Guide to Sensing Learning and Activity Through Practical Measures

Deb Morrison[1], Heena Lakhani[1], Kerri Wingert[2], Vanessa Wolbrink[3], Neelo Soltanzadeh[3], Robbin Riedy[2], Kristen Bergsman[4], Philip Bell[1] & William Penuel[2]
[1]University of Washington, [2]University of Colorado at Boulder, [3]WestEd, [4]Fred Hutch

Educational improvement requires timely strategies for sensing small shifts in learning or activity during improvement to identify opportunities in practice that can be leveraged or resourced further (Bryk, Gomez, Grunow, & LeMahieu, 2015). One way those involved in improvement science engage in such sensing is through the use of practical measures, a way of measuring specific steps and providing timely feedback to inform improvement work while it is in progress (Yeager, Bryk, Muhich, Hausman, & Morales, 2013).

What are practical measures?

Practical measures provide timely feedback that can inform instruction on *how* learning is occurring in a given context. Practical measures differ from other more traditional forms of educational measurement in that they:

- measure activity, engagement in activity, or identity with activity
- do not measure progress towards conceptual or practice based learning goals
- are not aimed at accountability or theory building
- use language embedded in the daily practices of classrooms
- are easily accessed and completed in everyday contexts
- produce data that is usable by practitioners
- can be used to understand progress, complexities, and hindrances in implementation

Practical measures fall into three main categories; measures of activity, of participants experiences with engagement in activity, and of participants identity development related to the type of activity. Measures of activity are designed to sense shifts in what a participant (learner or educator) thinks is occuring during the activity in terms of learning goals. An example of a practical measure of learner activity is a likert item such as “this week in class, I used models to understand science ideas”. The learner would then rate their agreement from strongly agree to strongly disagree. Such a measure is used to identify how much modeling a learner thinks is actually occuring during a defined window of science learning activity. Examples of a practical measure of educator activity are likert items such as “I used formative assessment to understand student thinking”, “My students have time to share their ideas with each other” or “Students have opportunities to argue using evidence”. These measures would sense learning goals such as formative assessment practices, classroom talk participation structures, or the teaching of science and engineering practices.

Measures of participants' experiences with engagement in activity include such things as interest in curricula or activity, and perceptions of relevance or connections activities have to everyday lives. Learner measures of engagement could be likert items such as “The work we did in class this week connected to my life”, “Going outside this week supported my learning”, “I found today’s activities interesting”. In terms of educators, such a measure could be an open ended question asking “What are you struggling with around formative assessment?” or a list of items that they rank order to indicate most challenging science and engineering practices to

engage students in. Items such as the likert measure “all students can learn chemistry” are also measures of engagement as help to identify educational biases and raise issues of equity that need support for teacher educators.

Measures of participants’ identity development related to the type of activity include such things as emerging science identities, feelings of belonging, and confidence with particular practices. These type of learner measures could include likert items such as “My ideas were heard and valued by my peers”, “I felt respected in class today”, or “We engaged in work similar to that of scientists”. These measures would sense constructs such as classroom talk participation structures, interpersonal interactions, or emerging science identities. Educator measures like this could include “I feel confident in engaging my students in modeling activity”, “I understand my grade level climate science content”, or “Students need to understand climate science to better participate in democratic decision making”.

Using Practical Measures to Sense Learning Experiences

To engage in practical measures work it is necessary to consider the nature of the work and the steps within the process. Practical measures work is best developed in collaboration with all involved parties (eg. teacher educators, teachers, students, parents) in order to ensure the questions asked are addressing relevant constructs of interest (clusters of questions) and that language is appropriate for the audience of the measures. In addition, practical measure development should be an iterative process that arises from a local need of practice.

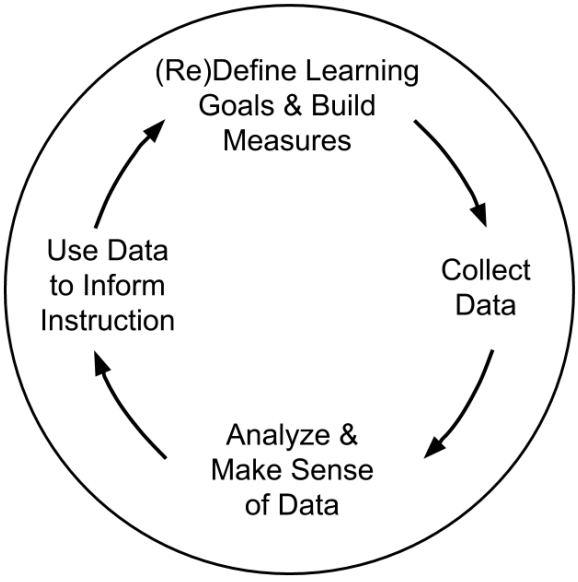

Figure 1: Cycle of Practical Measures

The cycle of practical measures (Figure 1) has been organized into four steps which is similar to that of the Formative Assessment Design Cycle (Ruiz-Primo & Furtak, 2007). These steps include:

- Step 1: Define Learning Goals and Build Measures
- Step 2: Collect Data
- Step 3: Analyze and Make Sense of Data

- Step 4: Use Data to Inform Instruction

Step 1: Define Learning Goals and Build Measures

The first step of the cycle of practical measures is to clearly identify what you want to know to inform instruction over a specific timeframe. As such, it is important to clearly define a goal that is to be measured which will guide the design of specific measures. For example, a group of teachers might decide to explore how participation structures are being experienced differentially by students in their classrooms. To sense this, teachers could regularly ask about how students' voices and ideas are being heard and taken up by peers in classroom conversations as they implement different interventions around classroom talk structures. Keeping a goal small and iterating on it over time provides the power around measures in practice for teachers to argue with evidence that a particular intervention is better in their specific context for achieving a clearly defined outcome.

Practical measures should be designed to ensure that they are timely and responsive to the learning goals defined. They should be short so as not to interfere with instruction, but given at regular intervals to make it more possible to sense smaller shifts in learning interactions or practices. While likert items are helpful in that they provide quick and easily interpretable data, it is also important to include some open ended questions for learners to provide more detail about their experiences. If an educator collects demographic data about learners (eg. gender, race, language) be sure to clearly communicate that you are using this information to interrogate your own potential biases and not essentialize the experiences of learners. To avoid having to collect demographic data, educators may want to collect non-anonymous data so that they can address the concerns of learners more specifically, but this can only be done if there is trusting and respectful relationships between learners and educators, otherwise, students will not clearly share about their experiences in such situations.

Designing Practical Measures

When designing measures consider:

- What questions will help you track how activity is unfolding in your context?
- Do the questions help you understand what activities are going on and how they are being experienced?
- Are the questions aligned with your goals and context? Have a couple of teachers review them and provide you with feedback.
- Do they make sense for your teachers/students? Is the language appropriate?

Expect to spend time choosing questions that will capture the range of participant experience. Questions have to be worded in a way that doesn't indicate a "right" or "wrong" answer. Some guidelines that we have found helpful are below.

- Measuring "confidence" can be a useful proxy for competence or identity. Asking people to report their competence (e.g., "I can design 3D assessments") is asking for over-reporting of positive outcomes. Instead of asking about their competence, "I can," try framing it in terms of confidence, such as "I am confident designing 3D assessments."
- Measuring an "amount" of practice is a good way to show growth over time. For example, see the following item:

  In the last week I asked students about their interests.

Often, Sometimes, Rarely, Not at All

- Don't indicate a "right" answer. Teachers in a hurry will respond that they do things that sound positive, even if they're not actually doing them. To avoid over-reporting, it is critical to word questions in as neutral a way as possible. This will also help you capture the range of things that teachers are doing. This is harder than it sounds, but one strategy is to avoid overly positive language. For example, take the comparison below:

Too positive	Just right
My assessments this week were high quality.	My assessments this week were three-dimensional.
I collaborate with my peers.	My peers and I planned a lesson together this week.

- In the case of educators, use practical measures to remind participants where they can go for help (Frerichs, Fenton, & Wingert, 2018) and what they should be trying out. Including a question about their practice and a link to a resource you used in the learning (a slide, an image, a planning document) might be just the reminder a participant needs to try out something new.

There are a number of different domains of interest for sensing learning and activity; however, each educational improvement endeavor has its own specific goals so the constructs of interest must be developed collaboratively within any effort. Some resources to consider constructs of interest include:

- *A Framework for K-12 Science Education* (NRC, 2012); including enterprise of science in appendices
- NGSS standards (http://www.nextgenscience.org/)
- STEM Teaching Tool resources on thinking about student experiences and engagement (http://stemteachingtools.org/)
- CSSS Science Professional Learning Standards (http://www.csss-science.org/downloads/SPLS.pdf)

Example Measures

The following example measures help to provide vision of what measures could look like for different audiences, across various constructs, and using multiple forms to address activity, engagement, and identity goals (Table 1).

Table 1. Example practical measures.

Audience, Construct, Type of Measure	**Example**
Teachers, Climate Science Learning, Measure of Identity with Activity - Likert	(Strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, strongly disagree)

	• I think it is important for students to learn about climate change in science class
Teachers, Climate Science Learning & NGSS, Measure of Activity - Likert	(Strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, strongly disagree) • My students reason with evidence about climate change • My students use large data sets to understand climate change • My students model how climate change is impacting their local area
Teachers, Climate Science Learning & NGSS, Measure of Engagement in Activity - Likert	(Strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, strongly disagree) • I engage my students in learning about climate change impacts in our local area • I engage my students in learning about climate change solutions • My students learn how different people are affected differently by climate change
Teachers, Climate Science Learning, Measure of Activity - Open Ended	What else would you like to learn about teaching climate change? What else would you like to learn about climate change?
Students, Climate Science Learning, Measure of Identity with Activity - Likert	(Strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, strongly disagree) • I think it is important to learn about climate change in science class
Students, Climate Science Learning & NGSS, Measure of Activity - Likert	(Strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, strongly disagree) • My teacher helps me reason with evidence about climate change • I used large data sets to understand climate change in my learning • I modeled how climate change is impacting my local area
Students, Climate Science Learning & NGSS, Measure of Engagement in Activity - Likert	(Strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, strongly disagree) • I like learning about climate change impacts in our local area • I appreciate learning about climate change solutions • In class we learn how different people are affected differently by climate change
Students, Climate Science Learning, Measure of Activity - Open Ended	What part of our learning experiences has most impacted your understanding of climate change?

Step 2: Collecting Data

The second step in the cycle of practical measures is focused on the question of how you get your data. Practical measures can take the form of electronic or paper surveys, or quick check-ins done on index cards as exit tickets. When designing measures consider:

- How will you collect the data from your target audience?
- Is the platform used to collect data accessible to your target audience?
- Does the platform provide you with tools to analyze the data?
- How are you ensuring the privacy of your data?

Electronic surveys

Electronic surveys are a common way to administer practical measures. Several considerations for the use of such surveys are cost and available features. Many survey platforms are free for a limited number of responses. For free or in a "trial version," some will gather a certain amount of data, use a limited number of question types, e.g., multiple choice or fill-in-the-blank, or a limited number of questions. Others are very expensive, and often require a license per user that is often not feasible for a department to pay for by itself. However, it may be worth it; some survey engines include a vast range of options that save time and result in much higher response rates, such as scheduling automatic email reminders for respondents who have not taken the survey yet and the ability to import spreadsheet data for drop-down items. If it is important that you get high response rates, it may be worth it to have automated reminder messages. Some platforms even include bot reporting algorithms and complex display and branch logic if you want to develop a complicated survey. It is a good idea to fully plan out what features are needed when implementing practical measures, and to reach out to the companies for pricing discounts.

As an open-source option, Google Forms has a minimalist interface that is free and highly shareable. A form and its results can be easily shared within a team of leaders. Importing a long list of questions can be quite time-consuming, and the question types are limited, but overall it offers cost-effectiveness and easy sharing.

Getting people to take your survey

One of the biggest issues in practical measurement is getting enough responses to inform instruction. A few different issues arise here.

Is a one-time survey enough? Usually not. If you really want to use practical measures to understand practice, you will want to see change over time. This introduces some complications to the method of surveying. For example, if you sample 50 teachers in August and get 23 responses, and again survey the same 50 teachers in May, you might get 23 totally different respondents. While both of these sets of responses can inform instruction, they are not representative necessarily of all participants so efforts should be made to embed time in the learning context to have respondents answer surveys or understand how to improve practical measures response rates.

How do you make sure you are sampling all participants? We recommend two approaches to try to ensure a sample that is complete, in which all participants respond to each survey over the entire time. Strategy 1: ask participants to enter their names into the survey. Strategy 2: Use email distribution from within your survey software, and turn on "automatic email collection," sometimes called "metadata." This method ensures the survey software collects the email of the respondents as they take the survey. Both strategies work well, but it is important to be transparent with participants about the reasons practical measurement are being used. Telling participants up front about the purpose of the survey can engender more trust among participants, and the items may perform better if teachers understand that they are not being evaluated by the surveys. Instead, surveys should have front matter that says they are designed to measure the success of the investment in their learning; is the work paying off? In what ways?

Privacy agreements: Am I allowed to collect this data? The short answer is that it depends on your arrangements with individual schools and districts. Districts all have their own guidelines and codes of conduct regarding collecting, storing, and sharing data. You should familiarize yourself with the policies your district recommends in regard to these questions. Your district has research specialists who can help if you are not familiar with these procedures or terms. Things to ask your district data team:

- Can I ask teachers to share their names with me on a repeated survey?
- What can I do with survey data I collect from teachers? Can I:
    - share it with external partners
    - share it with internal district teams
    - share it back to teachers in aggregate
    - store it on the cloud
    - use it to apply for grants
    - use it to write papers in journals.
    - under what circumstances must I get institutional review board approval for doing this survey?

Step 3: Analyzing and Making Sense of the Data -

The third step of the cycle of practical measures addresses the question of how to support sense making around the data produced. Consider making multiple representations of data and having participants comment on what works best for their sense making. When looking at data or visual representations of data, here are some reflection questions to consider:

- What does this tell you about [this aspect] of classroom experience?
- When looking at open ended questions, what stands out to you most? What gives you a visceral reaction, and why?
- What does this tell you about student engagement and motivation?

When leading teachers in professional learning around data, consider:

- **Presenting an anonymous set of data.** This might encourage teachers to pick out things that aren't stellar in student responses and reflect without feeling defensive.
- **Produce student profiles ("Student X")**: these student profiles could be an amalgamation of anonymous student data at certain snapshots in time (one at the beginning, middle, and end). This could be particularly useful when student surveys are anonymous and you can't track student data over time.
- **Visual representations**. Could show graphs of generalized data for groups. Allow a five minute period where nobody talks and educators just write their thoughts.

The design of the questions and the analysis of the data go hand-in-hand. There should be sets of complementary questions that are multiple choice/multiple select in nature as well as open-ended responses, so that one can understand the nature of the activity and how it is experienced, as described earlier. The multiple choice/select questions tend to give one a general sense of what is happening, while the open-ended questions will give one a deeper understanding of how implementation is unfolding. Analyses of these data can take multiple forms. Below, are some general guidelines for ways analyze that will help one analyze their data.

- Multiple choice/multiple select. For these types of questions, it will be easy to aggregate the number of responses to these questions. They will give one a quick overview of the trends that are occurring in activity. One can analyze the data in the following ways:
    - Descriptive statistics. This includes averages, frequencies, and the range of values for your data. This will give you an overall sense of how often an event is occurring, how many respondents are answering in specific ways, and the distribution of responses over a period of time.
    - Analyzing by other dimensions. The participants may be interested in breaking the data down into other dimensions such as pre-selected time intervals (e.g., weeks, months), demographics (e.g., students of color), etc. This data can help one understand how activity is changing over time and for whom it is changing.
- Open-ended responses. For these types of questions, there will be a range of responses to questions. This can be daunting at first, but there are some ways that can help one code the responses that are meaningful and productive. Coding is a way to sort and organize the data, which will then allow one to analyze the data.
    - Sorting & Organizing Data - Do a generous read through the data so that you can get an understanding of the lay-of-the land. These responses will be rich in detail, so multiple themes will be emerging from the data. As you are reading through the responses, it is important to start making notes so that you can go back and make sense of the data. Some questions that you can use to guide your notes include: *What do I notice? What surprises you? Are specific groups of students experiencing learning in different ways?* These types of questions will allow you to identify patterns in the data and to learn about the unique experiences of individuals. It is critical to attend to equity, thus strategically sorting responses by key demographic groups (e.g., students of color) will help you to understand how different groups of individuals are experiencing activity.
    - Analyzing Data - The notes and themes that were generated will be the foundation for analyzing the data. The goals are to understand which themes are coming up and their frequency. During the analysis phase, you will apply the

themes to your data; it is very likely that you will see more than one theme in a particular response. As you are applying the themes to the data, note the responses that are illustrative, so that you can have an example of what this theme looks like in context. Once you have applied the themes to the responses, a simple count of how often these themes come up will give you a sense of how representative or common that experience is. It is also important to attend to the themes that are less common, as they allow one to understand unique situations and will help you ensure that you are listening to all voices at the table.

Additional sensemaking tools to support this work will be added to this resource as developed.

Step 4: Use Data to Inform Instruction

The final step in the cycle of practical measures, prior to iterating back to step 1, addresses the question of how to use data to inform instruction. Based on the data, participants should consider:

- What are the opportunities for altering instructional design in your context? Operate in your own sphere of influence. If you cannot change the curriculum resource and that is identified as problematic, then consider how to enrich such resources with community involvement or digital resources.
- Who do you need to support instructional shifts? Is there a resource out there already to help inform instructional shifts (eg. www.stemteachingtools.org)? If not, who could you collaborate with to improve instruction as defined by your data?
- What professional learning should come next for the group? Do patterns in the data highlight specific areas for learning and ongoing shifts in teaching practice?

Implementation Challenges and Opportunities

What if data isn't valid because students don't take the surveys seriously?

Ways to address this concern:

- **Make surveys anonymous**. Anonymous surveys can encourage students to answer truthfully. This can be particularly important at first, but as teachers and students gain trust in the process, teachers may choose to ask for student names eventually.
- **Engage authentically with the data in class.** Students are often blunt and honest. But, if they don't see that a survey is contributing to real changes in the classroom, they won't put effort into it. Therefore, if a teacher over time shows students that they care about the results and are adjusting based on what they say, students will start to build what they're showing you and the responses will become more rich and complex. The bottom line is: If teachers don't use data to inform instruction, their students won't trust them and won't provide valuable answers. If they don't engage authentically with the data in class, it will backfire.
- **Make it clear to students you are learning together**. Deb Morrison has shared her experience when implementing these surveys, saying, "I had students come down and talk to me, saying, "I have a concern about your instruction". They become metacognitive about learning. They are thinking about the learning environment because I was, and I made it clear I was thinking about it. This is a life skill – we learn together with differences, we are coming together with different lived experiences, but we're all learners always. The more you show that vulnerability to your students and that you

want to develop the learning environment. I have been humbled by the desire to learn from students. These surveys help address implicit biases you have."

Teachers feel like these surveys are being forced on them and feel defensive.
Ways to address this concern:

- **Incorporate teacher feedback**. Provide an opportunity for teachers and school leaders to review the surveys and provide feedback on the questions.
- **Put teachers in control**. If teachers have sole access to their data and can use it in ways with which they're comfortable, they may be more likely to embrace the process. It is important to make clear to all involved that the data is not being used for monitoring teacher practice from a management / oversight perspective; the data is to support teacher improvement in their practice.
- **De-identify data**. For leaders using data for professional learning or to learn more about how to improve instruction school- or district-wide, make sure data is de-identified. This is particularly important for anybody in an evaluative role.

Teachers are worried about evaluation vs. assessment for improvement.
There may be security or other technological concerns with google docs.
Ways to address this concern:

- Be sure not to include student names if it's not a secure google format for educators.
- Consider engaging teachers in making sense of the data before bringing any of their educational leaders into the process.
- Consider providing technology professional learning for teachers and other leaders who will be accessing the data
- Consider who owns the data. Teachers should own the surveys and share them with whomever they feel comfortable.

Teachers feel like the survey takes up too much valuable instructional time. Survey fatigue.
To address this concern, make the following message clear:

- Practical measures can help teachers reach each and every student in the classroom and be more effective. The information you gather will have an impact on other learning goals because you're learning how to reach students and whether you're reaching each and every student.
- If your participants are being over surveyed, consider how to reorganize your survey schedules

References

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

Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of research in science teaching*, *44*(1), 57-84.

Jackson, K., Henrick, E., Cobb, P., Kochmanski, N., & Nieman, H. (2016). *Practical measures to improve the quality of small-group and whole-class discussion* [White Paper]. See https://www.education.uw.edu/slmi/practical-measures/ for information around this paper.

Kochmanski, N, Henrick, E & Cobb, A. (2015). On the Advancement of Content-Specific Practical Measures Assessing Aspects of Instruction Associated with Student Learning. National Center on Scaling Up Effective Schools Conference, Nashville, Tennessee, October 8, 2015.

Morrison, D, Stromholt, S., Lakhani, H., Bell, P., & (In Draft). Using practical measurement to sense district-level educational implementation and guide improvement in a research-practice partnership. *Journal of Learning Sciences.*

Yeager, D., Bryk, A., Muhich, J., Hausman, H., & Morales, L. (2013). *Practical measurement.* Palo Alto, CA: Carnegie Foundation for the Advancement of Teaching.