Toward a Common Research Framework for Professional Learning

Insights from RPPL Network Organizations on Data Use, Standardization, and Infrastructure

From the University of Pennsylvania's Penn Center for Learning Analytics and RPPL

Table of Contents:

Introduction and Goals

This report is based on the summaries and analysis of professional learning (PL) organizations within the RPPL network. Its purpose is to better understand the research infrastructure of the professional development tools used by the network. The interviews were conducted by members of the University of Pennsylvania team to comprehend better how organizations within the RPPL network use and organize data to work towards the following overarching goals: (i) Create standardized set of measurement tools across PL organizations; (ii) Broaden access among researchers and PL organizations to shared, standardized data; (iii) House shared data in a common platform that allows for secure researcher and RPPL member access and; (iv) Develop and validate innovative measures that will enable lower-cost, higher-quality research on PL.

The projects aim to build several interconnected strands of work through standardization measures of practice, collective use, secure data infrastructure, and the development of long-term innovative measures.

Scope of Review

Who we talked to: As part of the process, four (4) researchers were interviewed, and focus groups were conducted for thirteen (13) RPPL members and affiliates.

Interview Protocol: The same interview protocol was used for each interview and focus group. The first stage of the interviews was to understand better how each organization utilizes data within the network. Discussions opened with introductions to better understand everyone's positionality. Interviews and focus groups began with a PowerPoint presentation, attached as Appendix B. The PowerPoint slides were intended to aid the conversation and encourage emerging ideas. The interviewers' main objective was to assess the landscape of professional development measurement tools used within the RPPL network. Discussions focused on gathering information on how each organization analyzed and identified what was measured within their organization. Conversations were focused on what organizations perceived as the most effective factor, construct, or element in assessing the effectiveness of professional development.

Interviewers began by asking organizations how they utilized the Instructional Practice Guide (IPG) in their measurement procedures. In addition, interviewers prompted organizations as to how they were currently measuring or desired to measure the following areas: (i) expectations and beliefs; (ii) teacher measures; (iii) student measures; (iv) system measures, and (v) partnership outcomes. Specific organizational measures and outcomes are discussed in depth below.

The discussion then moved to the organization's current data infrastructure regarding questions around (i) the organization's current data collection and analysis; (ii) does the organization collect data from schools and districts; (iii) what platform does the organization use for surveys and; (iv) what support is currently in place for organizations data collection. Support within this context was intended for organizations to think about their personnel and, more broadly, how third parties, clients, providers, and school administration, assist with providing necessary data to conduct their work.

The interviews closed with questions about prosperous areas and challenges within the organizations' data management systems.

Where Are Organizations Today?

Each organization is a unique outlier taking its own approach to meet the organization's objectives and client/partner needs. There are several critical areas of measurement that organizations currently collect, measures regarding teachers and measures regarding students.

Measure of Teachers – Current:

Some organizations measure **teacher participation** in training by collecting attendance data and product usage data. Additionally, some organizations have created evaluation systems that rely on data gathered during professional learning sessions. The most commonality among the organizations is regarding the use of survey data. **Nearly all organizations administer a satisfaction or feedback survey.** Some organizations administer the survey immediately after PD training, while others administer multiple rounds to test satisfaction over time. A participant states that their survey, "Goes out to everyone who attends one of our workshops. It's designed to be taken immediately after the workshop concludes. It … has a pretty high level of participation from folks who enter in the workshop. That data comes back to us immediately" (Interview Participant). The participant discussed that this survey data is used to get immediate feedback and make adjustments. At the same time, they also administer another survey: "Designed deliberately not to be taken immediately after a workshop but to be taken by all folks who have been to a workshop at set points in the year" (Interview Participant). Organizations reported challenges in measuring progress over time through survey data and measuring the effectiveness of their services.

The second survey the interview participant mentions is intended to measure "Actual impact on instruction rather than just the impact that people hope for in the five minutes after they've left the workshop" (Interview Participant). Organizations spoke about the challenges of measuring impact over time, specifically around factors that might be difficult to quantify. For example, teachers are returning to their schools and finding barriers that impact them in deploying new knowledge learned. The barriers could be due to various factors, including school conditions, climate, or other obstacles leading to challenges in measuring satisfaction over time.

Many organizations report that collecting survey data immediately after a session is easiest for a better response rate. Some discussed that responses become more reliable when anonymous and that the value of the survey goes up as more respondents leave detailed descriptions in the comment sections. One interviewee remarked that they administer "anonymous student survey[s] because students said they would be more honest with us" (Interview Participant).

Organizations also measure teacher behavior. Some organizations have created their own scales for measuring perceptions, beliefs, and growth mindsets. The collection of these measures is done through surveys. Organizations have also created culturally responsive pedagogical tools to measure teachers and students about specific behaviors that teachers engage in. An organization recently reevaluated how they were thinking about teachers' expectations, beliefs, and mindsets and the role it played in the teachers' instructions. The organization conducted a literature review and met with several experts, which helped them develop a custom measure called the "teacher and student survey, sometimes also called the mindset and practices survey" (Interview Participant). Before the literature review, this organization used the "teaching tolerance tool, the common beliefs" (Interview Participant). The organization concluded from their literature review that "it's not so much teachers beliefs and mindsets or expectations that matter, or those things only matter to the extent that they change teachers practice, and so we sort of shifted how we're thinking about beliefs and mindsets to measuring more directly teachers behavior in the classroom and how those might connect to or be manifestations of or demonstrate their beliefs and their expectations" (Interview Participant).

The organizations discussed the value of these surveys. They saw the surveys providing insight into how students perceive their teachers' actions, which can often differ from teachers' self-perceptions. Measures of teachers' behaviors highlight the two-way nature of the teacher and student relationship and the importance of monitoring and managing it. An organization administers a survey to both teachers and students regarding specific behaviors teachers engage in the classroom. The organization reports that teachers have found it helpful because they do not often get information about the students' perspective about their experience as a student.

As discussed in this section, nearly every organization utilizes survey data, but the tools used to collect survey data vary considerably. Many software tools are utilized: Google Forms, Survey Monkey Alkimer (formerly Survey Gizmo), Questioner, or Qualtrics. It is challenging to unify at this level as organizations have created their own surveys on their own software systems. Standardization should not be to get everyone to use one tool but rather that all tools are compatible across each organization's software system.

Key Measurement and Measurement Approaches – Teachers:

Though the Instructional Practice Guide (IPG) is nominally used widely across organizations, nearly every organization using the IPG to measure teachers is customizing it in some form for their own needs. Organizations are changing the subjects of questions, adapting, deleting, or adding questions. These adjustments are mainly made because organizations are adopting the tool to meet the needs of their clients and what the client wants to see. Some organizations would consider developing a more flexible, upgraded version of the IPG that would include additional core sets of standards and testable indicators if it aligned with the organization's focus. Nearly every organization relies on some form of observational tools to collect outcome data. Many organizations create their own rubrics for observational collection.

While some organizations are adopting the IPG, many say it is not a helpful tool for them. One organization does not think the IPG would benefit their organization as "we're not taking a look at the actual instruction that's happening in classrooms, we're taking a look at the professional learning that prepares for the instruction" (Interview Participant). Another organization states, "We do not use the IPG as an observation tool because we have our own implementation reflection tool that we use when we have small visits" (Interview Participant). An organization has considered using the tool, "We have, at times, kind of opened consideration to using the IPGs in places where it makes sense based on what the focus of the PL is" (Interview Participant). While another participant sees value in the tool, but it does not meet their needs, "… I think the IPG is an exceptional element, is an exceptional instrument for measuring the rigor of mathematics and literacy instruction. It does not by design touch on many of the elements that we end up going into in ..." (Interview Participant).

Measure of Students – Current:

The most considerable variation among organizations was how and what measures for students are being used. Not all organizations can collect student-level data due to data sharing agreements or deem this level of data unnecessary to their specific client needs. Measures regarding learning, some organizations collect classroom information to match teachers to students. Some have created their own custom evaluation systems to measure student achievement through portfolio measures. Organizations that measure SRL and SEL do so mainly by portfolio measures through the submission of anonymized samples.

An organization has seen increases in student engagement using a gamified survey. The gamified survey measures student engagement with student learning, but they "realized, right, if this is a formative assessment of student engagement, it's gonna be given more than just like three times a year. So we realized ironically, for students to take it seriously, it had to be fun" (Interview Participant). Other organizations utilize surveys to measure student and teacher engagement, "We have also developed a student engagement and classroom experience survey that is aligned to our framework …." (Interview Participant). While another organization only measures student engagement when it is specific to client needs, "The student engagement and student experience is actually maybe the exception to that. … it would be very circumstantial, like we would only do that where it's the context of a particular engagement" (Interview Participant).

There is also variation in how organizations utilize student test scores. One organization states that they do not regularly collect this level of data, "maybe once or twice a year we collect information like that from districts is [test scores] … anonymous data as you previously said yes we have data sharing agreements and we generally ask them to de-identify the data …" (Interview Participant). Another organization has access to test score data but does not seem to have the staffing capacity to analyze it, "… we have the test scores around. Neither [redacted] nor I are necessarily experts at kind of the data management and organization procedures there but [redacted] has done some work with one of our large partners kind of drawing on some of those assessment data … We've brought it in again kind of where it makes sense and trying very hard with partners to make the caveat about where applying it doesn't make sense and try to avoid drawing too many conclusions" (Interview Participant).

Organizations that have access to collect administrative data and test scores seem to be more easily able to link teachers and students. An interview participant states that they collect "student IDs and other demographic factors to report on, but we also connected within the classroom. So that allows us to match teacher to student classroom information" (Interview Participant). The collection of student-level data is dependent on client needs. Some organizations are interested in summative data, while others utilize aggregate outcomes. An organization discusses that aggregate data might be easier to manage as they do not need to worry about the logistics and time required in executing a data sharing agreement. The participant states, "Where we need to access assessment data, we have on you know, at least advice of our council not needed to execute data sharing agreements for the aggregate data we collect, I think because it's not identifiable, so that's been again another opportunity to prioritize the work instead of prioritizing the logistics" (Interview Participant). Data collection decisions are based mainly on the client's need and the organization's capacity to collect, i.e., data sharing agreements and staffing. Student measurements are also collected via portfolio measures and surveys.

Measures of Systems:

Again, a running theme throughout the interviews is that what gets measured mainly depends on who the organization is partnering with and what the client needs. Some partners will provide data about curriculum usage to organizations, such as the number of users over a given academic year. Data that is provided to the organization by clients could aid organizations in measuring their effectiveness, but this is not consistently provided.

State and federal policies also impact teacher evaluations and how teachers might implement new training in the classroom, which could challenge an organization in measuring the impact/effectiveness of their tools. Some organizations provided examples that what is being learned in the PL is not always being evaluated in a teacher performance review. An Interview Participant states, "We're working with [redacted] who passed legislation to provide a substantial amount of performance-based compensation based on effectiveness. So, they had to then create a reliable measure of effectiveness in order to do that" (Interview Participant). The Interview Participant continues that effectiveness measurements are tied to teacher incentives and evaluation plans. Another participant remarks, "… I think every state has some policy that regulates and mandates that there be some way that the performance of teachers is, is assessed and recorded, right? So there's, this is a legal thing, if you will" (Interview Participant).

This creates additional hurdles for organizations as there might need to be an incentive for participants of a PL to practice what is learned. This also leads to not knowing the potential effectiveness of the PL if it's not continued to be practiced, evaluated, and implemented. School climate can also impact teachers implementing training from PL organizations.

Determining What to Measure:

There are three key ways that organizations determine what to measure:

I. **Partnership-based:** Determined by what the partner/client wants to be answered. Partners will have specific questions they want to be answered.
II. **Is the data available** to answer client/partner questions? Organizations discuss challenges around data sharing agreements. Data sharing agreements determine greatly what types of data organizations have access to, particularly at the student level. Data sharing agreements will determine if organizations can utilize classroom, school, or district-level data.
III. **Staffing capacity** also determines what an organization can analyze. Even if there is more data to analyze, human capital/funding is not always there to evaluate additional measures. This suggests an opportunity – if we can make it easier and cheaper to analyze data, PL organizations may be able to consider a broader set of measures.

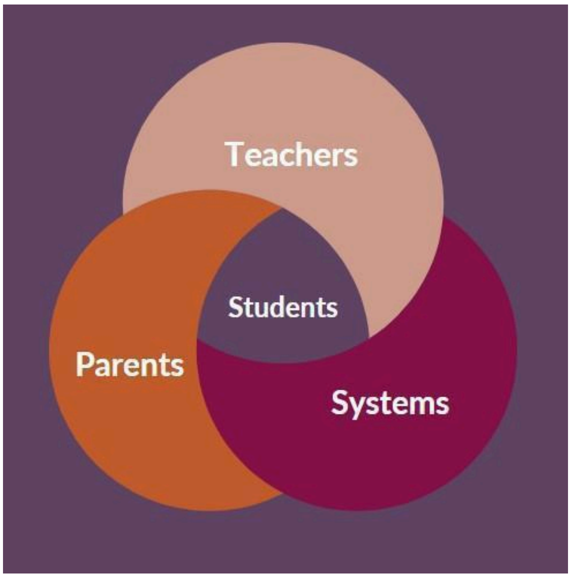
Key Findings

Universal Across Network:

Throughout the interview process, we noted several similarities across organizations. First, the desire to quantify the impact of their PL. Many organizations discussed a need and want to be able to make connections between the effectiveness of their professional learning organization and specific outcomes. However, this becomes a challenge when their partnerships determine much of the data they have access to. Second, every organization uses survey data in some form to measure satisfaction/feedback. As discussed above, organizations have trouble utilizing surveys to measure effectiveness over time. Third, organizations see standardization as challenging and unnecessary to their day-to-day operations. Fourth, staffing is a significant hurdle in the organizations' ability to collect and analyze current and new data. Five, with staffing already a challenge, organizations stated concerns around the ability to take on more. While these concerns were expressed, many discussed how they would benefit from having some processes automated for data analysis.

Key Findings:

While each organization is an outlier and has its own unique systems, they all face similar pain points. The image here helps represent what was heard among the organizations, the network's collective nature, and the ultimate goal of supporting students, teachers, communities, and parents. Each organization within the network has its own area of focus, with the ultimate goal of supporting teachers/staff, students, parents, and communities. The network is interconnected, policies and systems impact teachers, and teacher outcomes influence student outcomes. Each organization does not have the capacity to study all of these areas to gain a complete picture of what's going on. But within the RPPL network, there is an opportunity for organizations to utilize newly available data to answer more questions.



The image represents the interconnective nature of the RPPL network. The ending of one circle is the continuation of the next to support students at the center. Each area within the image, teacher, student, parents, and systems influences one another and are an extension of each other. Throughout the interview process, organizations wanted to answer more questions with data and time they did not have. For example, at the parent level, organizations discussed including student and parent perceptions in their work. At the teacher level, organizations spoke of a desire to better measure the effectiveness of their PL over time. At the system level, organizations wanted to examine further what system-level barriers are leading to teachers not implementing new PL practices and what supporting conditions could be implemented to aid teachers. At the student level, organizations addressed their desire to incorporate student voices and engagement better.

The above points are only a sample of ways organizations wish to enhance their work. Each of the issues an organization wishes to improve above is taking place by another organization to some degree within the RPPL network. The collaboration and pairing of organizations' shared research, ideas, and data could boost each organization's work and strengthen the entire network.

Near Future – What's Possible:

Standardization at any stage can help streamline workflow, provide an ability to answer more questions with additional data available, better determine the effectiveness of client services and collaborate within the network to serve client needs better.

More standardization and comparability of measures:

The organizations in the network use a wide range of multiple tools. There is no one best tool that should be focused on, and the tools used in common are frequently adapted. The organizations seem unwilling to stop modifying instruments (even when doing so invalidates evidence around the instruments' psychometric validity). However, by working with organizations to create an "unchanged subset" of items, we may be able to increase comparability across organizations. In addition, often, the same data is collected in different formats through different tools. It would be difficult to get organizations to switch to different survey collection tools, but data could be more comparable by creating data format converters. Increasing the compatibility of these tools across the network would provide the potential to share information and data more effectively and facilitate comparison of what is working in which situations.

Data requirements:

Many organizations spoke about two challenges around data that go in opposite directions: either an organization has too much data that they are not utilizing, primarily due to staffing capacity, or the organization does not have enough data to answer questions they would like to explore. (Sometimes, the same organization has both challenges!) The first challenge could be addressed by creating tools to automate analyses -- this will be considerably facilitated by increasing the standardization of instruments across organizations. Addressing the second challenge, we could meet again with organizations to give them recommendations about their data collection, discuss plans, and bring together sets of organizations with common interests/goals to plan future data collection together.

Standardization of Observational tools across the network:

Nearly every organization utilizes some type of observational tool within their practices to measure outcomes. Each organization uses multiple tools for data collection. Organizations have different procedures for running reports, analysis, and feedback depending on if the organization is partnered with districts, schools, or teachers. Organizations, too, are using different dashboards and data visualization tools.

Increasing the degree of standardization of observational tools across the network could help organizations learn from one another and save time and additional resources. There is no one observational instrument that will support all organizations, but there are clear commonalities, so subsets of organizations may be able to work together -- with scaffolding from RPPL and Penn -- to create common subsets within their observational methods.

Data Infrastructure:

We can identify two types of data infrastructure that could benefit organizations. First, once instruments are more standardized, we can create data converters (so that organizations can continue to use existing data collection systems) and automated reports (to address staffing limitations). Even without standardization of instruments, this is possible but will involve considerably more effort within the Penn team (since reports would then have to be customized for each organization).

Second, creating a shared database with de-identified data would support external researchers conducting analyses across organizations (or even using a single organization's data). Ideally, such a database would have (i) controlled access for researchers to help measure outcomes; (ii) Organizational access to enable organizations to learn and collaborate and; (iii) an easy and efficient design that could help organizations run simple reports and analysis. Challenges around privacy-protected analysis of student data can be addressed, for instance, by adopting a data enclave such as the MORF platform to allow controlled, limited access to data. Alternatively, it is now possible to largely automate the de-identification of text/transcript data based on Large Language Models, reducing human effort and cost considerably (though not entirely -- checking is still needed) and making it possible to share data more directly. Our interviews with researchers indicated that they strongly preferred downloading data directly to using contemporary data enclave technology due to the challenges and time costs of learning to use data enclaves; automated de-identification could enable this.

Slightly more distant future:

Within the slightly more distant future, the network can utilize richer additional forms of data from PL sessions, coaching sessions, and classrooms by using data from video, audio, transcripts, student work, and location badges. There have been advancements in machine learning and artificial intelligence that increase what can be obtained from these data sources, from methods for better video and audio processing (including audio-to-text processing) to technologies using large language models to speed and enhance qualitative coding and de-identify textual data.

More advanced detector technologies could also be utilized to capture constructs such as class and online participation, engagement, and uptake. It is also possible to automate the evaluation of student work through technologies that use pen-stroke analysis and image-to-text technologies to analyze portfolio works.

By developing, validating, and studying new measures, organizations can better serve their clients/partners. The network could pilot studies to test the impact and leverage the potential of the RPPL network and large-scale deployment.

Conclusion

Organizations generally appear excited about working within the RPPL network and collaborating on a standardization effort. The most significant challenge to overcome for organizations is the potential disruption of their day-to-day tasks that overhauling their data infrastructure, changing measures, or adopting new measures would cause. Therefore, finding solutions that make transitions maximally smooth (such as adopting converters rather than changing data collection tools) will increase the probability of success. In addition, continued conversation with organizations during development and deployment will be necessary to prevent unexpected or undesired disruptions.

Appendix A: Interview Protocol

RPPL Research Infrastructure Interviews

Research infrastructure

- Gates Funded Project
- Research Infrastructure for Professional Development
- Key Partners
- RPPL staff and RPPL's member organizations
- Annenberg Institute at Brown University
- University of Pennsylvania

Goals

1. Create standardized sets of measurement tools across PL organizations
2. Broaden access among researchers and PL organizations to shared, standardized data
3. House shared data in a common platform that allows for secure researcher and RPPL member access
4. Develop and validate innovative measures that will enable lower-cost, higher-quality research on PL

Next steps

1. Overview of measures used within the network
2. Survey: which measures you already use
3. Survey: standardization of measures
    a. Which of these measures would you like to be standardized (and available for quick deployment) and be able to compare across peer organizations?
4. What is missing?
5. Data infrastructure needs
6. Data sharing agreements

Inside the network

- Area: IPG (Instructional Practice Guide) & additional observation tools
    - ELA/Literacy: Grades K-2 Reading/Listening Comprehension
    - ELA/Literacy: Foundational Skills Observation Tool
    - ELA/Literacy: Grades 312
    - Math: Grades K8
    - Math: High School
    - Organization Specific Measure
- Area: Expectations & Beliefs
    - Teaching Tolerance Common Beliefs
    - Any other organization-specific measure?
- Area: Other teacher measures

    - PD Attendance Data
    - Satisfaction Survey
    - Product Usage
    - Interactions Data
    - Teacher Content Knowledge
    - Teacher Evaluation Data
- Area: Student measures
    - Summative Data
    - Student Level SEL (Social Emotional Learning)
    - Interim Assessments, Assignments or Tasks
    - Student Engagement & Experience
- Area: Systems measures
    - Partnership Outcomes
    - Partner/System Satisfaction
    - Supporting Conditions

Data infrastructure

- Org Data Collection & Analysis Needs
    - Does your organization have the capacity to collect or analyze new data? Are there specific areas where you need support?
- Data Collection from Districts
    - Do you regularly receive administrative data (student test scores, teacher-student linkages) from districts?
- Survey Platform
    - What systems/platforms are used for surveys?
    - How do you leverage LMS or other platforms?
- Survey Administration Support
    - What strategies do you use for data collection in general? (e.g., what's worked to get survey responses)
    - What strategies do you use to track participants across PL courses and/or years, student data, etc?

What else?

- What types of data do you collect? What is the most important?
- How satisfied are you with the measures you use?
- Are there other domains, data sources or measures you'd like to use?
- Innovative measures
    - Video/audio transcripts
    - Anonymization of transcripts
    - Detectors (uptake)
- “Right to be forgotten"
- Data sharing agreements
- Any other issues related to data collection/sharing?
- What are your strengths you'd be willing to share with the network?