International Society of the Learning Sciences

Supporting High School Science Teachers in Developing Pedagogical Content Knowledge for Data Literacy

Katherine Miller, Susan Yoon
kmiller@gse.upenn.edu, yoonsa@upenn.edu
University of Pennsylvania

Abstract: As the world becomes increasingly awash in data, there is a growing need for greater focus on data literacy at the K-12 education level. Since high school teachers are often trained in one specific subject, they need additional support to teach data literacy as an integrated STEM topic. In addition to providing subject matter knowledge of data literacy, this support must focus on pedagogical content knowledge (PCK) which connects the content to the pedagogy for teaching it. As there is a major dearth of research on PCK for data literacy this study seeks to apply existing strategies and design mechanisms that can support teachers in developing their PCK to the subject of data literacy to better understand how a new conception of PCK can be surfaced and understood.

Introduction

While the field of data science has expanded over the last few decades to address the explosion of data and the way it permeates so many aspects of life, K-12 education has not yet caught up, resulting in strong calls for greater focus on data literacy at all levels (e.g., Wolff et al., 2016). As the need for data literacy grows, the way we seek to teach students to interact with data also needs to change to promote a broader relationship with data that prepares students for working with the big, messy data sets that imbue research, industry, and society (Kjelvik & Schultheis, 2019; Lee & Wilkerson, 2018). Teachers need additional support to develop knowledge and tools for integrating that type of data into their classrooms, including both subject matter knowledge about data and knowledge about how to teach with it (Lee & Wilkerson, 2018). Pedagogical content knowledge (PCK) encompasses the professional knowledge that teachers hold about how to teach particular topics to particular groups of students (Shulman, 1987). While research on PCK for data literacy is mostly lacking, research on PCK for STEM integration has shown that in order to successfully integrate STEM learning in their classrooms, teachers need to understand how students learn and apply ideas in topic-integrated contexts and be able to ground the varied STEM concepts in the learning context and prior knowledge of their students (Vossen et al., 2020). This suggests that pedagogical strategies for data literacy implementation, are unique from teachers' existing PCK for the subject they teach. Therefore, any professional development (PD) that seeks to improve teachers' data literacy implementation should explicitly teach and model these strategies (Aydin-Gunbatar et al., 2020). However, the lack of research on PCK for data literacy makes it difficult to determine how to support teachers in developing it. As such, this project explores how to support teachers in developing their understanding of PCK for data literacy. Through a semester-long workshop with four in-service high school science teachers aimed specifically at examining their PCK for data literacy, this study asks, how did the components of the PD workshop support teachers' surfacing and development of PCK for data literacy?

Background

Data literacy is an emerging concept without a clear definition, and the overlap between data science, data literacy, computational literacy, and statistical literacy is still nebulous (Wolff et al., 2016). In an attempt to distinguish data literacy from some of the other fields, Kjelvik and Schultheis (2019) chose to define data literacy as existing in the overlap between quantitative reasoning and data science where data is grounded in an authentic context. This framework is useful in outlining some of the important components of data literacy, namely that it involves applying mathematical principles, working with computers and other technologies, and understanding the context of the data. The need to ground data literacy in real-world contexts and focus on *using* data rather than simply analyzing it is a common component across data literacy research (Rubin, 2020; Wolff, et al., 2016). However, authentic real-world data is often complex, messy, and unlike most of the data typically used in secondary school classrooms and teachers haven't always been taught how to engage with this data or the pedagogical strategy that is required to teach it (Kjelvik & Schultheis, 2019).

In order to engage teachers with pedagogical practices specific to teaching with and about data, we applied the framework of PCK to better support teachers' knowledge and learning. PCK is a distinct form of knowledge separate from subject matter knowledge (Gess-Newsome et al., 2019; Vossen, 2020). In the case of data literacy, the subject matter knowledge needed in the development of PCK draws from multiple disciplines

across science, statistics, and technology. For this project, the subject matter focus included a knowledge of the context of data, the role and purpose of data visualizations, and how to build inferences from data (Rubin, 2020). While most teachers have extensive PCK for subjects they received training in and have experience teaching, considering data literacy as explicit subject matter rather than a tool for teaching math or science is often new to teachers and they may need to develop additional PCK for data literacy in order to enact it successfully in the classroom. While PCK can develop through experience alone, studies have shown that PCK development can be guided and enhanced through PD that aligns with teachers' learning needs and goals (Aydin-Gunbatar et al., 2020; Gess-Newsome et al., 2019). However, there has been little to no empirical research published on PD specifically for developing knowledge for teaching data literacy. The theoretical research on potential strategies for teaching data literacy (e.g., Lee & Wilkerson, 2018) have been developed without engaging teachers with those strategies during PD. As such, this study relied on a method shown to develop PCK within PD interventions (e.g., Loughran et al., 2012) and applied those strategies to the development of knowledge for teaching with and about data.

Methods

This research engaged in an early-stage or exploratory study in which the goal was to examine how the knowledge teachers held about teaching for data literacy could be surfaced and developed.

Designing for PCK development

The intervention relied on a common method for engaging teachers in building and refining their PCK: Content Representations (CoRes, e.g., Loughran et al., 2012). CoRes were originally developed by Loughran and colleagues (2012) as a way to guide teachers to think about their practice in creating PCK around a particular big idea. A CoRe is a template which prompts a teacher or group of teachers to think about teaching a particular subject through big ideas with questions such as: *Why is it important for students to know this? What else do you know about this idea that you do not intend students to know yet? What is your knowledge of students' thinking which influences your teaching of this idea?* Loughran and colleagues (2012) found that CoRes could be used to develop PCK by enabling teachers to make their practice, and their thinking on their practice visible in a way that allowed them to then reflect on that knowledge. While teachers often struggle to construct big ideas for a subject and answer the questions in a CoRe individually, when building knowledge collaboratively they have more success (Aydin-Gunbatar et al., 2020). Using a CoRe as an artifact allows teachers to develop a shared language around the content knowledge and the PCK for teaching a particular topic (Loughran et al., 2012).

Context and participants

This study took place within the larger context of an intervention designed to support teachers in implementing a STEM-integrated unit on bioinformatics in their existing secondary school science classrooms (Yoon et al., 2022). An extension workshop was conducted with a subgroup of four teachers in which participants met virtually for about 20 hours over the course of five months while implementing the bioinformatics unit in their own classrooms. The workshop sessions relied on a number of different strategies, including explicit discussion of PCK, implementation discussions, and review of data literacy content components. During the workshop, four CoRes (and framing for a fifth) were discussed. The four participants all taught in the same large urban school district in the northeastern U.S. Three of the teachers (Hallie, Will, Manisha) had over ten years of teaching experience, while the fourth (Mary) was only in her second year in a formal classroom but possessed over 15 years of experience teaching informal science education. Two of the teachers (Hallie & Manisha) were teaching biology, two were teaching environmental science (Manisha & Will), and one (Anna) was teaching agriculture.

Data sources and analysis

The data for this paper came from transcripts of the workshop sessions and post intervention interviews. The nine sessions of the extension workshop series were conducted over Zoom and each session was recorded using Zoom's built in recording feature. A total of 15 hours, 20 minutes of video was captured and transcribed. Additionally, at the end of the workshop series, a semi-structured interview was conducted with each participant. The interviews asked teachers about teaching data literacy (e.g., *In your view, is data literacy important to teach in a science class, why or why not?*) as well as asking them to reflect on the components of the workshop series and how they affected their knowledge and learning (e.g., *What parts of the workshop series did you find most supportive of your growth as a teacher?*) The interviews ranged in length from 40 minutes to 58 minutes with an average length of 49 minutes and a total time of 3 hours and 19 minutes.

The transcripts of the workshops and interviews were organized in Delve Tool, a qualitative data analysis software and a constant comparative analysis (Glaser, 2008) was conducted to generate themes related to participants' perceptions of the intervention and its design components.

Findings

While the larger project that this paper is a part of focused on understanding what the components of PCK for data literacy are (Miller, 2022), this paper focuses on the process by which they were developed. The most valuable and tangible shift in teachers' PCK was simply becoming more aware of the knowledge they already held and being able to articulate it in a way that made it easier to apply to their implementation. To that end, the primary theme from the design of the workshop is that the CoRes, and especially the development of big ideas provided a successful framework for surfacing and organizing PCK for data literacy.

The CoRes served as a framework that allowed space for reflection. All four participants thought the CoRes were a useful component of the workshop series. They primarily reflected in their interviews about how the CoRe Template provided them a specific framework that allowed for more nuanced and focused reflection. Hallie and Will both reflected during their interviews about the usefulness of the conversations engaged around the CoRes. Hallie spoke to how being forced to articulate her own thinking to fit into the framework of the CoRe led to deeper thinking for her. She said in her interview, "So, I'll say just the brainstorming around what to do and how to do a thing because it's a lot easier for me to think deeply about a thing when I'm trying to express it to someone else than it is when I'm just trying to turn it around in my own head." Similarly, Will focused in his interview on how listening to other people articulate their thoughts in response to the CoRe framework was useful for him, reflecting during his interview, "I was more looking at what the other teachers were talking about. Seeing what other teachers were focusing on and what they knew opened my eyes: oh, I hadn't thought of it that way. So that was very beneficial to change my thinking in how I was going to present this." In this reflection, Will was talking about how the discussions around the CoRes and the ideas that the other teachers came up with and the knowledge they surfaced helped him to grow his own knowledge. They opened his eyes to new concepts and new strategies that he hoped to employ in his own classroom going forward.

While the CoRe template was useful for guiding discussions and reflections on teaching with and about data to be more focused and nuanced, perhaps the most useful part of completing the CoRes was identifying the big ideas for data literacy. Teachers were not used to thinking about data literacy as subject matter that needed to be explicitly taught rather than as a tool to be used to teach other concepts in science. There was extended time in the workshop sessions for discussing and identifying big ideas before building each CoRe, yet it was often still not enough. Simply coming up with big ideas that were appropriately specific but also appropriately generalized was a challenge for the teachers but a challenge that led to a lot of growth and recognition and articulation and refinement of knowledge that the teachers already held. One example of this is in attempting to develop big ideas for the concept of data in context, teachers landed on a big idea of interest: *interrogating data is useful and necessary*. The identification and refinement of this big idea led to the surfacing of PCK in that teachers then connected this to what they knew about how students' perceive data and strategies for teaching it. Specifically, they identified that students are not used to questioning data, and that therefore students would need to be explicitly taught how to interrogate data through tools such as employing text analysis to data representations.

The teachers came up with many more big ideas than they ended up focusing on in the CoRes they built, however having the CoRe as a framework allowed them to decide which of the big ideas were the most important or salient. So, while the brainstorming process for the big ideas was a crucial piece of surfacing knowledge about teaching data literacy, the CoRes allowed the participants to focus on specific big ideas and connect those to knowledge about students' understanding and specific teaching strategies, important components of PCK.

Discussion and implications

While the exploratory nature of this study did not allow for an exploration of the extent to which teachers developed additional PCK as a result of this intervention, it was clear that they became more aware of the concept of PCK and more inclined to reflect on their practice within the PCK framework. Engaging in development of big ideas for data literacy and subsequent completion of CoRes for those big ideas supported teachers in developing a better sense of the strategies they were using, language to describe them, and tools to reflect on them more productively. Unpacking the components of data literacy in order to determine the big ideas led teachers to, if not increase their subject matter knowledge of data literacy, certainly approach it with a different perspective. This supports previous research that shows CoRe development can support teachers' knowledge growth (e.g., Aydin-Gunbatar et al., 2020) However, most previous studies were conducted within disciplines that had well established big ideas and frameworks for PCK. The field of data literacy does not have that. So, while one implication of this research is to suggest that these tools continued to be used to help teachers develop knowledge and confidence for teaching data literacy, another implication is that the field needs to engage further with teacher knowledge for teaching data literacy. As Hallie reflected in her interview it would be nice if teachers had an established set of knowledge about students' preconceptions for data and teaching strategies to engage those preconceptions, but the field is not there yet, partly because there has been so little research conducted on teacher knowledge.

The use of CoRes in PD can guide teachers towards best practices for using authentic complex data in their classroom, but first a set of best practices needs to be established so that teachers can be supported in using those best practices to teach data literacy. Previous research has found that teachers often lack the content knowledge to engage in integrated STEM activities such as the use of complex authentic data (e.g., Aydin-Gunbatar et al., 2020) and this study supports that research. However, the larger struggle for the teachers in this study was not the subject matter knowledge of data literacy but the strategies for how to teach it. While the teachers already had knowledge of strategies for teaching science and engaging with data in the context of a science class and were able to reframe that knowledge in a way that applied to data literacy, most of them still felt at the end of the intervention that they were in need of additional support to grow their strategies for teaching with and about data.

References

Aydin-Gunbatar, S., Ekiz-Kiran, B., & Oztay, E. S. (2020). Pre-service chemistry teachers' pedagogical content knowledge for integrated STEM development with LESMeR model. *Chemistry Education Research and Practice*, *21*(4), 1063–1082. DOI:10.1039/d0rp00074d

Gess-Newsome, J., Taylor, J. A., Carlson, J., Gardner, A. L., Wilson, C. D., & Stuhlsatz, M. A. M. (2019). Teacher pedagogical content knowledge, practice, and student achievement. *International Journal of Science Education*, *41*(7), 944–963. DOI:10.1080/09500693.2016.1265158

Glaser, B. (2008). The constant comparative method of qualitative analysis. *Grounded Theory Review, 3*(7).

Kjelvik, M. K., & Schultheis, E. H. (2019). Getting messy with authentic data: Exploring the potential of using data from scientific research to support student data literacy. *CBE Life Sciences Education*, *18*(2), 1–8.

Lee, V. R., & Wilkerson, M. (2018). Data use by middle and secondary students in the digital age: A status report and future prospects. Commissioned Paper for the National Academies of Sciences, Engineering, and Medicine. Washington, D.C.

Loughran, J. J., Berry, A., & Mulhall, P. (2012). *Understanding and developing science teachers' pedagogical content knowledge* (2nd ed.). Sense Publishers.

Miller, K. M. (2022). *Developing Pedagogical Content Knowledge for STEM Integration Through Data Literacy: A Case Study of High School Science Teachers* [Doctoral Dissertation, University of Pennsylvania]. Proquest.

Rubin, A. (2020). Learning to reason with data: How did we get here and what do we know? *Journal of the Learning Sciences, 29*(1), 154–164.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*, 1–22.

Vossen, T. E., Henze, I., De Vries, M. J., & Van Driel, J. H. (2020). Finding the connection between research and design: the knowledge development of STEM teachers in a professional learning community. *International Journal of Technology and Design Education*, *30*(2), 295–320. DOI:10.1007/s10798-019-09507-7

Wolff, A., Gooch, D., Montaner, J. J. C., Rashid, U., & Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics, 12*(3), 9–26.

Yoon, S. A., Shim, J., Miller, K., Cottone, A., Noushad, N., Yoo, J., Gonzalez, M. V., Urbanowicz, R. U., & Himes, B. E. (2022). Professional development for STEM integration: Analyzing bioinformatics teaching by examining teachers' qualities of adaptive expertise. In A. C. Superfine, S. R. Goldman, & M. M. Ko (Eds.), *Teacher learning in changing contexts: Perspectives from the learning sciences*. Routledge.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. #1812738