Catalyzing Teachers' Evidence-Based Responses to Students' Problem-Based Learning in STEM

Nicole M. Hutchins, Vanderbilt University, nicole.m.hutchins@vanderbilt.edu
Gautam Biswas, Vanderbilt University, gautam.biswas@vanderbilt.edu

Abstract: This paper examines the processes middle school STEM teachers employ to interpret student learning and problem-solving activities during a problem-based learning unit and then design evidence-based lesson-plan customizations. Utilizing inductive and constant-comparative analysis of teachers' think-aloud data, we identify *catalyzing links* that support the transition from interpretation to enactment. We provide a contrasting case between an experienced and a novice teacher, and discuss how the results can inform STEM PBL professional development and teacher-support technology development.

Introduction

Prior research has demonstrated the importance of teacher engagement as students develop ideas and use of strategies to support their STEM learning in problem-based learning (PBL) environments that emphasize student-centered learning. This can pose challenges to classroom teachers, who have to interpret and respond to student progress in ways that target their learning and problem-solving needs while also adhering to the intent of the PBL learning design (e.g., not always addressing a specific knowledge gap through direct instruction) (Chen et al., 2021). Technology-enhanced approaches can assist in handling these challenges by logging student activities in a computer environment, and then using learning analytics to help teachers visualize student learning and problem-solving behaviors over time. This information can then be combined with orchestration technologies, such as teacher dashboards (Wiley et al., 2020) to develop subsequent instructional plans. However, more research is needed to target the complex task of translating our findings as researchers into a language that classroom teachers can interpret and convert to actionable information (Wiley et al., 2020).

Understanding how teachers use dashboards to develop and apply their evidence-based teaching practices in technology-enhanced curricula is critical for improving teacher support and preparation (Campos et al., 2021; Farrell and Marsh, 2016). This serves as the context for our research. In the study presented in this paper, a researcher (i.e., author 1 of this paper) interacted with eight teachers to understand how they used the Responsive Instruction for STEM Education (RISE) dashboard to assess and respond to students' learning and use of strategies as students worked on building computational models of water runoff in the C2STEM block-based coding environment. Teachers completed a series of five *Planning Period Simulations* that leveraged the co-designed dashboard to track students' learning progress and behaviors. We implemented think-aloud protocols using semi-structured interview questions enabling the teachers to verbalize their thought and evaluation processes (Charters, 2003). Our analyses targeted the following research question: *What processes do teachers use when making decisions by analyzing current student data for customizing subsequent lesson plans?* We conducted inductive analysis and constant-comparative analysis (Charmaz, 2006) to provide initial, exploratory patterns in the reasoning processes teachers used when transitioning from interpretations of student results to selecting evidence-based pedagogical responses, i.e., lesson plan customizations grounded in their domain-specific and pedagogical content knowledge and the student results visualized on RISE. A contrasting case between an experienced and a novice teacher explores and highlights these processes.

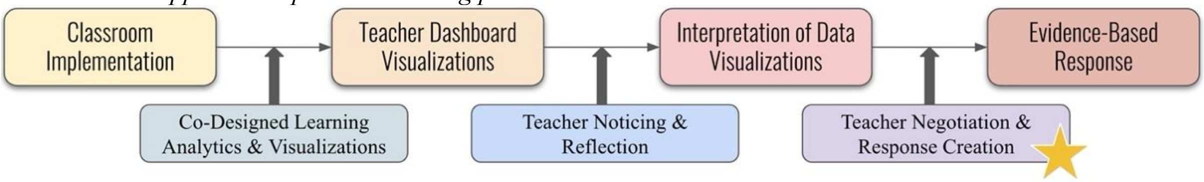
Background

A careful analysis of prior research models representing dashboard-supported responsive teaching identifies key research opportunities identified in Figure 1 (adapted from Campos et al., 2021). The first research area (Figure 1, in green) targets the impact of an educational event in the classroom by visualizing student data on a dashboard. Research in the area has targeted co-design methods for using teacher insights into developing and presenting such visualizations (Wiley et al., 2020), improving transparency in algorithm development (Holstein et al., 2019), and supporting teacher agency in the representations shown on the dashboards (Ahn et al., 2021). In our research, we have implemented a multi-step co-design process for the creation of RISE (Hutchins & Biswas, 2023).

Another research opportunity involves a deeper understanding of how teachers make sense of the information provided on dashboards and how we can support their interpretation processes (Figure 1, in blue). Campos et al. (2021) conducted a study with teachers and educational coaches to examine sensemaking processes and developed a typology of responses to different data visualizations. Others found sensemaking heuristics, which include comparing, monitoring, and exploring by teachers as they leveraged tools to support technology-supported collaborative learning (Voyiatzaki and Avouris, 2014). Molenaar et al. (2019) investigated how teachers make data visualizations actionable. Specific to our work, Chen et al. explored teacher dashboard use to support problem-based collaborative learning at the college level (Chen et al., 2021). Simultaneously, research on supporting teachers' interpretation process for improved decision making has recently received attention. For instance, researchers have evaluated the impact of different interpretive aids on teachers' sensemaking to support collaborative learning (van Leeuwen et al., 2019). Although we do not focus specifically on this research opportunity, we believe a deeper understanding of how teachers transition from learning analytics-based data visualizations to pedagogical decisions can aid in the future development of such tools and resources.

Finally, there is a need to understand how resulting teacher interpretations of dashboard visualization facilitate evidence-based pedagogical actions (Campos et al., 2021). To our knowledge, limited research exists that explores what teachers notice from classroom data and then generate evidence-based responses (identified with a star in Figure 1) to support students' learning and problem-solving strategies for K-12 science PBL curricula. As such, this research provides novel, exploratory findings on example pedagogical responses resulting from the noticing, interpretation, and reasoning about student data during a problem-based, middle school science curriculum.

Figure 1
A dashboard-supported responsive teaching process model



Instructional context

Science Projects Integrating Computing and Engineering (SPICE; spiceprojects.org) is a three- week, NGSS-aligned curriculum unit that challenges students to redesign their multi-functional schoolyard using different surface materials to minimize the amount of water runoff after a storm, while adhering to a set of design constraints (McElhaney et al., 2020). The problem-based learning curriculum consists of five core units. These units include: physical experiments, conceptual modeling, paper-based computational thinking tasks, computational modeling of the water runoff phenomenon, and engineering design, in which students use their computational models to design their schoolyard.

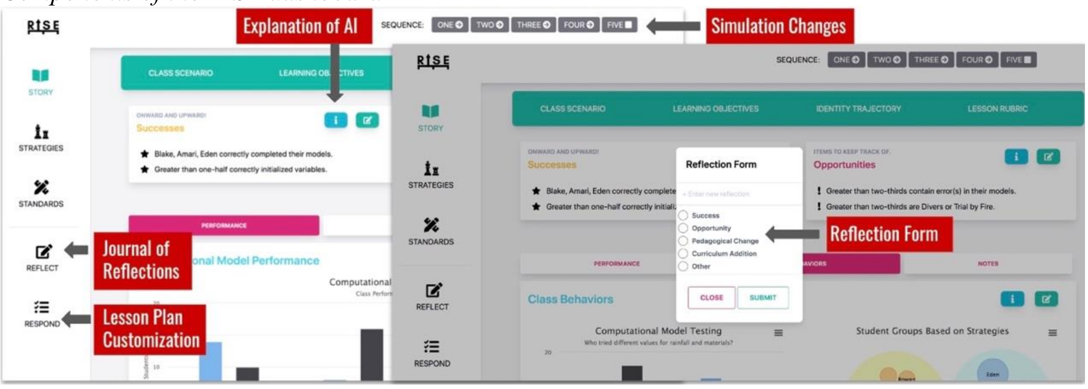
The dashboard simulations in this paper focused on supporting lesson plan customizations during the computational modeling task in SPICE, which were implemented in the C2STEM block-based coding environment (c2stem.org). In this task, students needed to initialize total rainfall and the absorption limit of the material they would test with their computational model. Students were required to construct their computational model with three conditional components corresponding to situations where the total runoff would be greater than, equal to, or less than the absorption limit of the selected material. Students tested their computational models as they built them by clicking on the green flag (similar to Scratch; scratch.mit.edu). Analysis of prior SPICE implementations highlighted the importance of (1) linking students' work during computational modeling to their prior science and unplugged CT work and (2) using productive computational modeling strategies (e.g., testing their models with different values of rainfall and materials) (Biswas & Hutchins, 2022). These findings were used to highlight relevant analytics during co-design of the RISE dashboard with teachers.

Co-designed dashboard components

The co-designed RISE dashboard (Hutchins & Biswas, 2023) consists of three core student result pages. The Story provides an overview of the class performance based on key immediate feedback needs recommended by teachers. This included text-based feedback highlighting class successes and opportunities for improvement based on performance measures (items scored by pre-defined rubrics) and strategies used (productive and unproductive strategies pre-defined based on the impact on student learning results). We included interactive data visualizations, such as the grouping of students based on strategy use with additional performance-based results (bottom right of Figure 2), based on teacher recommendations. The Strategies page provided a progression of student performance over the course of the curriculum (e.g., up to the "day" simulated in each planning period simulation) and their grouping by current strategy use. The strategy groups, e.g., Divers implemented unproductive, depth-first model construction strategies (c.f., Grover et al. 2016), were derived by previous analysis (Biswas & Hutchins, 2022). Finally, the Standards page provided a data table of all students with their scores on each completed curriculum task and identified strategy groups. All data visualizations based on artificial intelligence and machine learning methods include an explanation of the analysis approach to give teachers insight into the underlying mechanisms being visualized (e.g., a modal pops-up with the information when the blue button with an "i" is clicked).

The RISE dashboard is equipped with a Reflection Tool in which teachers can add reflections as they review the results (identified as the "Reflection Form" in Figure 2) and select categories for the type of reflection. Submitted forms were populated on the Reflection page based on the category selected (the page link is identified on the left-side menu bar in Figure 2). In the Reflection page, teachers can re-order and reorganize reflections as they see fit to support teacher engagement. Finally, teachers are also provided with a Response page. This page includes the current class plan for the next class day and tools to plan for any adjustments they deem necessary based on student performance. Teachers have access to a number of curriculum resources that include learning objectives and lesson plans relevant for the "day" to aid in their evaluation process.

Figure 2
Components of the RISE dashboard



Methods

Eight middle school STEM teachers (5 female, 3 male) participated in the study. The teachers were from varying urban and rural locations in Tennessee, Illinois, Virginia, New York, Wyoming, and the US Virgin Islands. While all teachers had prior experience with teacher dashboards, four teachers also had prior training and experience in implementing a computational modeling curriculum in their science classrooms, so we labeled them as experienced teachers, and four teachers had no prior training or experience with computational modeling in science, therefore, we labeled them as novice teachers. All teachers consented to participate in the Vanderbilt University IRB-approved study.

Planning period simulations

We focus this paper on five Planning Period Simulations in which teachers enacted five 15 minute "planning periods" by utilizing the RISE dashboard to review and reflect on student, group, and class performance and then developed lesson plan customizations for the "next" class day. These simulations were inspired by the Teacher Moments research at MIT (Benoit et al., 2021). Student data used for each simulation was pulled from prior SPICE studies using an approach similar to the Replay Enactment protocol (Holstein et al., 2019). Student data from the prior implementations were de-identified and students were given gender-neutral names. The five simulations were selected based on class averages in summative assessment performances in science and CT (e.g., one simulation used data from a class that had an above average pre-test performance in science, but below average pre-test performance in CT).

Each teacher first completed a 90-minute professional development session led by the research team in which they learned about the SPICE curriculum. For each simulation, a research team member first described the class scenario, including the class performance on the pretest and other class results prior to the simulation "day" (e.g., performance on the science conceptual models). Using a think-aloud protocol, teachers reviewed student results and feedback provided on the RISE dashboard, interpreted what they saw, and customized class lesson plans for the next day (as they saw fit). Prior research has noted the benefits of think-aloud protocols on tasks involving building interpretations (Charters, 2003; Campos et al., 2021). To obtain verbalizations that accurately reflected the cognitive processes teachers implemented during responsive teaching, we refrained from providing detailed instructions or interpretation of results. Instead, we utilized prompts such as "what possible actions would you take with this group?" and answered questions about technology that did not impact class evaluations (e.g., describing how to use the reflection form). This helped minimize issues concerning bias in data if researcher support or feedback impact teachers' responses (Sherin and Russ, 2014).

Finally, researchers completed an observation sheet during the simulations. The observation sheet consisted of a table for researchers to identify the (1) discussed idea (e.g., computational model scores), (2) visualization targeted, when applicable (e.g., bar graph of class performance), and (3) keywords used or links made (e.g., poor initialization of science variables score during computational modeling relating to prior science performance). These observations were used to support our analysis approach, discussed below.

Data collection and analysis

All Planning Period Simulations were conducted virtually and recorded on a video conferencing platform. In total, we had approximately 12 hours of video data, which we transcribed with an online transcription service. For this paper, we formulated the base unit of analysis by segmenting the transcripts into episodes of pedagogical reasoning (Horn and Little, 2010). An episode of pedagogical reasoning started after the researcher's opening statement about the class scenario and ended when the teacher submitted their customized lesson plan.

We used methods of inductive coding and constant-comparative analysis (Charmaz, 2006) as opposed to theoretically developed codes due to the exploratory nature of the work and a dearth of prior research examining how the resulting interpretations facilitate pedagogical actions. This approach led us to identify catalyzing links that teachers used to transition from their interpretations of AI-generated data visualizations to evidence-based lesson customizations. We develop conjectures about these links and their implications on teacher responses. Team members met to discuss episodes and the links that teachers applied in these episodes. We created analytical memos (Hatch, 2002) to help us then compare teachers' catalyzing links. In the discussion of the links identified, we reviewed the literature on processes that support learning in integrated domains to refine our understanding of the links to help us define the emerging patterns. In the context of the full picture of all pedagogical episodes, we noticed the recurrence of similar patterns as catalyzing links (e.g., supporting learning through multiple, linked representations) and planning period simulations, which suggested patterns exist in the relationship between interpreted classroom needs and class performance distributions.

Results

Utilizing the exploratory analysis process described above on episodes of pedagogical reasoning from all eight teachers, the team identified five key catalyzing link patterns that combined students' learning performance and use of strategies. The teachers utilized these links to support the decision-making processes that transitioned from interpretation of AI-based analysis of students' results to evidence-based lesson plan customizations. These links include: **Supporting Student Understanding Across Multiple, Linked Representations** (MLR; implemented by 4 experienced and 1 novice teacher), **Leveraging Student Successes** (LSS; implemented by 3 experienced and 2 novice teachers), **Representing, Addressing, and Leveraging Productive Failure in PBL** (PF; implemented by 3 experienced and 3 novice teachers), **Weighing responses for different levels of social interactions** (LSI; implemented by 3 experienced and 4 novice teachers), and **Integrating real-world contexts** (RWC; implemented by 2 experienced and 2 novice teachers). We describe each and provide example teacher quotes in Table 1.

Table 1
Identified catalyzing links supporting teacher transitions from noticing to response creation

Link	Description	Example Quote
MLR	Supporting students' understanding across multiple, linked representations (e.g., science, computing, engineering) based on students' performance across multiple assessments and their use of strategies during the computational modeling task	[reviewing Strategy Grouping Figure] *"I'm still thinking about the materials. How to get them to transfer that original [engineering] grid you'd set up so that they have to have different values for the materials. Because it's still more than half [that aren't testing]. I think those overt connections between the lab experiment [in science] and the [computational] model. We make those [connections] implicitly as adults, but I think it needs to be more obvious for a younger brain to connect the model to the real thing."*
LSS	Planning lesson plans to promote future success and/or motivating and engaging students who were not known to be computing enthusiasts or were new to	[reviewing Domain-Specific Score Figure] *"Well, I think in two, you probably need to reflect on the success of the [initializing variables]. Because they did. They were successful for the majority, but I think it would be good to reflect [on that] because that may push them to do better on this, the equal*

	computing (leveraging the "Successes" dashboard feature)	*to condition. Does that make sense? So it's important to have reflection on initializing variables as a class. Because if you focus on success, it drives success. Rather than say, Oh, y'all did a good job, let's go on to the next one."*
PF	Understanding, addressing, and utilizing productive failure during PBL to motivate and engage students in the difficult problem-solving process	[reviewing "Opportunities" Text Feedback] *"The other thing that I look at a lot is normalizing mistakes. And so if I would, it becomes tricky. And it really you have to normalize it from the first day. We all make mistakes, but mistakes can help us get better at it and have someone share a mistake related to the materials where everybody looks at it together to figure out."*
LSI	Determining the optimal pedagogical response targeting different levels of social interactions (e.g., lecture, group activity, or individual support)	[reviewing Strategy Grouping Figure] *"When you look at the dashboard, and you see those bigger amounts of needs, that's when you have to go triage time, and you got to think about okay, I'm gonna have to do something much different here. Because I've got a lot of misconceptions. Yeah, that's where a bigger [class-level] action will occur."*
RWC	Connect the real-world context of the curriculum's problem to create customizations that support collaborative work and allow for students to reinforce their knowledge	[reviewing Domain-Specific Score Figure] *"Because the whole framing of the problem is, you are a project manager and you're designing a playground. So you're gonna work with this team. And you're, you know, it's okay that you (a student that finished the computational model first) are a consultant. You are a part of the team that knows computing really well"*

The identified patterns in our exploratory analysis provide an initial framing to understand how teachers transitioned from their data interpretations to evidence-based pedagogical responses. In this section, we explore these transitions in more depth by comparing two episodes of pedagogical reasoning: one by an experienced teacher and the other by a novice teacher. Both cases involve an evaluation of a class that was low performing on the pretest in science, but high performing on the computing pretest.

This first example represents an experienced teacher. After the researcher provided the Simulation Scenario, the teacher began their think aloud process:

> TEACHER: *I always like looking at graphs first…I'm gonna go to that* [reviewing Domain-Specific Score Figure]. *So they had this same problem with initializing their variables. They did better at equal to, and then less than and greater than so not too bad. And this is probably reflecting that they were [science conceptual model group]…So yeah, knowing that, I would say because they did a not so great job here but they did a better job here. So they are getting some of this computing stuff. I would add in a quick physical demo showing the difference between absorption for a sponge and a paper. So they can just understand why they need to test more materials. So let me put in that.*

In this first segment, the teacher linked students' prior performance in science with issues concerning the initialization of science variables in the computational model. The teacher then checked student behaviors to identify if any other data was available to explain this issue. They then clicked to add a Reflection (a technical issue arose) and the teacher then added a reflection noting they would need to conduct another science demonstration to show the difference between the absorption of different materials to help students understand why they would need the science variables in their computer models. As such, the teacher utilized the process of ***Supporting Student Understanding Across Multiple, Linked Representations*** to generate a possible evidence-based response. At this point, the teacher had only viewed bar graphs illustrating the number of students who got initial variables and the conditions correct and the number of students who tested more than 2 materials. The teacher continued.

> TEACHER: *TEACHER: Let me look at behaviors* [reviewing the Computational Model Testing Figure seen in Figure 2]. *So yeah, I think that I'm going to add [a reflection]. So this is an opportunity [for students[to see more physical examples. So I think that's really important. Because they did a better job here. And yeah [reviewing the Student Groups Based on Strategy figure in Figure 2], we've got a lot of* [bad CT strategy group]. *So we need to fix that again. And I'm going to add an activity. Let's do Palmer. So they're gonna show and talk through their*

code. I think this is the nicest thing. I'm going to tell you why I like this…what I'm trying to do as a teacher is I am trying to get kids to be more [good CT strategy group] *and getting comfortable with coding… So seeing who's doing those techniques is really going to help me and then seeing who changes because sometimes in the moment, I'm only picking on the kids that I know are strong in CT, to show examples. And I think that can be a bit demoralizing for other students. So like as this goes on, let's say Kendall, all of a sudden jumps* into [good CT strategy group] *or something like that, that's like a great thing. But if I'm able to see like, someone made the jump from here to here, I can then highlight them and hopefully give them as you know, some nice positive praise, reinforcement kind of thing that I think would be really helpful.*

In this segment, the teacher further acknowledged the need to understand why testing different materials, now from the perspective that a lot of students did not test their computational models. The teacher connected science practices and computational practices, and in addition to conducting the science activity, she elected to do a class presentation in which a student was selected to demonstrate their code and testing practices. Using the ***Leveraging Student Successes*** approach, the teacher selected a student who changed to a more productive strategy group to promote the students' good work and improvement. The teacher concluded with their customized lesson plan:

TEACHER: *All right. So we still have the initializing variables, and the two thirds are [bad strategy group]. Okay, so I think let me go to reflect. Yeah, so I think having that physical example is really important for this class. And then maybe Palmer you know, maybe I bring in Palmer towards the end of class instead, for this group, and it because maybe the physical demo will help more. And then I can add Palmer in to wrap up.*

To address a conceptual issue regarding initializing variables and poor strategy performances by the class, this teacher used the processes of ***Supporting Student Understanding Across Multiple, Linked Representations*** and ***Leveraging Student Successes*** to determine and finalize an evidence-based lesson plan customization.

In the second example, we discuss a novice teacher presented with the same class simulation. The teacher utilized the links ***Weighing Responses at Multiple Social Levels*** and ***Integrating Real-World Contexts*** to create an evidence-based customized lesson plan for the class. The episode began with the teacher reviewing the Computational Model Testing Figure (seen in Figure 2) and identifying a class issue:

TEACHER: *Would it help to have samples of those materials on display in the classroom? Are they already doing that?*
RESEARCHER: [researcher describing that materials are available]
TEACHER: *I mean, it looks to me that the biggest need for the next day is to address the materials portion. And I like that this makes it clear.*

The teacher began with a curriculum question about the availability of physical materials. The teacher identified that the lack of testing the computational model with different materials was a problem. The teacher continued:

TEACHER: [reviewing the Student Groups Based on Strategy figure, Figure 2] *And I always worry about those students in the classroom who are done. Okay, now, what do you do now that you're done? I would be assigning them to work with someone who needed more support, so day Reese is going to be an expert. You can consult an expert with your work and bring in a consultant. And you can ask them three questions.*
RESEARCHER: ... [researcher agreement]
TEACHER: *You got to explain. But I would put a constraint on it. I think that's a challenge, that's why this could be really valuable to know that they're completed… But kind of put a constraint on it, like, you can only ask a consultant a question, they can't just tell you stuff. Yeah. And it can't be a question like, how do I write the code? Yeah.*

In this segment, the teacher applied Weighing Responses at Multiple Social Levels to reason about a potential lesson plan customization. The teacher also utilized pedagogical content knowledge (the consultant activity was described by the teacher in prior discussions during SPICE training) to determine a productive response. In discussing the idea further, the teacher said:

> TEACHER: *And this I think, I like also, because you can see who needs the most support. Yeah, you know, that they're kind of stalled … I like looking at the written feedback that you're pinpointing, you know, where there are opportunities, but that can also help you target the consultant. Yeah. I like that, because the whole framing of the problem is, you are a project manager and you're designing a playground. So you're gonna work with this team. And you're, you know, it's okay that you're a consultant. You are a part of the team that knows the computing really well.*

In this segment, the teacher described the opportunity presented by this pedagogical approach - being able to listen in on what consultants (students that finished the code) and project managers (in-progress students) were saying during this paired activity to identify students that may need additional support. The teacher reasoned about the choice of lesson activity by ***Integrating Real-World Contexts***. The teacher concluded:

> TEACHER: *Yeah, I think this is kind of invaluable. Alright, so the lesson plan is [coming together] I think the lesson plan for this is more, we'll have the consultants and the students that need work will create the three questions. And yeah, to have more complete testing behavior. Make sure everyone has like I also call them experts. calling you an expert. I do let them [talk to me] if they're stuck, I do become a spy but, don't do that often there.*

The teacher reiterated the response reasoning and acknowledged the underlying, student-centered design of the curriculum by noting that they should intervene when necessary but limit the number of interventions.

Comparison of the two teachers demonstrates key catalyzing links implemented by the experienced versus the novice teacher and highlights the differences in their approaches. For this simulation, the experienced teacher utilized the catalyzing links of ***Supporting Student Understanding Across Multiple, Linked Representations*** and ***Leveraging Student Successes*** to determine their lesson plan customizations, while the novice teacher utilized ***Weighing Responses at Multiple Social Levels*** and ***Integrating Real-World Contexts*** in their lesson plan customization. This seems to indicate that the experienced teachers leveraged their experience and comfort with the curriculum and technology (i.e., domain-specific knowledge) to notice and focus on multiple dimensions of student learning, while the novice teachers may have compensated for their lack of SPICE experience by leveraging their pedagogical content knowledge in PBL to discuss curriculum changes (e.g., discussing normalizing mistakes in Table 1, link PF), deciding on social level responses as gaps grow between groups of students, and leveraging the real-world context of the problem. We believe these results extend on Campos et al.'s (2021) findings by pointing to potential future research in (1) exploring how the provision of different types of teacher-created responses utilizing different catalyzing links post-simulation can help experienced and novice teachers critically reflect on their response choices and discuss how they might change, if at all, and (2) developing tools to aid in teachers' noticing by interpreting the complex learning analytics (e.g., van Leeuwen et al. 2019) that target their background and experience.

Results

This research presents a novel exploration into the processes teachers take from noticing and interpreting learning analytics from a co-designed dashboard to reasoning and enacting evidence-based pedagogical adjustments through lesson plan customizations. In particular, this exploratory work provides a preliminary framework for identifying and evaluating catalyzing links teachers implement to decide and create evidence-based pedagogical adjustments based on AI-based analyses of student learning and problem solving. Distinct from prior work in technology-enhanced responsive teaching (e.g., alerting teachers of individual student errors or disengagement) (Holstein et al., 2019; Van Lehn et al., 2021), these examples demonstrate that the dashboard supported teachers in implementing class activity that (1) increased class or group discussions and (2) supported the development of productive problem-solving strategies, both key for supporting PBL curricula.

We recognize limitations in our work. On the one hand, the low number of teacher participants in this study resulted in analyses focused on depth instead of breadth. Future work should increase the participant cohort to further validate our results and to ensure that teacher preparation is inclusive and supports equity in PBL applications. In addition, in terms of the selection of classes for each simulation, we recognize a limitation in the use of a high- vs low-performing dichotomy in the selection of classes as that approach may not fully represent the nuances of learning and problem-solving behaviors from a classroom context. Future work in selecting data for simulations (and co-design) can evolve more nuanced approaches to evaluating classes, groups within classes, and individual students. Finally, we aim to complete a full, iterative cycle in which the participating teachers will implement SPICE (supported by the RISE dashboard) in their classrooms to holistically examine the dashboard impact on lesson plan customizations and implementations.

References

Ahn, J., Nguyen, H., and Campos, F. (2021). From visible to understandable: Designing for teacher agency in education data visualizations. *Contemporary Issues in Technology and Teacher Education*.

Biswas, G. and Hutchins, N.M. (2022). Towards a Deeper Understanding of K-12 Students' CT and Engineering Design Processes. In Ouyang, F., Jiao, P., McLaren, B.M., Alavi, A.H. (Eds.) Artificial Intelligence in STEM Education: The Paradigmatic Shifts in Research, Education, and Technology. CRC Press.

Benoit, G., Slama, R., Moussapour, R. M., Reich, J., and Anderson, N. (2021). Simulating more equitable discussions: using teacher moments and practice-based teacher education in mathematical professional learning.

Campos, F., Ahn, J., DiGiacomo, D. K., Nguyen, H., and Hays, M. (2021). Making sense of sensemaking: Understanding how k–12 teachers and coaches react to visual analytics. *Journal of Learning Analytics*, 8(3):60–80.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage.

Charters, E. (2003). The use of think-aloud methods in qualitative research an introduction to think-aloud methods. *Brock Education Journal*, 12.

Chen, Y., Hmelo-Silver, C., Lajoie, S., Zheng, J., Huang, L., and Bodnar, S. (2021). Using teacher dashboards to assess group collaboration in problem-based learning. *Educational Technology Research and Development, 15*(2).

Farrell, C. C. and Marsh, J. A. (2016). Contributing conditions: A qualitative comparative analysis of teachers' instructional responses to data. *Teaching and Teacher Education, 60*:398–412.

Hatch, J. A. (2002). *Doing qualitative research in education settings*. SUNY Press.

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*(3):235–266.

Holstein, K., McLaren, B. M., and Aleven, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher–ai complementarity. *Journal of Learning Analytics, 6*(2):27–52.

Horn, I. S. and Little, J. W. (2010). Attending to problems of practice: Routines and resources for professional learning in teachers' workplace interactions. *American Educational Research Journal, 47*(1):181–217.

Hutchins, N.M., Biswas, G. (2023). Using Teacher Dashboards to Customize Lesson Plans for a Problem-Based, Middle School STEM Curriculum. In LAK23: 13th International Learning Analytics and Knowledge Conference (LAK 2023). Association for Computing Machinery, New York, NY, USA, 324–332.

McElhaney, K.W., Zhang, N., Basu, S., McBride, E., Biswas, G., & Chiu, J.L. (2020). Using Computational Modeling to Integrate Science and Engineering Curricular Activities. In Proceedings of the International Conference of the Learning Sciences (ICLS), Nashville, TN, USA.

Molenaar, I. and Knoop-van Campen, C. A. N. (2019). How teachers make dashboard information actionable. *IEEE Transactions on Learning Technologies, 12*(3):347–355.

Sherin, M. and Russ, R. (2014). *Teacher Noticing via Video: The Role of Interpretive Frames*, pages 11–28. Routledge.

van Leeuwen, A., Rummel, N., and van Gog, T. (2019). What information should cscl teacher dashboards provide to help teachers interpret cscl situations? *International Journal of Computer-Supported Collaborative Learning, 14*, 1–29.

VanLehn, K., Burkhardt, H., Cheema, S., Kang, S., Pead, D., Schoenfeld, A., & Wetzel, J. (2021). Can an orchestration system increase collaborative, productive struggle in teaching-by-eliciting classrooms? *Interactive Learning Environments, 29*(6), 987-1005, DOI: 10.1080/10494820.2019.1616567

Voyiatzaki, E. and Avouris, N. (2014). Support for the teacher in technology-enhanced collaborative classroom. *Education and Information Technologies, 19*(1):129–154.

Wiley, K. J., Dimitriadis, Y., Bradford, A., and Linn, M. C. (2020). From theory to action: Developing and evaluating learning analytics for learning design. In Proceedings of the Tenth International Conference on Learning Analytics Knowledge, LAK '20, page 569–578, New York, NY, USA.

Acknowledgments

We gratefully acknowledge the support of the National Science Foundation award DRL-1742195 and the National Science Foundation AI Institute for Engaged Learning (EngageAI Institute) under Grant No. DRL-2112635. Any opinions, findings, and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

