Title
A human-centred learning analytics approach for developing contextually scalable K-12 teacher dashboards

Journal
British Journal of Educational Technology, 55(3)

ISSN
0007-1013

Authors
Wiley, Korah
Dimitriadis, Yannis
Linn, Marcia

Publication Date
2024-05-01

DOI
10.1111/bjet.13383

Received: 1 March 2023 | Accepted: 16 August 2023

DOI: 10.1111/bjet.13383

ORIGINAL ARTICLE

British Journal of Educational Technology BERA

A human-centred learning analytics approach for developing contextually scalable K-12 teacher dashboards

Korah Wiley[1] | **Yannis Dimitriadis**[2] | **Marcia Linn**[3]

[1]Digital Promise, Center for Inclusive Innovation, Washington, D.C., USA

[2]Department of Signal Theory and Communications and Telematics Engineering, ETS de Ingenieros de Telecomunicación, Universidad de Valladolid, Valladolid, Spain

[3]Graduate School of Education, University of California at Berkeley, Berkeley, California, USA

Correspondence
Korah Wiley, Digital Promise, 1001 Connecticut Avenue NW, Suite 935, Washington, D.C. 20036, USA.
Email: kwiley@digitalpromise.org

Funding information
European Regional Development Fund; National Research Agency of the Spanish Ministry of Science and Innovation; National Science Foundation

Abstract
This paper describes a Human-Centred Learning Analytics (HCLA) design approach for developing learning analytics (LA) dashboards for K-12 classrooms that maintain both contextual relevance and scalability—two goals that are often in competition. Using mixed methods, we collected observational and interview data from teacher partners and assessment data from their students' engagement with the lesson materials. This DBR-based, human-centred design process resulted in a dashboard that supported teachers in addressing their students' learning needs. To develop the dashboard features that could support teachers, we found that a design refinement process that drew on the insights of teachers with varying teaching experience, philosophies and teaching contexts strengthened the resulting outcome. The versatile nature of the approach, in terms of student learning outcomes, makes it useful for HCLA design efforts across diverse K-12 educational contexts.

KEYWORDS
design-based research, human-centred learning analytics, K-12 education, knowledge integration, learning outcomes, teacher dashboard

Practitioner notes

What is already known about this topic

- Learning analytics that are aligned to both a learning theory and learning design support student learning.
- LA dashboards that support users to understand the associated learning analytics data provide actionable insight.
- Design-based research is a promising methodology for Human-Centred Learning Analytics design, particularly in the K-12 educational context.

What this paper adds

- Leveraging a longstanding, yet fluid, research-practice partnership is an effective design-based research adaptation for addressing the high variation in instructional practices that characterize K-12 education.
- Using both quantitative and qualitative data that reflects students' developing knowledge effectively supports teachers' inquiry into student learning.
- Teachers' use of learning analytics dashboards is heavily influenced by their perspectives on teaching and learning.

Implications for practice and/or policy

- Impact on student learning outcomes, alongside usability and feasibility, should be included as a necessary metric for the effectiveness of LA design.
- LA dashboard developers should both leverage learning data that reflect students' developing knowledge and position teachers to take responsive pedagogical action to support student learning.
- LA researchers and developers should utilize a long-term, yet fluid, research-practice partnership to form a multi-stakeholder, multidisciplinary design team for Human-Centred Learning Analytics design.

INTRODUCTION

Learning analytics (LA) can help teachers to understand and optimize the learning process and context (Siemens, 2012). Given the centrality of human involvement in learning (whether as teacher or student), if LA are to function as envisioned, they must attend to the needs of those engaged in the learning process. Additionally, LA must be sufficiently robust to effectively support the understanding and optimization of learning within increasingly complex educational contexts, especially in K-12 education (Kovanovic et al., 2021). Moreover, significant advances have been made so that LA dashboards can provide actionable insights (Jørnø & Gynther, 2018) to the main stakeholders involved in the teaching and learning process (Kasepalu et al., 2022; Verbert et al., 2013). While Human-Centred Learning Analytics (HCLA) design approaches have made strides in this direction, further expansion is needed (Buckingham Shum et al., 2019; Dimitriadis et al., 2021).

First, HCLA design approaches should be expanded to position usability and feasibility as starting rather than endpoint metrics for LA effectiveness. Second, new HCLA design approaches should utilize a relevant learning theory to align the LA with the desired learning outcomes (Gašević et al., 2015, 2017; Wiley et al., 2020)—a theory that accounts for the social aspects of learning in addition to individual student behaviours, as measured through clickstream data, eye movement or other indicators (Krumm et al., 2021; Ochoa et al., 2017).

Third, HCLA design approaches should generate LA that are effective in supporting the achievement of the desired learning outcomes across diverse users and contexts (ie, contextually scalable). This paper proposes an HCLA design approach that attends to *all three* of the above expansion areas. Overall, this paper shows that the proposed Human-Centred Design approach in the context of Design-Based Research, led by a longstanding Research-Practice Partnership (RPP), may achieve the desired learning outcomes across diverse users and contexts.

Since the human-centred design landscape is complex enough, some definitions of important terms are present. We adopt this definition of participatory design: "Participatory design in our definition includes activities where users, designers and researchers collaborate towards shared goals" (Bødker et al., 2022, p. 3). In participatory design, mutual learning among these groups is fundamental. We characterize co-design, as "a highly-facilitated, team-based process in which teachers, researchers, and developers work together in defined roles to design an educational innovation, realize the design in one or more prototypes, and evaluate each prototype's significance for addressing a concrete educational need" (Penuel et al., 2007, p. 51). Moreover, co-designers are not only end-users who provide feedback for the design team they are also decision-making members of the design team, from ideation to use in practice. Overall, we adopt the categorization of the human-centred design landscape (Sanders, 2006, Figure 1), in which approaches are categorized according to whether they follow expert versus participatory mindset, or are led by design or by research. We also acknowledge that such definitions are still under extensive discussion in the human-centred design research field.

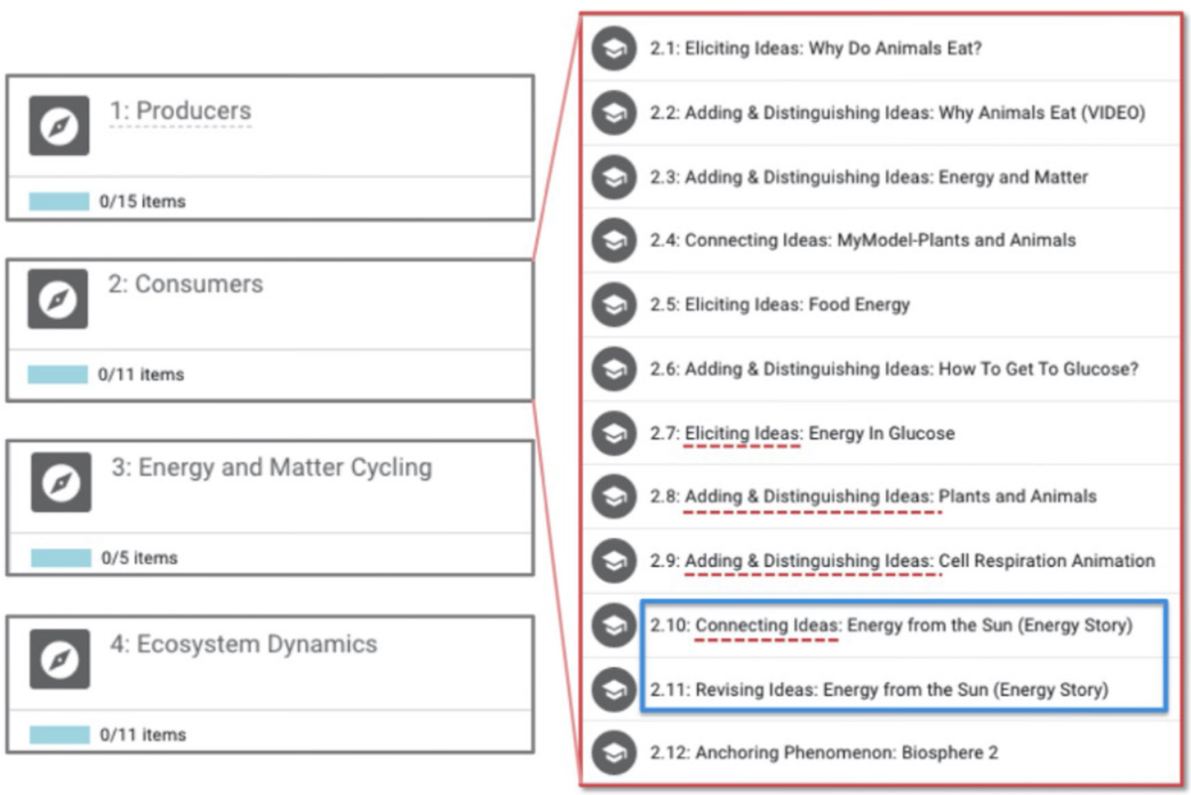


FIGURE 1 Schematic diagram of the WISE photosynthesis unit. Red dashed lines indicate the steps of the Knowledge Integration process being targeted by the unit activity. Blue box indicates the steps associated with the Initial Explanation item (2.10) and the Reflection Explanation item (2.11).

Background and rationale

Designing in the K-12 context presents unique challenges that warrant the use of HCLA design approaches. With the focus on supporting students to meet standards, it is important to involve teachers early in the design process to ensure that the type of data collected is going to help them understand their students' developing understanding. When teachers are not involved in this part of design decision-making, LA designers who are unfamiliar with the most relevant type of data may base data use decisions on convenience of access (Siemens & Baker, 2012). For example, while it may be convenient to create LA using readily available data that give insight into the platform navigation and engagement behaviours of the user and thus changes in student *activity* (eg, clickstream data), this type of data does not provide teachers with the insight they need to understand changes in student *knowledge* or to support their learning (Baker et al., 2020). Moreover, given the complexity of designing LA in the K-12 context, it is important to have the users (eg, teachers) involved in the design process as well as learning sciences researchers who bring experience in pedagogy. Combining the expertise of learning sciences researchers and teachers ensures that both the experiences of the teachers and the insights from pedagogical research are represented on the design team. Such researchers attend to the learning process in a way that can ensure that the LA solution can support the learning of students outside the experience of the teachers in the design team. Learning sciences researchers can identify relevant learning theories to support the achievement of the targeted learning outcomes (ie, meet the curriculum standards; Expansion Point: utilize a relevant learning theory to align the LA with the desired learning outcomes).

Furthermore, it is important to have designers on the team to support the process of the hallmark characteristics of design-based research methodology (Sandoval & Bell, 2004). We argue that when used to develop K-12 LA, DBR methodology is an effective strategy for centring human needs. To illustrate, in DBR, the primary focus is on understanding the context, particularly the users' experience and needs in the context (Collins, 1992). Additionally, in DBR, the goal is to develop an effective solution, a solution that will achieve the desired outcomes within the authentic context and that those outcomes are meaningful for those using and impacted by the solution (Barab & Squire, 2004). DBR also has the goal of refining the theory that informed the design of the solution (Collins et al., 2004). Thus, when DBR is used for HCLA development, it positions the HCLA solution as a tool for learning theory refinement. As such, the HCLA solution can support both researchers and users (eg, teachers) to gain a better understanding of how learning occurs in context. For the researcher, such data can either corroborate or challenge conceptions about how humans learn. For teachers, such data can either corroborate or challenge conceptions about how their students learn. Whether the learning theories are on the macro-level (for the researcher) or micro-level (for the teacher), the HCLA-mediated refinement to pedagogy can improve learning. Given the importance of designers in this process, we characterize the design approach present in this paper as DBR-based HCLA.

Thus, using DBR as the frame for an HCLA design approach can shift the research component from the LA development phase to the LA implementation phase, where teachers become researchers. This shift can work well within the K-12 context because teachers are expected to successfully intervene to support student learning. A DBR-based HCLA design approach, that is, an approach that employs DBR as a relevant theoretical framework, would position the LA as the tool through which teachers can iteratively support the learning of students as they get more data related to their students' learning. Rather than form theories that are generalizable to the larger field of education, they could form “micro-theories” about the individual learning processes of their students and how to redesign the learning environment to improve that learning (Mor et al., 2015).

Although K-12 education provides a rigorous HCLA design context, most of the research and development has taken place in and for postsecondary education (Kovanovic et al., 2021; Mangaroska & Giannakos, 2018). The reliance on standardized tests in K-12 education reflects a high-stakes ethos that K-12 education is foundational to students' future engagement in society. This perspective of K-12 education often leads to mandates for student success, rather than creating opportunities for it. For educational contexts, like K-12 classrooms, that also prioritize the achievement of specific learning outcomes, it is imperative that the HCLA design approach is grounded in a relevant theory of learning. With greater priority placed on integrated knowledge development, as evidenced in standards such as the US-based Next-Generation Science Standards (NGSS; NGSS Lead States, 2013) or world-wide assessments like Programme for International Student Assessment (PISA; https://www.oecd.org/pisa).

By focusing on the use of data (both qualitative and quantitative data) to provide teachers with rich and relevant insight into the learning process, in particular changes in student knowledge, an effective dashboard can contribute to student progress by including agreed upon indicators, aligned with the learning theory used by the partners. For example, in accordance with a learning theory that identified content-related reflection as important for learning, Cloude et al. (2021) developed LA based on the quantity and quality of students' learning reflections. By using an HCLA design approach that requires the use of a learning theory in conjunction with the evaluation of student learning as an integral part of the design process, we aim to contribute to the current dearth of related research (Expansion Point: utilize a relevant learning theory to align the LA with the desired learning outcomes).

Most K-12 content is driven by local, state/regional or national curriculum standards, so much of the complexity associated with K-12 education comes from *how* content is taught rather than what is taught. While lecture remains the primary instructional practice in postsecondary classrooms, the instructional practices in K-12 classrooms vary greatly, both across and within schools. These variations directly impact the way LA are implemented and ultimately their effectiveness for supporting teaching and learning. Recognizing the variability in K-12 teachers' use of LA dashboards, van Leeuwen et al. (2021) conducted a study to understand the factors that influence K-12 teachers' use of LA dashboards. They found that general teacher characteristics such as their teaching experience, age, gender and technology self-efficacy were not the primary influence. They hypothesized that complex factors including teachers' pedagogical knowledge, professional routines and data literacy may be more influential in dictating how teachers use LA dashboards.

To mitigate the impact of data literacy on teachers' LA dashboard use, Echeverria, Martinez-Maldonado, Buckingham Shum, et al. (2018), Echeverria, Martinez-Maldonaldo, Granda, et al. (2018) offer an HCLA design approach, Educational Data Storytelling (EDS), that forefronts educational outcomes while attending to important design aspects, such as usability and feasibility. These studies demonstrate how the EDS principles can enable the design of LA dashboards that guide teachers towards pedagogical actions that are aligned to specific educational outcomes. Indeed, co-designing learning analytics (LA) solutions with the intended users can produce a final design product that is highly effective in the implementation context. However, such co-design or broader participatory approaches can also lead to design products with limited utility for scaled contexts. (Expansion Point: position usability and feasibility as starting rather than endpoint metrics for LA effectiveness). Alternatively, design processes that prioritize scalability often shift towards a “designed for” rather than “designed with” approach, resulting in a product that may not meet teachers' relevant instructional needs based on their teaching context. Given the high variability in K-12 teaching practices and the tendency of classroom technology to be adopted at the department or school level, it is important that the HCLA design approach yields LA that can scale across such contextual variation (Expansion Point: generate LA that are effective in supporting the achievement of the desired learning outcomes across diverse users and contexts, ie, contextually scalable). When the design team and process include teachers with varying backgrounds, student populations and teaching practices, there is greater potential for the resulting LA to be able to function as compatible partners for classroom teachers. To develop LA that can scale to and be relevant for teachers in different teaching schools and with different pedagogical backgrounds (ie, have contextual relevance and scalability), some researchers (Krumm et al., 2021; Rodríguez-Triana et al., 2021) have used Design-based Research (DBR) to gain design insight from diverse teachers across multiple design cycles. DBR functions well as an HCLA design approach (Alvarez, 2020), given its focus on usable and feasible, yet theory-grounded solutions, iteration and multidisciplinary, multi-stakeholder design teams (Reimann, 2016; Sandoval & Bell, 2004), all of which can contribute to contextual scalability. As HCLA design approaches, like DBR, expand, they should not only position teachers as valuable for the LA design process but also for the LA refinement process associated with implementation and effectiveness (Expansion Point: position usability and feasibility as starting rather than endpoint metrics for LA effectiveness).

Another central element of the HCLA design approach described in this paper is the reliance on a K-12 educational technology (EdTech)-focused RPP. Using a K-12 EdTech-focused RPP can resolve numerous issues that commonly arise during the HCLA design process, especially in the K-12 context. A primary issue is having access to student data. Oftentimes, data sharing agreements are established during the process of forming the RPP or when new practitioner members (eg, districts, schools or teachers) join (He et al., 2020). Many times, these agreements are longstanding which significantly shortens the turnaround time during iterative design-test cycles. Ready access to student data allows each design-test cycle to include the evaluation of direct impact on student learning outcomes that are of great importance to school administrators when making decisions about technology adoption, namely the demonstrated improvement of student outcomes (eg, grades). Design HCLA within the context of a K-12 EdTech-focused RPP has the additional affordance of readily available, goal-aligned practitioner partners who can serve as co-designers and participants (He et al., 2020). More than just a sample of convenience, RPP practitioner members represent a pool of ideal and likely users. They are invested in the viability and effectiveness of the HCLA solution because the problem being solved was co-determined by them and relates directly to their practice. The requirement for addressing problems of practice that are of mutually beneficial value to both practitioners and researchers is another way the resulting HCLA is not overly specified for a particular set of users (Expansion Point: generate LA that are effective in supporting the achievement of the desired learning outcomes across diverse users and contexts, ie, contextually scalable). As mentioned previously, researchers attend to issues that hold promise for contributing to the larger understanding of the field so any problem addressed within an RPP must have potential for broader impact (Farrell et al., 2021). The potential for broader impact also comes when RPP members engage their networks to share the LA solution, making uptake and scale reasonable outcomes of a successful design. Since practitioner involvement in an RPP is often sanctioned by key administrators, in the event of a successful HCLA solution, the administrators are more inclined to adopt it because it would have already been vetted during the iterative design cycles (Penuel et al., 2021; Penuel et al., 2022). The practitioners chosen for membership in the RPP possess deep and relevant knowledge of the problem of practice and tend to represent diverse contexts. This diversity can be leveraged when selecting co-designers and participants for the design cycles. The fact that an RPP operates over the long-term makes it especially well-suited for a DBR-based HCLA design approach because the learnings from each iteration are retained as "institutional knowledge" within the RPP (Penuel et al., 2021). This affordance helps to mitigate the negative impact of member changes, such as when educator or system developer attrition rates are high or when researchers change institutions.

Although, collectively, recent HCLA design approaches attend to the areas of needed expansion, no single approach addresses all three areas. In this paper, we describe an HCLA design approach that addresses each of the aforementioned expansion areas and discuss the outcomes and learnings that can inform future research. Even though the expansion points and expected contributions have wider relevance in the general Learning Analytics and Learning Sciences fields, this paper aims at pointing out the specific connections to and implications for the field of HCLA.

Research questions

The DBR-based HCLA design approach described in this paper was driven by socio-constructivism, a learning theory that attends to both the social and cognitive aspects of learning. We leveraged data related to students' developing science ideas to create an LA dashboard in accordance with EDS principles that enables middle school science teachers, with diverse teaching backgrounds and in varying contexts, to investigate and develop interventions to help their students develop standards-aligned, multidimensional, integrated science knowledge. We evaluate the viability of this approach and effectiveness of the resulting LA dashboard by addressing the following questions:

1. In what ways can a DBR-based HCLA design process enable LA implementation within K-12 classrooms to *support teachers with varying teaching backgrounds, practices and perspectives in effective implementation*?
2. In what ways can a DBR-based HCLA design process enable LA implementation within K-12 classrooms to *support students in developing integrated, multidimensional content knowledge*?

In the following sections, we describe the specific features of the DBR-based HCLA design approach that we used (Methods), we present the effectiveness of this approach as defined by our research questions (RQs, Results), and we discuss the implications of our findings in terms of the expansion areas identified above and relevant research (Discussion). We conclude the paper with a statement regarding the contribution of our approach and findings to the HCLA design field and suggestions for future research.

STUDY DESIGN

Design iterations

The learning resources and LA dashboard developed in this study were the product of a design team consisting of researchers, teachers and system developers, and resulted from multiple cycles of design. All of the members of the design team, including the teachers, were members of the same K-12 EdTech-focused RPP, which consisted of the entire research group of the last author and over 100 middle and high school teachers from over 10 local school districts. The RPP members were oriented around the same goal of developing standards-aligned middle and secondary inquiry-based science curricula for use in the Web-based Inquiry Science Environment (WISE; https://wise.berkeley.edu/). Through the WISE platform, teachers across the world use the curricula not just those in the RPP. This study was motivated by the need that teachers expressed for support when using the RPP-developed WISE curriculum, namely in helping students meet the new science standards which called for multidimensional learning. While the specific standards targeted in the curricula were developed in the United States, they reflect science reform priorities that are shared across national contexts (Forbes et al., 2020). Thus, creating a technological solution to help teachers in supporting multidimensional learning represented a problem of practice with mutually beneficial outcomes, for teachers, researchers and system developers. The researchers and system developers formed a stable core of the RPP while teacher and administrator membership varied, with some teachers being members for over 20 years and others being first year members. The teaching experiences of the teacher members also varied independently of membership length. For this study, we co-designed with teachers who were new and longstanding RPP members, and who were new to and experienced in teaching (Table 1).

The determination of which unit to use for this study was based on whether the unit had an autoscored open response item and if it was being taught by a local RPP member teacher at the time of the study. Using a unit being taught by an RPP member teacher had logistical significance because it guaranteed access to student data, since RPP membership required participation in research studies. Particularly in contexts where access to student data is tightly regulated, designing HCLA within an RPP solves this logistical challenge of working within a K-12 context.

Design Phase 1 consisted of two design cycles that took place across two consecutive academic years. The goal of this phase was to design and test an HCLA solution to the critical problem of practice that the RPP members identified during the annual meeting in the school term prior to Phase 1. The consensus across teachers for the problem of practice to address was to develop technology that could help them support their students in developing multidimensional learning as called for by the new science standards.

For Phase 1, we invited a set of five local middle school science teachers in the RPP, who taught at the same school, to join the co-design team. Included in this set was the department leader, who had extensive professional training on how to support students in developing multidimensional knowledge in accordance with the science standards (ie, NGSS). The inclusion of this teacher was critical in the first design cycle as she was able to provide valuable guidance for the type of student data that teachers would need to support student learning. The perspectives provided by the other teachers informed recommendations for action in response to the LA data presented on the dashboard. Collectively, the teachers in the first design cycle provided guidance for designing an LA dashboard that would be effective and accessible to teachers with varying levels of teaching experience.

Although described in detail elsewhere (Wiley et al., 2020), we provide a brief description of the Phase 1 design process and outcomes here. Based on the teacher feedback from the pre-design phase, the co-authors and RPP system developers created several options

TABLE 1 Participant professional background.

Design phase	District ID	Teacher ID	Years teaching	Years in RPP
1	District-1	Teacher-A	15+	0–5
		Teacher-B	10–15	5–10
		Teacher-C	5–10	0–5
		Teacher-D	0–5	0–5
		Teacher-E	0–5	0–5
2	District-2	Teacher 1	15+	0–5
		Teacher 2	0–5	0–5
	District-3	Teacher 3	15+	15+

for technological solutions that could be developed using the logged data from the unit that the co-design teachers were teaching at the time. These options presented to the co-design teachers to stimulate ideation around a technology solution that would meet their need and that they would likely use given the constraints and opportunities of their teaching context (ie, feasibility). The researchers and system developers considered their ideas in terms of what would be executable within the timeframe of their use and what, if effective, could be usable across other contexts, particularly those represented within the RPP. Over the course of multiple synchronous and asynchronous conversations with the co-design teachers, we developed a prototype design that was subsequently implemented and tested for usability and perceived effectiveness in addressing the need of helping teachers to support students' multidimensional learning of the targeted content. After implementation the teachers gave their feedback on the data visualization approach and what action they planned on taking in light of the provided data analysis.

To guide the phase 1 design process, we drew on the OrLA practice of inter-stakeholder dialogue (Prieto et al., 2018) and used the LAID principles of coordination, comparison, and customization (Wise & Vytasek, 2017) as a conversational frame for the design discussions. Given their focus on critical factors for LA solution implementation, the LAID principles additionally functioned as usability and feasibility metrics for the design solution. During Cycle 1 of Design Phase 1, the design team identified and validated learning-congruent data from a learning theory-aligned, global climate change-focused inquiry science unit embedded in the WISE platform. Specifically, teachers and researchers decided together the type of LA to create after identifying key multiple-choice assessment items in the inquiry unit that were targeted to the distinguishing ideas step of the KI process. The LA were validated as being predictive of students' performance on a subsequent explanation assessment item (for more details, see Wiley et al., 2020). Additionally, the co-design teachers provided guidance for how the data could be analysed to provide them with insight into students' thinking on the way towards a multidimensional understanding of the targeted ideas. For example, one teacher commented, "I'm not looking for did they get it right or wrong, I want to know that explanation … I [want to] see right away who missed it, … instead of me having to look through". Based on this and other design recommendations by teachers, the design team created and evaluated a minimally viable HCLA solution, during Cycle 2 of Design Phase 1, using the validated learning-congruent data to determine whether and how it attended to issues of coordination, comparison and customization (ie, determined its usability and feasibility). A primary finding from Design Phase 1 was that while the HCLA solution had the potential to help teachers support students' multidimensional learning, specific attention needed to be given to design features related to the logistical use of the solution and the variation of use by different teachers (ie, logistical coordination and customization respectively; Wise & Vytasek, 2017). The Design Phase 1 feedback that was used to inform design revisions for Design Phase 2 is summarized in Table 2.

Another primary finding from the Design Phase 1 testing was that in order for the HCLA solution to support all teachers to take pedagogical actions in response to the analytics it needed to help provide them understanding the learning design related to the targeted assessments in a timely manner. Using the design and implementation feedback from the Phase 1 co-design teachers, we revised the technology solution. The agreed upon technology solution was an LA dashboard that would be embedded as a teacher feature within the unit curriculum. However, the feature development for the new HCLA solution was constrained by the academic calendar as teachers would be using the curriculum units in the next school term, which gave the system developers in collaboration with the researchers six months to complete a minimum viable product for classroom testing. The development timetable is linked to the school calendar because in-context design testing is central to a DBR-based design approach. For this study, "in-context" meant when classes were in

TABLE 2 Alignment of Design Phase 1 feedback to Design Phase 2 redesign.

Feature	Feedback	Redesign for Design Phase 2
Maximum Number of Attempts (MNA), in general	Teachers: Increased student anxiety without proper contextualization of its purpose; could encourage students to more seriously engage with the units if they perceived the limited number of attempts as an indicator of importance and take their time to provide more thoughtful responses; a means to provide insight into student thinking Ex: "I'm not looking for did they get it right or wrong, I want to know that explanation". "I [want to] see right away who missed it, instead of me having to look through" Researchers: most students unaware of the feature	Provide teachers with learning analytics that would give them insight into why students might have selected a particular incorrect answer as well as data regarding general correctness
Number of assessment items with MNA feature (all vs. select few closed-ended items)	Teachers: Caused students to no longer recognize it as an indicator of item importance; potential for students engaging more seriously on all the items would be "worth the risk" of decreased perceived importance	Restrict the MNA feature to only the multiple-option items that were aligned with the "distinguishing ideas" step in the Knowledge Integration process (described in the "Grounding Learning Theory" section), a critical step in the development of integrated knowledge but a challenging step for students to engage and teachers to support
Amount of analytics	Teachers: Regarding the type and amount of analytics that they wanted, the teachers all agreed that "less is more"	Provide teachers with near-time analysis to answer the following questions: How many students/groups (absolute and relative frequency) chose the correct answer on their first attempt? Of those who needed multiple attempts, what was the most commonly selected incorrect answer? What were alternative incorrect answer patterns? (the student/group ID codes were listed with the respective alternative pattern)
Meta-analysis	Teachers: Analysis resonated with in-class experience of students' learning needs. Researchers + System developers: not scalable if not automated, need meta-analysis that can be automated	Add section on analytics report called "Researchers' Insight" which reflected the researchers' conjectures about student thinking based on specific response patterns and automated feedback guidance associated with the incorrect submission Limit the Researchers' Insight to Design Phase 1, and replace it with automated "Key Insights" in Design Phase 2
Individual or aggregated data	Teachers: Had one teacher known more students were struggling with the same issue, she would have brought the discussion up to the whole-class level; teachers wanted to know the specific students who were struggling	Provide analytics at both the aggregate and individual level

TABLE 2 (Continued)

Feature	Feedback	Redesign for Design Phase 2
Timing of receiving/ reviewing the analytics: real-time or near-time	Teachers: Near-time access to the learning analytics (ie, after school or the next morning prior to instruction) would be the most realistic given their constraints and current teaching practices; Having the analytics in real-time would conflict with their practice of walking around the classroom when students worked on the unit as it would require, or at least encourage, them to be tethered to their desktop or to use a cumbersome mobile device System developers: While all the necessary log data was available, generating real-time/in-platform analytics would require new platform architecture	In Design Phase 1, provide teachers with an email report with the analytics after 50% of students completed the items Postpone real-time/in-platform analytics to Design Phase 2 to allow for time needed to develop the graphical displays and interface for the analysed data (ie, develop new platform architecture and system)
Feedback on MNA-activated items	Researchers+system developers: No students revisited the recommended resources after being prompted to do so by the automated feedback guidance; numerous students were choosing answers in a top-down pattern, suggesting a “clickthrough till correct” selection strategy	Place the multiple-option assessment items on the same page as their related resources so that students might be more inclined to reference the model if it appeared on the screen alongside the assessment item; limit the maximum number of attempts from the total number of answer choices to one less than this value to motivate students to attend to the feedback to get the correct answer rather than just clicking through response till they made the correct selection

session and when teachers were using the units, which varied by teacher and was determined by the timing of their curriculum progress. Situating this design process within the context of an RPP is valuable for this reason, namely that all parties involved are used to operating in sync, a rhythm that is most often driven by the educational system that practitioners are operating within (Penuel et al., 2021).

The objective for Design Phase 2 was to create an in-platform HCLA solution, namely a teacher dashboard, that supported all teachers to understand the learning design, to understand their students' progress and needs related to developing multidimensional learning of the targeted concepts, and to take effective pedagogical actions toward supporting such learning. We identified the Educational Data Storytelling approach as a promising strategy for simultaneously supporting teachers' understanding of the analytics and the unit's learning design.

After testing the usability and feasibility of the LA dashboard in the first design cycle, we invited a second set of RPP teachers to join the design team, as participatory designers. These teachers taught at different schools and had different teaching practices and perspectives (elaborated further in the Methods & Materials section). In addition to providing their perspective on the usability and feasibility of dashboard design, the second set of teachers provided design guidance for how the LA dashboard could support teachers with teaching practices similar to their own. The full set of design teachers functioned as a representative sample of teachers who would use the LA dashboard, including teachers with all levels of teaching experience, various teaching practices, and different student populations. In this way, the resulting LA dashboard would be neither overly fit to a select few or primarily master teachers nor require precise implementation to be effective.

Grounding learning theory

To guide the development and evaluation of LA and a teacher dashboard, we draw upon the Knowledge Integration (KI) pedagogical framework, a social constructivist perspective of learning that provides guidance for supporting students to develop integrated science knowledge. The framework holds that students enter any learning environment with preformed ideas that were developed through interactions with their physical and social environments and that inform their understanding of new ideas (Clark & Linn, 2013; Linn & Eylon, 2011). The major claim of the KI framework is that pedagogy that engages students in the aforementioned steps will support the development of integrated science knowledge. In this study, we use the KI pedagogical framework to develop the learning design used to facilitate student learning. In a KI-based learning design, students' ideas are elicited and made available for development by supporting them to distinguish between their *ideas* and evidence related to consensus science ideas. This distinguishing step is critical for students to develop integrated knowledge as it is when students determine which ideas are most productive for understanding the phenomena under study. The distinguishing step is often the step in which students and teachers need the greatest support (Vitale et al., 2016; Wiley et al., 2019). Students are then supported to refine their initial ideas in light of this evidence. Our selection of data to develop the LA for this study was guided by the well-researched and evidence-based KI framework. As such the resulting LA was highly likely to provide insight into student thinking and changes in students' developing knowledge. In the following section, we describe the alignment between the KI learning design and the LA used for the teacher dashboard. We also detail the implementation of the EDS principles into the LA dashboard design.

MATERIALS AND METHODS

Curriculum and assessments

Although the first design cycle was associated with a unit on climate change, this design cycle was associated with a unit on photosynthesis and cellular respiration. To assess students' starting and ending levels of integrated knowledge about photosynthesis and cellular respiration, we used an assessment item in the unit's pre/posttest. This item prompted students to write a story using scientific evidence to explain how a rabbit gets and uses energy from the sun, specifically how energy moves and changes. To assess students' learning during instruction, we used a unit-embedded assessment item, which prompted students to write a story to explain their ideas about how animals get and use energy to survive, again focusing on how energy moves and changes. This item appeared in two consecutive steps (Figure 1), before teachers reviewed the Teacher Action Plan (TAP) (referred to as Initial Explanation) and after their intervention (referred to as Reflection Explanation). All assessment items aligned to the middle school life science NGSS performance expectation, MS-LS1-6 (NGSS Lead States, 2013).

TAP design features

TAP dashboard development

We designed the TAP in collaboration with the platform system developers to provide teachers with information regarding student performance on the Initial/Reflection assessment item (see Curriculum section). To attend to factors related to usability, feasibility and learning outcomes, we implemented the six EDS principles (Echeverria, Martinez-Maldonado, et al., 2018; indicated with red circled numbers in Figure 2) as follows.

EDS Principle 1: *Align the LA dashboard with a clear educational goal derived from the learning design*. We aligned the TAP to the NGSS curriculum standard for photosynthesis (MS-LS1-6, NGSS Lead States, 2013). This curriculum standard calls for students to demonstrate an integrated understanding of energy and matter transformation by constructing a scientific explanation of the role of photosynthesis in the cycling of matter and the flow of energy. It was articulated in numerous ways at the top of the dashboard. We included a link to the webpage detailing the curriculum standard in the description of the assessment item and the item's location in the photosynthesis unit. We also provided the exact prompt for the assessment item and summarized the lesson's learning objective to link the language used in the curriculum standard with the language used in the item prompt.

EDS Principle 2: *Choose appropriate visuals to effectively communicate the desired message to the target audience*. The autoscores for students' performance on the embedded assessment item was aggregated by class period and reported using bar graphs as the percentage of students scoring at each score level for each dimension score. We represented the LA using bar graphs given teachers' likely familiarity with this data format (Alverson & Yamamoto, 2016; Whitaker & Jacobbe, 2017). Using bar graphs also seemed the most appropriate data format given our desire for teachers to compare the percentage of students across the different score levels.

EDS Principle 3: *Use narrative text to summarize visual features and create meaning*. Next to the graphs, we included a narrative summary of the LA, called Key Insights, which represented the key message that we wanted teachers to get from the LA. This

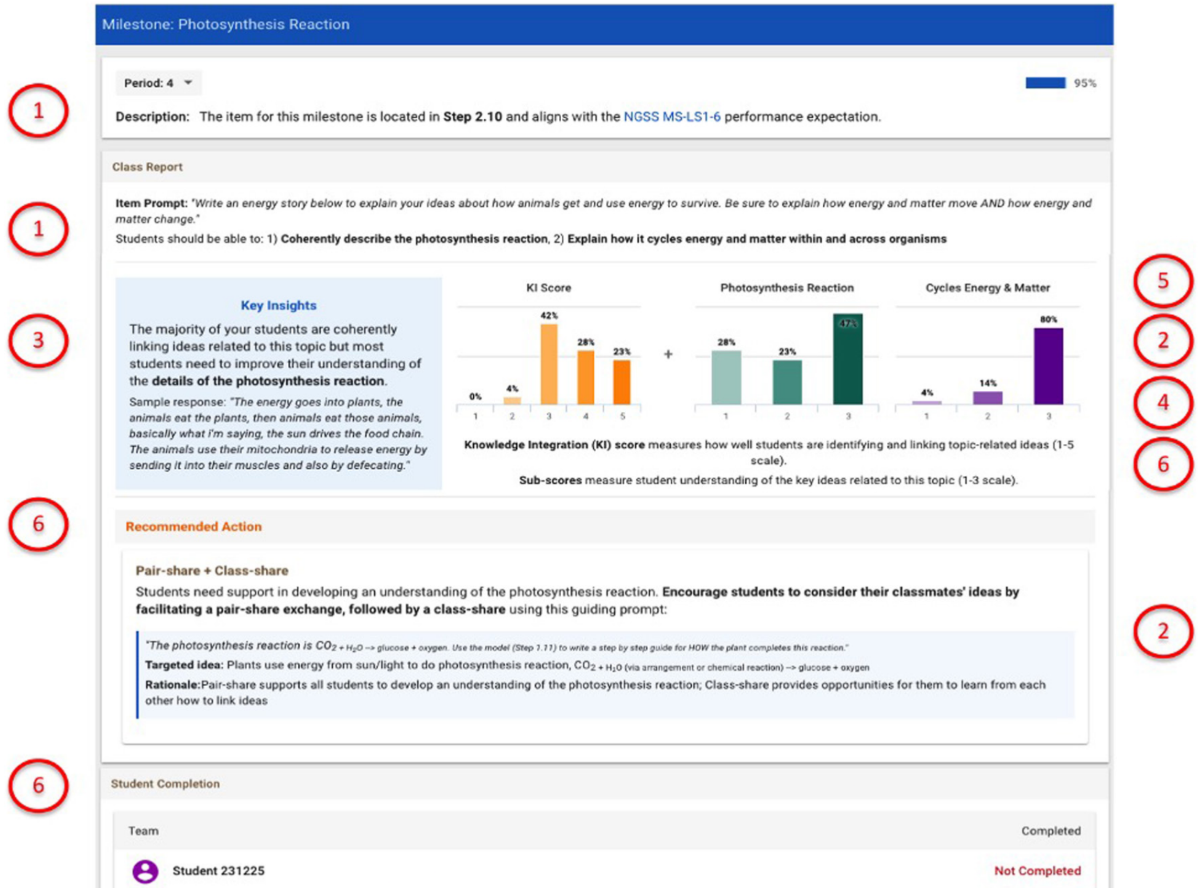


FIGURE 2 Annotated TAP displays. Red, circled numbers indicate the implementation of the correspondingly numbered Educational Data Storytelling Principle.

narrative content was written by the learning scientist on the design team, taking into consideration the KI-based scoring criteria. Specifically, the narrative described the aggregated class performance along two parameters, whether they expressed the targeted ideas (disciplinary core ideas and crosscutting concepts; DCI and CCC respectively) and whether those ideas were coherently linked (science and engineering practices/knowledge integration; SEP/KI), as determined by the algorithm-generated autoscore. Each sentence of the narrative was pre-written and in the form of a call to action for teachers to support students in the identified area of need. Thus, the final narrative summary was auto-generated based on the aggregated student scores for the class.

EDS Principle 4: *Use visual elements to direct visual attention and enhance sense making*. To support teachers in quickly understanding the LA data, we added data labels and major gridlines to the bar graphs. The data label provided teachers with specific percentages to take note of and the major gridlines would support teachers to make comparisons across score levels. We also emboldened certain words and phrases in the dashboard to draw teachers' attention to important aspects of the information being presented, such as the target learning objective, and the portion of the Key Insights that identified the target ideas for which students needed the greatest support to understand.

EDS Principle 5: *Use titles and captions to communicate the intent of the visual elements and explain relevant features in the data*. We used the assessment item topic for the title of the LA dashboard (eg, Milestone: Photosynthesis Reaction), and for each section of the dashboard, we included a short descriptive title (eg, Class Report, Recommended

Action). We captioned the graphs with the name and brief description of the scores being presented.

EDS Principle 6: *Maximize the data-ink ratio by excluding unnecessary headers, chart features, borders, grids*. Given our inclusion of data labels for the bar graph, we eliminated the *y*-axis line, deeming it redundant and unnecessary. We also chose not to have borders for the graphs.

TAP recommended actions

To support teachers in making evidence-based customizations to the learning design, the team created a recommended action for each possible data scenario across the three scoring dimensions, corresponding to each knowledge dimension of the targeted curriculum standard (ie, SEP/KI, DCI and CCC) and created binary score categories for each dimension. We designed the recommended actions in accordance with learning sciences research, to engage students in the full process of knowledge integration according to the KI framework (Figure 3). To support teachers in understanding the recommended action and encourage them to implement it, we included a statement of rationale. These recommended intervention activities make the TAP actionable by supporting teachers to take pedagogical action in response to the provided analytics, effectively "closing the loop" of the learning analytics cycle (Clow, 2012).

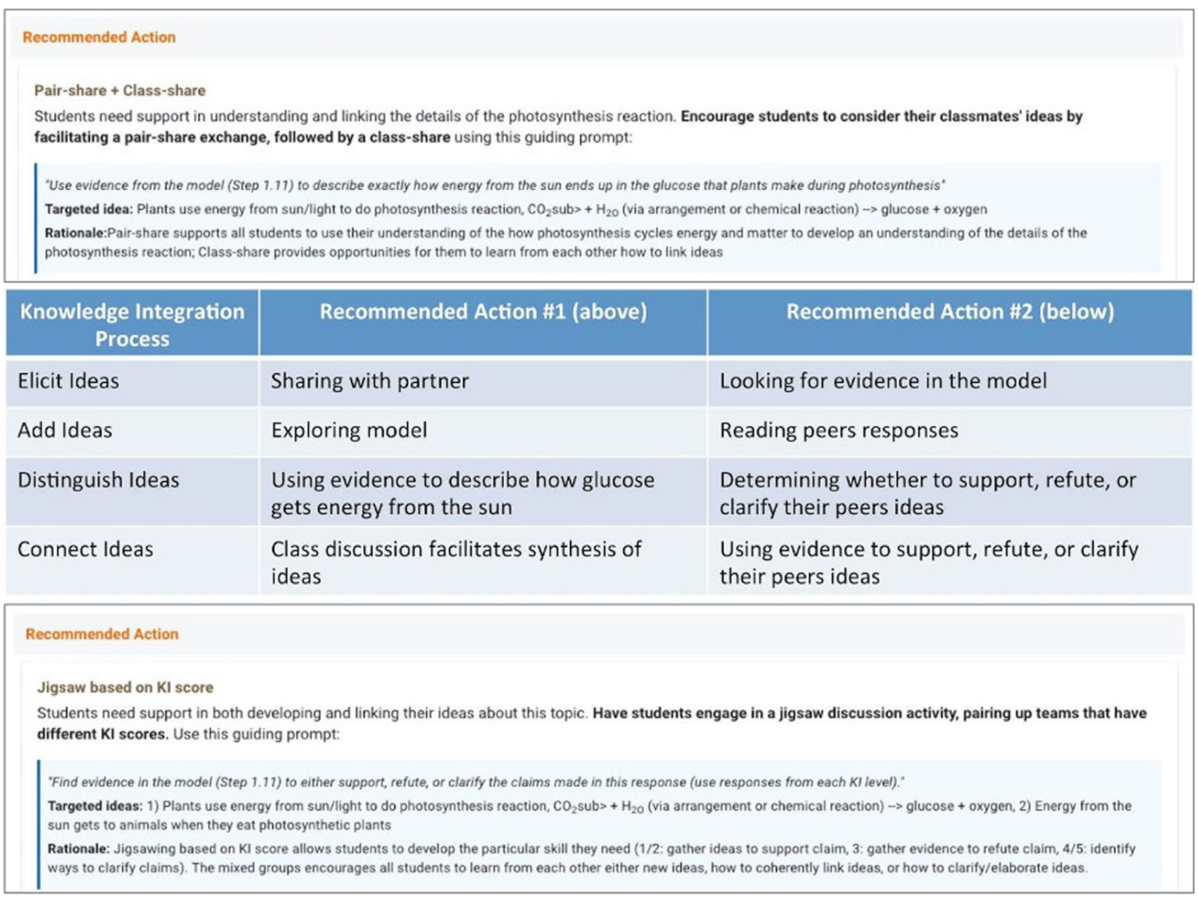


FIGURE 3 The knowledge integration alignment of TAP recommended actions received by teachers.

TAP availability

The TAP was located in the teacher interface of the learning platform and accessible to teachers via the Grading Tool feature after 75% of students completed the Initial Explanation item. It presented LA corresponding to class-aggregated data of students' autoscored responses to the assessment items and updated in real-time based on students' performance on the Reflection Explanation. The researchers and system developers set 75% as the threshold to ensure a sufficiently high sample size to yield a statistically valid aggregate data analysis while not delaying teacher action by forcing them to wait for 100% completion.

Design conjectures

Here, we describe our ideas (ie, our design conjectures; Sandoval, 2013) about how particular LA dashboard features would help teachers to provide their students with targeted learning support. A primary design conjecture was that linking LA to curriculum components aligned with the KI steps could provide teachers with insight into students' progress in the knowledge integration process through creating learning-congruent data rather than learning event data. In this way, KI-aligned LA could function as a "research" methodology for teachers (Reimann, 2016), helping to provide insight for supporting integrated, three-dimensional knowledge development. We further conjectured that using the KI framework to develop a LA dashboard would support teachers in taking pedagogical actions that help students develop an integrated, and three-dimensional understanding of targeted complex science concepts, which, for this study, was energy and matter transformation (see Figure 4 for a generalized illustration of these conjectures).

Participants

Three 7th grade science teachers in two different schools and districts as well as their 212, 7th grade students (ages 11–13) participated in this study. Teachers 1 and 3 had 20+ years of teaching experience, and Teacher 2 had less than one year of teaching experience and was in the process of getting his teaching licence. Although Teacher 3 had previous experience teaching with the lesson platform, unlike Teachers 1 and 2, the unit used in this study was new to all three teachers. All the teachers explicitly expressed a need for support in helping to assess their students' progress in developing integrated, three-dimensional science knowledge (ie, meeting the targeted curriculum standards). These teachers varied in terms of teaching experience, platform experience, length of time in the affiliated RPP, pedagogical knowledge and teaching practices. Their involvement in this design phase supported the refinement of the LA dashboard for contextual scalability.

Data sources and analysis

Researchers exported logged student data associated with the assessment items (ie, Initial/Reflection Explanation and pre/posttest item), which included automatically generated scores for students' assessment item responses. The automated scores were generated by a c-rater algorithm (Leacock & Chodorow, 2003) that was trained on 1000+ student responses that were human-scored using the curriculum standard rubric (see Table A1). The computer-human scoring agreement for this assessment item is approximately 70 + % for all three dimensions of the targeted curriculum standard (Riordan et al., 2020). To

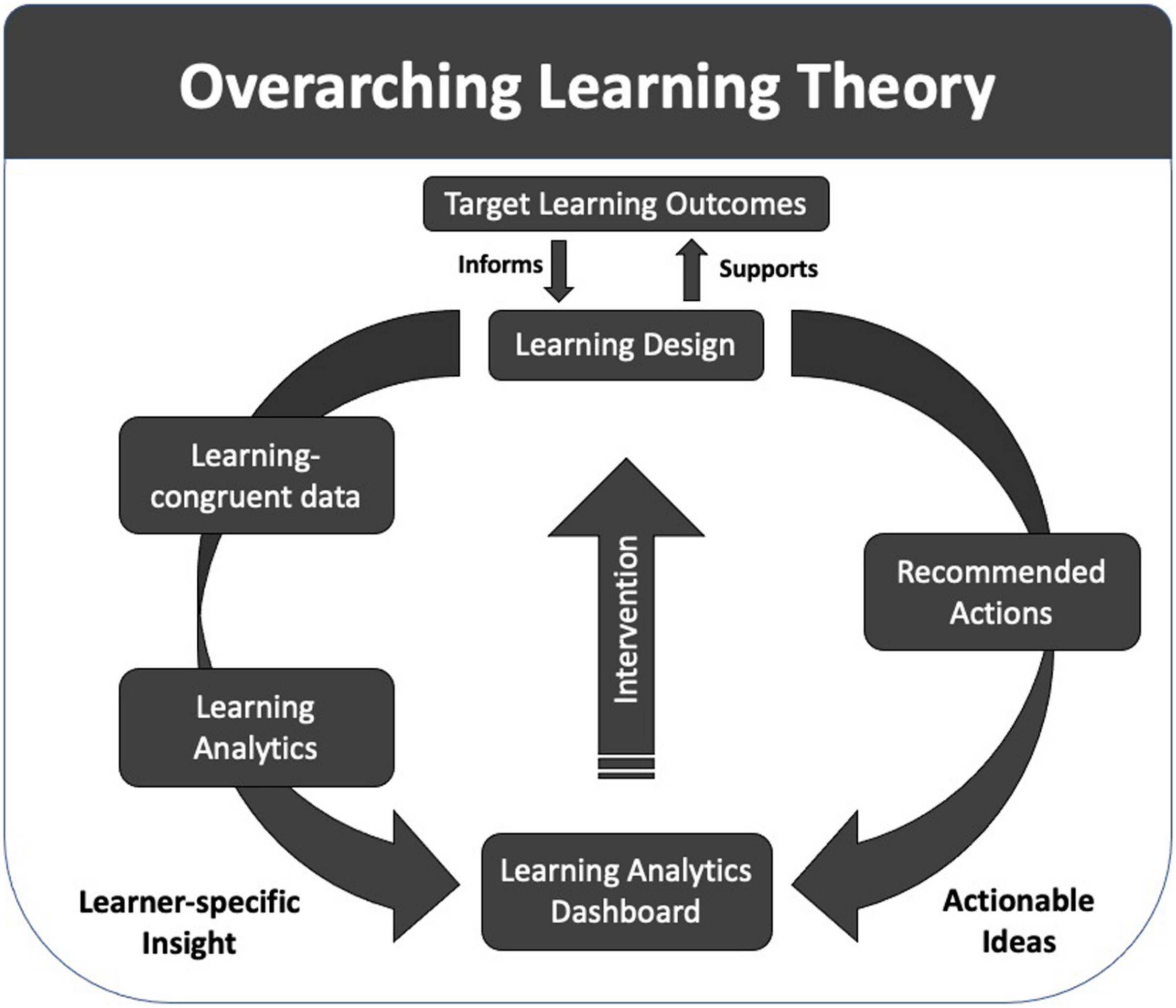


FIGURE 4 Generalized conjecture map that informed the TAP design.

determine whether students achieved learning gains after teachers' interventions, we performed McNemar's tests (Lachenbruch, 2014) on students' autoscored responses from the Initial to Reflection Explanation item. To determine if students experienced long-term learning gains, we conducted the same analysis on students' autoscores on the pretest and posttest item.

To understand teachers' pedagogy, we conducted semi-structured interviews (see Appendix B for the interview protocols) with each teacher before (pre-TAP) and after viewing the TAP generated from their class data (post-TAP). We also conducted classroom observations to document the pedagogical actions that teachers took after viewing data related to their students' performance on the Initial Explanation item. During the pre-TAP interview, we introduced the assessment items to teachers and solicited their feedback on the TAP features, using a sample TAP display. Also, during the pre-TAP interview, we solicited information from teachers regarding their typical assessment strategies and their expectations and interpretations of their students' performance on the embedded assessment item. The post-TAP interview was conducted either after the class period or after the school day. During the post-TAP interview, we asked teachers to describe their pedagogical actions in response to student data and provide a rationale for why they chose to implement them. Teachers were also asked how the TAP influenced their instructional intervention and subsequent instruction. The transcripts of these interviews were analysed for evidence of the teachers' pedagogy, specifically their beliefs and practices regarding teaching, student learning and assessments.

Interview coding and analysis

The transcripts of these interviews were analysed for evidence of the teachers' pedagogy, specifically their beliefs and practices regarding teaching, student learning and assessments, using the codes described in Table 3. We used these codes because we were interested in understanding the similarities and differences between teachers' pedagogy and the Knowledge Integration (KI)/constructivist pedagogy to determine relative alignment between teachers' interventions and the theory of learning (ie, constructivism-based KI process) that undergirded the HCLA solution. To mediate analysis of teachers' pedagogical practice in terms of its alignment to the theory-grounded pedagogical framework that informed the TAP design, namely KI, we applied all relevant KI codes to each coded instance of teachers' practice related to both student learning and assessment (ie, transcript segments coded

TABLE 3 Code descriptions and examples.

Code: Description (Cohen's kappa)	Example
Belief-teaching: held perspective related to teaching ($k=1.00$)	So I want them to make these connections instead of me guiding them per se … But, lay back and let them start to take ownership because moving seventh grade into eighth grade and high school, I believe that they have that ownership that they do not need some clown like me to show them
Belief-student learning: held perspective related to student learning ($k=0.86$)	Because when I tell them, it kind of gets lost. When they discover it on their own, I think it sticks better
Belief-assessments: held perspective related to assessing/assessments of student learning ($k=0.69$)	As long as you are trying to get it that's what I'm looking at. I look at the effort
Practice-student learning: specific behaviours/routines implemented when teaching related to student learning (0.80)	I passed it out and just talked about it a little bit. That's pretty much it
Practice-assessments: specific behaviours/routines implemented when teaching related to assessing student learning (0.83)	And then I'll check in with each group and they are going to be telling me their story
KI-elicit: teaching strategy/learning activity used to surface students' prior knowledge related to the target content/concept ($k=1.00$)	When I came around and asked them specifically how does the energy change, I'd have at least one kid say, "Oh, here's the kinetic and then chemical"
KI-add/discover: teaching strategy/learning activity used to increase students' knowledge base related to the target content/concept ($k=0.74$)	"Go back to the other pages and reference them and look at them"
KI-distinguish: teaching strategy/learning activity used to help students identify relevant ideas for understanding the target content/concept ($k=0.75$)	"Okay, yeah, that's right, but I want you to really think of, specifically, how does the energy and matter move through that system and how does it change?"
KI-connect/revise: teaching strategy/learning activity used to help students connect or revise their relevant ideas about the target content/concept ($k=0.67$)	They practice trying to say the science idea in a way that's making sense to them

with either a practice: student learning or practice: assessment code). Characterizing teachers' pedagogy as their beliefs and practices related to teaching and learning is consistent with the findings of numerous research studies (Mama & Hennessy, 2013; OECD, 2009; Pajares, 1992; Wladis et al., 2017).

To establish the reliability of our codes, 30% of the interview transcripts (which included both pre- and post-TAP interviews) were independently coded in five rounds by two coders, one of whom was the first author. At the end of each round, the coders reconciled all coding discrepancies to achieve at least 90% agreement. Prior to reconciliation during the final round, Cohen's kappa for interrater reliability was calculated for each code (Table 3). Once each code had a Cohen's kappa value greater than 0.61, corresponding to substantial agreement, the first author coded the remaining interview transcripts.

RESULTS

> RQ1 Finding 1—Giving teachers LA about their students' ideas with targeted recommendations for how to support students' specific learning needs is an effective motivator for diverse teachers to take pedagogical action.

During the pre-TAP interviews, we solicited teachers' feedback on the TAP features (ref. Figure 2). Teachers identified the TAP's alignment with an assessment targeted to a specific NGSS curriculum standard (*Description* section) as a key affordance. For example, Teacher 1 commented that having LA associated with a particular assessment item allowed her to redirect her time and energy towards extensively examining student performance on that item rather than giving cursory attention to all items. For Teachers 2 and 3, the TAP alignment with the targeted curriculum standard helped them feasibly determine whether their students were "getting it or not". Teacher 2 also identified the "Targeted Ideas" (*Recommended Action* section) as a valuable instructional tool as they would help him to specifically evaluate what information his students did and did not understand. Teacher 3 viewed the TAP as providing holistically valuable feedback consisting of "a real analysis of the data" (a reference to the score histograms in the *Class Report* section) that allowed him to support his students to engage in an important aspect of learning, namely revising their ideas. All teachers stated that the graphical presentation of the information facilitated their quick understanding of students' performance (*Class Report* section). In terms of the Key Insights and Recommended Action section, all teachers stated that they valued the recommended actions as they gave them specific ideas for how to address their students' learning needs. Teachers also offered suggestions for TAP redesign that might improve TAP implementation (Table 4).

To further assess whether and how the TAP supported teachers to understand their students' learning, teachers were also asked during the pre-TAP interview what expectations (ie, hypotheses) they had for students' learning, specifically in terms of their performance on the Initial Explanation item. All three teachers said that by the time students reached the Initial Explanation item they expected them to know the basics of photosynthesis since it was positioned at the end of instruction on photosynthesis and cellular respiration. Teacher 1 and 2's expectations were based on the instruction they provided their students prior to engaging the WISE unit. Even though Teacher 1 and 2 followed a similar lesson plan, what each teacher considered to be the basics differed. For Teacher 1, the basics meant knowing that plants take in carbon dioxide and water, and that they make glucose and oxygen. However, for Teacher 2 the basics meant knowing that animals get energy from plants, and plants get energy from the sun. Teacher 2 commented that he would expect students to struggle with the chemical reaction aspect of photosynthesis, recalling that his students had difficulty remembering the term "glucose". Although Teacher 3 did not articulate exactly what

TABLE 4 Alignment of Design Phase 2 feedback to TAP redesigns.

TAP feature	Feedback	Redesigns
Engagement metrics (proposed)	When asked about specific features that would be helpful for early career teachers, Teacher 2 said, "One thing that was kind of helpful was … It was on another curriculum…and they said how long [students] were on that page. The one who said, "I do not know it all," if I see they spent 10 seconds on the page, I do not even have to bother opening it, I know they did not … You can kind of address it then and be like, "All right, well I know you did not answer this very well, so let us …" That is one thing that I found kind of helpful"	These metrics are planned as a customization feature that users can turn on/off because previous RPP teachers, particularly more experienced teachers, said this would not be useful
Score descriptions	When asked about TAP features that felt extraneous, Teacher 2 identified the KI score description under the graphical display and said, "It's not really necessary, I do not think. It's pretty streamlined. I think it's not overwhelming with way too much information or anything. I think it's a good amount"	Score descriptions were placed in a hover box indicated by an information icon next to the graph subtitles
Recommended action	When asked to provide feedback on the recommended actions, Teachers 2 and 3 offered suggestions: Teacher 2 remarked, "Maybe the layout on this could be a little bit more easily accessible, because it's kind of a long sentence … so kind of the numbers broken up, because I was kind of looking, I'm like, "What is one two mean?" And then like, "Oh okay, so this is if you are in group one or two, you are going to do this. Maybe graphically it could just be a little more easily accessible" Teacher 3 said, "There definitely could be something like recommended action, and another link that you send the action to all of those in your group, all the kids right. Like how the old [platform] you could message them, message them out. So, "Check your messages first." Every day they get messages from me. So the first thing the kids go and do is check the messages, which are the prompts, if you will. So if you give them the chance to explore and then think that they, "I got this, I'm out, I'm on the next thing", thinking about playing the games later, or whatever it is. And then the next day you prompt them into going back and seeing and showing them the actual data and saying, this is where some of you are, this is where some of you aren't. And then give them that prompt. I think that lights a fire under them to want to improve it"	The recommendation actions were redesigned to reduce the text and provide a graphical representation of the activity structure. The whole-class comment feature from previous iterations of the platform were reinstated and provided as a recommendation action for teachers to use to give customized guidance to their students
Relevant auxiliary analytics (proposed)	In response to the idea of providing teachers with auxiliary data about students performance in the unit on questions/activities related to the target concept/content, such as maximum number of attempts data (similar to the data provided by the Phase 1 prototype, Teacher 2 said, "Yeah. I did notice on some of the earlier ones, it did say, "Attempt blank out of blank." And that's kind of good because I know some of my kids just select each one and do it until they get the right one. It's like, okay, did they really think about it and come to it, or did they just click until they got the right one? [Making] that information available, I think that would be helpful" When asked how the auxiliary LA data should be made available to teachers, Teacher 2 said, "Probably [at the] individual [student] level, and then maybe just something kind of out of the way, like a little footnote or something next to them. It does not have to be a huge (like a pop-out screen) … It's not the main thing, but if I wanted to look at it, just a quick footnote like, "Oh okay." You get a little better idea just quickly looking." He also suggested having an expand/collapse feature for the auxiliary data, stating, "I think because it can get hard to have to open, click every single thing open. It's nice, like the expand all on this, if I just want to go through and read, it's nice that it'll come through and expand"	The suggested redesigns are planned, but not yet implemented, as it will require a more extensive redesign of the platform

TABLE 4 Continued

TAP feature	Feedback	Redesigns
Rubric access	In regard to the rubric with criteria for the score levels (which was accessible via the gradebook, ie, outside the TAP), Teachers 2 and 3 provided the following suggestions: Teacher 2 said, "Maybe another link to the full metric page, like just a quick link so I could pull that up and then close it. That would probably be more helpful than this part down here (ie, the score descriptions under the graphs). Yeah, if it was easily available here … Because I have a general idea of what these mean, but if I wanted to look at the specific metrics it was rating it at, that would be a little more helpful" Teacher 3 shared similar suggestions, "Just a button that re-links it to that link. Right. Just put a little button, rubric button … Right now it's like the rubric is in one place. This is in another place. So maybe putting them together"	These suggestions were followed almost exactly, with a link to the rubric provided at the top of the TAP, near the graphs and information icon added near students' responses
Sample response	When asked about features that were not helpful, Teacher 2 identified the sample student response. Relatedly, Teacher 3 offered suggestions for how to make the sample responses helpful, namely to provide them to students in conjunction with the rubric rather than to teachers	The sample responses were removed from the TAP and a note was added with the rubric to let teachers know that if they saw fit they could provide their students with the rubric
Student scores	In regard to accessing individual student scores (which was available in the Teacher Gradebook section of the platform not the TAP), Teacher 2 said, "Having the whole grade under here (i.e. the Student Response section), just accessible there, would be helpful as well." He also suggested a student score/response grouping feature, saying "If you could tag [the student response] like, just needs everything, or … And then you could kind of have tagged groups. Like in Gmail, you have tags for emails that you can click on the tag and you could go and it'd be a group of emails like, oh, all my goals are … you know? I could see that being helpful. Okay, these are the kids who need to figure out the energy path, or these are the kids who need something else. And then just to be able to see the group of kids in sort of a broken down way. That would be pretty helpful"	Individual student scores were added alongside their response in the Student Response section, along with a sort feature for the score column. The re-activated whole-class comment feature (see the Recommended Action redesign) permits the teacher to assign comments based on score
Unit milestone	When discussing the contextual information provided at the top of the TAP, particularly the details of the milestone item, Teachers 2 noted that it would be particularly helpful if it were in closer proximity to the student responses and as a clickable link, saying, "I think that would be more relevant if I could access the questions here, because sometimes it's nice to have a reminder for what exactly did they ask them. It probably would not hurt to have a link to it … I could see having an easy link back to that on more detailed [milestone items] would be very helpful" When specially asked which phase in that learning process would it be most helpful for teachers to get TAP data, Teacher 3 recommended the following, "Middle, yeah. Because I want to figure out what we are struggling with at that point. So how fast can I move ahead versus going back and cycling other topics. And so obviously middle is really important"	The specific milestone text (eg, Step 2.10) was converted to a link that when clicked navigated the user to the actual location in the unit. A design rule was established that milestone items would be strategically situated in/developed for the middle of the unit learning experience

knowledge represented the basics, his expectation was that students would know the basics related to photosynthesis from their prior elementary school instruction.

> RQ1 Finding 2—Teachers enacted familiar practices and implemented LA products in keeping with their teaching philosophy.

We analysed the post-TAP interview transcripts for evidence of belief statements that reflect their teaching pedagogy, specifically, their perspective on teaching, student learning and assessment. With this analysis, we sought to contextualize teachers' pedagogical actions in relation to their teaching pedagogy to determine whether and how the TAP influenced it. The three teachers had different perspectives on teaching, student learning and assessment.

Teacher 1's pedagogy

Teacher 1 noted that her perspectives on student learning were developed from over 20 years of teaching experience. She believed that students "need more than just seeing stuff on a screen, they need to move stuff around and sort of make sense out of it". To foster more sharing of ideas among her students and to informally assess their ability to appropriately use science vocabulary, she developed a large repertoire of what she termed "kinaesthetic-verbal" activities, which require students to physically manipulate objects while simultaneously engaging in small group discourse to explain science ideas. She noted that these types of activities create space for students to productively struggle with, develop and articulate their understanding of the content in ways that suit them. She identified these kinaesthetic-verbal activities as necessary precursors for students being able to articulate their ideas in writing and summarized the process as "work-talk-revise".

Teacher 2's pedagogy

Similar to Teacher 1, Teacher 2 identified verbal explanations as a valuable means to gauge student learning. However, in contrast to Teacher 1, Teacher 2, who only had 3 months of teaching experience, viewed one-on-one, teacher-student conversations as the most effective assessment strategy rather than small group, student–student exchanges. Rather than relying only on written demonstrations of knowledge, he stated, "… if I sit with them or come over to them one-on-one and kind of talk it through with them, they usually reveal more of what they know." He viewed students' written responses as representing the ideas they were "confident enough to put down" and that this would increase with time and exposure. He expressed a belief that students truly understood a science idea if they can reproduce the correct answer several days later, which is also why he timed assessments to allow for information to "sink in a little longer".

Teacher 3's pedagogy

Like Teacher 1, Teacher 3 developed his perspective on learning and assessment over his 20+ years of teaching experience. He viewed learning as dictated by interest and intrinsic intellectual capacity, the former being the only factor that he, as the teacher, could influence. Teacher 3 believed that his actions as a teacher were more effective when they focused on inspiring and encouraging students rather than providing them with remedial support. He commented, "My job is to hook 'em. Hook 'em for life, right? Science for life." For the students that he recognized as having the intellectual capacity to learn, he saw his job as providing them with the skills to learn on their own and he expressed a view that assessments are a means of giving students credit for engaging the learning process, whether it be participation in the "hook 'em" activities or making use of their intellectual ability to critically think about and express science ideas. Similar to Teachers 1 and 2, Teacher 3 expressed a belief that learning took place through conversation and that regardless of whether students get assistance to complete formal assessments, like homework assignments and tests, "that, during that process, at least conversations are happening." Summarizing his perspective on assessment, Teacher 3 stated, "I'm not going to spend all my energy on assessment when I want to just give them the opportunity to explore and learn a lot."

When the pre-TAP transcripts were analysed to determine the alignment between teachers' expressed pedagogical practices and the KI process (Table 5), we found that Teachers 1 and 3's pedagogical practice was primarily characterized by supporting students to add/discover and connect/revise ideas. Consistent with his limited teaching experience, Teacher 2's KI practice was emerging, with the least evidence of supporting students to connect and revise ideas. All teachers expressed moderate engagement in the KI practice of supporting students to distinguish ideas, which is consistent with prior research on teachers' pedagogical challenges (Vitale et al., 2016; Wiley et al., 2019).

During the post-TAP interview, we asked teachers to share their interpretation of the LA presented. Teachers identified score levels of 4 and 5 as the most desirable in terms of student performance, treating the graphs as binary data (ie, either above a score of 3 or not). Teachers 1 and 2 expressed a need to read through student responses in the Grading Tool to gain a richer understanding of student thinking. For Teacher 3, on the other hand, the graphical data in the TAP provided him with a sufficient level of understanding regarding his students' learning to immediately implement the recommended instructional intervention.

During the classroom observations, we observed that all teachers took pedagogical actions in the form of instructional interventions in response to the TAP, albeit with varying degrees of alignment to the KI process, which the unit's learning design supported (Figure 5). Teacher 1 and 2 created their own interventions, while Teacher 3 relied upon the recommended action in the TAP. Each teacher's intervention is detailed below.

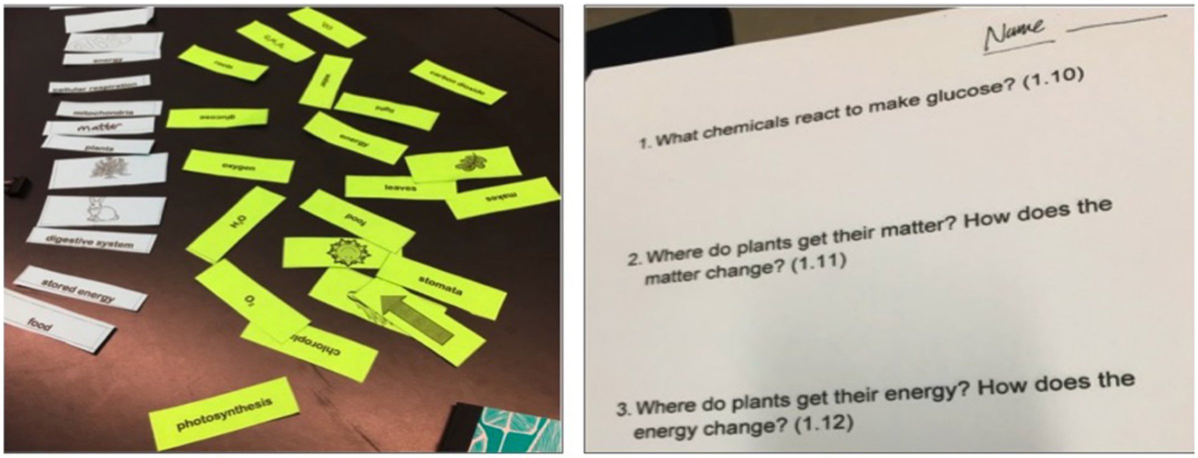
Teacher 1's pedagogical action

After reading through her students' responses to the Initial Explanation item, Teacher 1 noticed that, while many students discussed photosynthesis, few students discussed cellular respiration. In response to her students' performance, Teacher 1 decided to engage her students in a revised version of a storyboard activity they had previously completed. The storyboard activity is an example of what she termed a kinaesthetic-verbal activity, and her decision to use this activity reflects her perspective that learning is best achieved when

TABLE 5 Frequency of KI codes associated with teachers' pedagogical practices.

KI practice	Teacher 1	Teacher 2	Teacher 3
Elicit ideas	4	6	2
Add/discover ideas	19	5	16
Distinguish ideas	8	6	6
Connect/revise ideas	18	2	10

Note: Colors are gradated to reflect frequency amounts, with darker gradations reflecting higher frequency. <6 = red grading, 6 = white, >6 = green grading.



Knowledge Integration Process	Teacher 1 Storyboard Activity	Teacher 2 Worksheet Activity	Teacher 3 Peer-to-Peer Sharing
Elicit Ideas	Connecting the terms to known ideas	---------------------	Sharing responses
Add Ideas	Exploring ideas related to unknown terms	Looking for correct answer on specified unit step	Listening to responses
Distinguish Ideas	Determining which ideas to include in the story	---------------------	---------------------
Connect Ideas	Connecting the ideas to make a coherent and accurate story	Revising response to include ideas from worksheet	Revising response to include peers' ideas until satisfied with score

FIGURE 5 Knowledge integration alignment of teachers' interventions. Top left image: Teacher 1's storyboard activity materials. Top right image: A portion of Teacher 2's worksheet activity.

students do and talk while constructing their understanding. Students worked in groups (3–4 students) to organize strips of paper–some with words and images related to photosynthesis and cellular respiration and some blank, for them to use as needed—into a coherent narrative (Figure 5, top left image). Variation between groups was permitted by Teacher 1 and reflected her perspective that in science there are multiple ways of expressing the same ideas. The groups vocalized their narratives to Teacher 1, who provided feedback regarding its thoroughness in relation to the target ideas. Teacher 1 used this activity as a formative assessment to determine whether students were able to coherently connect their ideas, which was the learning goal. Afterwards, the groups completed the Reflection Explanation item.

Teacher 2's pedagogical action

After viewing the TAP report, Teacher 2 read all of his students' responses and constructed a worksheet of short-answer questions related to the assessment prompt (Figure 5, top right image), rather than implementing the recommended intervention. His intervention worksheet called for students to revisit specific steps in the WISE photosynthesis unit (the location of which was indicated alongside the question), an approach that aligned with his perspective that repeated exposure to the target information leads to learning. Students were expected to provide the "right answer" to the worksheet questions. He also engaged students one-on-one to support them in answering the questions and responded to students' clarifying questions by rephrasing the worksheet questions. Before the period ended, students were instructed to complete the Reflection Explanation item. Regarding his students' ability to engage in the activity, Teacher 2 stated that, in his estimation, few of his students had developed the skill “to go back and reference material,” and consequently were not going to benefit from his worksheet activity.

Teacher 3's pedagogical action

After viewing the TAP report, Teacher 3 read it to his class. He described the number of students at each score level and the scale for each score, read the Key Insights and Recommended Action sections to the class, and instructed them to complete the Reflection Explanation item. Per Teacher 3's request, the settings of the WISE platform were adjusted so that the students could see live updates to their autogenerated KI score as they worked on the Reflection Explanation item, which had the effect of gamifying the revision process. Teacher 3 expressed a belief that such motivation would encourage them “to cognitively understand what's going on” in terms of the science content. During the revision process, students queried each other regarding their scores and referenced each other's responses during conversations about how to improve their score. Teacher 3 was aware of this behaviour and identified it as valuable for learning, stating that it functioned to provide students with an “end goal”, an exemplar of how to construct their responses. Although this implementation of the recommended action was not completely consistent with the unit's design intention, it was consistent with Teacher 3's perspective on learning, in that it is incumbent upon students to facilitate the learning process themselves. Moreover, his approach to implementing the recommended action was consistent with his relaxed approach to assessment.

> RQ2 Finding—Students achieved targeted learning outcomes both in the short and longer-term.

Student Learning from Initial to Reflection. McNemar's test results revealed that for all three teachers' and for all three dimensions of the targeted curriculum standard, there was a statistically significant increase in the proportion of students expressing ideas that aligned with consensus science knowledge about photosynthesis and cellular respiration after the instructional intervention (Tables 6–8). Students improved the most along the dimension that their teachers' interventions focused on, which the TAP identified as the learning need.

Student Learning from Pretest to Posttest. We also conducted McNemar's tests on students' scores from the pretest and posttest item for each dimension to evaluate whether students retained the Initial to Reflection Explanation learning gains until the completion of the unit (Tables 9 and 10). Teacher 2's students did not complete the posttest, so their scores were excluded from this analysis. Of the students who completed both the pretest and posttest, 55 (56%) in Teacher 1's class and 18 (82%) in Teacher 1's class, only Teacher 1's students had statistically significant ($p < 0.05$) shifts in scores along the three dimensions (Table 9), although it is notable that most of Teacher 3's students began the unit expressing more normative and target ideas for all three dimensions (Table 10) than did Teacher 1's students.

In the next section, we provide a narrative account of the findings that provide an answer to each expansion need and the corresponding RQ, together with a more detailed description and discussion of the relevance of this finding for the HCLA field. Table 11 below provides a summary of the three HCLA expansion needs, together with the associated key findings of the study and relevant design features.

TABLE 6 Frequency of Teacher 1's students in each performance category of the Initial/Reflection assessment item, by teacher and learning dimension.

Teacher 1 initial	Reflection		
DCI $\chi^2(1)=6.00$[a] $p_{exact}=0.0312$*	Not present	Present	**Initial total**
Not present	64 (79%)	6 (7%)	70 (86%)
Present	0 (0%)	11 (14%)	**11 (14%)**
Reflection total	64 (69%)	**17 (21%)**	81 (100%)
CCC $\chi^2(1)=6.23$[a] $p_{exact}=0.0225$*	Not present	Present	**Initial total**
Not Present	45 (56%)	11 (14%)	56 (69%)
Present	2 (2%)	23 (28%)	**25 (31%)**
Reflection total	47 (58%)	**34 (42%)**	81 (100%)
SEP/KI $\chi^2(1)=6.40$[a] $p_{exact}=0.0215$*	Non-normative	Normative	**Initial total**
Non-normative	58 (72%)	9 (11%)	67 (83%)
Normative	1 (1%)	13 (16%)	**14 (17%)**
Reflection total	59 (73%)	**22 (27%)**	81 (100%)

Note: Colours are gradated to reflect frequency amounts, with darker gradations reflecting higher frequency. Percentage totals for the target performance category at the Initial and Reflection time points are bolded.

[a]McNemar's Chi-squared and exact significance probability.

*$p<0.05$

TABLE 7 Frequency of Teacher 2's students in each performance category of the Initial/Reflection assessment item, by teacher and learning dimension.

Teacher 2 initial	Reflection		
DCI $\chi^2(1)=7.00$[a] $p_{exact}=0.0156$*	Not present	Present	**Initial total**
Not present	70 (89%)	7 (9%)	77 (97%)
Present	0 (0%)	2 (3%)	**2 (3%)**
Reflection total	70 (89%)	**9 (11%)**	79 (100%)
CCC $\chi^2(1)=9.31$[a] $p_{exact}=0.0034$**	Not present	Present	**Initial total**
Not present	55 (70%)	12 (15%)	67 (85%)
Present	1 (1%)	11 (14%)	**12 (15%)**
Reflection total	56 (71%)	**23 (29%)**	79 (100%)
SEP/KI $\chi^2(1)=6.00$[a] $p_{exact}=0.0312$*	Unlinked	Linked	**Initial Total**
Unlinked	69 (87%)	6 (8%)	75 (95%)
Linked	0 (0%)	4 (5%)	**4 (5%)**
Reflection total	69 (87%)	**10 (13%)**	79 (100%)

Note: Colours are gradated to reflect frequency amounts, with darker gradations reflecting higher frequency. Percentage totals for the target performance category at the Initial and Reflection time points are bolded.

[a]McNemar's Chi-squared and exact significance probability.

*$p<0.05$; **$p<0.01$.

TABLE 8 Frequency of Teacher 3's students in each performance category of the Initial/Reflection assessment item, by teacher and learning dimension.

Teacher 3 initial	Reflection		
DCI $\chi^2(1)=8.00$[a] $p_{exact}=0.0078$**	Not present	Present	**Initial total**
Not present	11 (52%)	8 (38%)	19 (90%)
Present	0 (0%)	2 (10%)	**2 (10%)**
Reflection total	11 (52%)	**10 (48%)**	21 (100%)
CCC $\chi^2(1)=8.00$[a] $p_{exact}=0.0078$**	Not present	Present	**Initial total**
Not present	4 (19%)	8 (38%)	12 (57%)
Present	0 (0%)	9 (43%)	**9 (43%)**
Reflection total	4 (19%)	**17 (81%)**	21 (100%)
SEP/KI $\chi^2(1)=7.00$[a] $p_{exact}=0.0156$*	Unlinked	Linked	**Initial total**
Unlinked	10 (48%)	7 (33%)	17 (81%)
Linked	0 (0%)	4 (19%)	**4 (19%)**
Reflection total	10 (48%)	**11 (52%)**	21 (100%)

Note: Colours are gradated to reflect frequency amounts, with darker gradations reflecting higher frequency. Percentage totals for the target performance category at the Initial and Reflection time points are bolded.

[a]McNemar's Chi-squared and exact significance probability.

*$p<0.05$; **$p<0.01$.

TABLE 9 Frequency of Teacher 1's students in each performance category of the Pre/Posttest assessment item, by teacher and learning dimension.

Teacher 1 pretest	Posttest		
DCI $\chi^2(1)=7.36$[a] $p_{exact}=0.0117$*	Not present	Present	**Initial total**
Not present	42 (79%)	10 (19%)	52 (98%)
Present	1 (2%)	0 (0%)	**1 (2%)**
Reflection total	43 (81%)	**10 (19%)**	53 (100%)
CCC $\chi^2(1)=6.25$[a] $p_{exact}=0.0213$*	Not present	Present	**Initial total**
Not present	29 (53%)	13 (25%)	41 (77%)
Present	3 (6%)	9 (17%)	**12 (23%)**
Reflection total	31 (58%)	**22 (42%)**	53 (100%)
SEP/KI $\chi^2(1)=10.29$[a] $p_{exact}=0.0018$**	Non-normative	Normative	**Initial total**
Non-normative	33 (62%)	13 (25%)	46 (87%)
Normative	1 (2%)	6 (11%)	**7 (13%)**
Reflection total	34 (64%)	**19 (36%)**	53 (100%)

Note: Colours are gradated to reflect frequency amounts, with darker gradations reflecting higher frequency. Percentage totals for the target performance category at the Initial and Reflection time points are bolded.

[a]McNemar's Chi-squared and exact significance probability.

*$p<0.05$; **$p<0.01$.

TABLE 10 Frequency of Teacher 3's students in each performance category of the Pre/Posttest assessment item, by teacher and learning dimension.

Teacher 3 pretest	Posttest		
DCI $\chi^2(1)=1.00$[a] $p_{exact}=0.6250$	Not present	Present	**Initial total**
Not present	6 (33%)	3 (17%)	9 (50%)
Present	1 (6%)	8 (44%)	**9 (50%)**
Reflection total	7 (39%)	**11 (61%)**	18 (100%)
CCC $\chi^2(1)=1.00$[a] $p_{exact}=0.6250$	Not present	Present	**Initial total**
Not present	2 (11%)	3 (17%)	5 (28%)
Present	1 (6%)	12 (67%)	**13 (72%)**
Reflection total	3 (17%)	**15 (83%)**	18 (100%)
SEP/KI $\chi^2(1)=1.00$[a] $p_{exact}>0.9999$	Non-normative	Normative	**Initial total**
Non-normative	4 (22%)	1 (6%)	5 (28%)
Normative	0 (0%)	13 (72%)	**13 (72%)**
Reflection total	4 (22%)	**14 (78%)**	18 (100%)

Note: Colours are gradated to reflect frequency amounts, with darker gradations reflecting higher frequency. Percentage totals for the target performance category at the Initial and Reflection time points are bolded.
[a]McNemar's Chi-squared and exact significance probability.

TABLE 11 Alignment of expansion need, research questions, findings and design process features.

HCLA expansion need (aligned RQ)	Key finding	Relevant design feature
Generate LA that are effective in supporting the achievement of the desired learning outcomes across diverse users and contexts (ie, contextually scalable) (RQ1)	Co-designing with a longstanding multi-stakeholder, multidisciplinary design team led to an HCLA dashboard that effectively supported diverse teachers to support student learning	DBR: multidisciplinary, multi-stakeholder design team; Longstanding K-12 EdTech RPP
Position usability and feasibility as starting rather than endpoint metrics for LA effectiveness (RQ2)	Implementing EDS principles allowed teachers to attend to student learning	EDS Principles
	Incorporating the evaluation of student learning guided the development of an effective, HCLA solution	Longstanding K-12 EdTech RPP
Utilize a relevant learning theory to align the LA with the desired learning outcomes—a theory that accounts for the social aspects of learning in addition to individual student behaviours (RQ2)	Aligning the LA data and recommended actions with a learning theory positioned teachers as classroom learning researchers	DBR: theory-grounding; DBR: multidisciplinary, multi-stakeholder design team

DISCUSSION

In this section, we present a structured discussion around the two objectives of the paper focus (to both propose and test viability of an HCLA design approach), the two RQs, the three addressed expansion needs with respect to the current state of HCLA, and finally the design components of the proposed approach and findings of the study.

We argue that although HCLA solutions have effectively attended to the human needs within specific contexts, namely the classrooms of the initial design teachers, these products likely have limited use beyond those spaces (Li, 2021). Moreover, most HCLA design approaches culminate with the measurement of usability and feasibility rather than expanding the test of HCLA solution viability to include an evaluation of the impact on student learning (Mangaroska & Giannakos, 2018), which is arguably the ultimate human-centred education goal. To achieve contextual scalability and attend to student learning outcomes, we argue that HCLA need to be designed by a multidisciplinary, multi-stakeholder team that knows the educational contexts well and designed in a way that will centre the immediate needs of teachers and the ultimate needs of students. We hold that DBR, carried out within the context of an RPP and in conjunction with established HCLA design principles, like EDS, accomplishes this, as demonstrated by our findings.

> RQ1—In what ways can a DBR-based HCLA design process enable LA implementation within K-12 classrooms to *support teachers with varying teaching backgrounds, practices and perspectives in effective implementation*?

Addressed Expansion Need: Generate LA that are effective in supporting the achievement of the desired learning outcomes across diverse users and contexts (ie, contextually scalable).

Co-designing with a longstanding multi-stakeholder, multidisciplinary design team led to an HCLA dashboard that effectively supported diverse teachers to support student learning

HCLA designs centre the needs of the end-user by bringing them into the design process at the early stages of development (Revano & Garcia, 2021). In this way, the resulting LA solution can be *designed with* rather than *designed for* the end users. Common strategies for including the end user, particularly teachers, in the design process is to develop an LA solution to address the needs of the particular teacher involved (Dimitriadis et al., 2021). While this design strategy of making custom HCLA solutions may be sufficient for HigherEd teachers, who have the professional freedom to create "tailor-made" courses, a different approach is needed for the K-12 context. Since a key characteristic of K-12 education are courses developed based on a standards-based curriculum that are taught by teachers with varying backgrounds and perspectives, it is critical that an HCLA design approach for K-12 LA solution be of broad use and accommodate implementation variation without compromising student learning outcomes. To meet this need, we leveraged a longstanding K-12 EdTech RPP to build a multidisciplinary, multi-stakeholder design team that together held the necessary contextual and technical expertise. With teacher co-designers engaged throughout the entire design process, their needs as end users and those of their students (as the ultimate key stakeholder) were foregrounded and prioritized. A key benefit of designing within a longstanding K-12 EdTech RPP is the diverse pool of teachers from which to draw co-designers and participants. The fact that they are all from the same RPP ensures a shared problem of practice that the HCLA solution aims to address. Having practitioner engagement at the school- and/or district-level ensures variation in the professional and classroom contexts of the RPP teachers.

The finding that the LA dashboard supported teachers across contexts can be attributed to the use of DBR, specifically the work of a design team that was part of a longstanding RPP. The longevity of the partnership and inherent flux of membership resulted in an agile design team and process, which is important for the dynamic nature of the K-12 context.

Given that the specific expertise needed within each category of team member (researchers, software developers, teachers) varies throughout the design process, it is important that the involvement of different individuals, particularly different teachers, also varies. In this way, the composition of the design team can shift across the different phases of the process. The first phase is the design of a Minimally Viable Product (MVP) that will achieve the targeted learning outcomes. For the dashboard described here, this corresponds to the first round of design (Wiley et al., 2020). In this phase, the design team should be composed of members who thoroughly understand the major aspects of the problem space and implementation context, making contributions to the initial design based on their respective expertise. Initial testing of the MVP to verify functionality is critical. In the final phase of transition from MVP to wide-use product (the phase described in this study), the design team, which would include teachers from different schools/districts and with different pedagogical practices than those in previous design phases provides insight regarding how to refine and validate the functionality of the product. Intentionally involving teachers with key differences in complex teacher characteristics (eg, pedagogical knowledge and professional routines; van Leeuwen et al., 2021) across the various design iterations allowed us to design an LA dashboard that met diverse teachers' implementation needs. Including teachers with different perspectives and practices allowed us to see the important areas for teacher individuality/customization and the areas where additional support is needed to bring teachers' intervention designs into alignment with the underlying learning theory.

> RQ2: In what ways can a DBR-based HCLA design process enable LA implementation within K-12 classrooms to *support students in developing integrated, multidimensional content knowledge*?

Addressed Expansion Need: Position usability and feasibility as starting rather than endpoint metrics for LA effectiveness.

Implementing EDS principles allowed teachers to attend to student learning

Although usability and feasibility remain critical measures in determining the viability of an HCLA solution, in the K-12 context, where the stakes regarding student success are extremely high, these measures must be starting rather than ending points. As such, HCLA in the K-12 context must always centre the educational needs of students, whether they are the end user of the LA tool or not. EDS principles help to ensure usability and feasibility while centring students' educational needs (Echeverria, Martinez-Maldonaldo, Granda, et al. (2018)). Having been designed in accordance with EDS principles, the HCLA solution in this study (ie, the TAP) supported teachers to quickly assess students' learning needs and turn their attention to meeting those needs.

Specifically, providing teachers with subscores related to students' learning progress along each dimension of the curriculum standard supported them in taking evidence-based pedagogical actions, such as customizing the learning design, albeit in accordance with their pedagogical practices and perspectives (Campos et al., 2021). The three dimension scores presented graphically (quantitative data) and summarized in the Key Insights (qualitative data) motivated teachers' inquiry into student thinking (Mor et al., 2015), examining the relationship between students' integrated understanding and their knowledge of particular ideas and concepts targeted by the standards-aligned learning design. Dominant LA dashboards follow an exploratory visual analytics approach, present teachers with the data analysis and leave them unassisted in exploring the data to determine whether and how they can inform the pedagogical action needed to support student learning (Echeverria, Martinez-Maldonado, Buckingham Shum, et al., 2018; Echeverria, Martinez-Maldonaldo, Granda, et al., 2018). However, since the LA dashboard was designed following an explanatory approach using the EDS principles it was able to guide teachers, regardless of their data literacy skills, in not only how to make sense of the analysis but also how to take focused action to address student learning needs.

Incorporating the evaluation of student learning guided the development of an effective HCLA solution

As previously mentioned, usability and feasibility are the typical end measures for HCLA effectiveness. Rarely, if at all, is the effectiveness of the HCLA solution to support student learning measured, even in the K-12 context (Mangaroska & Giannakos, 2018). By measuring the impact of the HCLA solution across each phase of development, the needs of students can remain front and centre, even when they are not official members of the design team. Given the goal of all LA to support the understanding and optimization of learning and its context (Ferguson, 2012), it is especially important that K-12 HCLA design processes factor in the learning outcomes of students, the ultimate stakeholder in the K-12 context.

Gašević and colleagues contend that "the true test for learning analytics is demonstrating a longer term impact on student learning and teaching practice" (Gašević et al., 2015, p. 6). Our finding that, despite the differences among teachers' interventions, students of all three teachers had statistically significant improvement in scores across all three learning dimensions points to the effectiveness of designing HCLA using the described approach. It also suggests that, in terms of short-term learning, perhaps the most salient factor for supporting students is targeted intervention. This finding also highlights the value of creating LA from data that reflect students' *developing knowledge* rather than students' *behaviour* on the learning platform (ie, learning events), as it enabled teachers to implement interventions that address specific learning needs, something that learning event data are insufficient to do. Additionally, the finding that the LA motivated teachers to further inquire into student thinking by reading their students' written responses, suggests that facilitating this inquiry directly through the LA dashboard may be a valuable redesign.

Addressed Expansion Need: Utilize a relevant learning theory to align the LA with the desired learning outcomes—a theory that accounts for the social aspects of learning in addition to individual student behaviours.

Aligning the LA data and recommended actions with a learning theory positioned teachers as classroom learning researchers

In the K-12 context, a central responsibility (ie, need) of teachers is to assess and support their students in meeting the targeted curriculum standards. Thus, teacher-facing HCLA in K-12 should support teachers in meeting this need. However, if the HCLA solution does not provide teachers with relevant data on how students' learning is progressing (ie, changes in student knowledge rather than just changes in student activity), then perhaps this important need will go unmet. While principles like EDS can foreground the educational goals in the HCLA, they cannot change the relevance of the data nor overcome the limitations of data analysis that does not give targeted insight into students' learning needs. Only when the data used to create the HCLA are aligned to a theory of learning that is consistent with the targeted student learning outcomes can teachers' central need to assess and support student learning be met. When K-12 teachers are on the HCLA design team, they can help to identify relevant data and provide insight on the pedagogical actions that will be effective in classroom learning environments. However, most teachers do not have extensive training in learning theories, making it even more important that learning sciences researchers are included on the design team. In this way, the learning needs of humans in general and students in particular can be simultaneously centred. In this study, the socioconstructivism-based framework, Knowledge Integration, functioned to guide data selection and analysis and to inform the development of recommended actions so that the HCLA solution could provide teachers with the requisite insight into changes in student knowledge and thus help them to support students in achieving the targeted learning outcomes.

Given that the primary goal for LA in supporting classroom instruction is to provide teachers with actionable insight (Jørnø & Gynther, 2018), teachers must be positioned to take action toward optimizing student learning, which often entails customizing the learning design. The finding that students achieved long-term learning gains when their teacher made instructional interventions that aligned with the underlying pedagogical framework suggests that providing theory-aligned recommended actions in the LA dashboard is worthwhile, contrary to some perspectives on how best to support LA implementation (Wise & Vytasek, 2017).

Grounding all aspects of the LA and dashboards in a contemporary learning theory (eg, socioconstructivism) functionally positioned teachers as learning researchers, with the LA dashboard functioning as their methodological tool for exploring student learning. In conjunction with prior understanding of their students' needs, teachers queried the LA to gain real-time insight into their students' thinking. They developed hypotheses about their students' performance on the platform content, used the LA data to refine their theories and design an intervention, and revisited the LA dashboard to evaluate the need for additional interventions. Teachers' interventions were based both on the insight mediated by the LA dashboard and on their own pedagogy, which may or may not have aligned with the guiding learning theory. This finding highlights the need to develop LA (and LA dashboards) that support teachers in taking up ideas based on contemporary learning theory (ie, learning sciences research), especially when those ideas run counter to their pedagogy.

CONCLUSION

The findings in this study indicate that a DBR-based HCLA design process that leverages a long-term, yet fluid, RPP can effectively enable the development of an LA dashboard capable of supporting K-12 teachers with varying teaching backgrounds, practices and perspectives to develop and implement LA-informed interventions that help students meet targeted curriculum standards. Developing the teacher dashboard in accordance with EDS principles allowed us to simultaneously attend to human needs related to usability, feasibility, teaching and learning. In particular, the use of learning-congruent data that aligned with a contemporary theory of learning facilitated the repositioning of teachers as classroom researchers, capable of understanding and addressing their students' learning needs.

The aim of this study is to demonstrate the viability of a DBR-based HCLA approach that holds promise for designing LA solutions within the complexities of K-12 education. The design *approach* rather than the design solution is being proposed for use and further research by others. The particular design solution described in this study has limited use as it was designed to support students in meeting US science standards, standards that are not in use in all 50 US states and involved a rather small set of teachers from an RPP. However, we demonstrate that the design approach is one that can give rise to the type of LA solutions that can be effective within the K-12 context. The particular learning standards that are targeted are less relevant as demonstrated by the fact that an entirely different science topic and standards were used during the first design cycle. The learnings about the type of assessment that is useful for LA development was carried over to the development of the particular LA solution described here. Indeed, this approach has been used for a series of LA solutions that targets different science standards (Bichler et al., 2021; Gerard et al., 2020, 2022). The approach of developing a multidisciplinary, multi-stakeholder design team from a longstanding RPP to iteratively design a LA solution to support teachers in taking effective action in response to students developing understanding is the primary contribution of this paper to the field of HCLA design.

While the authors acknowledge the similarities between the proposed design approach and the LATUX framework (Martinez-Maldonado et al., 2016), such as the integration of rigorous evaluation as a critical component of the design process, the necessary precautions that should be taken when using student data and the thoughtful identification of relevant data, key stakeholders, and the problem of practice, there are substantive differences that set the proposed design approach apart. A critical distinction is the use of DBR as the design methodology. Using a DBR-based approach mitigates the tension articulated by Martinez-Maldonado and colleagues (Martinez-Maldonado et al., 2016) between authenticity and LA tool development. Since "in the wild" iterative design is a hallmark feature of the DBR methodology, many of the limitations associated with design in accordance with the LATUX workflow (eg, inability to test interventions and interactivity) are nullified. Another distinction is the use of an RPP as the design context. As a framework, LATUX does not recommend a particular design team structure or context, whereas the design approach proposed here contends that an RPP is ideal for HCLA design. Our findings demonstrate how the RPP design context in conjunction with the DBR methodology attends to all of the LATUX principles and mitigates or nullifies several of the framework's limitations.

In this paper, the study results have shown the potential of the proposed design approach to address three important expansion needs and its viability to achieve key outcomes, namely student learning and contextual scalability. The main components of the proposed design approach may also be valuable for the wider Learning and Learning Sciences field, although this paper aimed at discussing the shortcomings and challenges of existing HCLA design approaches. However, it is wise to carefully assess the implications of each component of the proposed HCLA design approach on the target expansion areas and key outcomes. While some factors may directly correspond to the proposed design components, due to the nature of the study, others may be derived from contextual aspects that emerge during this rather longstanding RPP. In this paper, we have tried to disentangle these relations and factors and show related evidence. However, future research in similar longstanding RPPs or with smaller scale projects should consider further validation of the proposed DBR-based, HCLA design components.

Future research should also explore extending the dashboard LA and recommended actions directly to students as a strategy for supporting self-directed learning. Recent research suggests value in supporting students to optimize their own learning (Buckingham Shum et al., 2019). Although these studies involved older students, it is worth investigating how to support younger students in effectively using LA to guide their own learning. The findings presented in this study raise other promising areas of research, namely using LA as a methodological tool to explore teaching and learning, and designing LA that can be an educative resource for teachers. Whether and how LA dashboards can be developed to influence teachers' philosophy for teaching needs further exploration. However, if such an LA solution can be developed, this could bridge the gap between theory and classroom action and open new frontiers for supporting teacher education, pre-service and in-service teacher training, as well as for supporting students during in-person and remote learning contexts.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the contribution of the participating teachers and students in the study, as well as the rest of the WISE research and development team at University of California at Berkeley. This material is based upon work supported in part by the National Science Foundation under Grant DRL-1813713: Supporting Teachers in Responsive Instruction for Developing Expertise in Science (STRIDES). Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The second author was partially supported by the European Regional Development Fund and the National Research Agency of the Spanish Ministry of Science and Innovation, under project grant PID2020-112584RB-C32.

FUNDING INFORMATION

Project grant PID2020-112584RB-C32 by the European Regional Development Fund and the National Research Agency of the Spanish Ministry of Science and Innovation. National Science Foundation Grant DRL-1813713: Supporting Teachers in Responsive Instruction for Developing Expertise in Science (STRIDES)

CONFLICT OF INTEREST STATEMENT

There is no potential conflict of interest in this work.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ETHICS STATEMENT

The study data are not open due to human subject protection policies. This study was conducted with the IRB approval of the University of California at Berkeley, USA.

ORCID

Korah Wiley https://orcid.org/0000-0001-5949-3506
Yannis Dimitriadis https://orcid.org/0000-0001-7275-2242
Marcia Linn https://orcid.org/0000-0003-2257-9474

REFERENCES

Alvarez, C. G. P. (2020). *Engaging stakeholders in the learning analytics design process*. Doctoral dissertation. University of Technology Sydney.

Alverson, C. Y., & Yamamoto, S. H. (2016). Educational decision making with visual data and graphical interpretation: Assessing the effects of user preference and accuracy. *SAGE Open*, *6*(4), 215824401667829. https://doi.org/10.1177/2158244016678290

Baker, R., Xu, D., Park, J., Yu, R., Li, Q., Cung, B., Fischer, C., Rodriguez, F., Warschauer, M., & Smyth, P. (2020). The benefits and caveats of using clickstream data to understand student self-regulatory behaviors: Opening the black box of learning processes. *International Journal of Educational Technology in Higher Education*, *17*(1), 1–24.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, *13*(1), 1–14.

Bichler, S., Gerard, L., Riordan, B., Lim-Breitbart, J., Bradford, A., Billings, K., & Linn, M. (2021). *Using learning analytics to assess students' ideas and generate reports for real time use and curriculum customization*. https://doi.org/10.13140/RG.2.2.12710.24649

Bødker, S., Dindler, C., Iversen, O. S., & Smith, R. C. (2022). What is participatory design? In *Participatory design. Synthesis lectures on human-centered informatics*. Springer. https://doi.org/10.1007/978-3-031-02235-7_2

Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-centred learning analytics. *Journal of Learning Analytics*, *6*(2), 1–9.

Campos, F. C., Ahn, J., DiGiacomo, D. K., Nguyen, H., & 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. https://doi.org/10.18608/jla.2021.7113

Clark, D., & Linn, M. (2013). The knowledge integration perspective: Connections across research and education. https://doi.org/10.4324/9780203154472

Cloude, E., Carpenter, D., Dever, D., Azevedo, R., & Lester, J. (2021). Game-based learning analytics for supporting adolescents' reflection. *Journal of Learning Analytics*, *8*(2), 51–72. https://doi.org/10.18608/jla.2021.7371

Clow, D. (2012). The learning analytics cycle: Closing the loop effectively. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge - LAK '12* (p. 134). ACM Press. https://doi.org/10.1145/2330601.2330636

Collins, A. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15–22). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-77750-9_2

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *Journal of the Learning Sciences*, *13*(1), 15–42.

Dimitriadis, Y., Martínez-Maldonado, R., & Wiley, K. (2021). Human-centered design principles for actionable learning analytics. In T. Tsiatsos, S. Demetriadis, A. Mikropoulos, & V. Dagdilelis (Eds.), *Research on E-learning and ICT in education*. Springer. https://doi.org/10.1007/978-3-030-64363-8_15

Echeverria, V., Martinez-Maldonado, R., Granda, R., Chiluiza, K., Conati, C., & Shum, S. B. (2018). Driving data storytelling from learning design. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge - LAK '18* (pp. 131–140). ACM Press. https://doi.org/10.1145/3170358.3170380

Echeverria, V., Martinez-Maldonaldo, R., Buckingham Shum, S., Chiluiza, K., Granda, R., & Conati, C. (2018). Exploratory versus explanatory visual learning analytics: Driving Teachers' attention through educational data storytelling. *Journal of Learning Analytics*, *5*(3), 72–97. https://doi.org/10.18608/jla.2018.53.6

Farrell, C. C., Penuel, W. R., Coburn, C., Daniel, J., & Steup, L. (2021). *Research-practice partnerships in education: The state of the field*. William T. Grant Foundation.

Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5/6), 304. https://doi.org/10.1504/IJTEL.2012.051816

Forbes, C. T., Neumann, K., & Schiepe-Tiska, A. (2020). Patterns of inquiry-based science instruction and student science achievement in PISA 2015. *International Journal of Science Education*, *42*(5), 783–806.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*, 64–71.

Gašević, D., Kovanović, V., & Joksimović, S. (2017). Piecing the learning analytics puzzle: A consolidated model of a field of research and practice. *Learning: Research and Practice*, *3*(1), 63–78.

Gerard, L., Bichler, S., Billings, K., Bradford, A., & Linn, M. C. (2022). A natural language processing (NLP) driven teacher dashboard to support responsive instruction. In J. Bywater & J. Littenberg-Tobias (Chairs), *Artificial intelligence and the future of STEM instruction: Designing new models to automate feedback to teachers* (structured poster session). 2022 AERA Annual Meeting, San Diego, CA, United States. https://doi.org/10.13140/RG.2.2.20272.02560

Gerard, L., Wiley, K., Bradford, A., Chen, J. K., Lim-Breitbart, J., & Linn, M. (2020). Impact of a teacher action planner that captures student ideas on teacher customization decisions. In M. Gresalfi, & I. S. Horn (Eds.), *The interdisciplinarity of the learning sciences, 14th international conference of the learning sciences (ICLS) 2020* (Vol. *4*, pp. 2077–2084). International Society of the Learning Sciences.

He, Y., Faircloth, B. S., Hewitt, K. K., Rock, M. L., Rodriguez, S., Gonzalez, L. M., & Vetter, A. (2020). Data management and use through research practice partnerships: A literature review. *Educational Research Review.*, *31*(1), 100360. https://doi.org/10.1016/j.edurev.2020.100360

Jørnø, R. L., & Gynther, K. (2018). What constitutes an 'actionable insight' in learning analytics? *Journal of Learning Analytics*, *5*(3), 198–221. https://doi.org/10.18608/jla.2018.53.13

Kasepalu, R., Chejara, P., Prieto, L. P., & Ley, T. (2022). Do teachers find dashboards trustworthy, actionable and useful? A vignette study using a logs and audio dashboard. *Technology, Knowledge and Learning*, *27*, 971–989. https://doi.org/10.1007/s10758-021-09522-5

Kovanovic, V., Mazziotti, C., & Lodge, J. (2021). Learning analytics for primary and secondary schools. *Journal of Learning Analytics*, *8*(2), 1–5. https://doi.org/10.18608/jla.2021.7543

Krumm, A. E., Boyce, J., & Everson, H. T. (2021). A collaborative approach to sharing learner event data. *Journal of Learning Analytics*, *8*(2), 73–82. https://doi.org/10.18608/jla.2021.7375

Lachenbruch, P. A. (2014). McNemar Test. In N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri, & J. L. Teugels (Eds.), *Wiley StatsRef: Statistics reference online*. https://doi.org/10.1002/9781118445112.stat04876

Leacock, C., & Chodorow, M. (2003). C-rater: Automated scoring of short-answer questions. *Computers and the Humanities*, *37*, 389–405. https://doi.org/10.1023/A:1025779619903

Li, Q. (2021). *Reflections on scaling up learning analytics*. Learning Analytics Research Network, NYU https://steinhardt.nyu.edu/learn/reflections-scaling-learning-analytics

Linn, M. C., & Eylon, B.-S. (2011). *Science learning and instruction: Taking advantage of technology to promote knowledge integration*. Routledge (OCLC: ocn176946367).

Mama, M., & Hennessy, S. (2013). Developing a typology of teacher beliefs and practices concerning classroom use of ICT. *Computers & Education*, *68*, 380–387.

Mangaroska, K., & Giannakos, M. (2018). Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies*, *12*(4), 516–534. https://doi.org/10.1109/TLT.2018.2868673

Martinez-Maldonado, R., Pardo, A., Mirriahi, N., Yacef, K., Kay, J., & Clayphan, A. (2016). LATUX: An iterative workflow for designing, validating and deploying learning analytics Visualisations. *Journal of Learning Analytics*, *2*(3), 9–39. https://doi.org/10.18608/jla.2015.23.3

Mor, Y., Ferguson, R., & Wasson, B. (2015). Editorial: Learning design, teacher inquiry into student learning and learning analytics: A call for action: Learning design, TISL and learning analytics. *British Journal of Educational Technology*, *46*(2), 221–229. https://doi.org/10.1111/bjet.12273

NGSS Lead States. (2013). *Next generation science standards: For states, by States*. The National Academies Press.

Ochoa, X., Lang, A. C., & Siemens, G. (2017). Multimodal learning analytics. *The Handbook of Learning Analytics*, *1*, 129–141.

OECD. (2009). Chapter 4 teaching practices, TEACHERS' beliefs and attitudes. In *Creating effective teaching and learning environments: First results from TALIS*. OECD Publishing. ISBN 978–92–64-05605-3.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, *62*(3), 307–332.

Penuel, W., Allen, C., Manz, E., Heredia, S., Superfine, A. C., Goldman, S. R., & Ko, M. L. M. (2022). Design-based implementation research as an approach to studying teacher learning in research-practice partnerships focused on equity. In *Teacher learning in changing contexts: Perspectives from the learning sciences*. Routledge.

Penuel, W., Furtak, E., & Farrell, C. (2021). Research-practice partnerships in education: Advancing an evolutionary logic of systems improvement. *DDS - the German School*, *113*(1), 45–62. https://doi.org/10.31244/dds.2021.01.05

Penuel, W. R., Roschelle, J., & Shechtman, N. (2007). Designing formative assessment software with teachers: An analysis of the co-design process. *Research and Practice in Technology Enhanced Learning*, *2*(1), 51–74.

Prieto, L. P., Rodríguez-Triana, M. J., Martínez-Maldonado, R., Dimitriadis, Y., & Gašević, D. (2018). Orchestrating learning analytics (OrLA): Supporting inter-stakeholder communication about adoption of learning analytics at the classroom level. *Australasian Journal of Educational Technology*, *35*(4), 14–33. https://doi.org/10.14742/ajet.4314

Reimann, P. (2016). Connecting learning analytics with learning research: The role of design-based research. *Learning: Research and Practice*, *2*(2), 130–142. https://doi.org/10.1080/23735082.2016.1210198

Revano, T. F., & Garcia, M. B. (2021). Designing human-centered learning analytics dashboard for higher education using a participatory design approach. In *2021 IEEE 13th international conference on humanoid, nanotechnology, information technology, communication and control, environment, and management (HNICEM)* (pp. 1–5). IEEE.

Riordan, B., Bichler, S., Bradford, A., King Chen, J., Wiley, K., Gerard, L., & C. Linn, M. (2020). An empirical investigation of neural methods for content scoring of science explanations. In *Proceedings of the fifteenth workshop on innovative use of nlp for building educational applications* (pp. 135–144). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.bea-1.13

Rodríguez-Triana, M. J., Prieto, L. P., Dimitriadis, Y., De Jong, T., & Gillet, D. (2021). ADA for IBL: Lessons learned in aligning learning design and analytics for inquiry-based learning orchestration. *Journal of Learning Analytics*, *8*(Article), 22–50. https://doi.org/10.18608/jla.2021.7357

Sanders, E. (2006). Design research in 2006. *Design Research Quarterly.*, *1*, 1–8.

Sandoval, W. A., & Bell, P. (2004). Design-based research methods for studying learning in context: Introduction. *Educational Psychologist*, *39*(4), 199–201. https://doi.org/10.1207/s15326985ep3904_1

Sandoval, W. (2013). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences.*, *23*, 18–36. https://doi.org/10.1080/10508406.2013.778204

Siemens, G. (2012, April). Learning analytics: Envisioning a research discipline and a domain of practice. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 4–8). LAK '12. Association for Computing Machinery. ISBN 9781450311113. https://doi.org/10.1145/2330601.2330605

Siemens, G., & Baker, R. S. D. (2012). Learning analytics and educational data mining: Towards communication and collaboration. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 252–254). LAK '12. Association for Computing Machinery. ISBN 9781450311113. https://doi.org/10.1145/2330601.2330661

van Leeuwen, A., Knoop-van Campen, C. A., Molenaar, I., & Rummel, N. (2021). How teacher characteristics relate to how teachers use dashboards: Results from two case studies in K-12. *Journal of Learning Analytics*, *8*(2), 6–21. https://doi.org/10.18608/jla.2021.7325

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, *57*(10), 1500–1509. https://doi.org/10.1177/0002764213479363

Vitale, J. M., McBride, E., & Linn, M. C. (2016). Distinguishing complex ideas about climate change: Knowledge integration vs. specific guidance. *International Journal of Science Education*, *38*(9), 1548–1569. https://doi.org/10.1080/09500693.2016.1198969

Whitaker, D., & Jacobbe, T. (2017). Students' understanding of Bar graphs and histograms: Results from the LOCUS assessments. *Journal of Statistics Education*, *25*(2), 90–102. https://doi.org/10.1080/10691898.2017.1321974

Wiley, K. J., Bradford, A., & Linn, M. C. (2019). Supporting collaborative curriculum customizations using the knowledge integration framework. In K. Lund, G. P. Niccolai, E. Lavoue, C. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), *A wide lens: Combining embodied, enactive, extended, and embedded learning in collaborative settings, 13th international conference on computer supported collaborative learning* (Vol. *1*, pp. 480–487). International Society of the Learning Sciences (ISLS).

Wiley, K. J., Dimitriadis, Y., Bradford, A., & Linn, M. C. (2020, March). From theory to action: Developing and evaluating learning analytics for learning design. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 569–578). Association for Computing Machinery. ISBN 9781450377126. https://doi.org/10.1145/3375462.3375540

Wise, A. F., & Vytasek, J. (2017). Learning analytics implementation design. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (1st ed., pp. 151–160). Society for Learning Analytics Research (SoLAR). https://doi.org/10.18608/hla17.013

Wladis, C., Offenholley, K., Lee, J.K., Dawes, D., Licwinko, S. (2017). A review of research on teacher beliefs and practices. In *Tenth congress of the European Society for Research in mathematics education CERME 10* (Vol. *38*, pp. 47–65). HAL Open Science. hal-01914615f. https://hal.science/hal-01914615

How to cite this article: Wiley, K., Dimitriadis, Y., & Linn, M. (2023). A human-centred learning analytics approach for developing contextually scalable K-12 teacher dashboards. *British Journal of Educational Technology*, *00*, 1–41. https://doi.org/10.1111/bjet.13383

APPENDIX A

TABLE A1 Rubric used to score the assessment items

Coherence score description	Score level	Idea score description
Off-task response	1	No discussion of target ideas
No link—Alternative or invalid ideas	2	Partial discussion of target ideas
Partial link—Normative ideas without any scientifically valid links between normative ideas	3	Full discussion of target ideas
Full link—One scientifically valid links between normative ideas	4	
Complex link—Two scientifically valid links between normative ideas	5	
Scientifically valid link description		Photosynthesis & Producer/Consumer Relationships (DCI)
Energy transformation: One type of energy can be transformed into another type of energy. Light energy is transformed into chemical energy during photosynthesis.		Idea #1: Full = $CO2 + H2O \rightarrow$ (via arrangement or chemical reaction) glucose + oxygen, uses energy from sun/light to do photosynthesis reaction Not good enough for partial credit to just say that the plant gets/uses energy from the sun to grow, need to say or describe photosynthesis
Matter transformation: Reactants, such as carbon dioxide and water, are transformed into products, glucose and oxygen		Idea #2: Full = Explicit statement about what specifically the PLANT uses glucose/food (not CO2, H2O, light) for: growth, energy, stored, repair, seed production
Energy–matter relationship: Energy is required to initiate a chemical reaction (bond breaking). Chemical energy is stored in glucose (matter)		
		Energy Transfer Drives Matter Cycling (CCC)
		Idea #1: Full = Energy (any type of relevant energy) and matter (any type of relevant matter) moves WITHIN THE PLANT during photosynthesis
		Idea#2: Full = Energy from the sun gets to animals when they eat photosynthetic plants, directly or indirectly. (ie, when animals EAT plants or plant-eating animals breathe in oxygen for food/energy)

APPENDIX B

B.1 | Plan for three interviews

- PRE: intro to research study, before unit implementation.
- DURING: after milestone, when viewing TAP data.
- POST: after implementation of instructional customization.

B.2 | PRE INTERVIEW PROTOCOL

Express Gratitude, Build Rapport

- Thank you for partnering with us for this research project.

Provide Overview of Research Study

- **T**he goal of this project is to provide teachers with NGSS aligned curriculum, give them data to know how well their students are meeting the three dimensional learning goals for the targeted PE.

Solicit Information about their Current Practice around Assessing student understanding.
Question Series 1

- Is this your first time teaching this content (PE)/specific curriculum unit?
- Why did you select this unit? What are your goals for what you want your students to get out of this unit?
- What are some of the ways that you currently find out how students are meeting the learning goals?
- Is there one way that you find to be the most effective at giving you that insight?
- Is that a method you are able to use often?

Introduce the New Milestone Feature of Units

- In response to the feedback we were getting from teachers, we have redesigned our units to 1.) focus on specific PEs, and 2.) put in automatically scored milestone assessments (learning checkpoints) in the unit.
  - Direct them to Milestone page in the platform
  - Show them where in the unit the milestone is situated.
  - Show them the scoring rubrics for the PE and curriculum resources targeting these ideas
- Each milestone will generate a report in the teacher grading tool that will give you a summary of how well your students are meeting those PE specific goals. (Assure teacher we will be there to help them find the report.)
- How do you think your students are likely to perform on this milestone?
- What do you think you will see in your student work on this milestone? Which ideas do you think your students will understand easily? Which do you think they will have the most difficulty understanding? Ask teacher about their predictions for patterns for embedded item, how much they think it'll vary by class period
  - Show sample summary report.
- This milestone that we will focus on for this part of the study targets MS-LS1-6 which calls for students to construct a scientific argument that explains how the photosynthesis reaction facilitates the cycling of energy and matter.
- So the report will give you feedback about how well your students coherently linking the details of the photosynthesis reaction with the concept of energy and matter cycling.

Solicit Information about their Impressions of the Report.
Question Series 2

- In general, what are your initial thoughts or impressions about the report?
- What do you find most useful or informative?
- Specific to the design, what aspects of the report draw you attention?
- Do you think the report complements in any way what you currently do?
- In what ways, if any, do you see the report interfering with your current practices?

- Another aspect of this project is to, as needed, provide teachers with support to make instructional adjustments in response to the information about how well their students are meeting the three dimensional learning goals so that students can better met them. (Space for teachers to use other resources.)

Question Series 3

- What are your thoughts about the recommended activities portion of the report?
- Ask elaboration questions based on their response

End Interview, Emphasis Partnership

- That was great information, thank you so much for taking the time to meet with us today and sharing your ideas and perspectives about this study.
- As you are going through the unit, document your ideas using the Teacher Notebook (will always be accessible).
- Our plan is to observe your class as usual, and once your report has been generated meet briefly with you, when it is most convenient to talk about the results and what your ideas are for how you might respond. Get details for when the best times will be. (Two additional interviews.)
- Go over teacher consent form (have teacher sign) and mention videotaping during interviews and during class (implementation of instructional interventions).
- Thank them again and leave.

B.3 | DURING INTERVIEW PROTOCOL

- What customizations did you make prior to implementation? (Or “Which instructional activity are you planning to implement?”—whether ours or theirs)
- Why did you decide to make a customization?
- How did you decide on this design for your customization?
- How did your customization respond to your students' ideas?
- Did you work with a partner to make customizations or by yourself?
- What was your sense of how students were understanding this idea from your experience teaching (before look at report)
- What do you notice in the report about your students' understanding? How does it reflect (or not) what you observed while teaching?
- Imagine yourself teaching this unit the next day without this report—how do you think this affects how you might teach differently tomorrow?
- What do you see in this—what does it tell you that has implications for teaching tomorrow?
- Do you plan to adjust your lesson tomorrow based on this information?
- What surprises you in this report?
- Is this different than what you'd predict/expect to see for this class?
- How do these recommended actions resonate with what you think you would do to respond to students' ideas? What changes would you make if any?
- Do you plan to use the report for each class period?
- Which customization do you plan to use and why?
- Based on the information—would you choose to customize it in other ways?
- As the recommendation action, do you want students to see their scores?
- Take note of what they are attending do, then in post ask more specifically about various features.

B.4 | POST INTERVIEW PROTOCOL

- How did the implementation of the unit go in general?
- What did you observe about students' science learning? Please give an example
- Did you customize the unit during implementation? If so, please give an example—Were you able to carry out the customization you had planned to implement, or did you have to tweak or make adjustments (and why)?
- How did students respond to your customization? Did it surprise you?
- Compare report 1 with report "revised". What do you notice about changes in student understanding?
- How did you feel like using the TAP and recommended actions supported your teaching and students learning versus when you have a unit without the TAP
- Were you oriented enough to content in the unit to be able to use the report, interpret the data and consider the recommended actions?
- What would you do differently next time when using the TAP in this unit—in terms of how you used the report (by period, what you attended to, how you followed the recommended action)?
- Did your use of the TAP (seeing student performance and the suggested recommendations) influence your teaching beyond the recommended actions, as you continued to teach the rest of the unit?
- Gather feedback on specific design features of the TAP: key insights, presentation of 3 graphs, student responses, recommended actions (other options they would suggest)
- How do you see this report fitting in the context of your teaching practice or assessment routines more broadly? Is the information in the TAP something you could see yourself wanting and using when teaching other lessons (if possible)?
- What did you observe was challenging for students during the unit?
- How do you want to customize the unit for next time?
- Did you find the Teacher Notebook helpful? Would prompts be useful for helping you to reflect on your instructional decisions after viewing the Teacher Report data?