JOURNAL OF LEARNING ANALYTICS
SoLAR SOCIETY for LEARNING ANALYTICS RESEARCH
Volume 8(2), 73–82. https://doi.org/10.18608/jla.2021.7375

A Collaborative Approach to Sharing Learner Event Data

Andrew E. Krumm[1], Jared Boyce[2], Howard T. Everson[3]

Abstract
This paper describes a collaboration organized around exchanging data between two technological systems to support teachers' instructional decision-making. The goals of the collaboration among researchers, technology developers, and practitioners were not only to support teachers' instructional decision-making but also to document the challenges and opportunities associated with bringing together data from instruction- and assessment-focused technologies. The approach described in this paper illustrates the potential importance of anchoring data products that combine data between two systems in the needs of teachers as well as aligning the content that students learn and are assessed on between systems. The increasing presence of data standards has made sharing complex data increasingly more feasible. The example collaboration described in this paper demonstrates the role that non-technical activities can play in supporting the exchange and use of learner event data.

Notes for Practice

- Exchanging and combining data from two or more technologies requires common understandings of how data are generated by students.
- There are few established strategies for making sense of learner event data across different technologies. This gap signals the potential importance of higher bandwidth interactions between and among researchers, technology developers, and school practitioners.

Keywords
Data interoperability, data products, decision-making, instructional support

Submitted: 01/10/20 — **Accepted:** 23/05/21 — **Published:** 03/09/21

Corresponding author [1]Email: aekrumm@umich.edu Address: University of Michigan, Medical School, 221 Victor Vaughan, 1111 E. Catherine St., Ann Arbor, MI 48109-2054, USA. ORCID ID: https://orcid.org/0000-0001-9714-1632
[2]Email: jared.boyce@sri.com Address: SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025, USA ORCID ID: https://orcid.org/0000-0002-7030-6836
[3]Email: howard.everson@sri.com Address: SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025, USA ORCID ID: https://orcid.org/0000-0002-8688-1819

1. Introduction

Imagine a school where students use multiple technologies to advance learning, data from these technologies are analyzed regularly, and educators use these data to improve student opportunities to learn. For many K–12 schools and districts, this vision remains underdeveloped despite the fact that students use a variety of learning technologies in their day-to-day activities and that these technologies can produce large quantities of data (Baker & Siemens, 2014; Koedinger, D'Mello, McLaughlin, Pardos, & Rosé, 2015). While technology use and quantities of data increase, a key challenge for practitioners in schools is accessing and visualizing data from multiple technologies in a single system or platform (e.g., DiCerbo & Korbin, 2016).

As the volume of available data has increased, a variety of data interoperability standards have emerged to support the "seamless, secure, and controlled exchange of data between applications" (State Educational Technology Directors Association, 2018, p. 8). While data standards have made storing and sharing data feasible, longstanding challenges remain in exchanging data to support specific work practices and routines (i.e., organizational interoperability), in communicating the meaning of data elements from one context to the next (i.e., semantic interoperability), and in facilitating the exchange and use of data from multiple systems (i.e., technical interoperability; Pagano, Candela, & Castelli, 2013).

Collaborations among researchers, technology developers, and practitioners offer a promising approach for addressing multiple challenges associated with collecting, sharing, and analyzing data from digital systems and platforms (Krumm, Means, & Bienkowksi, 2018). In this paper, we describe a collaborative project that joined data from an instruction-focused and an assessment-focused technology to support teachers' instructional decision-making. As a collaborative project, we sought to 1) learn directly from teachers how they could use data products that joined data from two technologies as well as 2) understand the challenges and opportunities associated with exchanging data in order to inform broader efforts directed at supporting data interoperability.

2. Data Interoperability and Learning Event Data

From the introduction of the earliest online learning technologies, there has been interest in developing common ways to share data between and among technologies. Some of the first successes in interoperability came from higher education and the increased use of learning management systems like Blackboard, Sakai, Desire2Learn, and, more recently, Canvas. The centrality of learning management systems to higher education institutions in the mid-2000s made them ideal for integrating multiple applications, such as through the Learning Tools Interoperability standard. Other interoperability successes have occurred with single sign-on (e.g., Clever and OneRoster) and administrative data system standards like the Common Education Standards, the Schools Interoperability Framework, and the Ed-Fi Data Standard in the United States. These initiatives have helped to solve real and immediate problems facing educators. Faculty using learning management systems needed ways to integrate a variety of learning tools into their courses; schools supporting multiple learning technologies needed to reduce the complexity of managing multiple student accounts; and districts and schools needed ways to move student records in less error-prone ways between schools and among various administrative systems (e.g., student information systems).

Along with specific standards, there are a growing number of organizations supporting various elements of data interoperability, such as technical assistance organizations like the Data Quality Campaign and the Privacy Technical Assistance Center, which offer resources related to both data sharing and data privacy. Project Unicorn, based out of InnovateEDU, also works to address multiple aspects of interoperability through partnerships with schools and technology vendors.

Through initiatives like Ed-Fi, progress has been made in sharing data between assessment-focused systems and student information systems. Assessment-focused systems administer tests and quizzes to students as well as provide analyses and reports on students. Student information systems collect and store demographic information about students along with longitudinal test scores, attendance information, and course grades. Sharing learner event data (i.e., system log data, telemetry data, digital exhaust, usage data, or system utilization data) with assessment-focused technologies or student information systems, however, still poses several challenges despite the successes of data standards specifically geared toward learner event data.

Learner event data are most often collected by instruction-focused digital environments designed primarily to help students acquire and practice new content and skills. Instruction-focused technologies are widely used in schools and can provide direct instruction to students, support homework activities, and/or assist in blended learning models (e.g., Murphy et al., 2014). These types of environments can provide an entire curriculum, or they can provide supplementary learning experiences. They can also consist of one or more different types of games, simulations, playlists, and intelligent tutoring experiences. For consistency, we define learner event data as *data that capture specific actions taken by a student at a particular time within a specific learning task*. Actions within and across environments can be numerous and can vary in terms of their granularity from “moved mouse to the left” to “logged in” to “accessed resource.” A central challenge associated with exchanging learner event data is the amount of context that needs to be shared with an event in order for that event to be useable by another system (e.g., Dietz et al., 2012; Sottilare, Long, & Goldberg, 2017).

A handful of interoperability standards have made sharing learner event data more and more feasible, such as Experience API (xAPI from Advanced Distributed Learning) and Caliper (from IMS Global). xAPI and Caliper, in particular, have addressed a key issue in working with learner event data, namely, the diversity of learning activities within instruction-focused technologies (e.g., games, simulations, puzzles, and quizzes). The variety of activities that students can engage in contributes to an almost overwhelming diversity in the number and types of events that can be collected across technologies (e.g., del Blanco et al., 2013). xAPI and Caliper have approached the problem by creating a standard that facilitates sharing events using the general categories of subject (e.g., “student”), object (e.g., “game level”), and verb (e.g., “started”). These standards, along with smaller scale efforts like the Assessment Data Aggregator for Game Environments (ADAGE) from the University of Wisconsin (Owen, Ramirez, Salmon, & Halverson, 2014) and proofs-of-concept from organizations like ETS (Hao, Smith, Mislevy, von Davier, & Bauer, 2016), are increasing in popularity.

Within the context of data interoperability, learner event data has increasingly been an area of research in the field of learning analytics. Bakharia and colleagues (2016), for example, examined the potential of xAPI to combine data from multiple social media platforms. In their work, they identified the critical role that contextual information within xAPI statements played in joining data to perform social network analyses. Similarly, Manso-Vazquez, Caeiro-Rodriguez, and Llamas-Nistal (2018) demonstrated the often-intense work involved in disambiguating learner event data in relation to specific self-regulated learning processes across multiple technologies (e.g., Winne & Hadwin, 1998). While xAPI offered multiple affordances, “the translation from the strategies to the profile was not trivial. Every action had to be tackled separately while trying to maintain coherence throughout the profile" (p. 42480). The Pittsburgh Science of Learning Center's DataShop and the data workflow tool LearnSphere have both provided researchers with access to complex, granular learner event data as well as analytical tools to carry out analyses like generating learning curves (e.g., Martin et al., 2011). DataShop and LearnSphere make it easy to work with data that have been collected and stored by specific technologies and can serve as the kind of "sandboxes" called for by Pagano, Candela, and Castelli (2013) when working to make sense of and combine data from multiple sources.

As the above examples imply, a central task in working with event data is understanding the multiple contexts in which students generate data. Context can be understood in at least two ways: the learning activities within an instructional-focused technology itself and the broader instructional activities in which the technology is used within a classroom (Krumm, Means, & Bienkowski, 2018). For many educational data elements, such as those related to student information systems and assessment-focused technologies, the importance of context is somewhat reduced. For example, when working with data from a test, those familiar with education have commonly shared references for "items" and an "overall score." Furthermore, many people who have spent time in schools intuitively understand that a test's purpose is to quantify what a student knows at a given point in time. In the case of learner event data captured by instruction-focused technologies, there are comparatively fewer shared understandings to draw upon. For example, a student can view a video, complete a problem set, be provided with feedback on correct or incorrect responses, engage in open-ended coding activities, manipulate an avatar, operate a pencil, input numbers into fixed problems, choose an activity to begin, and/or be directed to the next activity by the system itself. These example features of a digital learning task can be combined in unique ways, spread out over time, and be more or less salient depending on the organization of the content to be learned (e.g., lessons, units). Given the many ways in which these features of learning activities can be combined, a considerable amount of context must be communicated in order for an event to be appropriately used in a given analysis or data product.

Drawing on insights from the above research, we used a collaborative approach informed by Penuel and colleagues (2011) to directly intervene upon challenges associated with semantic and organizational interoperability. Our working hypothesis was that researchers, developers, and practitioners would need to work together in order to share data between two technologies, make sense of data originating from both systems (i.e., semantic interoperability), and prototype data products (i.e., organizational interoperability). The importance of developing prototypes represents an additional but subtle takeaway from the existing literature around data interoperability in education: the power of data standards is in their ability to support the exchange and use of data from multiple systems (i.e., technical interoperability), and only when data are actually brought together can challenges and opportunities be identified.

3. Data and Methods

This project sought to understand the benefits and challenges of data interoperability by developing prototype data products that joined data from two different technological systems. The collaboration included Renaissance Learning (Renaissance), the MIND Research Institute (makers of ST Math and hereafter MIND), SRI International, and one elementary school on the U.S. West Coast that was using Renaissance Star Assessments and ST Math. The role of the school was particularly important because a key assumption of our approach was that data products that combined data from multiple systems needed to address problems raised by teachers, and that issues of organizational, technical, and semantic interoperability could perhaps best be identified by working to combine data in the service of addressing teachers' data-related problems.

Within the collaboration, researchers were responsible for organizing focus groups, collecting data from the focus groups, and prototyping data products. In developing prototype data products, researchers helped to develop a shared understanding across the collaboration of each system's data by examining the contexts in which students generate data in Star Math and ST Math. To better understand Star Math, researchers in the collaboration worked with Renaissance to understand how different assessments were administered and how different scores were generated. As a computer adaptive test, Star Math provides scores of student abilities across a wide variety of mathematical skills in addition to overall scores and growth projections. To better understand ST Math, researchers used guest accounts to directly experience learning tasks while simultaneously consulting data models and data dictionaries (i.e., descriptions of data elements and valid values) from ST Math to make connections between collected data and student actions.

The basic learning task within ST Math is a puzzle. A puzzle presents students with a visual representation of a mathematical concept that can be solved by moving virtual manipulatives. Data on student puzzle and level performances can easily be summarized and joined with metadata related to the games and learning objectives of a level as well as the session to which a student's activity belongs. A high-level representation of student movement through ST Math is presented in Figure 1. A syllabus (not pictured) is comprised of multiple objectives for a grade level; objectives are comprised of games, levels, and puzzles.

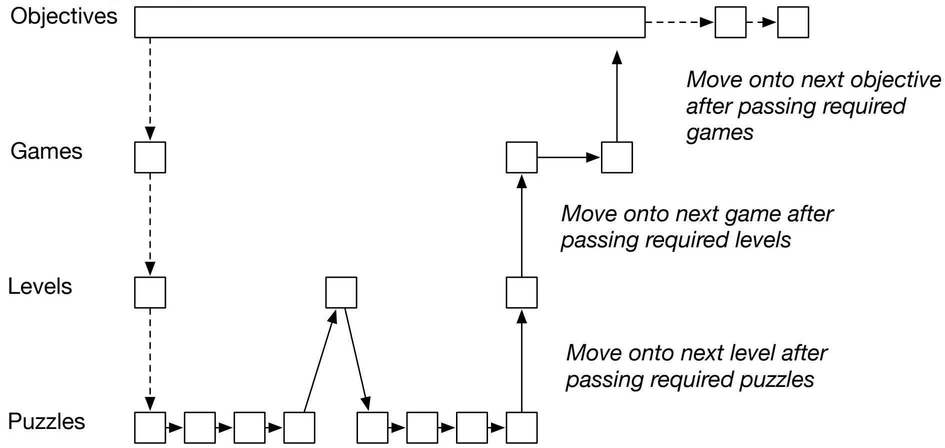

Figure 1. Movement across puzzles, levels, games, and objectives in ST Math.

Two focus groups were organized to learn directly from teachers. Five Grade 3 teachers and six Grade 4 teachers participated in the first focus groups. Three Grade 4 teachers participated in the second focus group. Two of the school's technology specialists also participated in both focus groups. The initial focus group for Grade 3 teachers was conducted in person; videoconferencing was used for the initial focus group for Grade 4 teachers. Based on feedback from the virtual focus group, the second focus group was conducted in person. Data from focus groups were collected using in-depth field notes collected by one researcher. Following a focus group session, the researcher generated a memo that was shared among members of the collaboration. Initial focus groups were organized around 1) how teachers used ST Math, 2) how they used ST Math and Star Math reports, and 3) design ideas for combining data between ST Math and Star Math. Figure 2 illustrates the organization of focus groups 1 and 2 along with example questions. It was important to understand how teachers used ST Math because their use of a given technology can play an important role in their motivation to use reports from that technology. Likewise, it was important to understand how teachers used each system's reports in order to understand what data they most attended to. To ensure that teachers had a common set of experiences with Star Math reports, collections of classroom- and student-level ST Math and Star Math reports were compiled for each teacher and distributed before the first focus group.

Focus Group 1
Distribute reports from ST Math and Star Math

Introductions and tone setting:
Since you've been at this school, how, if at all, have you changed the ways you teach math?

ST Math:
How do you use ST Math in your class? When is it used during the day and how often? Do all students use it in the same way? Do you use the reports provided by ST Math? If so, which ones? How often do you look at an ST Math report (e.g., daily, weekly)? For what purposes do you use a given report?

Star Assessments:
What is your experience with Star Assessments? Which have you used more, the reading or math tests? Thinking about whichever Start test you've used most, which reports do you find most helpful and how do you use them? Specifically thinking about the Star Math test, what reports are most helpful?

Design ideas and general discussion

Focus Group 2
Distribute prototype visualizations

Gather reactions to following sections:
Heading
Scatterplot of ST Math game recommendations (Figure 3)
Table of consolidated ST Math data (Table 1)

General reflections and discussion

Figure 2. Organization of focus groups 1 and 2.

In addition to making sense of data elements across both systems and organizing focus groups, researchers in the collaboration generated prototype data products (i.e., visualizations and tables) using data from both systems. Teachers' design ideas from the first focus group provided direction on what data elements to attend to and how to combine them. Combining data across systems provided the opportunity to surface challenges and opportunities associated with exchanging and merging data. Researchers and representatives from MIND and Renaissance met regularly to discuss teachers' design ideas and issues associated with merging data.

After finalizing a set of prototypes, a second focus group was held with teachers to gather feedback on the prototypes. For the second focus group, we hypothesized that teachers would find the prototype data products more meaningful if they could connect the information presented directly to their students, giving them the opportunity to question the information, use their classroom experiences to interpret the information, and decide how best to act on their interpretations. This second focus group consisted of a researcher walking teachers through the prototype data products one element at a time and gathering their feedback on each element.

4. Results and Discussions

4.1. Focus Group 1

The first focus group was organized around understanding 1) how teachers used ST Math, 2) how they used ST Math reports and Star Math reports, and 3) their ideas for combining ST Math and Star Math data. Teachers reported that when they used ST Math, they used it for 30–45 minutes on average. Teachers did, however, vary in their use of ST Math throughout the week, with some using it daily and others using it twice per week. Teachers had different reasons for using ST Math. Some teachers had all students use ST Math at the same time, with the teacher's role largely to assist students. Other teachers had half the class use ST Math independently while the other half received direct instruction from the teacher. As part of this approach, students switched midway through the class between ST Math and working with the teacher. In addition to these two general approaches, some teachers described different degrees of comfort in working with small groups of students while the rest of the class worked on ST Math. While ST Math offers the opportunity to modify what objectives students work on, only some teachers made modifications, and other teachers either were not aware of some of these features or did not feel comfortable modifying student experiences of ST Math.

Teachers stated that they typically only used ST Math's classroom-level progress report. For this report, teachers almost exclusively referred to two data elements: % Syllabus Progress (the proportion of the curriculum each student had progressed through) and Alert Buttons (indicators of struggle, low time on task, and low post-quiz scores). Teachers understood other data elements on the classroom-level progress report (e.g., a timestamp of a student's last login), although they did not connect them to specific actions they could or would take with students. Teachers said that their formal professional development work on ST Math had occurred some time ago and wondered whether another round of training might be valuable to help them better understand what reports are available and how best to use them.

Teachers had significant experience in using Star Reading reports because their district regularly used Star Reading as its universal screener and progress monitoring assessment for language arts as part of its Response to Intervention model. Based largely on their experience with Star Reading, teachers noted that Star Math reports were easy to understand and easy to use for assigning student groups. Additionally, teachers found the skills descriptions in Star Math reports easy to interpret and easy to develop potential instructional interventions around. Teachers reported using only a single Star report — the classroom-level summary report — which they found particularly useful for understanding the performance of individuals as well as groups of students. Teachers also appreciated that Star reports provide cross-grade-level information (e.g., the teacher of a Grade 4 student who is significantly below grade level will see that the student needs Grade 3 skills development). Teachers said that this was essential for working with their students given the variation that existed in their classrooms.

A clear vision emerged for how data from ST Math and Star Math could be combined. Based on their prior use of Star Reading and Accelerated Reader, a widely adopted digital reading platform from Renaissance Learning, teachers had experience with explicitly linking Star assessment results to recommended books in Accelerated Reader for students. This functionality helped teachers quickly and easily decide what students could or should work on. Teachers identified the potential value of aligning ST Math games and puzzles with the same standards used by Star Math. And while teachers noted that these types of combined data products could be useful, several comments addressed more than simply presenting data; teachers wanted additional functionality that let them assign individual as well as groups of students to ST Math activities based on Star Math results.

4.2. Prototype Data Product Development

Using the insights gained from the first focus group as well as researchers' developing understanding of ST Math's and Star Math's data elements, researchers began developing data products with regular input from Renaissance and MIND around a core design idea: helping teachers connect the skills that students demonstrated on Star Math with student progress in ST Math. Framed in question form, "If a student earned *X* on Star Math, what should he or she be working on in ST Math?" Aligning what students work on in ST Math with their performances on Star Math was intended to help teachers group students and/or modify what objective a student could work on next in ST Math (e.g., DiCerbo & Korbin, 2016). To make progress on this design idea, the collaboration recognized the importance of mapping the mathematical content that students were assessed on using Star Math and the content students were learning in ST Math.

4.2.1. Mapping Content Across Instructional and Assessment Uses

The mapping between ST Math games and Star Math domains took place over several weeks and required the insights of MIND, Renaissance, researchers, and teachers at the partnering school. For example, a school technology specialist worked with a group of teachers at the school to do initial mappings between Star Math and ST Math and reported significant challenges related to the scope of the task. The process involved organizing the skills and domains assessed in Grade 4 on Star Math (n = 75) and the content covered across ST Math games in Grade 4 (n = 145). Across multiple meetings in which mappings were iteratively refined, 95 ST Math games were mapped to 56 Star Math domains. Given the need to develop data products between the first and second focus groups, we stopped the mapping process once we had reached an adequate number of mappings to create necessary prototypes.

Domain-specific proficiency scores provided a way to align a student's performance on a Star Math test with what the student could be working on in ST Math. For example, a student with a Star Math scale score of 633 could work on "explaining why a figure is line-symmetric" based on the entering proficiency score for that domain of 635. This particular Star Math skill was aligned with the ST Math game "Ice Cave." Therefore, if a student had not yet already completed this game, it could serve as a reasonable recommendation for what that student could work on next.

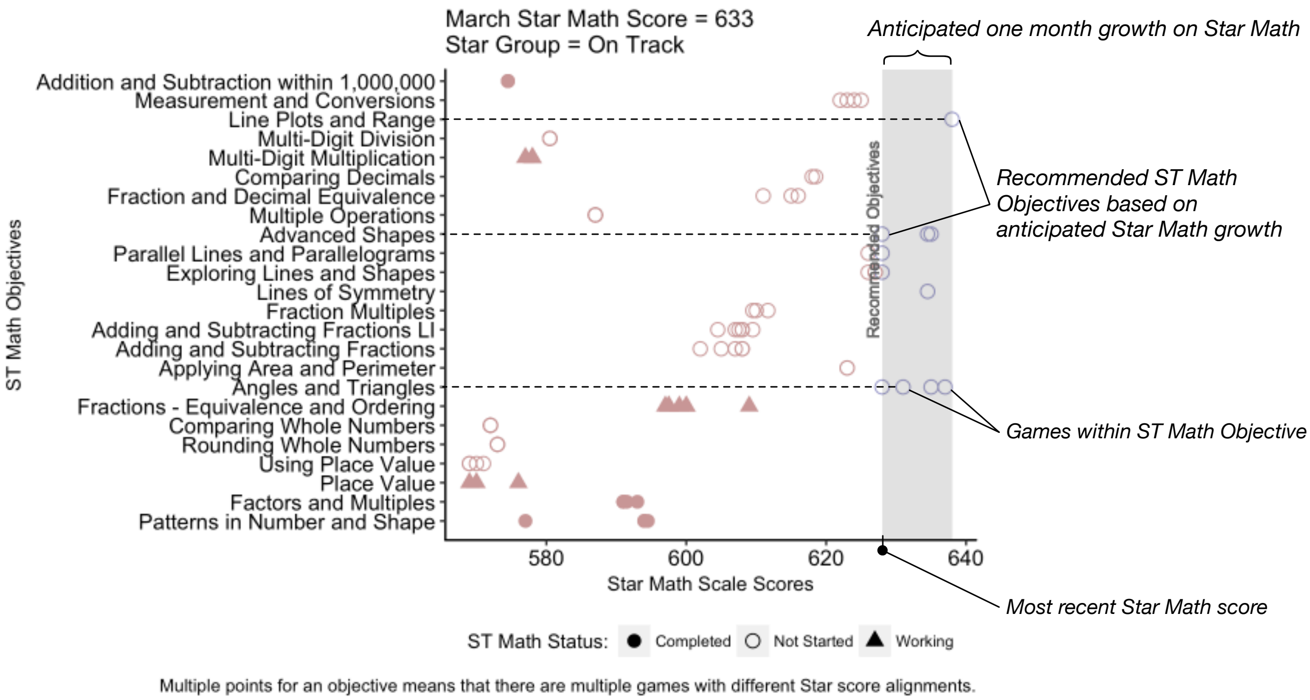

Figure 3. Game recommendations visualization.

4.2.2. Prototype Visualization

After developing the various content mappings between ST Math Grade 4 games and Star Math Grade 4 domains, we iteratively refined a series of visualizations that responded to teacher requests from the first focus group. Figure 3 is a sample visualization given to teachers in the second focus group. Prototypes were developed using R (R Core Team, 2019) and R Markdown (Allaire et al., 2019) files that were rendered in HTML. Figure 3 and Table 1 represent a single student's recent Star Math results and ST Math progress.

From bottom to top, the *y*-axis in Figure 3 organizes ST Math objectives in the order they are presented to students and teachers. We chose objectives over games as a way to organize ST Math content because teachers cannot assign individual games to students whereas they can modify the order of objectives. The *x*-axis presents Star Math scale scores. The scatterplot indicates whether a student has "Completed," "Not Started," or is currently "Working" on games within an ST Math objective. Based on a student's Star Math scale score and projected growth, games and objectives were recommended for that student to work on. This is illustrated by the vertical gray bar, which signified a month's worth of anticipated growth on the Star Math test. Therefore, the student represented in Figure 3 could potentially be working on six different objectives in ST Math based

ISSN 1929-7750 (online).

on her most recent Star score of 633. Beyond making recommendations, a teacher could also view whether a student was making adequate progress in ST Math based on his or her Star Math scores.

Along with Figure 3, we provided teachers with a table that summarized other ST Math-related information such as the pre- and post-quiz scores for an objective, a mean-level score for that objective (i.e., an average of puzzles passed per level, per game), whether or not a student triggered a "hurdle" (i.e., 10 consecutive unsuccessful level attempts), and a flag for whether a student could be working on a given objective based on his or her Star Math score. This extra detail was intended to further support teachers' instructional decision-making on what ST Math content to assign to individuals and/or groups of students.

4.3. Focus Group 2: Reactions to Data Products

Even though we initially had worked with Grade 3 and 4 teachers for the first focus group, we created mappings only for Grade 4 because of time restrictions; thus, only Grade 4 teachers participated in the second focus group. Teachers in this group appreciated having key information (e.g., Star Math score and Star Math grouping) organized together at the top of an individual student report (see Figure 3). While the teachers saw value in the summary information, they wanted "grade-level-equivalent" information attached to the Star Math scale score. Teachers noted that the Star Math reports contextualized scale scores (e.g., a traffic light system of green for "on-target" and yellow for "at risk") beyond what was presented in the partnership's prototype visualization. Thus, as we combined data from the two platforms, we realized that we had dropped information that the teachers valued from Star Math's original report.

Overall, teachers had mixed reactions to the scatterplot in Figure 3. For some teachers, the scatterplot validated their experiences of the complexities involved in aligning ST Math content and standardized assessment results. Other teachers found the visual confusing and struggled with the idea that ST Math and Star Math had different sequencing of content. Teachers stated that they liked the "bottom-to-top" visualization of ST Math objectives, and while teachers saw value in visually tracking student progress in ST Math, they wondered whether their desire to reorder objectives for students would make the visual difficult to use in the long run. Lastly, teachers noted the potential utility of "target lines" or "on-track indicators" that contextualized where students should be in ST Math throughout the year based on pacing guidelines.

Teachers had differing experiences understanding how students significantly above or below grade level were represented in the scatterplot. Students scoring in Grade 5 or higher according to Star Math were placed far off to the right of the scatterplot; students scoring in Grade 3 or lower according to Star Math were placed far off to the left. Some of the teachers understood this limitation of the visualization and found that it validated their prior experiences of some students needing something other than core Grade 4 math instruction. Other teachers, due in part to not understanding the limitation of the visualization, stated that these reports would not be helpful to them in working with students who were significantly above or below grade level. For students working at or near grade level, all teachers liked the gray "recommended" column for how students' Star Math scores could be translated into which parts of the ST Math curriculum they could be working on. Teachers believed that this type of visualization, combined with an "on-target" visualization could provide them with meaningful insight into how ST Math and Star Math related to each another.

For Table 1, all teachers liked the "all-in-one" presentation of ST Math data. They felt that this summarized student progress well and gave them the information they needed from ST Math reports in a single location. Teachers greatly appreciated the inclusion of the "hurdles" data. They reported that in class they relied on the "colour borders" displayed on their screens when students triggered a hurdle, i.e., 10 or more unsuccessful level attempts. However, teachers did not actively track this information over time. Hurdles data resonated with their experience of working with students in the classroom and with helping to identify students who may need additional support. Teachers also liked the "Recommended Objectives" column and understood how the recommendation information connected to the recommendation visualization in the scatterplot.

Looking across both Figure 3 and Table 1 and echoing feedback from the first round of focus groups, teachers consistently voiced a need for data products that could 1) help them form instructional groupings of students and 2) identify content that individuals, groups, or the whole class is struggling with. In general, teachers noted that Figure 3 and Table 1 could, with modifications, be used to help group students more intentionally. For example, teachers noted that using hurdles information could help them identify students who passed an objective but may still need reteaching to ensure mastery. In considering both data products, teachers voiced a desire for the prototyped data reports to be more directive. While the data products were intended to provide teachers with recommendations, teachers wanted more definitive conclusions from the visualizations and tables.

Table 1. Consolidated ST Math data with recommendations

ST Math Obj.	Pre Quiz	Post Quiz	Mean Level Score	Hurdles	Recommended Obj.
Addition and Subtraction within 1,000,000	100	60	59.6		
Line Plots and Range					<–
Multi-Digit Multiplication	0		95.0		
Advanced Shapes					<–
Parallel Lines and Parallelograms					<–
Exploring Lines and Shapes					<–
Lines of Symmetry					<–
Angles and Triangles					<–
Fractions - Equivalence and Ordering	100		33.9		
Place Value	20		79.9		
Factors and Multiples	40	60	63.5		
Patterns in Number and Shape	80	60	71.4		

Date of last login: 2018-02-28

5. Discussion

The organizing idea for this project was that a collaboration among researchers, practitioners, and technology developers could help to overcome semantic and organizational data interoperability challenges. To address semantic interoperability, members of the collaboration worked to understand data generated by each system both in terms of student actions and teachers' intended uses in order to combine data from each system. To address organizational interoperability, the collaboration worked to identify what a teacher would actually do with a data product that combined data from two systems in order to identify what data to attend to as well as any additional data alignments that needed to be developed, such as content mappings between the two systems.

To address teachers' needs related to providing ST Math recommendations based on Star Math scores, a relatively limited amount of data needed to be combined and shared. To develop the visualization depicted in Figure 3, for example, a student's Star Math domain scores and scale scores were combined with a student's progress on ST Math games and objectives. While limited in the volume of data that needed to be blended, data from ST Math needed to be summarized into the categories of "Not Started," "Completed," and "Working" to be combined with Star Math data, which amplifies the importance of understanding the underlying data in order to manipulate it accurately.

The task of using assessment data to make instructional decisions remains conceptually and methodologically challenging (Connor 2019; Halverson, 2010). The collaborative approach adopted for this project sought to address issues of semantic and organizational interoperability through high-bandwidth interactions (Daft & Lengel, 1986) among researchers, practitioners, and developers in order to support instructional decision-making. In combining data and developing prototype data products, several challenges emerged. Some of the challenges may be inherent to the task of making instructional recommendations from assessment data (e.g., Schifter et al., 2014). For example, we observed that Star Math scores do not align in a perfectly linear manner with ST Math games and objectives. In Figure 3, a score of 580 or lower is aligned with both the first and last objectives in the ST Math Grade 4 progression. While the collaboration's mapping between ST Math game content and Star Math domains could have contributed to non-linear alignments, these off-trend pairs may also indicate differences in how MIND and Renaissance view student progression through mathematics content. These differences may prove more challenging for other combinations of instruction- and assessment-focused technologies to overcome (Penuel et al., 2014).

While there are multiple choices that technology companies can make in terms of the standards (e.g., xAPI or Caliper) and infrastructure they use to support data exchange, one piece of infrastructure for future data interoperability projects going forward may be data analysis tools that support the kind of semantic interoperability work that was central to this collaboration. LearnSphere, for example, represents a set of tools for analyzing learner event data so that multiple individuals can develop shared mental models. TeamSpace, which was developed by researchers at Digital Promise, is a cloud-based analytics workspace that helps educational organizations securely and collaboratively analyze complex data coming from one or more digital environments (Krumm, Roschelle, & Schank, 2019). Given the importance of communicating the meaning behind events and drawing on multiple individuals' knowledge and expertise, LearnSphere and TeamSpace could play an important role in helping researchers, practitioners, and developers make sense of different data sources. Lastly, the kinds of collaboration described in this paper benefited from the lessons of previously successful interoperability initiatives — address real problems experienced by practitioners in schools and universities (e.g., Learning Tools Interoperability and single sign-on). The problem of joining data from Star Math and ST Math to support teachers' instructional decision-making helped in shaping what data to attend to as well as ideas for how to manipulate and visualize data. Therefore, the concreteness of addressing specific problems may help other systems overcome semantic and organizational challenges while taking advantage of technological interoperability advances (e.g., xAPI and Caliper).

The collaboration described in this paper provides an example of both the semantic and organizational work that may need to take place in order to further accelerate multiple technical advancements. The collaboration, however, addressed semantic interoperability more fully than organizational interoperability, and as a single case, the findings from this collaboration may not generalize to other partnerships or interoperability projects. Moreover, the importance of higher-bandwidth interactions and aligning content between systems may be unique to the instruction- and assessment-focused technologies described in this paper. While a single case, the problem-focused, collaborative approach used in this project may be worth replicating given the lessons learned from this collaboration and previous efforts to join data from across multiple systems.

6. Conclusion

This project set out to understand how two learning technologies could share data with each another in order to develop data products that could support teachers' instructional decision-making. In working to blend data across two technologies, this project identified important obstacles and the role that a collaboration among researchers, practitioners, and developers can play in tackling data interoperability challenges. While data interoperability is often framed in technical terms, this project highlighted the critical role that mapping content between instruction-focused and assessment-focused technologies can play. While standards for transporting data between systems continue to mature, developing the kinds of data products demonstrated in this paper required many individuals to make sense of multiple, distinct data streams and work with teachers to understand their needs for data products that combine data from multiple systems. Identifying ways to make these kinds of collaborations more efficient and effective could serve as an important problem to solve for those seeking to advance data interoperability.

Declaration of Conflicting Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The publication of this article received financial support from Renaissance Learning to SRI International who coordinated the collaboration, generated prototypes, and conducted focus groups.

Acknowledgements

The collaboration described in this paper was unique, requiring the extra — often unplanned — efforts of many at the MIND Research Institute and Renaissance Learning. Lastly, we wish to acknowledge the teachers and leaders at the participating school who provided their valuable time, energy, and insights.

References

Allaire, J. J., Horner, J., Xie, Y., Marti, V., & Porte, N. (2019). *Markdown: Render Markdown with the C Library "Sundown."* R package version 1.1. https://CRAN.R-project.org/package=markdown

Baker, R. S., & Siemens, G. (2014). Educational data mining and learning analytics. In K. Sawyer (Ed.), *Cambridge handbook of the learning sciences*, 2nd ed., pp. 253–272. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.016

Bakharia, A., Kitto, K., Pardo, A., Gašević, D., & Dawson, S. (2016). Recipe for success: Lessons learnt from using xAPI within the connected learning analytics toolkit. *Proceedings of the 6th International Conference on Learning Analytics and Knowledge* (LAK '16), 25–29 April 2016, Edinburgh, UK (pp. 378–382). New York: ACM. https://doi.org/10.1145/2883851.2883882

Connor, C. M. (2019). Using technology and assessment to personalize instruction: Preventing reading problems. *Prevention Science*, *20*(1), 89–99. https://doi.org/10.1007/s11121-017-0842-9

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science*, *32*(5), 554–571. https://doi.org/10.1287/mnsc.32.5.554

del Blanco, A., Serrano, A., Freire, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2013). E-Learning standards and learning analytics: Can data collection be improved by using standard data models? *Proceedings of the 2013 IEEE Global Engineering Education Conference* (EDUCON 2013), 13–15 March 2013, Berlin, Germany (pp. 1255–1261). Washington, DC: IEEE Computer Society. https://doi.org/10.1109/EduCon.2013.6530268

DiCerbo, K. E., & Kobrin, J. (2016). Communicating assessment results based on learning progressions. Paper presented at the American Educational Research Association Annual Conference (AERA 2016), 8–12 April 2016, Washington, DC, USA.

Dietze, S., Yu, H. Q., Giordano, D., Kaldoudi, E., Dovrolis, N., & Taibi, D. (2012). Linked education: Interlinking educational resources and the Web of data. *Proceedings of the 27th Annual ACM Symposium on Applied Computing* (SAC '12), 26–30 March 2012, Riva (Trento), Italy (pp. 366–371). New York: ACM. https://doi.org/10.1145/2245276.2245347

Halverson, R. (2010). School formative feedback systems. *Peabody Journal of Education*, *85*(2), 130–146. https://doi.org/10.1080/01619561003685270

Hao, J., Smith, L., Mislevy, R., von Davier, A., & Bauer, M. (2016). *Taming log files from game/simulation-based assessments: Data models and data analysis tools*. Princeton, NJ: ETS Research Report (No. RR-16-10).

Koedinger, K. R., D'Mello, S., McLaughlin, E. A., Pardos, Z. A., & Rosé, C. P. (2015). Data mining and education. *WIREs Cognitive Science, 6*(4), 333–353. https://doi.org/10.1002/wcs.1350

Krumm, A. E., Roschelle, J., & Schank, P. (2019). Analytics as a team sport: Using cloud-based tools to support data-intensive research–practice partnerships. Workshop at the 9th International Conference on Learning Analytics and Knowledge (LAK '19), 4–8 March 2019, Tempe, AZ, USA. https://circlcenter.org/events/teamspace-lak19/

Krumm A. E., Means, B., & Bienkowski, M. (2018). *Learning analytics goes to school: A collaborative approach to improving education*. New York: Routledge.

Manso-Vazquez, M., Caeiro-Rodriguez, M., & Llamas-Nistal, M. (2018). An xAPI application profile to monitor self-regulated learning strategies. *IEEE Access*, *6*, 42467–42481. https://doi.org/10.1109/ACCESS.2018.2860519

Martin, B., Mitrovic, A., Koedinger, K. R., & Mathan, S. (2011). Evaluating and improving adaptive educational systems with learning curves. *User Modeling and User-Adapted Interaction, 21*, 249–283. https://doi.org/10.1007/s11257-010-9084-2

Murphy, R., Snow, E., Mislevy, J., Gallagher, L., Krumm, A. E., & Wei, X. (2014). *Blended learning report*. Menlo Park, CA: SRI Education.

Owen, V. E., Ramirez, D., Salmon, A., & Halverson, R. (2014). Capturing learner trajectories in educational games through ADAGE (Assessment Data Aggregator for Game Environments): A click-stream data framework for assessment of learning in play. Paper presented at the American Educational Research Association Annual Conference (AERA 2014), 3–7 April 2014, Philadelphia, PA, USA.

Pagano, P., Candela, L., & Castelli, D. (2013). Data interoperability. *Data Science Journal, 12*. https://doi.org/10.2481/dsj.GRDI-004

Penuel, W. R., Confrey, J., Maloney, A., & Rupp, A. A. (2014). Design decisions in developing learning trajectories–based assessments in mathematics: A case study. *Journal of the Learning Sciences*, *23*(1), 47–95. https://doi.org/10.1080/10508406.2013.866118

Penuel, W. R., Fishman, B. J., Haugan Cheng, B., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher, 40*(7), 331–337. https://doi.org/10.3102/0013189X11421826

R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Schifter, C., Natarajan, U., Ketelhut, D. J., & Kirchgessner, A. (2014). Data-driven decision-making: Facilitating teacher use of student data to inform classroom instruction. *Contemporary Issues in Technology and Teacher Education, 14*(4), 419–432. https://citejournal.org/volume-14/issue-4-14/science/data-driven-decision-making-facilitating-teacher-use-of-student-data-to-inform-classroom-instruction

Sottilare, R. A., Long, R. A., & Goldberg, B. S. (2017). Enhancing the Experience Application Program Interface (xAPI) to improve domain competency modeling for adaptive instruction. *Proceedings of the 4th ACM Conference on Learning @ Scale* (L@S 2017)*,* 20–21 April 2017, Cambridge, MA, USA (pp. 265–268). New York: ACM. https://doi.org/10.1145/3051457.3054001

State Educational Technology Directors Association. (2018). *State education leadership interoperability: Leveraging data for academic excellence*. Retrieved from http://www.setda.org/master/wp-content/uploads/2018/05/State-Leadership-Interoperability.pdf

Winne, P. H., & Hadwin, A. F. (1998) Studying as self-regulated learning. In D. J. Hacker & J. Dunlosky (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ: Erlbaum.