I-SMART Goal 3 Technical Report: Teacher Dashboard Usability and Utility Study

Robert P. Dolan
Samantha M. Gilbert
Kim Ducharme
Cara Wojcik
Allison Posey
CAST

Preferred Citation: Dolan, R. P., Gilbert, S. M., Ducharme, K., Wojcik, C., Posey, A., CAST. (2022). *I-SMART goal 3 technical report: Teacher dashboard usability and utility study*. Accessible Teaching, Learning, and Assessment Systems (ATLAS), the University of Kansas.

The project described in this report was developed under a grant from the U.S. Department of Education. However, the content does not necessarily represent the policy of the U.S. Department of Education, and you should not assume endorsement by the Federal government.

Introduction

In this report, we describe a usability and utility study conducted with a prototype teacher score-reporting dashboard for supporting teachers' effective interpretation of—and instructional planning based upon—student performance on science testlets. The design and development of this dashboard is described in the "Goal 3: Teacher Dashboard Design" companion report, which documents the process of researching and designing the user experience, interaction, and interface for the dashboard through an iterative codesign process.

Project Background

Goal 3 of the I-SMART research project focused on developing a score-reporting dashboard to support teachers' interpretation and use of students' science testlet results and provide evidence for any gaps in students' NGSS-aligned knowledge, skills, and understandings (KSAs). These science testlets, per the broader goals of I-SMART, are designed to be instructionally embedded and based on learning map models as expounded by the Dynamic Learning Maps® (DLM®; Kingston et al., 2016) assessment approach.

While this dashboard has been designed primarily for use with alternate assessment of students with the most significant cognitive disabilities, we are also considering its utility for students with and without disabilities who fail to meet grade-level academic achievement standards in science. As such, the implications of this research apply to I-SMART, DLM, and instructionally embedded and learning map model-based assessments—both general and alternate.

Formative assessment can be most comprehensively defined as "a planned, ongoing process used by all students and teachers during learning and teaching to elicit and use evidence of student learning to improve student understanding of intended disciplinary learning outcomes and support students to become self-directed learners" (Council of Chief State School Officers [CCSSO], 2018, p. 2). Properly implemented, formative assessment can provide a means for effectively differentiating instruction to diverse learners in a science classroom.

Ideally, teachers receive comprehensive, pre-service formative assessment training to develop and use their own items, rubrics, scoring procedures, analyses, and decision support tools (Black & Wiliam, 1998; Heritage, 2010). However, the processes needed for effective formative data analysis are time consuming and difficult to develop (Bill and Melinda Gates Foundation, 2015). That time and energy might be better spent preparing lessons, interventions, and working directly with students—unless teachers are provided with coherent tools to support formative assessment processes that streamline the data-to-intervention process and offer better just-in-time support for real-time differentiation (Wylie & Dolan, 2013). In reality, teachers often rely on third-party content and tools to drive student data–based instructional decision-making. Where such assessment tools fall short is in providing teachers with interpretable and actionable information about their students' KSUs.

Appendix A: Terminology defines key terminology central to DLM and I-SMART projects.

Teacher Dashboard Design

If we are to support teachers and students in conducting effective formative assessment using student results from external—not teacher-developed—instructionally embedded instruments, student data must be presented in a manner that complements existing formative assessment practices. We hypothesize that use of I-SMART's learning map neighborhoods could enhance such formative use of student results data. Results of a needs assessment conducted through a series of focus groups with educators indicated a need to provide options and flexibility to interpret and engage with assessment results in both student and class aggregated views.

These findings were the basis of a rigorous codesign process with educators (Dolan et al., 2020). This iterative design process focused on applying Universal Design for Learning (UDL) principles to provide teachers with multiple means for engagement, representation, and action and expression (CAST, 2018) during dashboard use—rather than a one-size-fits-all design—as well as scaffolding teachers on more difficult processes, such as interpreting dynamic learning maps (Kingston et al., 2016).

Researchers enlisted a cadre of thirteen special and general educators to iteratively design the dashboard through a set of virtual design team meetings. Redesigns were based on data including thematic analysis of open-ended feedback and quantitative evidence of navigability and utility. The final dashboard design was created to provide teachers with a flexible environment for supporting data-driven instructional decision-making and includes graphic displays of skill mastery in learning map model and tabular formats. The architecture of the dashboard consists of six main functional areas (i.e., screens), with three providing roster-level information and three providing individual student-level information.

These final designs were used to develop a prototype dashboard within the Kite Suite (https://kiteassessments.org). Due largely to system constraints, there were some compromises to user interface design in the final prototype implementation, including (a) use of unfamiliar Essential Element (EE) codes rather than familiar unit names, (b) lack of visual indication of element clickability, (c) inability to reliably zoom in and out of learning map model views, and (d) lack of a key to explain graphic elements. The Learning Map views also incorrectly depicted which nodes were included in the testlets at each linkage level, as discussed later. Despite these differences, overall user experience functionality paralleled that of the original design.

The following sections describe the six screens. Two screens show the final versions implemented in the prototype.

Roster Overview Screen

The Roster Overview screen provides the landing page and navigational fulcrum for the teacher dashboard. It is designed to support teachers in spotting patterns among student performance and instructional status. Figure 1 shows the Roster Overview

screen as designed. Figure 2 shows the Roster Overview screen as implemented in the prototype.

Figure 1

Roster Overview Screen as Designed

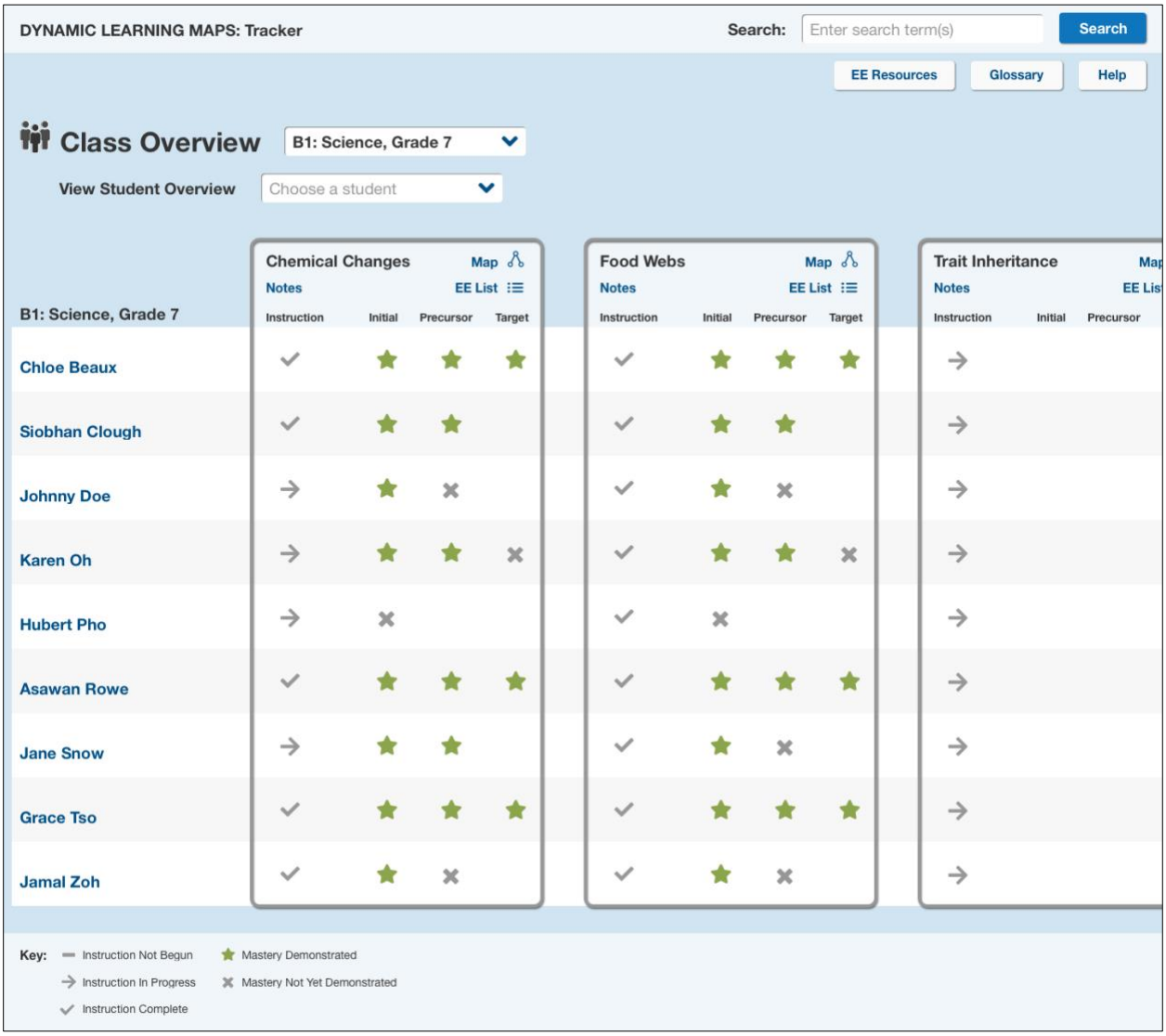
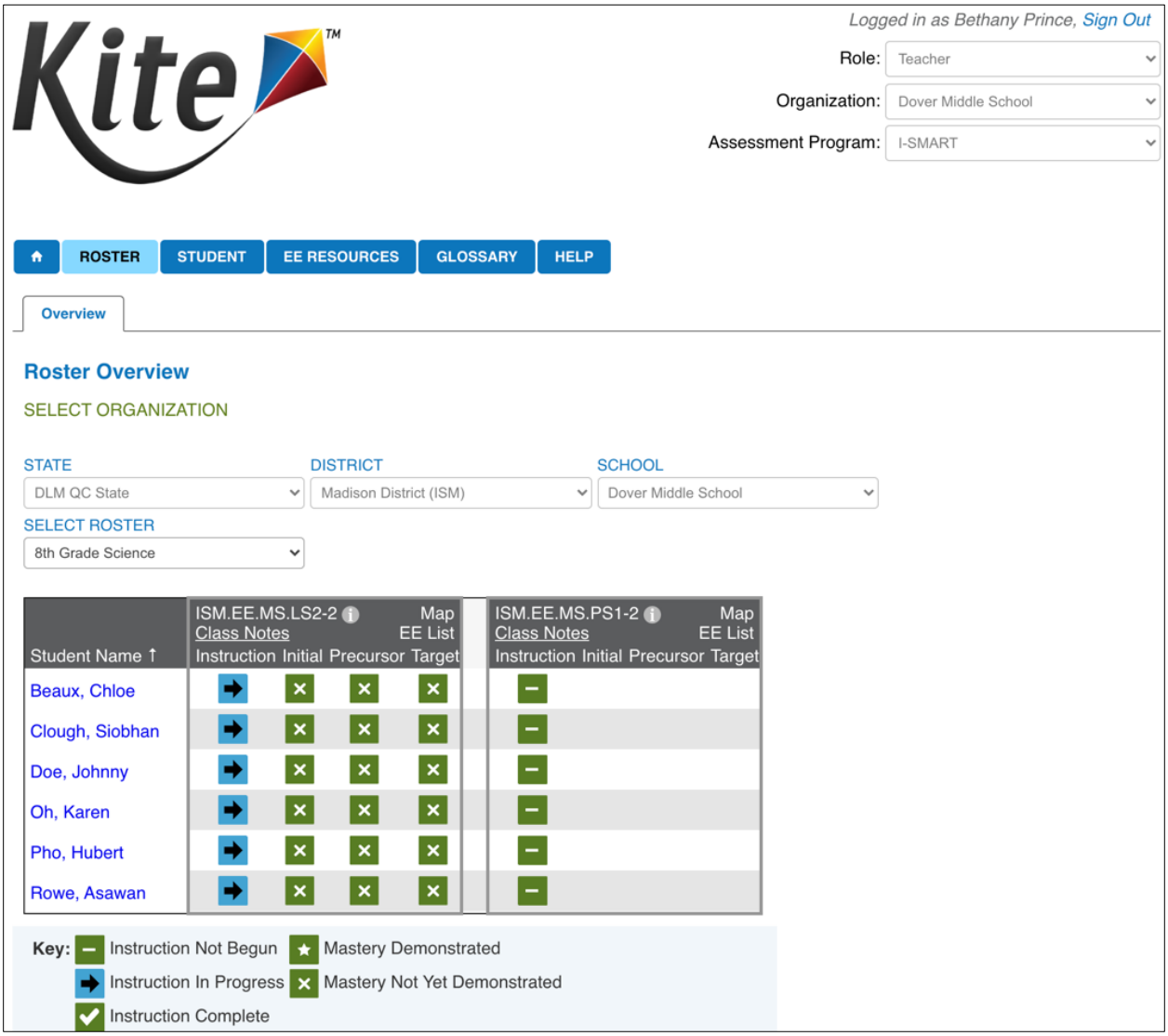


Figure 2

Roster Overview Screen as Implemented in the Prototype

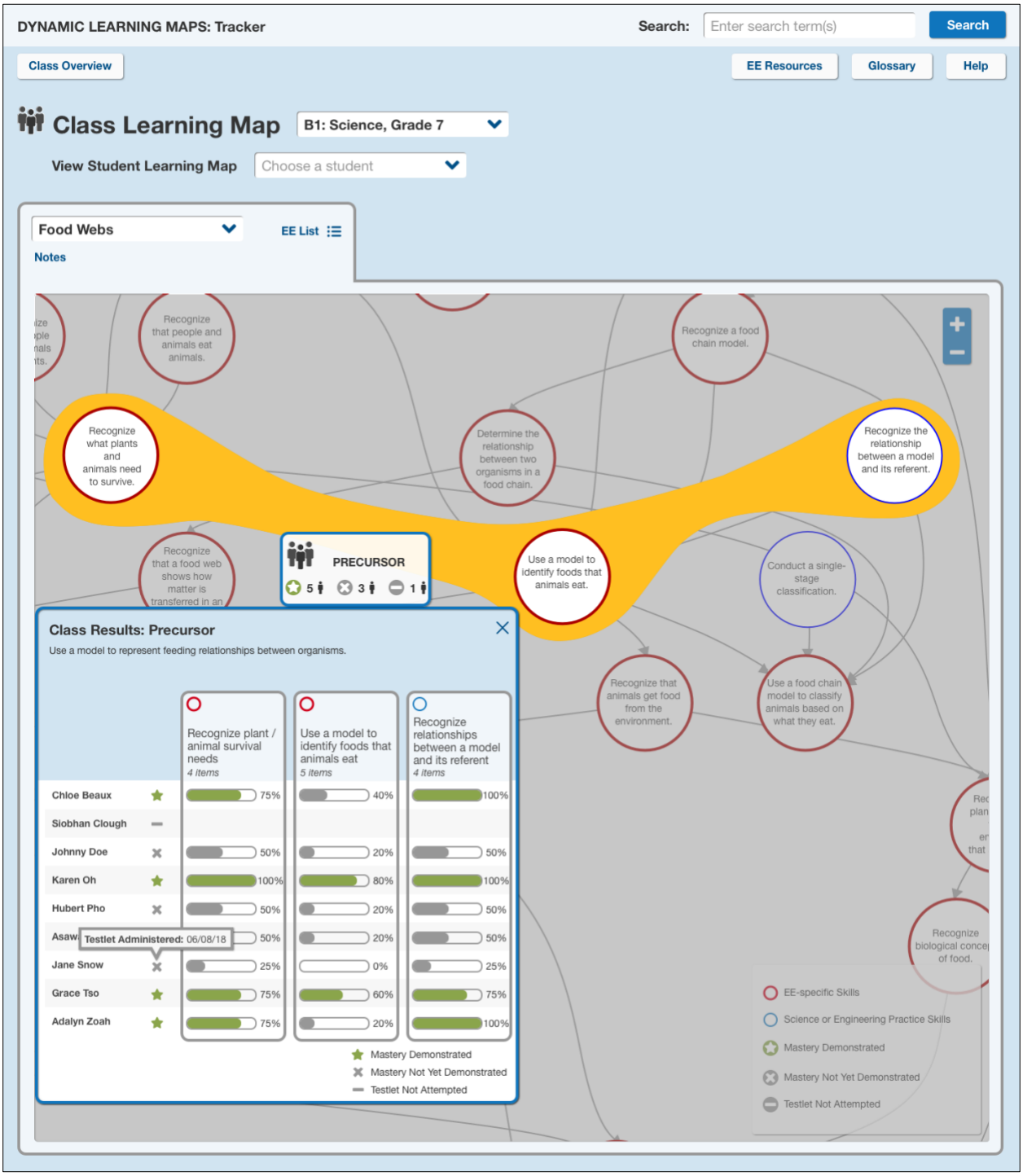


Roster Learning Map Screen

The Roster Learning Map screen (Figure 3) provides an interactive view of the learning map neighborhood for one EE, and depicts students' testlet performance at each linkage level, with the ability to click on nodes to obtain details for that node and drill down into a testlet for additional student- and node-specific performance details. The screen is designed to support teachers in whole-roster instructional planning.

Figure 3

Roster Learning Map Screen as Designed



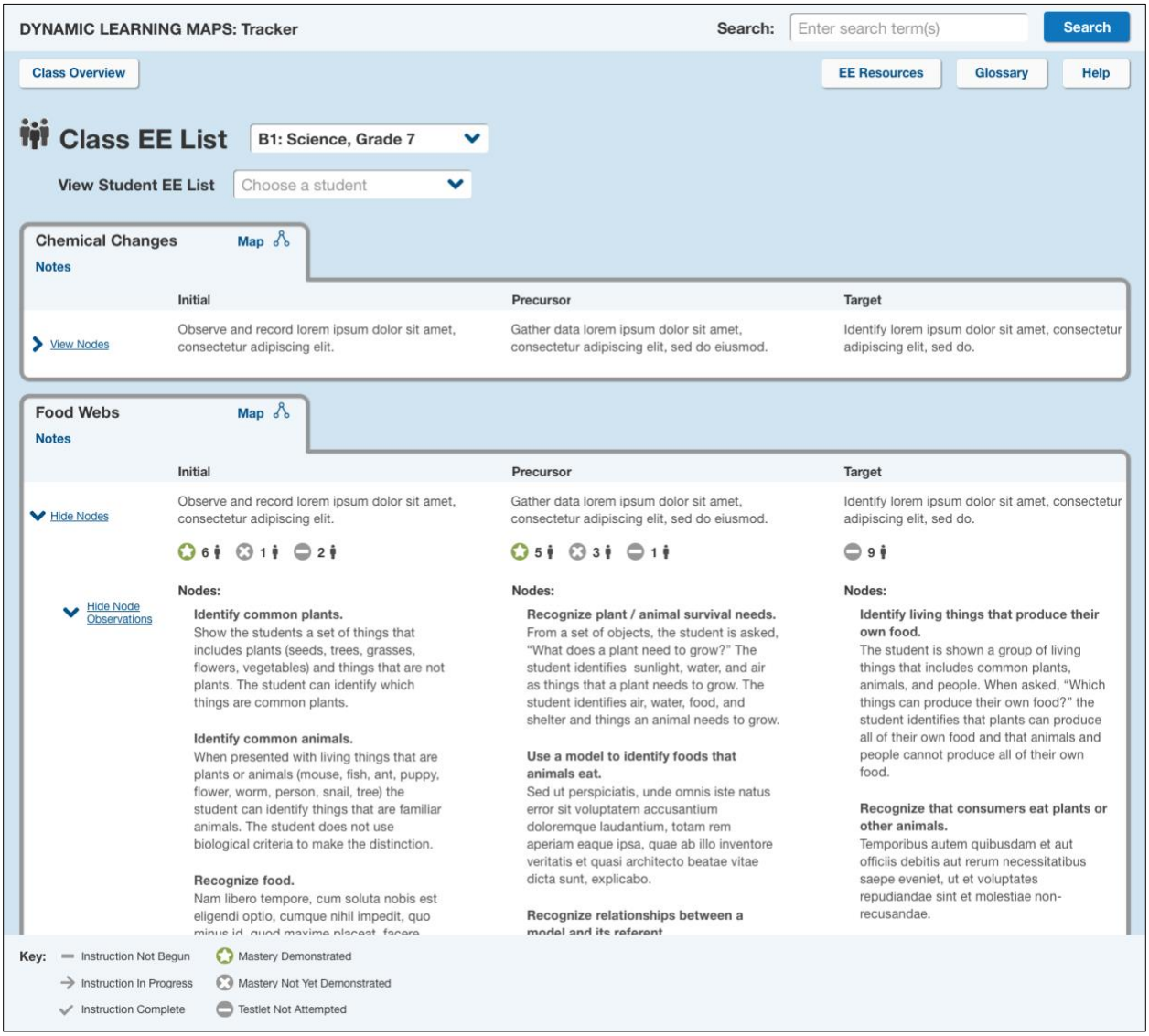
Roster Essential Element List Screen

The Roster Essential Element List screen (Figure 4; depicted as Class EE Screen in the dashboard design screens) provides teachers access to full information for each EE,

linkage level, and tested node, atop of which is layered students' testlet performance. It provides an alternative way than the learning map view to contextualize student performance. Show/hide (i.e., expand and collapse interface) helps decrease complexity to allow teachers to see "both the forest and the trees." As with the Roster Learning Map screen, it is designed to support teachers in whole-roster instructional planning.

Figure 4

Roster EE Screen as Designed

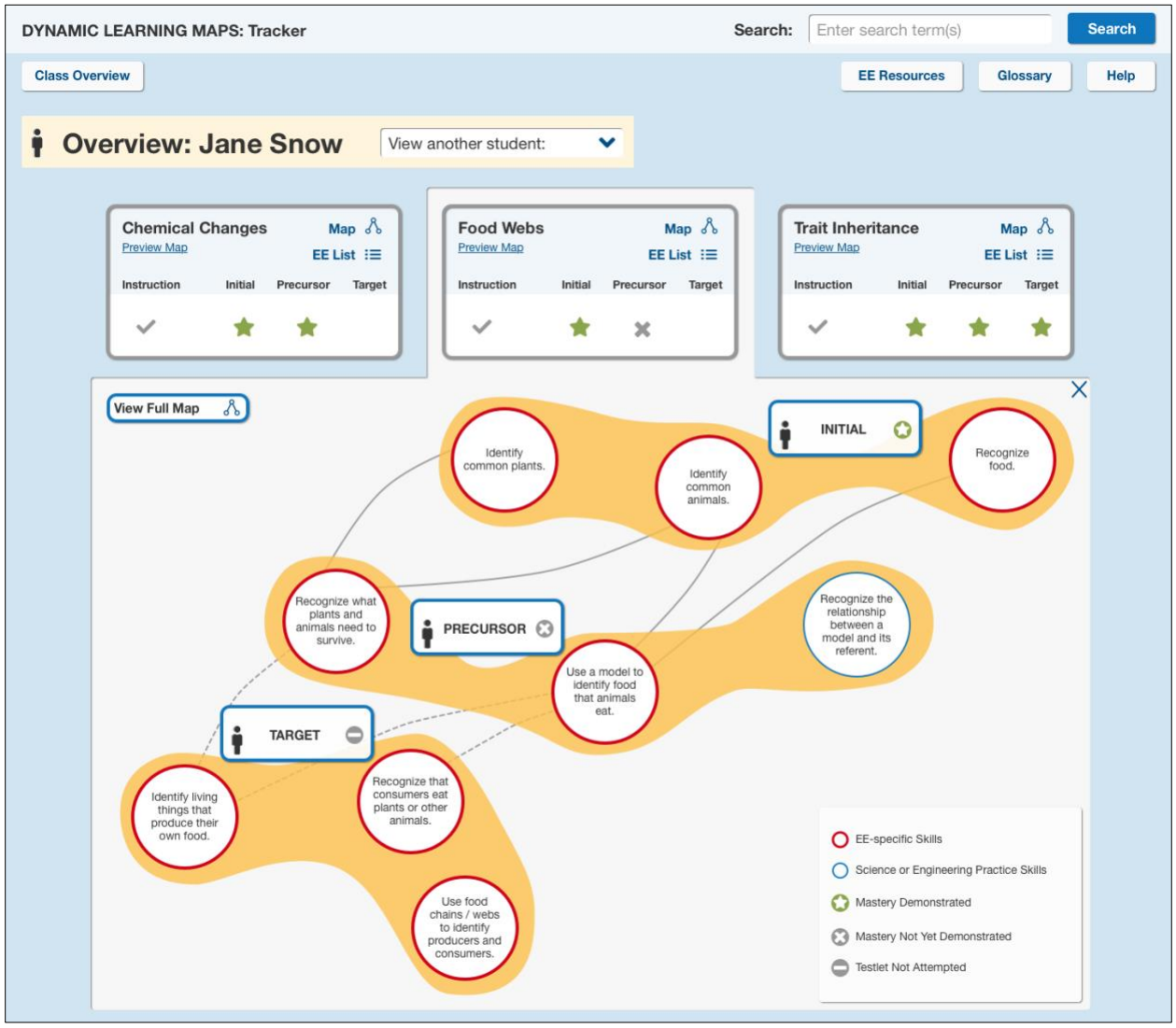


Student Overview Screen

The Student Overview screen (Figure 5) was designed to provide teachers with a broad overview of a single student's testlet performance and support teachers in identifying and understanding a student's achievement across all EEs in aggregate. The screen serves as a gateway to the Learning Map. A map preview featured a small section of the EE neighborhood map with visual indications of student performance by testlet and node.

Figure 5

Student Overview Screen as Designed



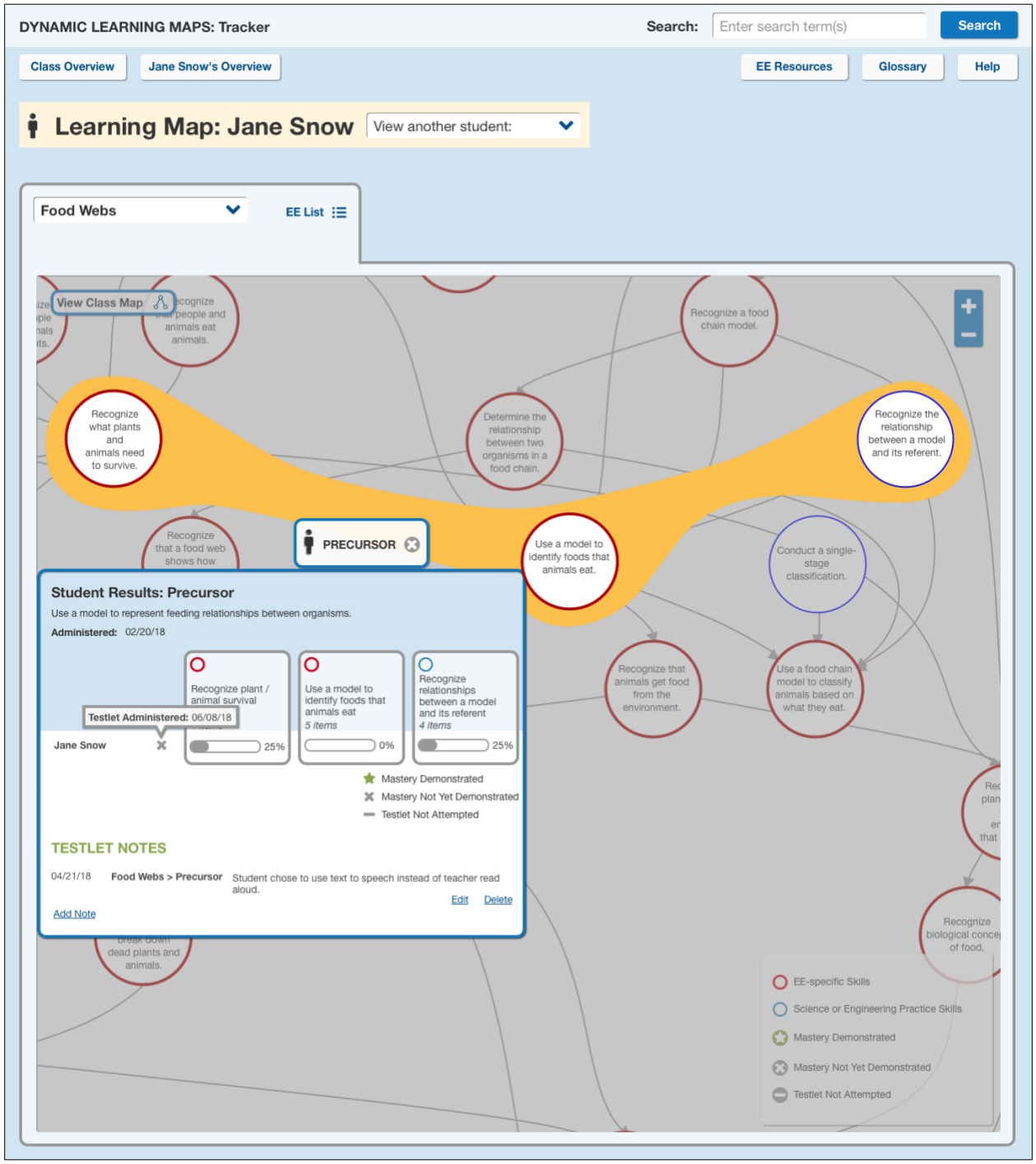
Student Learning Map Screen

The Student Learning Map screen (Figure 6) is similar to the Roster Learning Map screen but is limited to data from an individual student. By providing testlet administration details, such as the date of administration, node-specific performance,

and context notes recorded by the teacher, it is designed to support teachers in student-specific instructional planning.

Figure 6

Student Learning Map Screen as Designed



Student Essential Element List Screen

The Student Essential Element List screen is similar to the Roster EE List screen but limited to data from an individual student. As with the Roster EE List screen, it provides an alternative and more traditional view of testlet performance together with details of each EE to support teachers in student-specific instructional planning. Figure 7 shows the Student Essential Element List screen design and Figure 8 shows the Student Essential Element List screen as implemented in the prototype.

Figure 7

Student EE List Screen as Designed

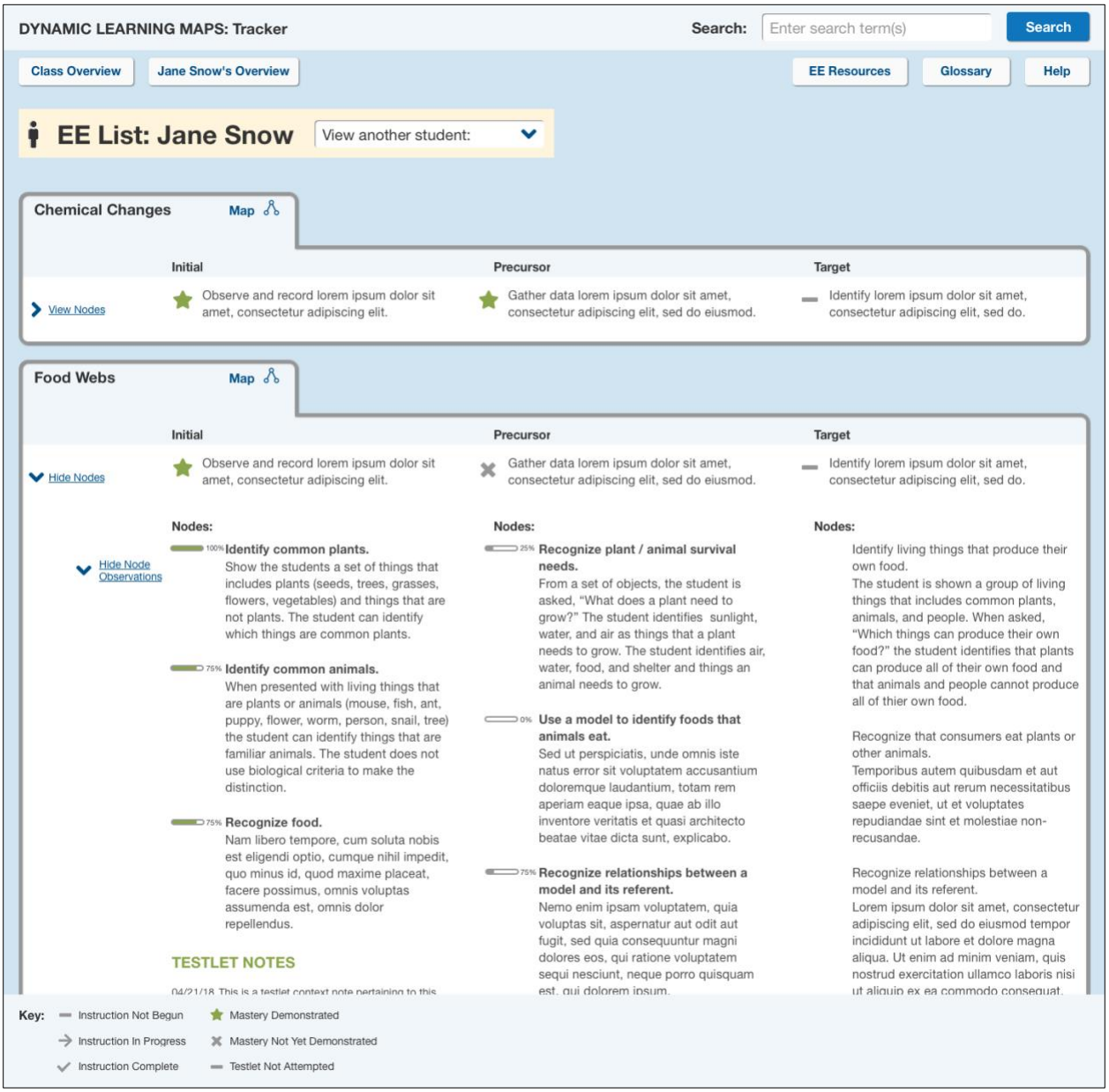
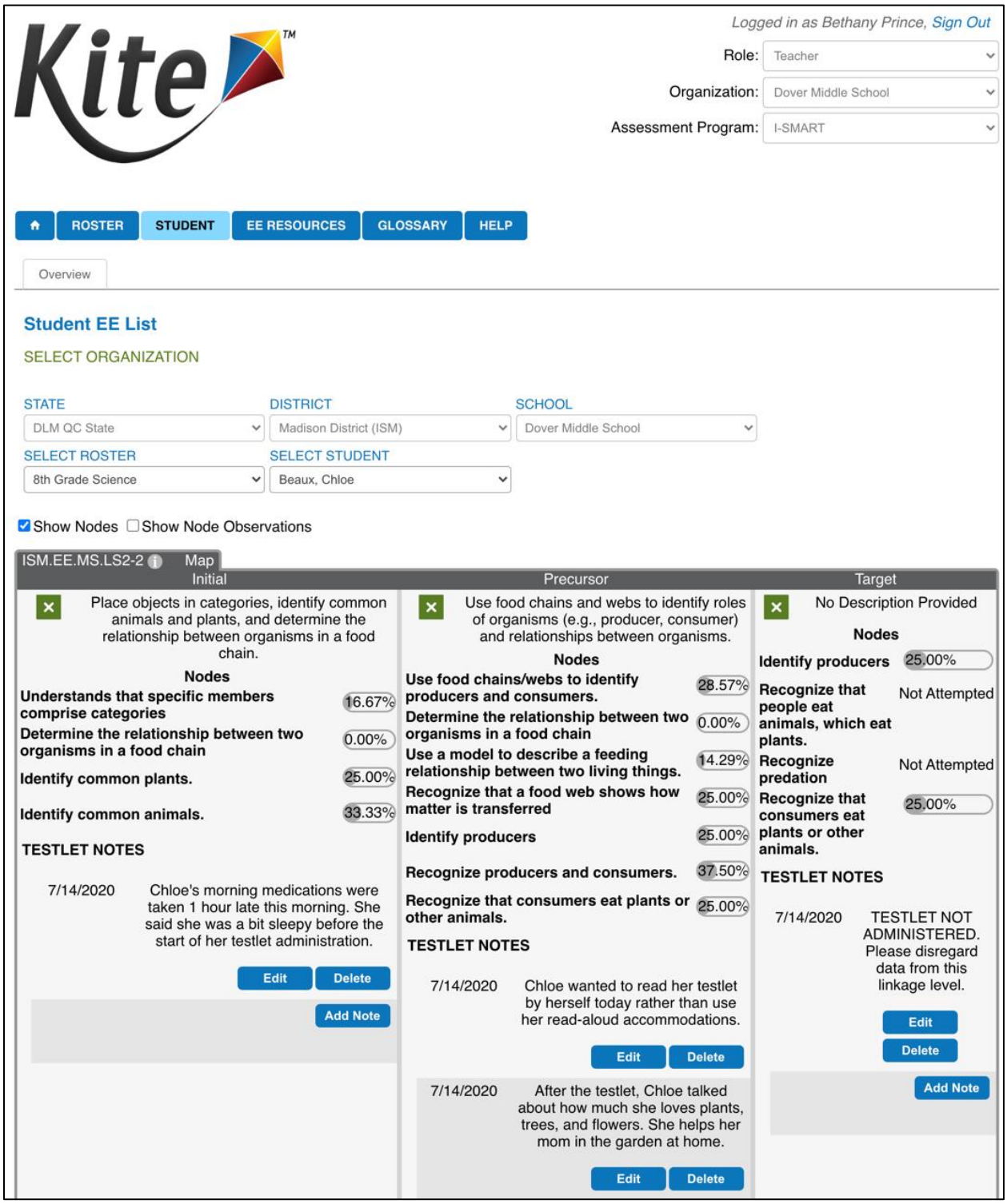


Figure 8

Student EE List Screen as Implemented in the Prototype



Current Study

To evaluate the potential for this teacher dashboard design to support effective formative assessment, we conducted a small-scale usability and utility study of a prototype teacher dashboard with elementary, middle, and high school special education teachers who teach science. Dashboard usability was evaluated in terms of teachers' ability to navigate the dashboard user interface and access student data. Dashboard utility was evaluated in terms of how teachers would be able to use the dashboard to support instructional planning and formative assessment by addressing the following research questions:

1. Do teachers believe the **dashboard provides an effective means for understanding learning map models**?
2. Do teachers believe **learning map models *in general* can effectively support instructional decision-making** by guiding choice of appropriate instructional pathways for students?
3. Do teachers believe the **dashboard conveys student data** in a means conducive to informing instructional decision-making?
4. Do teachers believe that the **dashboard would improve their ability to make instructional decisions**? If so, what kinds of decisions and how does the dashboard improve their ability to make them?
5. **Can teachers be trained** to make effective use of the dashboard in informing instructional decision-making? If so, what type of training?

Methods

Recruitment

Four of the five I-SMART partner states provided participants for this study. These states released a short summary and a one-page recruitment letter to potential candidates detailing a study overview and eligibility. Each state had their own recruitment system; one state recruited teachers via a listserv, whereas other states emailed individuals directly. Candidates emailed the I-SMART team directly with their interest and experience.

Researchers selected 17 teachers across four of the five I-SMART partner states and representing the three grade bands. As shown in

Table 1, all teachers had experience teaching science to students with significant cognitive disabilities, with most having also taught science in inclusive general education settings. All teachers had administered DLM alternate assessments, although with a range of experience. None had participated in the codesign process.

Table 1

Teacher Participant Demographics Showing Raw Counts for the 17 Participants

Demographic	*n*
State	
Maryland	3
Missouri	12
New York	1
Oklahoma	1
School setting	
Rural	7
Suburban	2
Urban	8
Primary grade band	
Elementary	6
Middle school	5
High school	6
Science DLM administration experience	
Advanced	3
Intermediate	9
Novice	5
None	0
General education teaching experience	
Yes	14
No	3
Teaching training experience	
Yes, formal	6
Yes, informal	4
No	3
Unknown	4

Study Components

The study was conducted with each teacher over a period of two to three weeks and consisted of the following six components conducted in order: (a) pre-study survey, (b) training video, (c) dashboard usability session, (d) self-guided exploration, (e) dashboard utility session, and (f) post-study survey.

Pre-Study Survey

Teachers were administered an online survey (see Appendix B: Pre-Session Survey) to evaluate their attitudes and beliefs about formative assessment, data-driven instructional decision-making, and the role that learning map models might play.

Training Video

Teachers independently viewed an eight-minute training video prepared as a refresher on DLM learning map model concepts and terminology.

Dashboard Usability Session

Teachers participated in the first of two virtual, one-hour recorded video sessions conducted using Zoom. The first session evaluated dashboard usability and was conducted according to a think-aloud protocol (Dolan et al., 2011; Leighton, 2017). Teachers had access to the dashboard prototype populated with mock student data at the grade bands at which they currently taught. Teachers were asked to complete a series of tasks (see Appendix C: Usability and Utility Sessions Protocol) and scored on their ability according to the following rubric:

- able to complete independently and readily;
- able to complete independently with some effort;
- able to complete with minimal prompting;
- able to complete with extensive prompting; or
- unable to complete, required researcher intervention.

Scoring reliability was established through spot checking by another member of the research team. In addition, teachers were asked to “think aloud” as they completed the tasks. Two members of the I-SMART research team conducted each session, with one leading the session and the other observing and taking notes.

Self-guided Exploration

Teachers were provided access to the dashboard prototype to explore independently and were asked to consider the following questions:

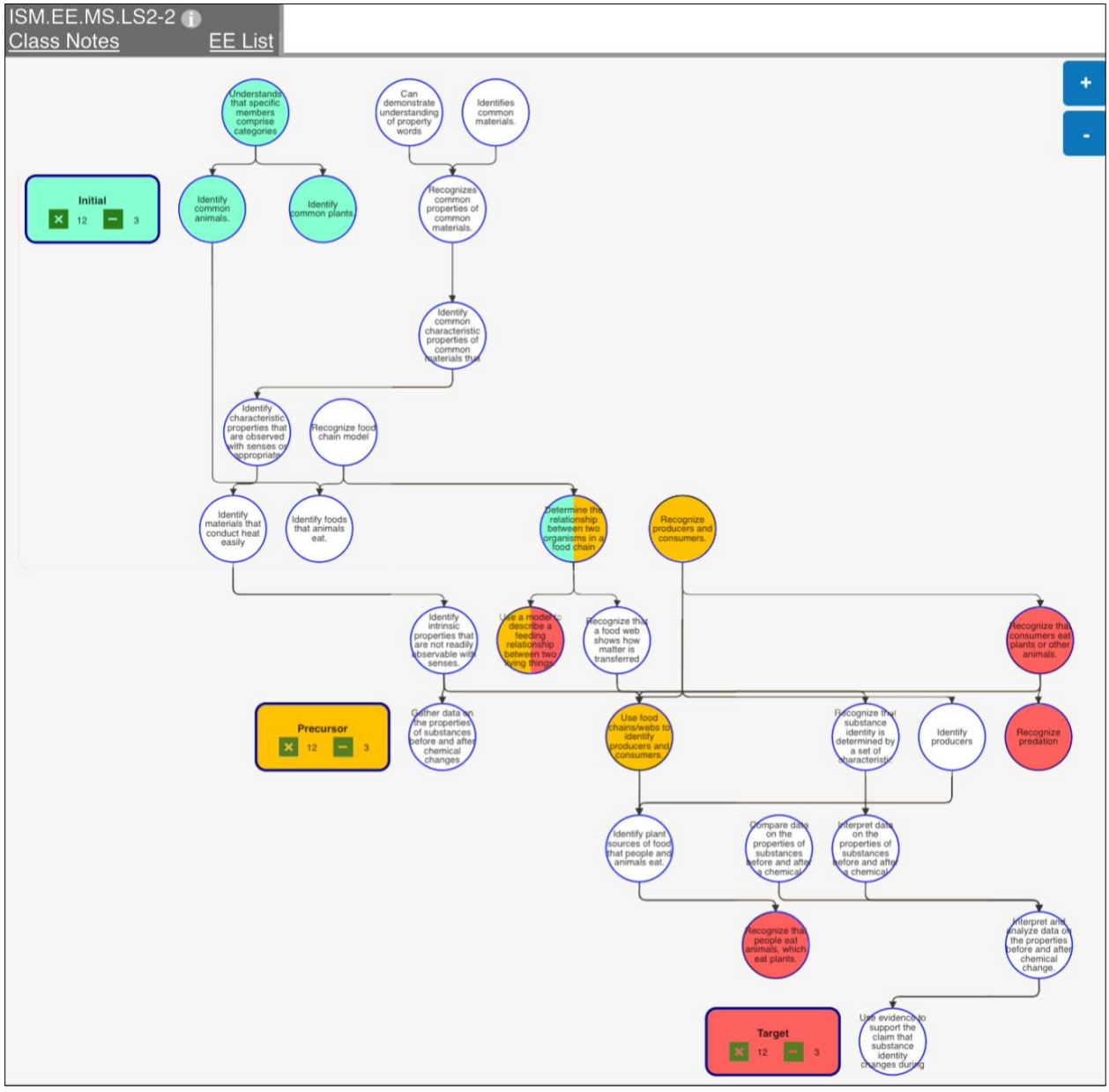
1. How might you use the dashboard to help you with instructional planning?
2. Would you use the dashboard to examine data from a student’s previous year for instructional planning this year?
3. Would you use the dashboard during meetings with students? Parents or guardians? Other educators?

4. How might the dashboard be useful in inclusive or general education settings, for students who take general or alternate assessments?

During the time between sessions, teachers were also shown a static image depicting a redesign of the learning map model (see Figure 9) and asked to consider its clarity compared to the original learning map models (that incorrectly depicted the nodes in each linkage level) in the dashboard prototype used in usability sessions.

Figure 9

Learning Map Model View Redesigned but Not Implemented in the Prototype



Note: The colors in the model represent tested nodes at the Initial (green), Precursor (orange), and Target (red) linkage levels. White nodes represent non-tested nodes.

Dashboard Utility Session

Teachers participated in the second session, conducted in a similar manner as the first, but in accordance with a cognitive lab protocol to evaluate dashboard utility (Dolan et al., 2011; Leighton, 2017; see Appendix C: Usability and Utility Sessions Protocol). During this session, teachers were asked to consider how they might use the dashboard in the classroom, and the extent to which its use could help them with instructional decision-making. They did this by logging in to the online dashboard prototype and completing a series of tasks using the dashboard with mock student data.

Post-Study Survey

Teachers were administered an online survey (see Appendix D: Post-Session Survey) to evaluate changes in attitudes and beliefs and their experience with the dashboard prototype.

Analysis

All survey, usability session, and utility session data were recorded and catalogued for analysis. For the usability analysis, usability session data were analyzed quantitatively to calculate critical error, error-free, and scaffolding rates. For the utility analysis, we applied a qualitative content analysis approach using survey, usability session, and utility session data, in which we developed a set of thematic codes based on our research questions and on additional emergent themes uncovered through the analysis process. Table 2 summarizes the sources of information used to provide evidence for each of the dashboard utility research questions.

Table 2

Sources of Evidence Used to Evaluate Research Questions Pertaining to Dashboard Utility

Research question	Source of evidence		
	Usability session	Utility session	Surveys
1: Dashboard support of understanding of learning map models	✓	✓	
2: Learning map model utility		✓	✓
3: Dashboard conveying of data to support instructional decision-making	✓	✓	✓
4: Dashboard improving instructional decision-making		✓	✓
5: Dashboard training		✓	

Results

Results will be described first in terms of dashboard usability, then in terms of dashboard utility by addressing each of the five research questions as well as emergent findings.

Dashboard Usability

Dashboard usability data were collected during the usability session. Teachers' ability to complete tasks independently in each of the dashboard's six main functional areas (Overview, Learning Map, and EE List; each for roster and student data) is shown in Table 3 with data aggregated across all teachers. Overall, teachers were able to complete 82% of tasks independently and another 10% of tasks with minimal prompting. In 8% of cases, teachers were unable to complete the tasks without extensive prompting or intervention by researchers. Since the nature and complexity of the tasks differed across roster and student views, and since learning transfer likely occurred across tasks, these findings provide only initial indicators of usability challenge areas.

Table 3

Usability Results Table Showing Level of Independence with Which Teachers Were Able to Complete Tasks

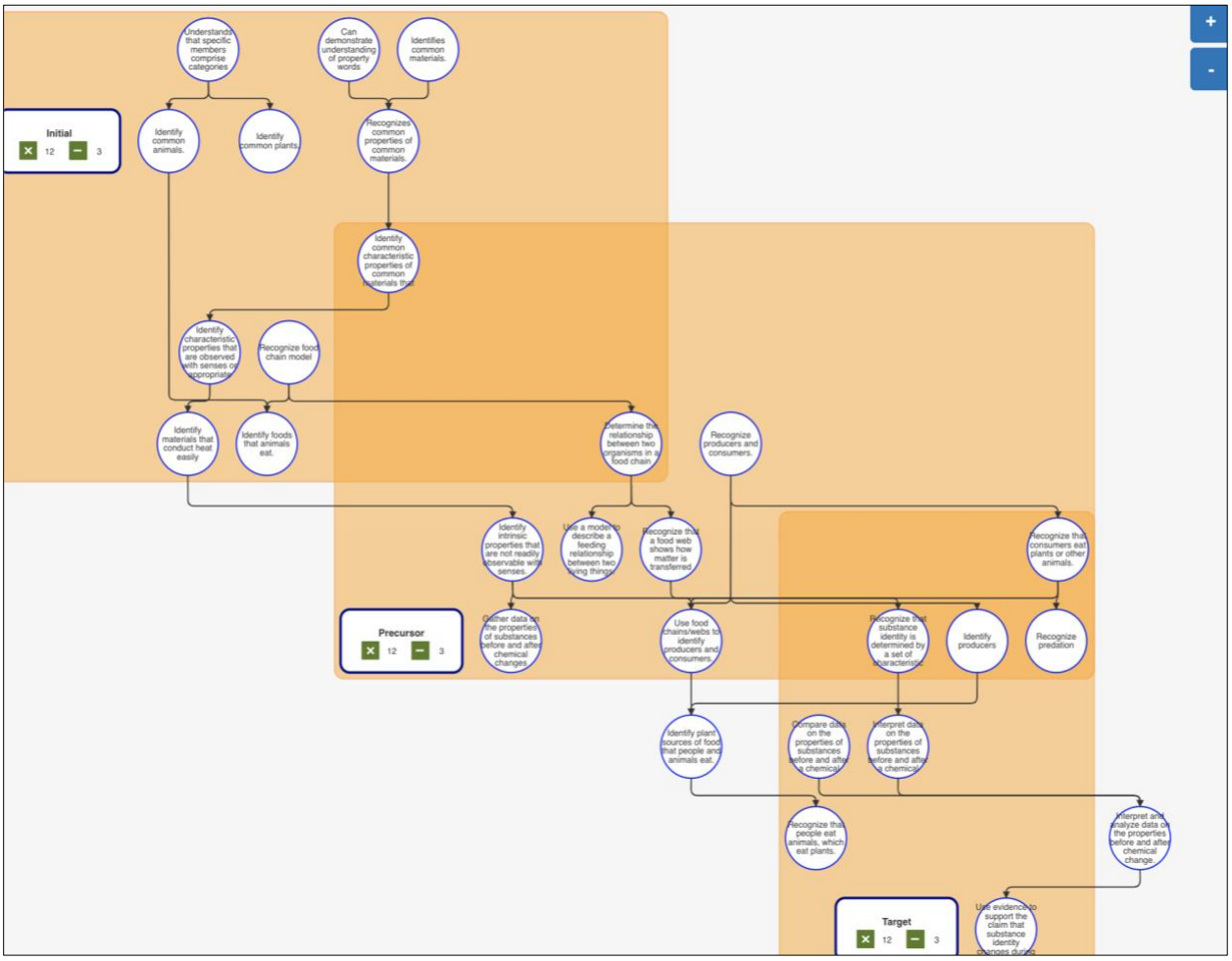
Ability to complete tasks	Roster overview, 7 tasks	Roster map, 4 tasks	Roster EE list, 3 tasks	Student overview, 3 tasks	Student map, 4 tasks	Student EE list, 4 tasks	Overall, 25 tasks
Yes, independently and readily	63%	70%	88%	94%	44%	61%	70%
Yes, independently with some effort	14%	12%	6%	0%	22%	20%	12%
Yes, with minimal prompting	14%	10%	0%	6%	28%	2%	10%
Yes, with extensive prompting	4%	6%	6%	0%	3%	15%	6%
No, required researcher intervention	6%	2%	0%	0%	3%	2%	2%

During the session, two concerns were frequently expressed by teachers regarding dashboard usability and served as sources of challenge in their ability to complete tasks independently. First, teachers often did not know that clickable elements were in fact clickable. Normally, clickability is indicated by visual appearance (e.g., underlining of links) or by changing the cursor (e.g., from arrow to pointing finger). Although both of these were included in the original dashboard design, neither was implemented in the final dashboard prototype. Second, the use of what teachers referred to as “assessment jargon” (e.g., “nodes,” “linkage levels”) rather than teacher-friendly terminology was cited as confusing, despite these terms being explained to teachers during DLM training and in the training video at the start of this study.

Two inconsistencies in the dashboard prototype implementation with respect to the design had a pronounced negative impact on teachers’ ability to use the roster and student map views independently. The first was a problem with zooming in and out of maps. As a result, we frequently prompted teachers to reload the web page in the browser to reset the map. Second, the way testlets were delineated on maps was incorrect and includes too many nodes (see Figure 10). As shown in Figure 9, we provided teachers a redesigned map model prior to the utility sessions as a supplement to the prototype, in which the nodes within each testlet were correctly identified using a different design approach.

Figure 10

Learning Map Model View from the Implemented Prototype



During usability sessions, several teachers—including some with prior DLM map review experience—demonstrated difficulty with interpreting the maps, describing them as overwhelming, complex, and abstract. All 17 teachers indicated that the redesigned maps improved clarity, with only two teachers continuing to view the map models to be too complex or visually abstract for interpretation.

Dashboard-based Understanding of Learning Map Models (Research Question 1)

Dashboard utility data to address the research question “Do teachers believe the dashboard provides an effective means for understanding the learning map models?” were collected during the utility session and post-study survey.

Utility Session Results

Just over half of teachers (*n* = 10; 59%) reported the dashboard interface improved their ability to understand learning map models compared with static representations. Five

teachers (29%) indicated that the dashboard interface did nothing to address the fact that the learning map models remained complex, jargon-filled, and unfamiliar. Two teachers (12%) provided neutral responses or insufficient information to address this research question.

Utility of Learning Map Models (Research Question 2)

Dashboard utility data to address the research question "Do teachers believe the map models in general can effectively guide appropriate instructional pathways for students?" were collected during the utility session and post-study survey.

Utility Session Results

Based on comments during the utility sessions, 65% (*n* = 11) of teachers indicated they believe learning map models in general can effectively guide appropriate instructional pathways for students. They reported the map models could serve as a visual aid for effectively guiding goal setting, communication, and choice of appropriate instructional pathways with built in "detour" or "re-route" options. By contrast, 24% (*n* = 4) of teachers indicated they do not believe map models effectively guide appropriate instructional pathways for students because they are complex, jargon-filled, and unfamiliar.

Two of these four teachers indicated the learning map models added an additional layer of complexity that provided little to no additional support compared to their familiar table-based models of presenting student data. They cited the use of "assessment-industry jargon," such as NGSS-based codes (e.g., "EE-MS-LS2-2") instead of unit names (e.g., "Food Webs") as well as "nodes" instead of "knowledge and skills," as contributing factors to this unnecessary complexity. The other two of these four teachers (the same ones identified in the usability study as believing the map models were too complex to interpret) described the learning map models as "conceptually and visually unfamiliar," indicating that their abstractness would serve as a barrier for guiding choice of instructional pathways.

Survey Results

Teachers were asked to indicate their level of agreement to the statement, "Learning map models can help with interpretation of student results from DLM assessments," on the pre- and post-study surveys. Results are shown in

Table 4 for the sixteen teachers who completed both pre- and post-study surveys. (For the one additional teacher, the post-study response was “Agree.”) Half of the teachers increased their belief that learning map models can help, with all teachers either agreeing or strongly agreeing by the end of the study.

Table 4

Teachers' Changing Level of Agreement with the Statement, "Learning Map Models Can Help with Interpretation of Student Results from DLM Assessments," on the Pre- and Post-study Surveys

		Post-study				
		SD	D	A	SA	T
Pre-study	SD					
	D			2	1	3
	A			6	5	11
	SA				2	2
	T			8	8	**16**

Note. SD = strongly disagree, D = disagree, A = agree, SA = strongly agree, T = total.

Dashboard Utility for Conveying Student Data (Research Question 3)

Dashboard utility data to address the research question "Do teachers believe the dashboard conveys student data in a means conducive to informing instructional decision-making?" were collected during the utility session and post-study survey.

Utility Session Results

Four teachers (24%) reported the dashboard did not effectively convey student data, noting the dashboard lacked conceptual "entry points" for using data for instructional planning. Three of these teachers indicated the dashboard would also need to include visual representations that summarize student progress—such as through depiction as a timeline showing when instruction began and when testlets were administered—in order to be fully effective in informing instructional decision-making.

Of the 13 teachers (76%) who believed the dashboard conveyed data in a means conducive to informing instructional decision-making, nine teachers (54%) indicated the use of the Roster Overview page would be conducive to grouping students for differentiation by providing a "quick glance" at student data. Seven teachers (41%) further indicated the Student EE List provided necessary in-depth results to complement the Roster Overview. Also, three teachers (18%) volunteered the Learning Map views were especially conducive to informing instructional decision-making by helping them connect students' KSUs with instructional pathways.

In response to the question of whether they prefer the graphic visualization of student data provided in the Learning Map view to the more traditional representation of data provided by the EE List View, seven teachers (41%) preferred the EE List view, two (12%) preferred the Learning Map view, and six (35%) preferred to use a combination of the two representations. Two teachers (12%) preferred neither view, stating that they both would instead prefer to use the Roster Overview page.

Survey Results

All teachers reported that the dashboard conveys student test performance data better than traditional score reports for informing instructional decision-making. 94% of teachers (*n* = 16) believed the dashboard to be much better at reporting individual student data than current DLM score reports, with one teacher (6%) considering them equivalent. For interpreting data from multiple students, all teachers believed the dashboard to be better than using individual DLM score reports, with 15 teachers (88%) stating much better and two teachers (12%) stating somewhat better.

Dashboard Utility for Supporting Instructional Decision-Making (Research Question 4)

Dashboard utility data to address the research question "Do teachers believe that the dashboard would improve their ability to make instructional decisions? If so, what kinds of decisions and how does the dashboard improve their ability to make them?" were collected during the utility session and surveys.

Utility Session Results

All 17 teachers indicated the dashboard would improve their ability to make instructional decisions—the dashboard offered flexible ways to view testlet results, such as aggregated views (individual student vs. roster views) and in multiple representations (i.e., Roster, EE List, and Learning Map views). Teachers identified four common types of instructional decisions they believe would be improved or supported by use of the dashboard: (a) content and lesson planning, (b) goal setting, (c) instructional grouping, and (d) testlet administration. Below, findings for each of the four instructional decisions are summarized and include evidence to support whether the dashboard effectively or does not effectively convey student data in a way that improves teachers' ability to make that instructional decision.

Content and Lesson Planning. The majority of teachers (*n* = 11; 65%) indicated that use of the dashboard would improve their ability to plan content and lessons. Common reasons included that the dashboard "indicates specific skills, knowledge, and areas of growth," and "breaks [them] down" so they can be used to "target lessons specifically." Teachers also indicated the dashboard could support not only the sequencing of specific skill instruction, but it also "gives [them] the ability to make better curriculum modifications" and "modify those skills that [they] were trying to teach in a general education" setting.

Goal-setting. During the previous needs assessment study, teachers reported they "need to know what students', parents', and other stakeholders' goals are for the students' progress, especially in science, in which goals are not set in the IEP," (Dolan et al.,

2020, p. 7). A similar need was reflected by teachers in this study, seven of whom (41%) indicated the prevalence of goal setting in their instructional decisions. One teacher stated the dashboard "would be helpful in discussing and setting goals for how we can show goals and progress." Furthermore, teachers believed the dashboard supported goal-directed instructional planning with students, which is a key part of the formative assessment process. For example, 18% of teachers (*n* = 3) noted they would use the dashboard to share learning map models with students to help them identify "where they are now, where they are going, and how to get there."

Instructional Grouping. 59% of teachers (*n* = 10) reported they would use the Roster Overview page in particular to group students for differentiated instruction. Responses indicated the dashboard effectively conveys data in a "clear and concise" way that supports the instructional act of grouping students, as well as grouping instruction itself. Teachers reported two common methods of grouping using dashboard results: homogeneously according to their current linkage levels to "hone in on the skills [they] may not have mastered yet," or heterogeneously across linkage levels (mixed-level) to enable student mentorship, where students who have mastered a given node(s) could help other students who have not yet demonstrated mastery.

Survey Results

Teachers were asked to indicate their level of agreement with the statement, "In general, results from instructionally embedded assessments can effectively inform instructional decision-making," on the pre- and post-study surveys. Results are shown in Teachers were asked to indicate their level of agreement with the statement, "DLM assessment results can effectively inform instructional decision-making," on the pre- and post-study surveys. Results are shown in Table 6 for the 16 teachers who completed both pre- and post-study surveys. (For the one additional teacher, the post-study response was "agree.") Almost half of the teachers (*n* = 7) increased their belief that results from DLM assessments can effectively inform instructional decision-making, with teachers either agreeing (*n* = 9) or strongly agreeing (*n* = 7) by the end of the study.

Table 5 for the 16 teachers who completed both pre- and post-study surveys. (For the one additional teacher, the post-study response was “strongly agree.”) Almost half of the teachers (*n* = 7) increased their belief that instructionally embedded assessments can effectively inform instructional decision-making, with all teachers either agreeing (*n* = 5) or strongly agreeing (*n* = 11) by the end of the study.

Teachers were asked to indicate their level of agreement with the statement, “DLM assessment results can effectively inform instructional decision-making,” on the pre- and post-study surveys. Results are shown in Table 6 for the 16 teachers who completed both pre- and post-study surveys. (For the one additional teacher, the post-study response was “agree.”) Almost half of the teachers (*n* = 7) increased their belief that results from DLM assessments can effectively inform instructional decision-making, with teachers either agreeing (*n* = 9) or strongly agreeing (*n* = 7) by the end of the study.

Table 5

Teachers' Changing Level of Agreement with the Statement, "In General, Results from Instructionally Embedded Assessments Can Effectively Inform Instructional Decision-making," on the Pre- and Post-study Surveys

		Post-study				
		SD	D	A	SA	T
Pre-study	SD					
	D			1		1
	A			4	6	10
	SA				5	5
	T			5	11	16

Note. SD = strongly disagree, D = disagree, A = agree, SA = strongly agree, T = total.

Table 6

Teachers Changing Beliefs in Response to the Statement, "DLM Assessment Results Can Effectively Inform Instructional Decision-making," on the Pre- and Post-study Surveys

		Post-study				
		S D	D	A	S A	T
Pre-study	SD					
	D			2		2
	A			7	5	12
	SA				2	2
	T			9	7	16

Note. SD = strongly disagree, D = disagree, A = agree, SA = strongly agree, T = total.

During the post-study survey, teachers were asked to compare the dashboard with standard DLM score reports (see Appendices Appendix E: Sample DLM Science Score Report: Performance Profile and Appendix F: Sample DLM Science Score Report: Learning Profile) in terms of a number of factors. Responses for the 17 teachers are shown in Table 7.

Table 7

Teachers' Beliefs About How the Dashboard Compared with Standard DLM Score Reports in Terms of Four Different Factors

Factor	Much worse	Some-what worse	About the same	Some-what better	Much better
Ease of use in interpreting individual student results			1, 6%		16, 94%
Ease of use in comparing across multiple students				2, 12%	15, 88%
Effectiveness for informing instructional decision-making for individual students			1, 6%	1, 6%	15, 88%
Effectiveness for informing instructional decision-making for multiple students				3, 18%	14, 82%

Dashboard Training (Research Question 5)

Dashboard utility data to address the research question, "Can teachers be trained to make effective use of the dashboard in informing instructional decision-making? If so, what type of training?" were collected during the utility session.

Utility Session Results

All teachers believed that teachers can be trained to make effective use of the dashboard in informing instructional decision-making. Common responses indicated that effective dashboard training should be flexible, consistently accessible, active, and focus on leveraging communities of practice. Their suggested training methods varied, including "hands on" simulations or case studies (47%, *n* = 8), training videos (18%, *n* = 3), and mentorship-style training opportunities for new teachers (24%, *n* = 4).

Emergent Themes

In addition to themes covered by the research questions, three themes emerged during analysis of utility session data: (a) use of the dashboard to improve communications with stakeholders (other educators, parents/guardians), (b) sharing of dashboard with students, and (c) use of dashboard in general education settings. Although short of being research questions, these themes emerged as important factors in considering the utility of the dashboard in practice. The following includes results aligned with these themes.

Communication with Stakeholders

All teachers reported the different views and representations within the dashboard would make it conducive to supporting communication with stakeholders, and they would use the dashboard as such. In particular, teachers suggested which views would be exceptionally conducive to communicating and conveying data to other educators and to parents/guardians.

Teachers reported the dashboard could facilitate communication and instructional planning with administrators to help communicate student progress. As an aid in improving communication with general educators within and across grade levels, the dashboard could help bridge a longstanding gap between general and special education. Finally, teachers suggested that information presented throughout the Dashboard could support teams at school and district levels to "figure out weaknesses in our curriculum and plan better," and hence serve as a curriculum and instructional design support tool.

Four teachers (24%) volunteered that additional data visualizations, such as line graphs that demonstrate student progress in mastering various linkage levels, would help in stakeholder communications. Two of these teachers further suggested there be a way to depict student individual learning goals when they differ from the location of the testlet linkage levels as predefined on learning map models.

Teachers reported they would use the Student Learning Map (n = 8, 44%) and EE List (n = 7, 41%) views to communicate and facilitate parents'/guardians' understanding of what their children are learning and in relation to their testlet performance results. For example, the Student Learning Map view was reported to provide parents/guardians a potential "visual representation of what their kids know," how they are progressing, and anticipation for subsequent areas of instruction. The EE List would provide specific mastery data. Additionally, it was commonly reported that this increased understanding might help parents/guardians become more actively involved in their child's education and aid in collaboratively setting clear and concrete goals for their students.

Three teachers (18%), each with administrative experience, noted the importance of maintaining confidentiality when displaying results to parents/guardians. Teachers preferred using the Learning Map views to show where a student is now and where they are going because they believed it could be easily understood by parents/guardians. Use of the Roster Learning Map view would provide a high-level picture of the students' progress with respect to peers but would need to be redesigned to not breach confidentiality. When asked if they would share their screen or print out individual views

for parents/guardians, many (*n* = 8, 47%) suggested they would print out a copy of their child's results from the EE List view.

Student Use of Dashboard

While no teacher suggested the dashboard might be used independently by students, many reported they would use the dashboard in meetings with their students as a way of communicating progress and creating, monitoring, and evaluating short-term instructional goals.

When asked with whom they would use the map view, seven teachers (41%) replied with no hesitation "my students," but emphasized the role that student variability and disability plays in deciding what to use with them. For example, students might need iconic or visual representations to best work with teachers and track their progress. In this case, the color-coded maps might work great with modifications to text and vocabulary. Furthermore, one teacher said, "I think the color coded one I could [use with students]—I like to have the kids talk about where they are goal-wise—if they see the colors, they can see they are working at that." On the other hand, one teacher said, "there's no way that this would be meaningful for them," as "most of the students in [their] classroom who qualify for alternative assessment cannot read."

Dashboard Use in General Education Settings

One of the most notable findings that emerged from utility sessions was that all 17 teachers believed the Dashboard would be very helpful in general education or in co-teaching environments. Not only did teachers believe the Dashboard could be *useful* in General Education classrooms, they thought it would be *important* or *improve* General Education classrooms. Based on utility sessions, teachers' beliefs encompassed multiple categories, including (a) curriculum and instructional planning, (b) aligning standards, and (c) improving communication and collaboration between general education and special education co-teachers.

Discussion

Usability

There are several aspects to the teacher dashboard that make it complex from a user experience design perspective. First, we provide both individual and aggregate views of data. Second, we provide access to rich descriptions of EEs and nodes. Third, we provide access to—and overlay student testlet results upon—learning map models. Finally, we provide access to testlet administration notes and instructional status. Teachers were able to accomplish most of our usability tasks independently and readily, and with minimal support from study personnel. There are two notable exceptions.

First, use of the Student Learning Map views came easily to teachers in fewer than half of the tasks, with about a quarter of all tasks requiring prompting, albeit minimal. This was primarily due to challenges the teachers faced with zoom and pan feature implementation. Zooming in and out was an important codesign feature as it allows teachers to both discern local map neighborhood nodes and performance and contextualize this information with respect to the entire EE. Also, incorrect depiction of

nodes within each testlet made it difficult for teachers to complete some tasks. Both of these challenges can readily be ameliorated in future iterations (our redesign suggestion for the latter was welcomed by teachers), and we believe are likely to increase map view usability significantly.

Second, about a sixth of the tasks involving the Student EE List view required extensive prompting or were not completable. This was due primarily to the inherent conceptual complexity of the DLM framework and compounded by what teachers referred to as "jargon." Based on our codesign sessions, we believe that use of familiar terminology (e.g., "skills" instead of "nodes"), use of EE names rather than codes, and consistent availability of a key will increase usability of the EE List views.

Overall, and from a general user experience perspective, our findings support the possibility of providing teachers a rich, full-featured dashboard they can use. That said, adequate training on the use of the dashboard would be critical. Proper usability design must be considered "necessary but insufficient" for many if not most teachers.

Instructional Use of Learning Map Models

The learning map models at the heart of DLM assessments were originally designed for test development purposes, not for teacher instructional planning. Learning map models tend to be large and complex, and are usually printed on large sheets of paper, something not typically available to teachers. A separate "node and connections" document—one per EE—provides further details used during learning map model external review. The dashboard was codesigned with teachers to combine these separate documents with less available real estate, and with features that facilitate interpretation, such as the ability to zoom in and out and pan, and the ability to click on any node for drill-down details. Also, the nodes comprising each testlet are also depicted.

Our findings generally support the notion that teachers can use map models to support interpretation of student testlet results and for instructional planning. Half of the teachers' belief in the utility of the map models in interpretation of student's testlet scores shifted positively during the study, with all teachers agreeing or strongly agreeing by the end. Furthermore, two-thirds of teachers believed map models can further help with instructional planning. We believe that teachers' belief in the utility of the dashboard was hampered by limited implementation of design features, such as zooming in and out and node details. When discussing these limitations, teachers speculated that proper implementation of these features might indeed improve the utility of the dashboard for learning maps interpretation and use. As noted previously in terms of usability, remaining challenges include the "assessment jargon-filled" nature of the information and interfaces, and lack of training.

Dashboard Use to Inform Instructional Decision-Making

Effective use of the teacher dashboard to support formative assessment practices requires proper depiction of student results and scaffolds for instructional decision-making. Our results indicate that teachers believed the dashboard would provide both of

these. The various and deeply interconnected interfaces provide flexible means to interpret student data (i.e., KSUs comprising the learning map models) and models may be viewed as maps or in more traditional tabular formats. The intentional codesign of using traditional formats as an "entry point" into understanding and using the less familiar maps seems to have worked well, with all but one teacher preferring the dashboard to traditional score reports, and all teachers believing dashboard use would improve their instructional decision-making processes of content and lesson planning, goal setting, instructional grouping, and testlet administration.

Almost half of the teachers' believed that results from instructionally embedded assessments can effectively inform instructional decision-making, with beliefs shifting positively during the study, and all teachers agreeing or strongly agreeing by the end. Furthermore, almost half of the teachers' believed that results from DLM assessments *specifically* can effectively inform instructional decision-making, with beliefs shifting positively during the study, and all teachers agreeing or strongly agreeing by the end.

Training

If teachers are to use this dashboard effectively for instructional decision-making, they would require adequate training on formative assessment practices in general, and on the use of assessment results, learning map models, and this prototype dashboard specifically. As such training was beyond the scope of this initial study, we are relying on teachers' perceptions about the potential for such training and how it would most optimally be delivered. As several of the teachers in this study have training experience, their opinions are particularly valuable. As we presented in the Results section above, all teachers believed that teachers can indeed be trained to make effective use of the dashboard in informing instructional decision-making despite the complexity of both the dashboard and the formative assessment process. What was clear is that such training would need to be as immersive and hands-on as possible, rather than the common "stand and deliver" method of providing professional learning. Furthermore, on-the-ground supports for teachers would be necessary to use the dashboard efficiently and effectively.

Improving Communication with Stakeholders

Communication is one of the pillars of successful instructional planning and decision-making. While the dashboard was designed primarily for use by and for teachers only, teachers commonly indicated the dashboard would be used for communication with other educators and with parents/guardians. As one teacher said, "I think anyone in education should be able to look at [the dashboard] and see, 'Oh that's where they need to work.'" Improved understanding and engagement by parents/guardians when discussing alternate assessment results would help ameliorate what one teacher described as, "If you're talking about an alternative assessment, you can just watch their eyes glaze over." Furthermore, clear communication and understanding of student data are especially critical during IEP team meetings, which consist of—at a minimum—special education teachers, general education teachers, and parents/guardians, as they rely on results for communicating action plans and goals.

One of the most notable findings was teachers' belief that the dashboard would be very useful in general education or in co-teaching environments, across curriculum and instructional planning, aligning standards, and improving communication and collaboration between general education and special education co-teachers.

Involvement of Students

As a "planned, ongoing process used by all students and teachers during learning and teaching to elicit and use evidence of student learning to improvise student understanding of intended disciplinary learning outcomes and support students to become self-directed learners" (CCSSO, 2018), formative assessment practices could be greatly enhanced by sharing dashboard with students, as strongly indicated by teachers in this study. Use of the dashboard in this way was not part of the original design criteria, and much could be done in terms of providing simplified or streamlined student views that teachers could show to individual or groups of students.

Study Limitations

It is important to note the current study did not attempt in-vivo evaluation of the dashboard. As such, we could not evaluate a number of actual factors that could impact its utility. For example, use of the dashboard could encourage teachers to become overfocused on discrete skills, or viewing nodes in isolation and drilling until students demonstrate mastery. To that end, five teachers (29%) did note the testlets themselves are only "one piece of the puzzle," and that additional factors need to be considered when using assessment data to support instructional planning.

Future Implications

The I-SMART Dashboard was developed with the intention of providing data in a way that supports teachers' understanding of science testlet results and allows them to use it effectively in their planning. Effective instructional planning requires at least three core components: (a) data that can be interpreted for use, (b) context to determine validity and reliability of testlet results, and (c) pedagogical supports for professional learning. This dashboard study has concluded the I-SMART Dashboard provides the access, clarity, and disaggregation of data needed to bridge the gaps between *having* assessment results and *using* assessment results to support effective formative assessment practices. That said, additional research is necessary to understand the potential value of the dashboard as an actual support to formative assessment practices in the classroom. We hope future studies consider how the dashboard might incorporate these pedagogical supports to improve teachers' instructional planning and support effective formative assessment. It would also be important to explore application of extended validity model approaches that explicitly include measures of prescription and implementation of data-driven interventions, as well as closing of achievement gaps (Way et al., 2009).

References

Bill and Melinda Gates Foundation. (2015). *Teachers know best—Making data work*. https://usprogram.gatesfoundation.org/news-and-insights/articles/teachers-know-best-making-data-work

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *80*(2), 139–148.

CAST. (2018). *Universal Design for Learning guidelines version 2.2*. http://udlguidelines.cast.org

Council of Chief State School Officers. (2018). *Revising the definition of formative assessment*. Council of Chief State School Officers. https://ccsso.org/sites/default/files/2018-06/Revising%20the%20Definition%20of%20Formative%20Assessment.pdf

Dolan, R. P., Goodman, J., Strain-Seymour, E., Adams, J., & Sethuraman, S. (2011). *Cognitive lab evaluation of innovative items in mathematics and English language arts assessment of elementary, middle, and high school students*. Pearson. http://images.pearsonassessments.com/images/tmrs/cognitive_lab_evaluation_of_innovative_items.pdf

Dolan, R. P., Wojcik, C., Ducharme, K., Starr, E., & Blackorby, J. (2020). *I-SMART Goal 3: Teacher dashboard design*. CAST, Inc.

Heritage, M. (2010). *Formative assessment and next-generation assessment systems: Are we losing an opportunity?* Council of Chief State School Officers.

Kingston, N. M., Karvonen, M., Bechard, S., & Erickson, K. (2016). The philosophical underpinnings and key features of the Dynamic Learning Maps Alternate Assessment. *Teachers College Record*, *118*(14). https://www.tcrecord.org/Content.asp?ContentId=21546

Leighton, J. P. (2017). *Using think-aloud interviews and cognitive labs in educational research*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199372904.001.0001

Way, W. D., Dolan, R. P., & Nichols, P. (2009). Psychometric challenges and opportunities in implementing formative assessment. In H. L. Andrade & G. J. Cizek, *Handbook of formative assessment* (pp. 297–315). Routledge.

Wylie, C., & Dolan, R. P. (2013, April). *The role of formalized tools in formative assessment* [Paper presentation]. American Educational Research Association Annual Meeting, San Diego, CA.

Appendix A: Terminology

The following terminology is central to DLM and I-SMART projects and will serve as an aid in reading this paper.

Essential Elements

Essential Elements are grade-level–specific expectations about what students with the most significant cognitive disabilities should know and be able to do. Essential Elements are related to College and Career Readiness standards (Next Generation Science Standards in the I-SMART project) for students without significant cognitive disabilities.

Linkage Levels

Linkage levels are small collections of nodes that are measured at different levels of complexity. Target linkage levels are at the highest level of complexity and are most closely aligned with the Essential Element and the NGSS standard. Initial and Precursor linkage levels are connected to the Essential Element at a reduced level of complexity.

Nodes

Nodes are individual knowledge and skill areas that can be represented by a single "point" or "circle" in a learning map model.

Node Observations

Node observations describe student behaviors that can provide evidence in evaluating their knowledge, skills, and understandings (KSUs) aligned with a given node.

Testlets

Testlets are short computer-delivered and instructionally embedded science assessments that can be used throughout the school year to ascertain students' progress with science content. They share a context and engagement activity and can be dynamically routed based on the difficulty level required to evaluate a student's construct-relevant KSUs.

Appendix B: Pre-Session Survey

I-SMART Teacher Dashboard Usability Study Pre-Session Survey

Note: all responses will be kept entirely confidential

* Required

1. Email *

2. What is your full name?

3. Preferred name:

4. What state do you teach in?

Mark only one oval.

- ◯ Maryland
- ◯ Missouri
- ◯ New York
- ◯ New Jersey
- ◯ Oklahoma
- ◯ Other: _____

5. What is the name of your district?

6. What is the name of your school?

7. What grade(s) do you currently teach? (2019-2020 academic year) (check all that apply)

Check all that apply.

- [ ] 3
- [ ] 4
- [ ] 5
- [ ] 6
- [ ] 7
- [ ] 8
- [ ] 9
- [ ] 10
- [ ] 11
- [ ] 12

8. In which grades have you administered alternate assessments in the past five years? (check all that apply)

Check all that apply.

- [ ] 3
- [ ] 4
- [ ] 5
- [ ] 6
- [ ] 7
- [ ] 8
- [ ] 9
- [ ] 10
- [ ] 11
- [ ] 12

9. In which of the learning environments have you taught in within the past five years? (check all that apply)

Check all that apply.

- [ ] Self-contained Classroom
- [ ] Resource Room
- [ ] Inclusive Classroom
- [ ] Other: _____

10. What devices/tools do you most typically use for instructional planning? (check all that apply.)

Check all that apply.

- [ ] Paper-Based Materials (notebooks, etc.)
- [ ] Desktop or Laptop computer
- [ ] Tablet
- [ ] Smart Phone
- [ ] Dedicated software
- [ ] Other: _____

Instructionally Embedded Assessments

Indicate the degree to which you agree with the following:

11. In general, results from instructionally embedded assessments can effectively inform instructional decision making.

Mark only one oval per row.

	Strongly Disagree	Disagree	Agree	Strongly Agree
-->	⬭	⬭	⬭	⬭

12. DLM assessment results can effectively inform instructional decision-making.

Mark only one oval per row.

	Strongly Disagree	Disagree	Agree	Strongly Agree
-->				

13. I use DLM assessment results to inform instructional decision making.

Mark only one oval per row.

	Never	Rarely	Occassionally	Moderately	Frequently
-->					

14. I find it easy to use DLM assessment results to inform instructional decision making.

Mark only one oval per row.

	Strongly Disagree	Disagree	Agree	Strongly Agree
-->				

15. Please explain your answer.

16. How do you use DLM assessment results to inform instructional decision making? (if applicable)

17. What barriers have you encountered during instructional decision making?

18. What is your level of expertise administering DLM assessments in ELA, Math, and Science?

Mark only one oval per row.

	None	Novice	Intermediate	Advanced
ELA	◯	◯	◯	◯
Math	◯	◯	◯	◯
Science	◯	◯	◯	◯

Learning Map Models

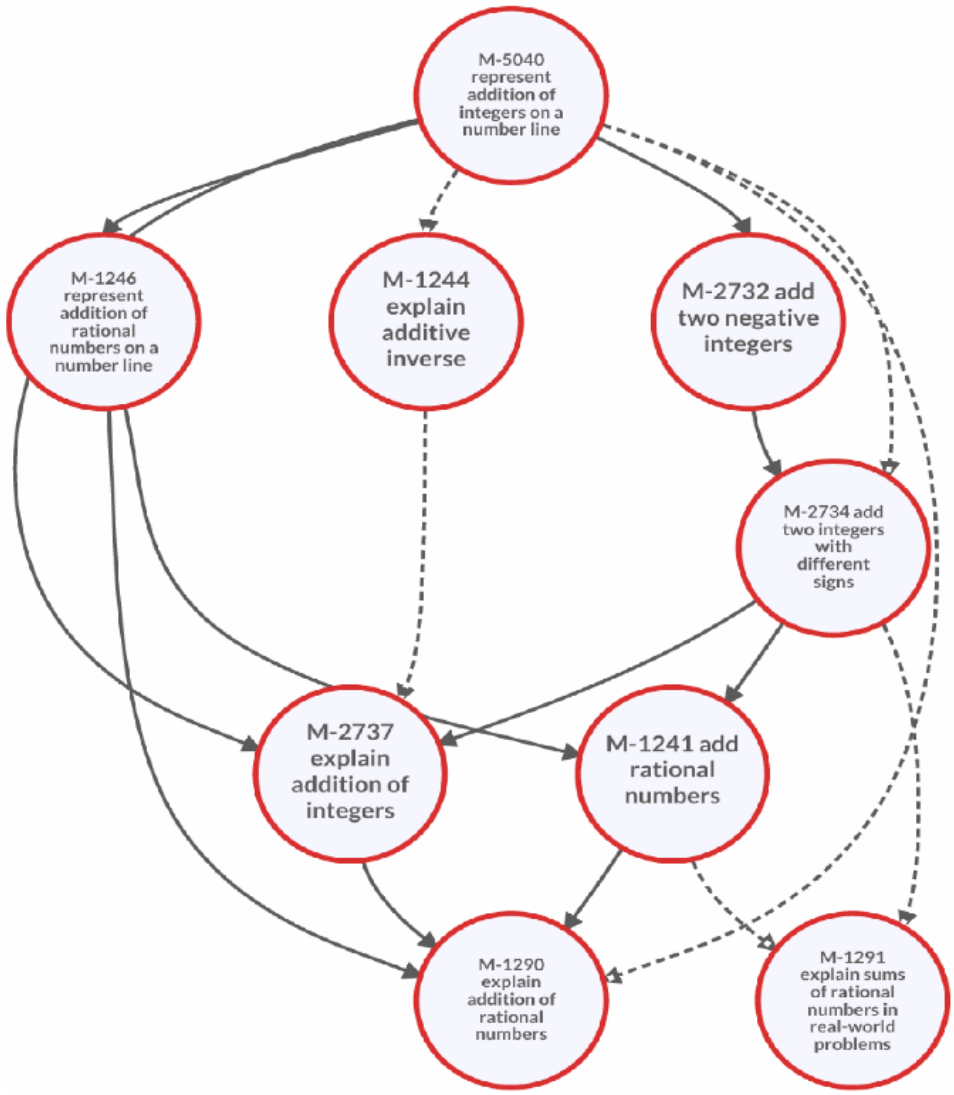
The following questions will help the research team understand your familiarity with learning map models.

19. How familiar are you with learning map models that underly ELA and Math DLM assessments? (see image below)

Mark only one oval per row.

	Not at all familiar	Slightly familiar	Moderately familiar	Extremely familiar
-->				

Example of Learning Map Model



20. Do you think learning map models can help with interpretation of student results from DLM assessments?

Mark only one oval per row.

	Strongly Disagree	Disagree	Agree	Strongly Agree
-->				

21. Do you think learning map models can effectively support instructional decision-making?

Mark only one oval per row.

	Strongly Disagree	Disagree	Not Applicable	Strongly Agree
-->				

22. What would it take for learning map models to effectively support instructional decision making?

Appendix C: Usability and Utility Sessions Protocol

Usability Study Protocol for KITE

Version 1.3, updated July 12, 2020 after KITE fixes and mock data complete.

- KITE dashboard: https://educator.kiteaai.org
- Responses spreadsheet
- Dashboard prototype (Design, Marvel app)
- Dashboard prototype (Development)
- Accessing the Dashboard (usernames, passwords for Production environment)

Overview

This document describes a plan for conducting a usability test of the I-SMART Reporting Dashboard user interface design.

Testing Goals

1. Evaluate the dashboard usability.
2. Evaluate the interpretability of score report contents.
3. Evaluate the utility of the score report contents and interface in terms of their ability to meaningfully inform instructional decision-making.
4. Inform the production of a summary of findings and recommendations for dashboard design refinements.

Methodology

Approximately 16 teachers from DLM partner states will participate in two 60-minute, individual remote cognitive interview sessions beginning in Spring 2020. The two sessions will be spaced out by three to seven days.

Training

Prior to the first session, teachers will view an eight-minute video explaining learning map models. At the start of the first session, teachers will receive interactive training on use of the dashboard. Existing ATLAS training materials (e.g., PowerPoints, scripts, videos) may be adapted for these trainings. We will assume teachers already have training on DLM fundamentals and use.

Participants

Sixteen teachers at the elementary, middle, or high school levels from two or more partner states (including Maryland) will be recruited for testing purposes. All teachers included in the study will have DLM science experience, but none will have taken part in previous dashboard design or map review processes. Approximately half of these teachers will have experience working with students who *do not* take the alternate

assessment, and who may or may not have disabilities, but struggle to meet grade-level standards in science.

Procedure

Pre-session Survey. In advance of the first session, teachers will be administered an online survey to evaluate their attitudes and beliefs about formative assessment, data-driven instructional decision-making, and the role that DLM assessments can play. The survey should take approximately 10–15 minutes to complete.

Pre-session Training. Following completion of the pre-session survey and in advance of the first session, teachers will be administered training on learning map models in the form of an eight-minute video.

Testing Sessions. Participants will take part in the usability test via remote screen-sharing software that will be recorded. Participants will be provided the following participation guidelines in advance of the session:

- must be at a desktop or laptop computer, not on a tablet or cell phone;
- must have a built-in or external video camera connected to their computer;
- must have Zoom video communication software installed on their computer;
- must have a stable, high-bandwidth internet connection;
- must be in a private, stationary location, free of distractions (no bells, announcements, teachers or students wandering around in the background);
- must have provided materials (if any) available, such as printouts; and
- must have both video and audio turned on in Zoom.

The facilitator will instruct the participant that they are evaluating the application rather than the facilitator evaluating the participant. The facilitator will ask the participant if they have any questions. The session will begin when all participant questions are answered by the facilitator.

At times, the facilitator will instruct the participant to “think aloud” so that a verbal record exists of their interaction with the I-SMART Reporting Dashboard. The facilitator or the notetaker will observe and record user behavior, user comments, and interactions with the application. At the start of each task, the facilitator will read aloud the task description then instruct the participant to begin the task. Participants will be reminded to think aloud periodically if they do not do so on their own.

Participants will be asked to optionally record notes of their ideas that they share with I-SMART staff after each session concludes, as a further means to document their thinking.

There will be a total two sessions, each lasting 60 minutes.

Second-Session Preparation. Following the first session, teachers will be asked to consider how they might use the dashboard in the classroom, and the extent to which its use could help them with instructional decision-making. Teachers will log in to the online dashboard and be provided instructions for completing a series of tasks using the dashboard with mock student data that will take approximately 30–60 minutes to complete.

Post-session Survey. Following the second session, teachers will be readministered the pre-session online survey to reevaluate their attitudes and beliefs about formative assessment, data-driven instructional decision-making, and the role that DLM assessments can play. Additional questions will cover impressions of the Reporting Dashboard. The survey should take approximately 10–20 minutes to complete.

Roles

Trainer: [Name]

- create an eight-minute training video on learning map models

Facilitator: [Name]

- provide an overview of the study to participants
- define usability and purpose of usability testing to participants
- provide interactive training on use of dashboard
- lead participants through testing protocols
- respond to participant's requests for assistance

Notetaker: [Name]

- record participant's actions and comments

Testing Metrics and Analysis

Usability performance measures to be evaluated:

- task completion rate
- error-free rate
- critical error count
- recoverable (non-critical) error count
- responses to embedded "temperature check" questions administered several times during each session, which probe current level of understanding vs. confusion
- time on task
- subjective measures (user-reported attitude, satisfaction with experience)

Performance measures will be prioritized by impact and dashboard function (e.g., individual student vs. aggregate views).

Data interpretability performance to be evaluated by coding teacher actions, utterances, and survey results using a to-be developed taxonomy based on previous score report interpretation research.

Instructional decision-making utility to be evaluated by coding teacher actions, utterances, and survey results using a to-be developed taxonomy based on formative and classroom assessment frameworks (e.g., Wylie).

Testing Protocol—Session One Script

Introduction

Facilitator: Hi, [tester's name]. My name is [facilitator's name].

[Name] will be joining us to take notes.

(If applicable: …and we have two other observers from CAST. They're part of the Research and Design team.)

Thank you for taking the time to participate in this study. Before we begin, I'm going to give you a brief overview of the test and how it will work. This session will last about 50 minutes. For this test, I'll be sharing with you a series of specific tasks to complete. The test will be split into two sections. Today we would like to know whether the dashboard environment is easy to use and if the information is clear. The next session will be more focused on whether the information helps you with instructional decision-making.

It's really important to know that we are only testing the design of the environment, not you. There are no wrong answers—we consider you the expert. Please be honest and let me know at any time if there's something you like, dislike, or are confused about. I promise you won't hurt my feelings.

I'd like you to "think aloud" as much as possible. By that, I mean that I'd like you to verbalize your thoughts as often as you can. For example, imagine you're hungry and walk into your kitchen to make yourself a snack. "Think aloud" right now what you'd do.

If at any point you have questions, please don't hesitate to ask. Do you have any questions so far?

As a reminder, we will videotape this session—the computer screen you'll share and what you have to say. The video will be used only to help us figure out how to improve the application, and it won't be seen by anyone except the people working on the project. Do we have your permission to record the session?

[Record.]

Okay, let's get started.

Scenario

Facilitator: Imagine you are a teacher in a/an [elementary | middle | high school] and have administered I-SMART testlets. These are like DLM science testlets but you can administer them whenever you want during the year and they provide information about students' knowledge, skills, and understandings in science at the time of the assessment. You would like to review student performance on these testlets and are open to the idea that these results can help you with instructional decision-making.

I'm going to ask you to share your screen and think aloud as you go through the questions I will ask.

Are you seeing this for the first time, or have you explored the environment a bit? Either is fine, just curious!

Sign in using the Kite Suite link, username, and password provided in the chat window, and click in the lower right on "LM Dashboard," then on the Roster button.

Questions to Determine the Usability and Basic Interpretability of Student Data Displayed through the I-SMART Reporting Dashboard

1. General question - no page [–]

    Did you watch the video on learning map models?
    Was it new? Old material? Was it useful?

2. Roster Overview page [Navigation/Usability]

    Let's start here. What do you think is the title of the page you're looking at?

3. Roster Overview page [Data Interpretability]

    You'll notice two columns, each starting with the label ISM.EE… Each column is for a different Essential Element neighborhood. The one with the label ending in (ES) LS2-1/(MS/HS) LS2-2 is about food webs.

    For the next several questions, let's look at the Food Web Essential Element, which is the column ending in LS2-1/LS 2-2. LS stands for life sciences.

    What is the status of instruction for [Johnny Doe]? In other words, has Johnny's instruction not begun, is it in progress, or is it complete?

4. Roster Overview page [Data Interpretability]

    How did [Johnny] do on the [precursor] Testlet?

5. Roster Overview page [Data Interpretability]

    Is the status of instruction the same or different for Johnny and Chloe? Is the status of instruction the same or different for Johnny and Anastasia / Leann / Hayden?

6. Roster Overview page [Navigation/Usability]

Can you change the status of instruction for ____ (ES)/____ (MS)/____ (HS) from “instruction not begun” to “instruction in progress” and set the date to today’s date?

7. Roster Overview page [Data Interpretability]

   How long was the period of instruction on the Food Web Essential Element for [Caitlin Verna / Neal Liberty / Danielle Celeste]?

8. Roster Overview page [Navigation/Usability]

   I’m going to ask you to go to a different page. Can you navigate to the Learning Map for the Food Web Essential Element?

   (*If they get stuck, direct them to click on the word “Map” located at the top of the Food Web column.*)

9. Class Learning Map page [Data Interpretability]

   To scroll within this page, you need to put your cursor outside of the map box. To move the map. If you lose your map, reload the page (and make sure you’re still on [LS2-1/LS2-1]

   First, are you looking at the correct EE? If not, please change it to [LS2-1/LS2-2].

   I’m going to ask you about that first box with the green icons, labeled “initial.” Zoom in until you can see the information clearly.

   Without clicking on the box, how is the class doing for the initial testlet?

10. Class Learning Map page [Navigation/Usability]

    On this same class map, can you open up the “Class Results Summary” Box for the Precursor testlet?

11. Class Learning Map page [Data Interpretability]

    In this summary box, in which node did [Johnny Doe] show the most mastery?

12. Class Learning Map page [Navigation/Usability]

    Okay, now I’d like you to close the class summary box.

13. Class Learning Map page [Navigation/Usability]

    Can you navigate to the student pages using the global navigation?

14. Class Learning Map page [Navigation/Usability]

    Can you change the view to Johnny Doe’s Student Overview?

15. Class Learning Map page [Navigation/Usability]

    Great. Now navigate to Johnny’s Food Web Learning Map. Remember, that’s the one ending in [LS2-1/LS2-2].

16. Student Learning Map page [Data Interpretability]

    Which testlet(s) has Johnny taken? In other words, which ones have been administered?

17. Student Learning Map page [Data Interpretability]

    Which testlet(s) did Johnny struggle with the least?

18. Student Learning Map page [Navigation/Usability]

    On this map, can you access the Node Observation for the…

    - [ES] …recognize food chain models node. It's one of the circles in the initial.
    - [MS] …identifies common materials node. It's one of the circles in the top row.
    - [HS] …recognize population node. It's one of the circles to the right of the initial box.

    (*It may take a few tries clicking on the circle node—a "beta" glitch. If they're thinking out loud, you'll know if they got it "right."*)

19. Student Learning Map page [Navigation/Usability]

    Can you navigate to Johnny's EE List View for this Food Web Essential Element?

    After they go to the listview, say: You will need to change the student pulldown back to Johnny.

20. Student EE List View page [Navigation/Usability]

    Click on the "Show Nodes" box. It's just under all the pull-down menus.

    Can you describe the information you are seeing in the Food Webs Essential Element neighborhood? That's the one ending in [LS2-1/LS2-2].

21. Student EE List View page [Data Interpretability]

    Great. Can you find the Node Observation for the precursor node "Use food chains/webs to identify producers and consumers?" (Determine the relationship between two organisms in a food chain.)

    (If they can't find it right away, ask them to scroll up and check the box that says, "Show node observations," then repeat the question.)

22. Student EE List View page [Data Interpretability]

    What is the testlet note that pertains to Johnny's Precursor testlet?

23. Student EE List View page [Navigation/Usability]

    Can you navigate to the Student Overview for Chloe Beaux?

(If they use the pull-down to get to Chloe, ask them to check on the title of the page. Hopefully they'll realize they're not on the student overview. They can click on the "overview" tab, or "student" in the global navigation to proceed.)

24. Student Overview page [Data Interpretability]

    Which Essential Element is Chloe currently receiving instruction in?

25. Student Overview page [Navigation/Usability]

    On what date did instruction begin for Food Webs Essential Element? (LS 2-1/LS2-2)?

26. Student Overview page [Navigation/Usability]

    Can you navigate back to the Roster Overview?

27. General question - no page [–]

    Great. Now I'd like to share my screen. Can you click on the stop sharing button? It's a red button at the top middle of your screen.

    (Share your screen and open the map view and list view of the dashboard side by side.)

    Did you find the map view or list view more useful? Why?

28. General question - no page [–]

    Thinking back, in general, what was the most difficult and what was easy for you?

29. Any other thoughts in general?

Facilitator: In preparation for our second session in [3–7] days, as early as Wednesday, July 21, we'd like you to consider how you might use the dashboard in your classroom, and the extent to which its use could help you with instructional decision-making. We will email you dashboard login information and instructions for completing a series of tasks using the dashboard with mock student data that will take approximately 30–60 minutes to complete.

Testing Protocol—Intersession Homework

Thank you again for participating in the first of the two usability testing sessions for our I-SMART Teacher Dashboard. Prior to the second session, we ask that you spend up to 15–30 minutes on the following activities:

1. Please watch this eight-minute video on learning map models (if you haven't already, or if you wish to review).

If you are familiar with learning map models and would like to review specific portions, please consult these time stamps:

- Essential Elements and Standards (0:12)
- General Education Standards versus Essential Elements—An Example (1:47)
- What Is a Learning Map Neighborhood? (2:49)
- Example of a Learning Map Neighborhood (3:32)
- Components of a Learning Map Neighborhood—Nodes and Connections (3:43)
- Node Descriptions and Observations (4:25)
- Characteristics of a Learning Map Neighborhood—Alternate and Multiple Pathways (4:55)
- Interpret a Learning Map Neighborhood (6:04)

2. Please look at the attached screenshot of a different version of a learning map neighborhood than the one you saw in the prototype. In this version, all the nodes assessed in a given testlet are colored the same. For example, four nodes are tested in the initial linkage level testlet, as shown in green. One of these nodes [NAME] is also purple because this node is also assessed in the Precursor/Distal linkage level testlet.
3. Please consider the following questions which we'll be discussing during the second session. Feel free to record any ideas for our discussion during the second session.
    1. How might you use the dashboard to help you with instructional planning?
    2. Would you use the dashboard to examine data from a student's previous year for instructional planning this year?
    3. Would you use the dashboard during meetings with students? Parents or guardians? Other educators?
    4. How might the dashboard be useful in inclusive or general education settings?

Testing Protocol—Session Two

Background

The second session will be less structured than the first. Teachers had the opportunity during the first session to understand the intended design of the

dashboard and its general usage and have been encouraged to consider (between sessions) how dashboard use could help them with instructional decision-making in the classroom or learning environment. We will now evaluate the potential utility of the dashboard in supporting teachers' instructional decision-making according to the following research questions.

Research Questions

1. Do teachers believe the dashboard provides an effective means for **understanding the learning map models**?
2. Do teachers believe the map models *in general* can effectively **guide choice of appropriate instructional pathways** for students?
3. Do teachers believe the dashboard conveys student data in a means **conducive to informing instructional decision-making**?
4. What **kinds of instructional decisions** do teachers believe the dashboard can help them make?
5. What **type of training** would teachers need in order to make effective use of the dashboard in informing instructional decision-making?

Teachers will be asked to explain specifically what they would do with specific mock students based on their data as made available in the dashboard. For example:

- When you come into class on Wednesday, what factors go into your decisions on what to teach [Sally]?
- At what point during instructional decision-making are students' test scores considered?
- How often would you use the dashboard?
- How might summative (i.e., end-of-year) use of the dashboard support student transitioning to the next year?

Teacher responses will be gathered both through questioning during the second session and in the post-session survey.

Protocol Script

Facilitator will share their screen.

Facilitator: Thanks again for participating in the previous session and for returning for this follow-up session. Were you able to do the "homework" in which you considered the role the dashboard might play in instructional planning? Before we begin with specific questions, do you have any overarching thoughts, reflections, or suggestions to share from doing that homework and since the previous session?

Remind me, did you watch the eight-minute video on learning map models prior to our previous session together?

Yes: Great

No: Did you watch it before **this** session? If so: Was it new? Old material? Was it useful?

As we discussed last time, the dashboard prototype is in its early stages and some of the information provided in the learning map neighborhood wasn't quite correct. This image better shows the nodes included in each testlet for one the Essential Elements…

- **Elementary School:** EE LS2-1 Food Chains
- **Middle School:** EE LS2-2 Food Webs
- **High School:** EE LS2-2 Food Webs

… by using different colors. If you were explaining this to a fellow teacher, how would you describe what you're seeing? What about the nodes that have two colors?

Facilitator: Today I'll be asking a series of questions about the dashboard and how it might support instructional decision-making and planning, that is, support you in determining instruction, interventions, remediations, or other factors directly impacting your students. We'll refer to all these collectively as instructional planning. I'll be sharing my screen and have screenshots of the KITE system we can look at as we discuss the dashboard.

1. How do you decide when to administer the science alternate assessment testlets and at what linkage levels?
2. Do you currently use results from the science alternate assessment for instructional planning?
    a. If so, how?
    b. If not, would availability of student data through such a dashboard change your practices?
    c. How about for ELA and math?
3. Do you currently use learning maps for instructional planning in science, ELA, or math?
    a. If so, how?
4. How might you use the dashboard to help you with instructional planning?
5. How would you use the following to help with instructional planning [show appropriate screenshots]…

a. Roster Overview?

b. Student EE List?

c. Map View?

6. Would you use the dashboard to examine data from a student's previous year for instructional planning this year? Why or why not?

7. Would you use the dashboard during meetings with…

    a. Students

    b. Parents or guardians?

    c. Other educators?

    Note: Give teachers the opportunity to mention **unprompted** the idea of actively sharing the dashboard screen with these stakeholders. If they don't, then ask. And make sure they consider what that would mean for students.

8. If you have taught in an inclusive or general education setting as well for students struggling in science, how do you think this dashboard could be used or modified for that environment?

9. Would students in general education be better able to make sense of the dashboard if you shared it with them? Would it be useful for them?

10. What might you add to or change in the dashboard to make it more useful for instructional planning?

    Note: Allow teachers to mention the following **unprompted**. If they don't, then ask.

    a. Is there additional information that the dashboard could provide?

    b. links to Next Generation Science Standards

    c. direct access to students' IEP plans

    d. integration to Student Information Systems

    e. links to instructional resources (optional: Who do you imagine would curate such resources, you as a teacher, a community of teachers, the University of Kansas/KITE folks?)

11. How do you think educators might best be trained to effectively use this dashboard to guide instructional planning?

    a. Do you train teachers yourself?

12. What aspects of the dashboard use might be challenging to teach or learn? What about interpreting and leveraging the learning maps themselves?
13. That’s all our questions. Is there anything else you'd like to share?

Thank you again for your participation in these sessions. Your feedback is critically important as we explore new directions for using DLM assessments to support instructional planning. We’ll send you a link to the post-session survey, which will include us asking for your address so we can send you your stipend check.

Testlets Completed by Mock Students

Table C1

Testlests and Linkage Levels Completed by Each Student

Student	Elementary	Middle school	High school
Chloe Beaux	Initial, Precursor	Initial, Precursor	Initial, Distal
Siobhan Chlough	Initial, Precursor	Initial, Precursor	Initial, Distal
Johnny Doe	Initial, Precursor	Initial, Precursor	Distal, Precursor
Karen Oh	Precursor, Target	Precursor, Target	Distal, Precursor
Hubert Pho	Precursor, Target	Precursor, Target	Precursor, Target
Asawan Rowe	Precursor, Target	Precursor, Target	Precursor, Target

Appendix D: Post-Session Survey

I-SMART Dashboard Usability Study Post-Session Survey

Note: all responses will be kept entirely confidential

* Required

1. Email *

2. What is your full name?

Instructionally Embedded Assessments

Indicate the degree to which you agree with the following

3. In general, results from instructionally embedded assessments can effectively inform instructional decision-making.

Mark only one oval per row.

	Strongly Disagree	Disagree	Agree	Strongly Agree
-->				

4. DLM assessment results can effectively inform instructional decision-making.

Mark only one oval per row.

	Strongly Disagree	Disagree	Agree	Strongly Agree
-->				

5. I use DLM assessment results to inform instructional decision-making.

Mark only one oval per row.

	Strongly Disagree	Disagree	Agree	Strongly Agree
-->				

6. I find it easy to use UDL assessment results to inform instructional decision-making.

Mark only one oval per row.

	Strongly Disagree	Disagree	Agree	Strongly Agree
-->				

7. Please explain your answer.

8. How do you use DLM assessment results to inform instructional decision-making? (if applicable)

9. What barriers have you encountered during instructional decision-making?

10. What is your level of expertise administering DLM assessments in ELA, Math, and Science?

Mark only one oval per row.

	None	Novice	Intermediate	Advanced
ELA	◯	◯	◯	◯
Math	◯	◯	◯	◯
Science	◯	◯	◯	◯

Learning Map Models

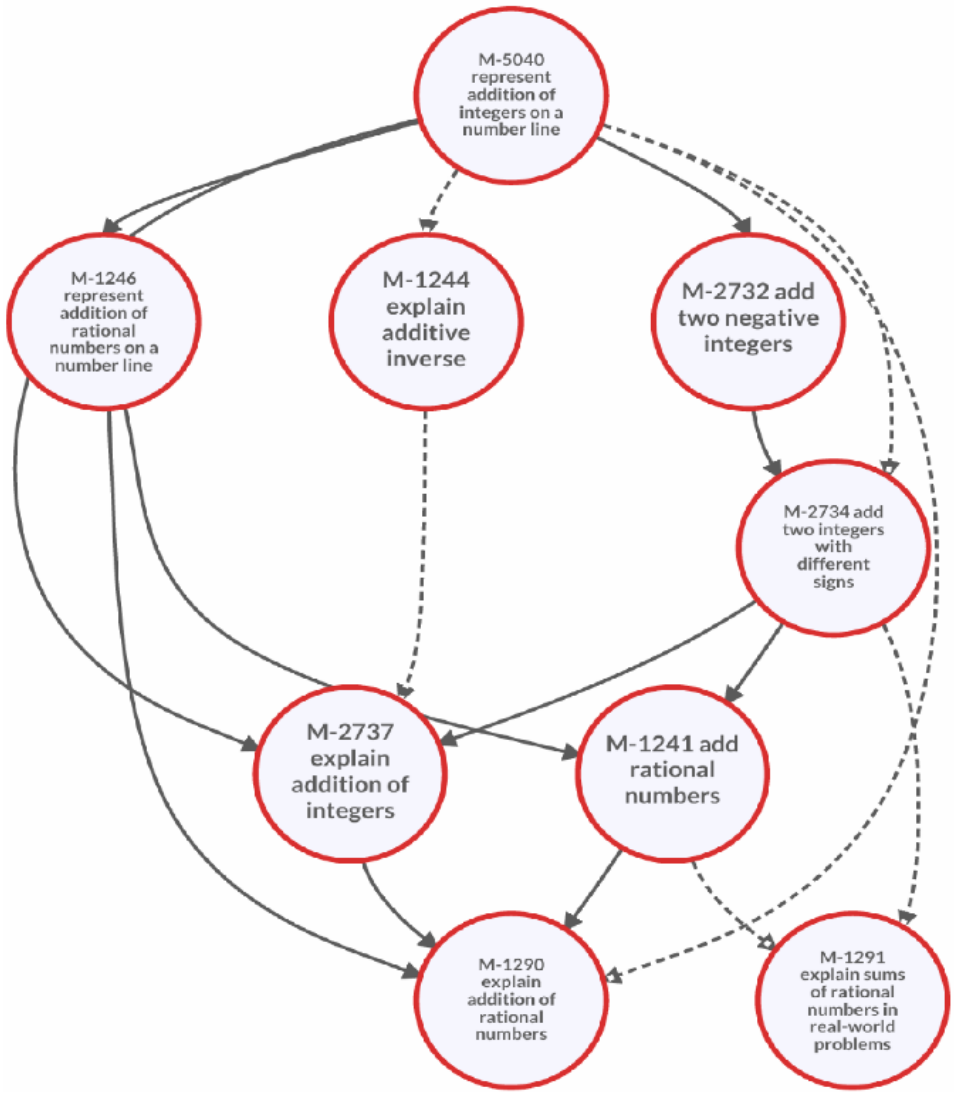
The following questions will help the research team understand your familiarity with learning map models.

11. How familiar are you with learning map models that underly ELA and Math DLM assessments? (see image below)

Mark only one oval per row.

	Not at all familiar	Slightly familiar	Moderately familiar	Extremely familiar
-->	◯	◯	◯	◯

Example of Learning Map Model



12. Do you think learning map models can help with interpretation of student results from DLM assessments?

Mark only one oval per row.

	Strongly Disagree	Disagree	Agree	Strongly Agree
-->				

13. Do you think learning map models can effectively support instructional decision-making?

Mark only one oval per row.

	Strongly Disagree	Disagree	Agree	Strongly Agree
-->				

14. What would it take for learning map models to effectively support instructional decision-making?

I-SMART Teacher Dashboard Feedback

15. How does the I-SMART Teacher Dashboard compare with standard DLM score reports in terms of ...

Mark only one oval per row.

	much worse	somewhat worse	about the same	somewhat better	much better
ease of use in interpreting individual student results	◯	◯	◯	◯	◯
ease of use in comparing across multiple students	◯	◯	◯	◯	◯
effectiveness for informing instructional decision-making for individual students	◯	◯	◯	◯	◯
effectiveness for informing instructional decision-making for multiple students	◯	◯	◯	◯	◯

16. How does the I-SMART Teacher Dashboard compare with other (non-DLM) methods you use for accessing score reports?

17. How do you envision the I-SMART Teacher Dashboard might be used in your classroom?

18. Are there any features that you wish the I-SMART Teacher Dashboard had that were not available to you today?

19. Is there any other feedback that you would like to provide that could improve the I-SMART Teacher Dashboard?

Address for Stipend

Please provide your mailing address for us to mail your stipend.

20. Where would you like us to mail your stipend? (Please include street address, city, state/territory, and zip code) *

Appendix E: Sample DLM Science Score Report: Performance Profile

REPORT DATE: 03-20-2017
SUBJECT: Science
GRADE: 5

Individual Student Year-End Report
Performance Profile 2016-17

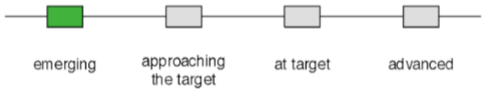
DYNAMIC® LEARNING MAPS

NAME: Student DLM
DISTRICT: DLM District
SCHOOL: DLM School

DISTRICT ID: DLM District ID
STATE: DLM State

Overall Results

Elementary science allows students to show their achievement in 27 skills related to 9 Essential Elements. Student has mastered 7 of those 27 skills during the 2016-17 school year. Overall, Student's mastery of Science fell into the first of four performance categories: **emerging**. The specific skills Student has and has not mastered can be found in Student's Learning Profile.



EMERGING:	The student demonstrates **emerging** understanding of and ability to apply content knowledge and skills represented by the Essential Elements.
APPROACHING THE TARGET:	The student's understanding of and ability to apply targeted content knowledge and skills represented by the Essential Elements is **approaching the target**.
AT TARGET:	The student's understanding of and ability to apply content knowledge and skills represented by the Essential Elements is **at target**.
ADVANCED:	The student demonstrates **advanced** understanding of and ability to apply targeted content knowledge and skills represented by the Essential Elements.

A student who achieves at the **emerging** performance level typically can recognize changes in state of matter, match properties, observe the effects of gravity, distinguish living from non-living things, identify human needs, order daily events, and anticipate routines.

In physical science, the student can

- recognize melting and freezing
- match materials with similar physical properties
- recognize the direction objects go when dropped
- identify models that show plants need sunlight to grow

In life science, the student can

REPORT DATE: 03-20-2017
SUBJECT: Science
GRADE: 5

Individual Student Year-End Report
Performance Profile 2016-17

DYNAMIC® LEARNING MAPS

NAME: Student DLM
DISTRICT: DLM District
SCHOOL: DLM School

DISTRICT ID: DLM District ID
STATE: DLM State

Performance Profile, continued

- distinguish things that grow from things that do not grow
- identify common human foods

In earth and space science, the student can

- order events in daily routines, including sunrise and sunset
- identify routines to follow when it is raining

Domain

More information about Student's performance on each of the Essential Elements that make up the Domains is located in the Learning Profile.

Appendix F: Sample DLM Science Score Report: Learning Profile

REPORT DATE: 03-20-2017
SUBJECT: Science
GRADE: 5

Individual Student Year-End Report
Learning Profile 2016-17

DYNAMIC® LEARNING MAPS

NAME: Student DLM
DISTRICT: DLM District
SCHOOL: DLM School

DISTRICT ID: DLM District ID
STATE: DLM State

Student's performance in elementary science Essential Elements is summarized below. This information is based on all of the DLM tests Student took during the 2016-17 school year. Student was assessed on 9 out of 9 Essential Elements expected in elementary science. Student was assessed on 3 out of 3 Domains expected in elementary science.

In order to master an Essential Element, a student must master a series of skills leading up to the specific skill identified in the Essential Element. This table describes what skills your child demonstrated in the assessment and how those skills compare to grade level expectations.

Green shading shows levels mastered this year. Blue shading shows Essential Elements with no evidence of mastery. Gray shading indicates the Essential Element was not assessed this year.

	Level Mastery		
Essential Element	1	2	3 (Target)
SCI.5.ESS.1.2	Order events including sunrise and sunset	Recognize patterns in the length of day	Show seasonal patterns in the length of day
SCI.5.ESS.2.1	Anticipates routine to follow when it is raining	Recognize how water affects people	Model how water affects the living things
SCI.5.ESS.3.1	Identify one way to protect a resource of Earth	Compare methods that help protect the Earth's resources	Describe how to protect the Earth's resources
SCI.5.LS.1.1	Distinguish things that grow from things that don't grow	Provide evidence that plants grow	Provide evidence that plants need air and water to grow
SCI.5.LS.2.1	Identify common human foods	Identify a model that shows matter moving from plants to animals	Model matter moving through living things
SCI.5.PS.1.2	Recognize melting and freezing	Compare weight before and after melting and freezing	Compare weight before and after heating, cooling, or mixing
SCI.5.PS.1.3	Match physical properties	Classify materials by physical properties	Identify materials based on properties

Levels mastered this year | No evidence of mastery on this Essential Element | Essential Element not tested

Page 1 of 2

REPORT DATE: 03-20-2017
SUBJECT: Science
GRADE: 5

Individual Student Year-End Report
Learning Profile 2016-17

DYNAMIC® LEARNING MAPS

NAME: Student DLM
DISTRICT: DLM District
SCHOOL: DLM School

DISTRICT ID: DLM District ID
STATE: DLM State

Essential Element	Level Mastery		
	1	2	3 (Target)
SCI.5.PS.2.1	Recognize the direction objects go when dropped	Predict the direction objects go when dropped	Demonstrate that gravity is directed down
SCI.5.PS.3.1	Identify models that show plants need sunlight to grow	Model plants capturing energy from sunlight	Model energy in food coming from the Sun

Levels mastered this year | No evidence of mastery on this Essential Element | Essential Element not tested