

## **Promises and Challenges in Teacher Implementation of Instructionally Embedded Assessments**

Amy Clark, Jennifer L. Kobrin, Megan Mulvihill, and Ashley Hirt

Accessible Teaching, Learning, and Assessment Systems (ATLAS), University of Kansas

### Author Note

Paper presented at the 2023 annual meeting of the National Council on Measurement in Education, Chicago, IL. Correspondence concerning this paper should be addressed to Amy Clark (akclark@ku.edu). Do not redistribute this paper without permission of the authors. The authors would like to acknowledge Elizabeth Kavitsky for her contributions to this work.

## **Abstract**

Many state education agencies (SEAs) have begun to develop and implement new, innovative assessment systems to meet statewide accountability requirements. To support these efforts, it is essential that existing programs share lessons learned about the development and implementation of innovative assessments. However, most new assessment systems are still under development and there is limited information available. The Dynamic Learning Maps (DLM) Instructionally Embedded Alternate Assessment System has been operational since the 2014-2015 school year and provides a unique opportunity to learn and share valuable lessons about using an innovative instructionally embedded assessment system for statewide accountability purposes. Unlike traditional summative assessments, instructionally embedded assessments are administered throughout the school year and provide real-time information about students' performance of specific, measurable skills. We conducted focus groups with thirty teachers from 6 states about the implementation of the DLM instructionally embedded assessment system. Findings from the focus groups demonstrate the flexibility and utility of instructionally embedded assessments, which teachers use to plan instruction, inform individualized education programs, evaluate students' progress, and share information about students' achievement with parents. The focus groups also revealed that some teachers face challenges to accessing and using assessment results effectively and that teachers would benefit from targeted training in these areas. Teachers' perceptions and suggestions for the interface used in the online assessment system are also discussed. These findings represent important considerations for developers of new, innovative assessment systems.

## **Promises and Challenges in Teacher Implementation of Instructionally Embedded Assessments**

State education agencies and test developers have begun to develop new, innovative assessment systems to meet statewide accountability requirements, including under the Innovative Assessment Demonstration Authority. As additional states explore options for flexibility available under the Every Student Succeeds Act (ESSA), it is important to share lessons learned from existing programs. However, as most new assessment systems are still under development, there is limited information available about the process of implementing innovative assessment models in schools and how teachers view them.

The Dynamic Learning Maps (DLM) Instructionally Embedded Alternate Assessment System (DLM Consortium, 2022) provides an example of an innovative assessment system. Unlike traditional summative assessments, instructionally embedded assessments are administered during the school year, following instruction, and results provide teachers with actionable data to inform classroom instruction (Clark & Karvonen, 2021). Instructionally embedded assessment results can also be combined to produce summative scores (Clark et al., 2017), useful for state accountability purposes. As a fully operational assessment system since 2015, DLM instructionally embedded assessments provide a unique opportunity to explore teacher perceptions of an innovative, instructionally embedded assessment system and share lessons learned with other programs considering similar innovative testing models.

### **Innovative Assessment Models**

Since the No Child Left Behind Act of 2001 (2002) and later Race to the Top (2010), state education agencies have been required to use statewide academic assessments to meet accountability requirements. While summative assessments adopted by states have been useful for reporting overall achievement, they have received criticism from teachers, parents, and policymakers concerned about the impact of high stakes testing on instruction (Olson & Jerald, 2020) as well as unintended consequences such as test anxiety and students' ability to fully show their knowledge and skills (Spoden et al., 2020). Assessment results from traditional summative assessments are typically released during the summer, when students are out of school, and report students' subject-level achievement rather than specific skills or abilities, which lack utility for teachers, parents, and students who want to use test results to improve instruction and learning (e.g., Marion, 2018).

Many state education agencies have implemented balanced assessment systems to provide timely, actionable data for instruction while still meeting state accountability requirements (e.g., Gong, 2010). However, because balanced assessment systems are comprised of multiple types of assessments (e.g., formative, interim, and summative), they can increase total testing time, which can detract from instruction (O'Keefe & Lewis, 2019). Additionally, the different types of assessments that make up a balanced assessment system may lack coherence with one another (Marion, 2018), creating challenges for educators in preparing students for assessments and connecting results to instruction.

The Every Student Succeeds Act (ESSA) provides states with options for flexible assessment systems. For example, five states have pursued innovative assessment systems as part of the Innovative Assessment Demonstration Authority (IADA) program (Evans & Marion, 2021). Among the five programs, three states (New Hampshire, Louisiana, and North Carolina) have shortened their end-of-year assessments and supplemented summative scores with a combination of interim content-based (Louisiana

Department of Education, 2018) or performance assessments (O’Keefe & Lewis, 2019). One of the remaining states, Georgia, is currently piloting two new assessment systems (GMAP and NAAVY), while Massachusetts continues to refine the process for its multi-part assessment system. For further discussion of these and other innovative state accountability assessments, see O’Keefe and Lewis (2019) and Evans and Marion (2021).

### **Dynamic Learning Maps Instructionally Embedded Assessments**

Several states have also pursued innovative testing models outside the IADA. DLM alternate assessments measure academic achievement for students with the most significant cognitive disabilities in English language arts (ELA), mathematics, and science. While most states adopt a year-end administration model, six states administer instructionally embedded assessments as their state accountability assessment. States adopting the year-end model have access to instructionally embedded assessments for optional use ahead of their spring administration window, though use has historically been limited. DLM instructionally embedded assessments are designed to integrate assessment into instruction, whereby teachers select content, provide instruction, assess student learning, and use the results to inform subsequent instruction.

DLM instructionally embedded assessments prioritize flexibility, accessibility, and student outcomes, allowing teachers to select the content, complexity, timing, and frequency of testing that is most appropriate for each student (within constraints). During fall and spring testing windows, teachers select academic content standards from an assessment blueprint based on state and local testing requirements. These requirements ensure construct representation across each subject (e.g., the teacher should select at least two standards within each of four subdomains). Each standard is available at five complexity levels to provide all students access to grade-level academic content. The levels for assessment were determined from an underlying research-based learning map (Thompson & Nash, 2022), which specifies the order of acquisition of knowledge, skills, and understandings spanning from early foundational representations to college and career ready skills. Small sections of the maps (“mini-maps”) are made available to teachers as instructional resources.

DLM instructionally embedded assessments are administered using an online platform (see screen shot in Figure 1). In the online platform, teachers create instructional plans for the standards and complexity levels of their choosing to meet blueprint requirements. The online system recommends a level for each standard based on a survey the teacher completes specifying the student’s academic and expressive communication skills, and the teacher has the option to accept the recommendation or assign a different level. Each assessment, called a testlet in the DLM system, is for a single standard and level and consists of three to nine items. Testlets are delivered after the teacher indicates instruction is complete.

The online system provides results on demand, reflecting mastery results for testlets completed to that point. Assessments are scored using a diagnostic classification model, which provides mastery information for each assessed standard. At the end of the year, mastery results from both the fall and spring windows are compiled into summative performance levels describing overall student achievement in the subject (Clark et al., 2017). These end-of-year results are used for state accountability purposes. Technical documentation summarizes validity evidence for intended uses of the system (Clark & Karvonen, 2021; DLM Consortium, 2022).

Because DLM instructionally embedded assessments have been operational since 2014-2015, the system provides a unique opportunity to learn and share valuable information about teacher perceptions of of an instructionally embedded assessment system. Other programs and agencies interested in adopting new, innovative assessment models to meet both instructional and state accountability needs can use information from the current study to inform their decisions. We specifically explored the following research questions:

1. What are teachers' perceptions of the instructionally embedded assessment model?
2. What are teachers' perceptions of instructional uses of instructionally embedded assessments?
3. How do teachers use instructionally embedded assessment results?
4. What are teachers' perceptions of the online system used for embedded assessment?

### **Method**

We conducted semi-structured focus group interviews in November 2021 with teachers who administer DLM instructionally embedded assessments to collect feedback about their perceptions of the assessment model. A total of 30 teachers participated in 11 focus groups.

#### **Recruitment**

We aimed to recruit a range of educators to share diverse perspectives on instructionally embedded assessments. Using system extracts, we identified students in grades 3-8 and high school who took DLM instructionally embedded assessments in the 2020-2021 school year, along with their teacher. To sample a range of teachers, not just those who used the system extensively or might provide positive feedback, we assigned each student a blueprint coverage category (i.e., did not meet, met, or exceeded blueprint requirements). Next, we reduced the sample to include teachers who had administered at least two instructionally embedded testlets, to maximize their likelihood of remembering the system and how they had used it.

We created a sampling frame of 162 teachers by randomly selecting teachers from each blueprint coverage group. We prioritized selection of educators who had a roster of four or more students so they would have a range of student experiences to share; we also sought to include teachers from a range of DLM states (including teachers from states adopting the instructionally embedded model for accountability and those who used the optional instructionally embedded assessments in the year-end model). We invited teachers to complete a Qualtrics survey to indicate their interest in focus group participation, and sent consent forms and information about honoraria prior to the focus group sessions.

#### **Participants**

Thirty teachers from six states participated in the focus groups. All but one participant (29, 97%) represented states using the DLM instructionally embedded model for state accountability purposes, while one participant was from a state adopting the DLM year-end assessment model but had administered optional instructionally embedded assessments during the year. Participating teachers had a range of teaching experience. The median ELA and mathematics teaching experience was 15 years, while the median for science and teaching students with significant cognitive disabilities was 12 years. This amount of experience is consistent with the broader DLM population of educators, based on an

annual teacher survey (DLM Consortium, 2022). Teachers in the focus groups had a median 8 years of experience administering alternate assessments. Teachers’ caseloads ranged from one to eight students.

Nearly all ( $n = 28$ ; 93%) participating teachers identified as female, and two (7%) identified as male. Most ( $n = 27$ ; 90%) of the participants were White, 29 (97%) were non-Hispanic/Latino, one (3%) participant indicated Hispanic/Latino ethnicity, and three participants (10%) were African American. Ten (33%) participating teachers taught in rural settings, 12 (40%) teachers taught in town or suburban settings, and eight (27%) taught in a city.

### Data Collection

Teachers participated in focus groups online through the Zoom platform. Each focus group began with a review of the informed consent materials, then followed a semi-structured interview protocol (Appendix A). Following completion of the sessions, participants received a \$50 honorarium. The focus group transcripts were professionally transcribed for review and de-identified for analysis. Because testing requirements can vary by state, some location data (e.g., state) were retained for analysis.

### Data Analysis

The research team used content analysis with a combination of deductive and inductive approaches to identify themes and patterns in the focus group data related to the research questions. We developed a codebook (see Appendix B) containing codes to apply to the transcripts. We identified initial codes based on the research questions and focus group interview protocol. We then conducted a microanalysis (i.e., detailed line by line analysis) of the transcripts (Strauss & Corbin, 1998) and added codes where needed. As shown in Table 1, the final coding protocol contained four broad categories aligned to the four research questions: perceptions of instructionally embedded assessments, use of instructionally embedded assessments to inform instruction, use of instructionally embedded assessment results, and perceptions of the online assessment system. Within those categories, 29 codes were identified and applied to the transcripts.

*Table 1: Categories developed for coding analysis*

<b>Coding Category</b>	<b>Description</b>
Perceptions of Instructionally Embedded Assessments	Positive and negative perceptions of instructionally embedded assessments, including comparisons with other types of assessments, flexibility in assessment timing, and state/district guidance influencing instructionally embedded use
Use of Instructionally Embedded Assessments to Inform Instruction	Comments related to the impact of instructionally embedded assessments on instruction, including informing individualized education programs (IEPs) and how embedded assessments are incorporated into instructional practice
Use of Instructionally Embedded Assessment Results	Comments related to use of results, including how fall results impacted spring instructional planning, and use of results from the previous year
Perceptions of the Online Assessment System	Positive and negative appraisals of the assessment system interface, including comments related to the use of system planning functions to scaffold instruction and assessment, and suggestions for improving the online system

Each focus group transcript was coded by two of the four members of the research team, with pairings counterbalanced across all transcripts. Research team members reviewed the transcripts, applied codes to segments of text, then compared codes and discussed discrepancies to come to consensus. After all coding was completed, the research team used thematic analysis procedures derived from Silverman (2021) to identify themes in the data, relying on the quoted excerpts of text for each research question and reports of code frequencies and code occurrences. Following initial analysis, each research team member reviewed the themes for a different research question to confirm the themes present in the data and determine if any were missing or required revision. Quotations are included for transparency (Eldh et al., 2020).

We used various methods to demonstrate rigor of the study (e.g., credibility, trustworthiness, transferability; Amankwaa, 2016). We triangulated perspectives by including diverse participants (Patton, 1999). During focus group sessions, facilitators used clarifying questions and paraphrasing to confirm intended meaning. The research team collaborated on all aspects of the study, including use of multiple researchers for interviews and coding. A codebook promoted coding consistency across researchers.

## Results

Results are organized according to the four research questions.

### Perceptions of Instructionally Embedded Assessments

Overall, teachers shared positive perceptions of instructionally embedded assessments. Several teachers liked the proximity of assessment to instruction and the ability to assess students throughout the year to track progress.

*Well, it just fits right into where you're teaching. I mean because when you give an assessment and you kind of see how they did, you can just know where to focus on the next time. . .*

*I like being able to do it at different times throughout the year using it at the beginning and end of year. . . So just a constant check on exactly where they're at is important to me and I guess I like that.*

Teachers described how having more frequent testing spread out over time was better for their students' recall and reduced test anxiety.

*I think this [testing throughout the year] is a better option just for our students because waiting until the end of the year, they probably wouldn't remember. I don't even know if a big, cumulative end of the year assessment, I don't even know if that's a good measure of what our students can do.*

*When we just break it up, and my student has just a little 10-minute testlet, it doesn't cause them anxiety. They don't get overwhelmed.*

However, some teachers felt that the instructionally embedded assessments required too much testing and could be stressful for students.

*To me, it doesn't really matter [taking a summative or instructionally embedded assessment]. I would say that I know testing can be stressful for them. One of the students is very aware, and so*

*he would say, "I'm taking a test. I'm taking a test," and I'd have to kind of talk him down. I feel like more for him would be more stressful rather than the end-of-the year, but I don't know.*

Teachers liked the flexibility to select assessment content and timing for individual students and expressed that it allows them to meet their individual students' needs.

*It's really nice that I get to pick to kind of tailor to the student . . . They're totally different students, and they have different levels of where they're at. It's really nice that I can pick what meets his needs and what meets his ability level so I can target those in the assessments.*

*That flexibility to teach where the kid is at, not where the whole world wants him to be and help him get there in a logical way is what really works for me.*

Teachers described several different approaches related to the timing of assessment administration for their students. Some teachers start assessing early in the window. Others start midway through the assessment window or spread out their testing throughout the entire window.

*I kind of take a few at a time. I'll go in and kind of teach it. Okay, we're going to assess these couple [standards]. And then we take a little break. And we assess a couple more. I never do it kind of all at once. We do a few at a time.*

*I usually try to teach several of the concepts and then provide the assessment, so I'm doing a couple [testlets] at a time usually that kind of correlate together.*

A few teachers chose to assess a standard when it's covered in their curricula (e.g., corresponding to established pacing guides), while others base the timing of assessment on their students' moods.

*I don't want him to take a test on a day when he is just not having it. It wouldn't be a good day for me to have a test on a bad day even though sometimes I have to. But I want him to be able to show what he knows, so I look for if it's a good day, and I have that flexibility. . .*

Most teachers described using the same approach for choosing assessment content across both fall and spring windows, but several described selecting higher levels in the spring in order to "push harder" to "see how close we can get to [the grade-level target]."

A few teachers noted that they choose to assess students beyond the blueprint minimum requirements to challenge the student. One teacher described the careful balance necessary to maintain high expectations for students while providing opportunities for the student to feel motivated and successful.

*I challenge. I do test above and beyond [blueprint requirements], just because at the high school level, I feel like we need to kind of push them a little bit, just to kind of see, and I know parents like that feedback. They like for us to tell them that their kid, you know, could do a little bit more than expected.*

Some teachers felt that the DLM instructionally embedded assessments required more work and more time for the teacher than traditional summative assessments. A few expressed concern about completing all required testing in the allotted time, especially for states or districts that have a shorter assessment administration window. One teacher shared, "I have the test window and as soon as she says, 'Go,' well, I start then and just kind of go forward in crunch time to get it done by the deadline." Another shared, "I'm always terrified I won't get it done."



Teachers also expressed that the DLM instructionally embedded model requires more organization and monitoring than a summative assessment. Some teachers reported challenges when they had multiple students to assess, such as selecting content for administration and coordinating the testing (e.g., pulling students out of class).

When asked what advice teachers would give to a new teacher starting instructionally embedded assessments, several said to “*start early.*” One teacher said, “*I think you just have to know that it’s something you check every week,*” while another suggested making a calendar and timeline to follow to “*keep on track and not get behind.*”

### **Use of Embedded Assessments to Inform Instruction**

Teachers described how they used the instructionally embedded assessments to inform their instructional practice. Teachers found the complexity levels and learning maps useful, using them to identify skill progressions related to the content standards and to guide instruction.

*One of the standards was, I don’t know what the actual full name is, but recognizing separateness, set, and subset, and the goal is to explain rates and ratios. So we will start off at that basic level and work through the map because, again, if they don’t have that, they can’t recognize separateness. They’re not going to even understand what a rate or ratio is.*

*I like how it [use of levels for standards] focuses on what’s next. When you say, they’re at pre-proximal, I like that they are separated more. And that you can look at, okay, so they’ve got this one, so now I need to be able to work up to this one.*

Teachers particularly liked being able to identify areas of focus for individual students.

*It just fits right into where you’re teaching. I mean because when you give an assessment and you kind of see how they did, you can just know where to focus on the next time.*

*I think it just again tells me where they are and what we need to really hone in on.*

Teachers also described using the learning maps to create their lesson plans.

*I like the [mini] maps so that if you assign something and you’re going to test it then you kind of know how to create your lesson plan and what needs to be covered before you test that subject. It really helps in creating my lesson plans.*

*DLM takes the math and the science and the literature and gives you a logical ladder with lesson planning kind of canned for you already. Here’s where he is. Here’s what we’re working on. Here’s what we’ve mastered. Here’s what we’re doing next.*

However, because of the level of detail in the maps, some teachers indicated they can be confusing. One teacher stated, “*Sometimes I look at those math or those ELA mini maps and I sometimes do not know how to get started.*”

Teachers described how the assessment blueprint informs their instruction.

*It gives me more of a focus ... to kind of plan what I’m focused on as compared to giving me a list of things that we have to do and just try and get it all done at the end of the year. It can be kind of like a pacing guide in a way.*

*I just look at where we have to assign the testlets, which ones we're going to teach and I pick that and I just kind of look at that and then I look at the map and see what was coupled before that and that's kind of what I use as my pacing guide.*

Teachers also used the DLM instructionally embedded assessments to inform students' IEPs.

*We use it on our IEPs and our present levels of performance and stuff like that. So, you know, to indicate how are they doing, and what goals are we going to make and what objectives.*

However, because DLM instructionally embedded assessments are administered at fall and spring intervals, teachers described needing more frequent progress monitoring data to evaluate IEP goals.

*It doesn't give me data because I don't do it enough. You know, I need weekly data for that. But it does give me just another layer of data I can look at and make sure things are aligning and agreeing, and if there's glaring differences, then I need to spend some time kind of considering why and looking at where else I need to go with it.*

### **Use of Instructionally Embedded Assessment Results**

Teachers described using assessment results to plan instruction in both the fall and spring academic terms.

*[When planning fall testing,] I will typically look back at how they did the previous year and I'll kind of consider where they were at [then].*

*If they didn't master it the first semester, then we will redo it the second semester on the same level; but if they master at that level [in the fall], I try to see how far they can go up [in the spring]... We just push until we no longer master it for the second semester.*

Teachers also described using the results to monitor students' progress during the school year and across grade levels.

*I think it's good to see if they did okay in the fall, how did they do in the spring, and if they really maintained it [that skill level].*

*I've had the same kids the last three years and now they're eighth graders... I still have a core group of six that I started [DLM] with, so they are the ones that I'm able to have seen how much they've grown. They've done really well.*

Many teachers described sharing the assessment results with parents. They stated that the DLM instructionally embedded assessment results made it easier to demonstrate students' ongoing progress than previous portfolio-style assessments. One teacher reported using the results to give parents materials to work with the students at home.

*I like how that [the score report] has that progression [of linkage levels]. Because then you can tell parents, they can do this, but now we're working toward this.*

*You can build all the portfolios you want, but progressive data within one standard was very hard to establish before I was introduced to the DLM and I think if anything, being able to show a parent beyond a shadow of a doubt the progress that their child is making...has made a big difference to me.*

However, some teachers experienced barriers that prevent them from using DLM instructionally embedded results as intended. For example, a few did not know they could access their students' results in the online interface during the year.

*I just note what I see [correct responses for] during the testing session. I didn't know that we could pull up the results at the end of each assessment window.*

*I've just gotten the end of the year report [from school administrators]. I guess I've never learned how to do the others.*

Other teachers shared that they desired additional training specific to accessing and using results to support instruction and learning. They emphasized that the DLM annual training prioritizes information on assessment administration and the test format and does not include enough on understanding and using score reports.

*I didn't know that you could pull up reports because all the DLM training that is provided is how to give the test and what the test is about and all those things. It doesn't train us on how to get reports and how to use the reports.*

*It would be far more beneficial [for the teacher/administrator training] to say this is what we do, this is how you use it, this is how you read it, this is where you find your information so you can make those [instructional] decisions... I want to be able to use that information.*

### **Perceptions of the Online Assessment System**

Teachers reported that the layout of the online system (Figure 1) supported them in planning instruction and meeting instructionally embedded blueprint requirements for content coverage across areas.

*You're able to know exactly, you know, how many [standards] you have to put, from what areas, you know, and then as you complete them and all, you know, it's like showing you how many you have completed, you know...it's way more user friendly than what it used to be.*

*I love now that I can just quickly go on there and when you first log in, you can see [testing] requirements met and I love that...that's one of the things that I like. It's just right there for you.*

The layout also supported them in selecting the appropriate complexity level for each content standard.

*It's much easier to see what the levels are, see the [content standards]...and pick what we – if we don't think that the computer has picked appropriately or if it's the first time, picking the [complexity] level where we think our students will be able to show mastery. It's much easier that way, which helps us a lot.*

When asked how the assessment system interface could be improved, teachers offered several suggestions, ranging from adjustments to layout and design, to incorporating additional visual indications of student status, to a streamlined process for generating reports. Many teachers expressed a desire for system interface changes that would provide them with more information at a glance, while others described approaches to data visualization and reporting that would assist them in monitoring student progress and planning instruction across grade levels.

*If I'm going to my instructional [sic] and assessment planner, maybe if there was some type of caution-looking symbol for each student whenever [the test blueprint is] not completed.*

Project staff will consider these points of feedback for informing future updates to the online system.

### **Discussion**

As states increasingly adopt flexible, embedded, and through-course assessment designs, it is important to collect stakeholder feedback from users in the field that can inform design, development, and use of such systems. Instructionally embedded systems, like DLM, offer unique benefits over traditional testing formats, but also introduce unique challenges. For instance, assessment administration is individualized to the student, within constraints. While teachers value and appreciate being able to tailor each student's experience to better meet individual learning needs or instructional goals, they also shared that having too many elements of choice or flexibility could potentially introduce challenges with larger caseloads. States and districts considering adoption of instructionally embedded or other flexible assessment administration models should consider the amount of flexibility they want to support and collect educator feedback on what level of individual student customization would be feasible to implement given other factors, such as length of test administration windows.

When adopting innovative test designs, programs should also consider the amount and scope of training needed for educators to be able to administer the assessment as intended and use results effectively. The teachers we interviewed emphasized that training is important, and they also desired additional information about how to access and use instructionally embedded assessment results. Because DLM training prioritizes information about administration, teachers felt prepared to administer assessments, but not necessarily on how to use results. Several disclosed they did not know they could access results during the window. This finding led to important changes to the online system layout to make finding and accessing results clearer. This finding also underscores the importance of evaluating training and resource materials and their relevance to educators. Because educators have limited time to devote to training, not only is length of the training a primary consideration, but also criticality of information. Programs should collect stakeholder feedback and make adjustments so that educators have access to the most important information, and also know where to find additional resources outside of the training should they want additional information.

Usability of the online interface is another important factor in teachers' implementation and use of instructionally embedded assessments. The DLM online interface was specially designed to support teachers when administering the instructionally embedded assessments, informed by an educator cadre group that met regularly. Yet, even with an educator-focused design process, the present focus groups identified additional areas for improvement. As such, we encourage other testing programs to leverage educators in providing feedback not only in early design phases of the work, but periodically over time, so that online systems meet user needs as those needs may evolve or change.

Across the major theme categories, we also noticed variability in state and local policies that impacted user experience with the system. For instance, some states or districts have additional blueprint constraints beyond consortium-level requirements; states and districts also vary in the length of their test administration windows and the amount of training and support they provide teachers. As such, educator perceptions of and experience with instructionally embedded assessments varied based on how much time they perceived they had to complete requirements, how stringent their local

requirements are, and how much support they perceived having. Some described the benefits of having strong building and district support, while others described being the only one in their area or not having any support in their building or district. Recognizing that these elements can impact educator use of the system, programs adopting innovative testing models should evaluate how their decisions may impact implementation, including potentially unintended negative consequences.

While the present study identified numerous strengths of flexible, instructionally embedded assessments, we also note limitations of the study. While we sought to identify a representative sample of teachers across states, teaching locations, and demographic groups, participating teachers may not fully represent all users in the population or the full range of perspectives. Regardless, we believe this study provides an important contribution to the field on educator perspectives of instructionally embedded assessments that can be useful to informing design and implementation of other innovative assessment programs.

### References

- Amankwaa, L. (2016). Creating protocols for trustworthiness. *Qualitative Research Journal of Cultural Diversity*, 23(3), 121–127.
- Clark, A. K., & Karvonen, M. (2021). Instructionally embedded assessment: Theory of action for an innovative system. *Frontiers in Education*, 6(724938).
- Clark, A. K., Nash, B., Karvonen, M., & Kingston, N. (2017). Condensed mastery profile method for setting standards for diagnostic assessment systems. *Educational Measurement: Issues and Practice*, 36(4), 5-15.
- Dynamic Learning Maps (DLM) Consortium. (2022). 2021 – 2022 Technical Manual: Instructionally Embedded Model. University of Kansas, Accessible Teaching, Learning, and Assessment Systems (ATLAS).
- Eldh, A. C., Årestedt, L., & Berterö, C. (2020). Quotations in qualitative studies: Reflections on constituents, custom, and purpose. *International Journal of Qualitative Methods*, 19, 1609406920969268. <https://doi.org/10.1177/1609406920969268>
- Evans, C. & Marion, S. (2021). *Summary of IADA Programs Approved as of Sept. 2021*. Center for Assessment.
- Gong, B. (2010). *Using balanced assessment systems to improve student learning and school capacity: An introduction*. Council of Chief State School Officers and Renaissance Learning. <https://silo.tips/download/using-balanced-assessment-systems-to-improve-student-learning-and-school-capacity>
- Kennedy, C. A., & Wilson, M. (2007). *Using Progress Variables to Interpret Student Achievement and Progress* [Technical Report No. 2006-12-1]. Berkely Evaluation & Assessment Research (BEAR) Center, University of California, Berkeley.
- Kingston, N. M., Karvonen, M., Bechard, S., Erickson, K. A. (2016). The philosophical underpinnings and key features of the Dynamic Learning Maps Alternate Assessment. *Teachers College Record*, 118.

- Louisiana Department of Education (2018). *Louisiana Believes: Louisiana's Application for Innovative Assessment Demonstration Authority under Section 1204 of the Elementary and Secondary Education Act (ESEA)*. US Department of Education, <https://www2.ed.gov/admins/lead/account/iada/laiadaapplication.pdf>.
- Marion, S. (2018). The opportunities and challenges of a systems approach to assessment. *Educational Measurement: Issues and Practice*, 37(1), 45–48. <https://doi.org/10.1111/emip.12193>
- O'Keefe, B., & Lewis, B. (2019). *The State of Assessment: A Look Forward on Innovation in State Testing Systems*. Bellwether Education Partners.
- Olson, L., & Jerald, C. (2020). *The big test: The future of statewide standardized assessments*. Future Ed.
- Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Services Research*, 34(5 part 2), 1189–1189.
- Silverman, D. (2021). *Qualitative Research* (5th ed.). Sage.
- Spoden, C., Fleischer, J., & Frey, A. (2020). Person misfit, test anxiety, and test-taking motivation in a large-scale mathematics proficiency test for self-evaluation. *Studies in Educational Evaluation*, 67, 100910. <https://doi.org/10.1016/j.stueduc.2020.100910>
- Strauss, A. and Corbin, J. (1998): *Basics of qualitative research: techniques and procedures for developing grounded theory* (2nd ed.). Sage.
- Thompson, W. J., & Nash, B. (2022). A Diagnostic Framework for the Empirical Evaluation of Learning Maps. *Frontiers in Education*, 6, 714736. <https://doi.org/10.3389/educ.2021.714736>
- Wilson, M. (2009, December). Assessment for learning AND for accountability. Keynote presentation at the Exploratory Seminar: Next Generation K-12 Assessment Systems. Educational Testing Service, Princeton, NJ.

Figure 1

Screenshot of Instruction and Assessment Planner Tool

SETTINGS
MANAGE TESTS
REPORTS
HELP

Fall Window

District 6 / School 11 / English Language Arts

[Tran, Anne](#)

State ID: 10000536

First Contact Survey

PNP Profile

Credentials

Themes

Fall Performance

Select an Essential Element and linkage level.

After completing all blueprint requirements, you may receive a field test testlet. To access the Testlet Information Page for the field test, go to the [Test Management](#) tab.

Between the close of the Fall window and the opening of the Spring window, you may still view student plans and print the Fall Essential Element Status Report. No other work may be completed.

Instruction In Progress

Testlet Assigned

Testing In Progress

Complete

Recommended Linkage Level

Mastery Demonstrated

Mastery Not Demonstrated

Results Not Available

Choose one EE in C1.2 (L, RL or RI). Complete

Claim: ELA.C1 Students can comprehend text in increasingly complex ways.  
Conceptual Area: ELA.C1.2 Construct Understandings of Text

Essential Element	Initial Precursor	Distal Precursor	Proximal Precursor	Target	Successor
<b>ELA.EE.RL.9-10.1</b> Determine which citations demonstrate what the text says explicitly as well as inferences drawn from the text.	Identify concrete details in a familiar story	Answer questions by referring to a text  <span style="background-color: #0070C0; color: white; padding: 2px;">Testlet Assigned 01/13</span>	Cite textual evidence for explicit information in text	Discriminate between explicit and implicit citations	Determine a narrative's explicit meaning
<b>ELA.EE.RL.9-10.2</b> Recount events related to the theme or central idea, including details about character and setting.	Identify the forward sequence in a familiar routine	Identify main idea  <span style="background-color: #0070C0; color: white; padding: 2px;">Instruction In Progress 01/13</span>	Identify details related to the theme of a story	Recount events contributing to the theme using details	Recount main events related to the theme
<b>ELA.EE.RL.9-10.4</b> Determine the meaning of words and phrases as they are used in a text, including idioms, analogies, and figures of speech.	Identify descriptive words	Identify the words or phrases to complete a literal sentence  <span style="background-color: #4F81BD; color: white; padding: 2px;">Complete: Mastered 01/17</span>	Determine the meaning of idioms and figures of speech	Determine the meaning of words and phrases	Determine the meaning and impact of words and phrases

**Note.** Teachers use this interface to select content standards (called Essential Elements in the figure) and complexity levels for instruction and assessment (indicated as Initial Precursor, Distal Precursor, Proximal Precursor, Target, Successor in the figure). The status indicates instructional plans in progress, testlets ready for administration, and mastery of completed assessments.

## Appendix A

### Focus Group Protocol

During the focus groups, participating teachers will be asked to provide some background on their teaching experience, caseload, and how many students participate in DLM instructionally embedded assessments; and then will be asked to describe how they use the instructionally embedded assessments. The focus group questions are framed around the logic model for instructionally embedded assessments.

#### Background

- 1) Tell the group a little about yourself and your background.
  - a) How long have you been teaching/working with students with significant cognitive disabilities?
  - b) What subjects do you teach? Grades?
  - c) How long have you been administering DLM instructionally embedded assessments?
  - d) For how many students do you typically administer instructionally embedded assessments?
- 2) What types of assessment do you use in your daily instruction? (Listen for whether they are doing any types of formative assessment)
- 3) What kind of information do you currently use to plan your instruction?

#### Plan and Instruct

- 4) How do you currently use the Instruction and Assessment Planner?
  - a) Probe for how they use for instructional planning/monitoring/adjustment, planning across/within students, using results to inform next steps, whether they receive any state/district guidance regarding use of the system
- 5) What features of the Planner work well when planning instruction?
  - a) Probe for any specific features of the system that support instructional decision making or areas for which they recommend improvements (including how it has improved practice from prior years, for teachers who previously used ITI)
  - b) Listen for whether they like being able to select EEs, adjust LLs, create plans on their own timeline, etc.
  - c) What would you change about the Planner if you had a magic wand?
- 6) How do you make plans for your students? How do you know when you need to change your approach?
- 7) What role does the DLM instructionally embedded assessments have in your instructional planning?
  - a) Does it help you determine what to teach? If so, how?
    - i) How do you determine your plans for covering the blueprint?
    - ii) How do you select which EEs to instruct and assess?
    - iii) Do you typically choose recommended LLs or adjust, and why?
  - b) Does it help you determine how to teach/what instructional methods to use? If so, how?
  - c) Does it help you to determine how to group students for instruction?
- 8) What role does the DLM instructionally embedded assessments have in your IEP goal setting?

#### Assess

- 9) Do you typically aim to meet exact blueprint requirements, or do you assess more than the required number of EEs, and why?



- 10) How often do you re-assess EEs?
  - a) What is your primary reason(s) for re-assessing EEs?
  - b) How do you determine which EEs to re-assess?
  - c) How do you determine the linkage level for EEs that you re-assess?
- 11) How do instructionally embedded assessments compare to other assessments you have administered to these students (e.g., old portfolio assessments, etc.)?
  - a) What works better about instructionally embedded assessments? Worse?
  - b) Do you find DLM assessments to be instructionally relevant? Does the content of testlets look like activities you do during instruction?
  - c) Probe for whether they see benefit to conducting assessments throughout the year, whether other students not currently taking DLM assessments would benefit from access, potential barriers to implementing with a whole class

### Evaluate

- 12) To what extent do you believe the DLM assessment results reflect what your students know and can do?
- 13) To what extent do you review student results after the student has completed a testlet? (either in main Planner interface or via the progress report?)
  - a) Probe for frequency, what they do with the information
- 14) What types of decisions do you make based on assessment results?
  - a) How do results inform instruction?
  - b) Do you use results to determine what EEs and LLs to assess next? If so, how?
  - c) What decisions would you like to be able to make that the system currently does not support?

### Overall system use

- 15) Do you use the system similarly in the fall and spring assessment windows? If you use it differently, please describe how fall and spring use differs.
- 16) How well do the instructionally embedded assessments fit into your daily routine in your classroom?
- 17) Do you see benefits in spreading out testing throughout the year compared to a traditional end-of-year assessment? Why or why not?
- 18) Could you see using this type of assessment approach for a whole class?  
How could the system be improved to better support you in your instruction? What additional types of support would be helpful?

## Appendix B

### Codebook for DLM instructionally embedded Focus Groups Analysis

Code	Subcode	Label	Definition
<b>IE Model</b>			
PIE+/-		Dis/likes IE Assessments	Positive and negative perceptions of IE assessment
	FLX	PIE-Flexibility	Perceptions of the IE assessment flexibility
EE+/-		Experiences with EE	Comments related to EEs including positive and negative appraisals
LL+/-		Experiences with LL	Comments related to LLs including positive and negative appraisals
AT+/-		Assessment timing	Factors that influence when students are assessed on EE/LL (i.e., logistics of assessment administration, preparedness of the student, etc.) including positive and negative appraisals of choosing when to assess each student
DIFS		Difference between Fall and Spring	Discussion of any elements that differ between the Fall and Spring administrations
IPMD		Instructional plans multiple DLM	Strategies for creating instructional plans, implementing instruction, and assessing multiple DLM students
GUID		State/district guidance	State/district guidance impacting IE use
SUG	Other: Suggestions	Comments that include suggestions for improvements/modifications	SUG
AD		Advice	Advice to new users of the IES
<b>Use Results</b>			
DATA		General use of result data	Any mention of how result data was used
	TRU	Testlet result usage	How were testlet results used
	LY	Last year	Use of last year's results
	DNU	Did not use results	
	FR	Fall results	How the results from the Fall administration are used (if at all) in determining the Spring administration

Code	Subcode	Label	Definition
<b>Planner Use</b>			
BP		Blueprint	Comments related to the Blueprint
PT+/-		Dis/likes Planner tool	Positive and negative appraisals of the Planner tool
	PTS	Planner tool suggestions	Suggestions for improving the Planner tool in meeting assessment requirements and connecting assessment to instruction
IIP		Instructional info preferences	Preferences for receiving instructional info on IEAS and Planner
<b>Instructional</b>			
IEP		IEP	Comments related to IEPs
	REC	Other: Resources	Comments addressing resources specifically for following through with IE implementation
IOI		Impacts on instruction	Impacts of having instructionally embedded assessment on instruction including how embedded assessments are incorporated into instructional practice
	CP	Other: Conversations with Parents	Refers to discussions with parents
<b>Other</b>			
O		other	Other comments
	NS	Other: New Students	
	PS	Other: Prior Students	Comments referring to work with prior students
	CTXT	Other: Context	Comments related to context for instruction or assessment
	NQ	Other: Notable quotes	Comments of significance related to IE
	GNQ	Other: General Notable Quotes	Important comments that may not be directly related to IE