DYNAMIC® LEARNING MAPS

PROJECT BRIEF

The Impact of COVID-19: Validity Considerations and Scoring and Reporting in Flexible Scenarios

AUTHORS: Amy K. Clark, W. Jake Thompson, Jennifer Kobrin, Elizabeth Kavitsky, Meagan Karvonen

KEY TAKEAWAYS

» Disruptions to instruction and assessment due to the COVID-19 pandemic introduce risks to claims made in the DLM validity argument.
» These risks and their implications for scoring and reporting can be evaluated by identifying and exploring various instruction and assessment scenarios.

INTENDED USES

Results from the DLM alternate assessment may be used for

» instructional planning,
» reporting summative performance in state accountability systems,
» guiding resource allocation,
» and reporting student performance to teachers and parents.

We use the DLM validity argument to evaluate the extent to which intended uses are supported.

INSTRUCTION AND ASSESSMENT SCENARIOS

» **Scenario 1:** School resumes normally.
» **Scenario 2:** School requires alternative scheduling (assessments are administered in school when students are in the building).
» **Scenario 3:** School resumes with multiple disruptions throughout the year due to COVID-19 outbreaks (assessment administration depends on length and timing of disruptions).
» **Scenario 4:** School does not reopen in person at all.
» **Scenario 5:** Testing is halted.

We expect that there will be combinations of scenarios, and that scenarios will change over time within and across states.

BACKGROUND

In a typical administration year, DLM assessments provide results at two levels. The Performance Profile section of the score report describes student performance in the subject overall and supports the intended uses of results in accountability systems, and for resource allocation and reporting student performance to teachers and parents. The Learning Profile section provides detailed reporting of student mastery of individual skills and supports the intended use of results for instructional planning and for reporting student performance to teachers and parents.

The COVID-19 pandemic introduced variability in instruction and assessment that has implications for defensible reporting of results. We proactively evaluated several potential scenarios to inform decisions about scoring and reporting for the 2020-2021 year.

VALIDITY RISKS

We used claims from the DLM theory of action to evaluate potential validity risks under five instruction and assessment scenarios. The risk to claims differs depending on the various scenarios and have implications for the level of reporting that is defensible. The risks also differ depending on the assessment model (Instructionally Embedded, which is a through-course model, or Year-End, which are administered only in the spring).

We focused on claims related to assessment administration, instruction and scoring. Scenario 1 presents no risks because it represents normal instruction and assessment. The other scenarios each have different associated risks that are described below.

	Claim G	Claim H	Claim I	Claim J	Claim K	Claim L	Claim M
Scenario 1	●	●	●	●	●	●	●
Scenario 2	◒	●	●	?	● *	● *	● *
Scenario 3	◒	◒	◒	?	◒	◒	◒
Scenario 4	◒	◒	◒	?	◒	◒	◒
Scenario 5	○	◒	◒	?	◒	○	◒

Note. ● = No risk, ◒ = Partial Risk, ○ = At Risk, ? = Unknown risk.
* Conditional on amount of quality instruction received.

Administration and Implementation Claims

Claim G: "The combination of administered assessments measure knowledge and skills at the appropriate breadth, depth, and complexity." "Appropriate complexity" is potentially at risk for states using the Instructionally Embedded Model in Scenarios 2, 3, and 4, if a lack of regular direct contact with their students hinders teachers choosing the most appropriate level for instruction and assessment. "Appropriate breadth" is potentially at risk in Scenarios 3, 4, and 5 for both models due to a possible lack of time for students to complete all blueprint requirements. Scenario 5 presents the strongest risk to this claim.

VALIDITY

» The DLM validity argument collects evidence to evaluate claims in the underlying theory of action.
» The theory of action is a logic model that details the claims in the validity argument and how they inform one another.
» Different levels of evidence are required for each claim to support the validity of the inferences made from assessment results for intended uses.

ADDITIONAL CONSIDERATIONS

» Each state will choose their reporting option based on their own instruction and assessment scenario(s). States may also decide not to receive any DLM results in 2020-21.
» Hypothetical scenarios may not all be realistic depending on how the year unfolds.
» State education agencies have been encouraged to collect other data, including evidence of opportunity to learn, before making decisions about reporting options or using results for accountability. For more information, see the "Additional Resources" section below.
» Federal guidance may also inform state education agency decisions about the appropriate level of reporting for their state.

ADDITIONAL RESOURCES

» Clark, A. K., & Karvonen, M. (2020). Constructing and evaluating a validation argument for a next-generation alternate assessment. *Educational Assessment, 25*(1), 47-64. https://doi.org/10.35542/osf.io/n6j7w
» Domaleski, C., Boyer, M., & Evans, C. (2020). Accountability interrupted: *Guidance for collecting, evaluating, and reporting data in 2020–2021 (Restart and Recovery)*. Council of Chief State School Officers. https://753a0706.flowpaper.com/CCSSORRAccountabilityv3/
» Marion, S., Gong, B., Lorié, W., & Kockler, R. (2020). *Restart & Recovery: Assessment Considerations for Fall 2020. Council of Chief State School Officers.* https://ccsso.org/sites/default/files/2020-07/Assessment%20Considerations%20for%20Fall%202020.pdf

VALIDITY RISKS CONTINUED

Claim H: "Students interact with the system to show their knowledge, skills, and understandings"
There is potential risk in Scenarios 3, 4, and 5, depending on students' technology access outside of school as well as the quality of accessibility supports available during off-site administration by a trained administrator.

Claim I: "Educators administer assessments with fidelity"
If in-school administration of the assessment is not possible due to disruptions, assessments are administered off-site by a trained test administrator or through in-school testing appointments. Other remote administration is not supported for DLM assessments. There is potential risk in Scenarios 3, 4, and 5, depending on the level of off-site administration by a trained administrator.

Claim J: "Educators provide instruction aligned with content standards and at an appropriate level of challenge"
There is an unknown risk in Scenarios 2, 3, 4, and 5, since it is unclear how the quality of remote learning may compare to in person learning. Disruptions may limit the data normally used as evidence to evaluate this claim.

Scoring Claims

Claim K: "Mastery results indicate what students know and can do"
Claim L: "Results indicate summative performance relative to alternate achievement standards"
Claim M: "Results can be used for instructional planning, monitoring, and adjustment"

These claims feed into each other, and Claim K assumes a full year of instruction. Under Scenarios 2, 3, 4, and 5, Claim L may be at risk, which has implications for use of DLM results in state accountability systems. Scenario 5 presents the strongest risk for Claim L. Under these scenarios, performance results need to be interpreted with caution and cannot be compared to other years' assessment data. In Scenarios 3, 4, and 5 there is potential risk for Claims K and M due to previously stated risks to Claim G, H, and I (if results do not represent what a student knows and can do, teachers may struggle to use results to make important instructional decisions).

SCORING AND REPORTING POSSIBILITIES

The appropriate level of reporting will be based on the risk to the claims. Risks to Claims K and M generally impact the Learning Profile, while risks to Claim L impact the Performance Profile.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
Performance Profile	●	◔	◔	◔	○
Learning Profile	●	◔	◔	◔	◔

Note. ● = Standard, ◔ = Modified, ○ = Not Provided.

Possibility 1 - Full scoring and reporting: The score report would include both profiles and is likely only feasible under Scenario 1.

Possibility 2 - Modified scoring and reporting: Modified reporting requires sufficient evidence for Claims K, L, and M and would likely be seen in Scenarios 2, 3, and 4. Score reports would include both profiles, but have caveat language added to help guard against potential misinterpretations.

Possibility 3 - Limited scoring and reporting: Due to the risks in Scenario 5, the score report would be limited to the Learning Profile, showing student mastery based on content assessed before testing was halted. This would be used as one source of evidence to support instructional planning in the subsequent year. The score report would also include caveat language.

CAVEAT LANGUAGE FOR SCORE REPORTS

The 2020-2021 academic year was significantly impacted by the COVID-19 pandemic. Results may reflect the unusual circumstances for instruction and assessment this year. Use results with caution.