

Promoting Accurate Score Report Interpretation and Use for Instructional Planning

Meagan Karvonen, Russell Swinburne Romine, Amy Clark, Jennifer Brussow & Neal Kingston

University of Kansas

Abstract

The purpose of the study was to evaluate teachers' interpretations and use of individual student score reports that describe results from Dynamic Learning Maps alternate assessments. This paper describes results from two studies on the interpretation and use of alternate assessment (AA-AAAS) score reports. The first study focuses on usability of report contents for communication with parents and instructional planning based on individual and paired interviews with teachers from two states. The second study examined the impact of interpretation resources on educators' understanding of report contents. Findings suggest that teachers are able to use score report contents as intended to inform instruction and develop Individualized Education Programs (IEPs).

Promoting Accurate Score Report Interpretation and Use for Instructional Planning

Alternate assessments based on alternate academic achievement standards (AA-AAAS) are relatively new large-scale assessments for students with significant cognitive disabilities (SWSCD), a small but extremely diverse population. Past limitations of AA-AAAS score reports have included unfamiliar terminology, unclear scoring methods, a focus on deficits, and limited information to guide changes in instruction or supports (Nitsch, 2013). There is also evidence that teachers have not systematically used AA-AAAS results or considered content standards when assessing progress or deciding what to teach after students have mastered academic skills (Karvonen, Wakeman, Moody, & Flowers, 2013).

Score reports for Dynamic Learning Maps (DLM) Alternate Assessments are designed to address these challenges by providing actionable information to guide instructional decisions while also being appropriate for accountability purposes. Assessment results are based on mastery classification (i.e., whether a student demonstrated mastery of skills aligned to learning map nodes) using cognitive diagnostic modeling rather than a scaled score on a latent trait. Summative score reports consist of a fine-grained learning profile and a performance profile that aggregates information across content standards. Each report was first developed by staff based on research literature and refined after multiple rounds of focus groups with educators and parents. Previous research has documented interpretability of the final prototypes (Authors, 2015) and preliminary evidence of how teachers evaluate score report contents (Authors, 2016). The current paper builds on previous research and presents results from two studies.

Background

States participating in DLM assessments follow one of two blueprint testing models: the integrated model, which includes instructionally embedded testing throughout the year in addition to a spring window, and the year-end model, which only includes a spring testing window. Because the

integrated model provides more opportunities for students to demonstrate their knowledge, skills, and abilities, the individual student score reports include high-level summary information in the Performance Profile along with additional information at a more fine-grained level in the Learning Profile, whereas score reports for the year-end model only include the Performance Profile portion of the reports.

DLM Score Reports

Three score reports have been developed at the individual student level: 1) a progress report, 2) an end of year performance profile, and 3) an end of year learning profile. This paper focuses on the second and third reports, which together comprise the individual student summative score report. An example based on the 2014-15 assessment year is provided in the appendix.

The **performance profile** aggregates linkage level mastery information for reporting on each conceptual area and for the subject overall. It contains three main sections. The first section (Overall Results) includes a text summary of the student's performance, including the total number of Essential Elements mastered during the year and the student's final performance level. The second section (Conceptual Areas) reports the percent of skills, or linkage levels, within each conceptual area that the student mastered. This value is calculated as the number of linkage levels mastered (as reported on the learning profile) out of the total number of linkage levels possible for the grade and content area. The third section contains bulleted lists of skills mastered and skills assessed but not mastered. Lists are organized by Conceptual Area.

The **learning profile** shows rows for each Essential Element and columns that correspond to the five linkage levels (initial precursor, distal precursor, proximal precursor, target, and successor). The Essential Elements are grouped by Conceptual Area. Shading is used to distinguish between linkage levels the student mastered, levels assessed but not mastered, and levels not assessed that year.

The reports depict student “mastery” of “skills.” Mastery here is determined using cognitive diagnostic modeling with thresholds set for linkage level mastery. If the student’s probability of linkage level mastery is greater than this threshold, the student is considered a master of the linkage level.

“Skill” refers to a linkage level for an Essential Element.

Table 1 summarizes the components of the performance profile and learning profile that make up the individual student score report. These components were part of the coding scheme used for data analysis and are referred to by number throughout the results section.

Table 1. Components of the DLM 2014-15 Individual Student Score Report

Performance Profile	Learning Profile
1) Overall performance level:	4) Learning profile narrative
a) narrative	5) Conceptual Area and Essential Element codes
b) graphic	6) Mastery information:
c) performance level descriptors	a) Mastered (green)
2) Conceptual areas: bar graphs with subtitles	b) No evidence of mastery (blue)
3) Mastery list:	c) Untested (no shading)
a) Conceptual area headings	
b) Introductory statement	
c) Bulleted statements	

To support correct interpretation and use of reports, a PDF interpretation guide was made available to teachers on an annual basis. However, due to the complexity of the assessments and unique nature of the score reports, the PDF interpretation guide may not provide teachers with enough support. This study evaluated the extent that a video tutorial helped support teacher interpretation of individual student score report contents.

Methods

Two studies were conducted to support interpretation and use of individual student score reports. The studies were designed to answer following research questions:

1. How do participants read and interpret the information in reports?
2. How do participants explain results to parents?

3. How do participants use report contents for educational planning and instruction?
4. Does an online tutorial support teacher interpretation of score report contents?

Research questions 1-3 were addressed via individual and paired interviews with teachers.

Research question 4 was addressed via an online score report tutorial.

Data Collection

Results for the first study described in this paper are based on individual interviews and paired interviews conducted with teachers in two states. Protocols were slightly different for individual and paired interviews but both versions were semi-structured.

The individual interview protocol began with general questions about the participant's background with DLM assessments and previous experience with the score reports. Then the participant was presented with a score report and asked what it said about the student. Participants were asked to think aloud while they read the contents. Probes were used for clarification of responses and to ensure participants attended to each part of the report (e.g., to point them back to a section they skipped). After interpreting each section of the report (i.e., performance profile and learning profile), the participant was asked how they might explain the report to a parent. The same process (initial interpretation and reinterpretation for a parent) was followed for a second, contrasting report. The interview concluded with an opportunity for the participant to make recommendations about resources that other teachers would need to support their interpretation and use of DLM score reports.

The paired interview began with the same general background questions as the individual interview but also included a question about the participants' history of collaboration. The pair was then presented with a score report and asked to talk aloud about their interpretation of its contents. The primary focus of the interview was the use of the report to plan for instruction, including long-term educational planning and for mid-year adjustments to instruction. Participants engaged in unstructured dialog about the contents and in vivo probes were used as needed for clarification and elaboration to

cover both major categories of use (instruction and IEP planning). After repeating the process with a second, contrasting report, the interview concluded with an opportunity for recommendations about resources to support score report interpretation and use.

Both types of interviews used score reports from the 2014-15 academic year with realistic student results but fictitious student identifiers. Sample score reports were prepared in both subjects (ELA and mathematics) and across elementary, middle, and high school grades. Samples were also selected within each subject/grade band to provide contrasting patterns of student performance.

Each interview incorporated two sample reports. The choice of specific reports for each interview were based on the participant's familiarity with the grade band and subject. For example, a middle school teacher who was responsible for both ELA and mathematics might be presented with an ELA grade 6 report for a high-achieving student and a math grade 7 report for a low-achieving student. There was no intentional sequence in which report was presented first.

Results for the second study described in this paper are based on individual participation in an online on-demand score report tutorial. Because of the differences in the individual student score reports for each blueprint testing model, two tutorials were created. Each tutorial included an informed consent portion, followed by pre-test items, the training video, evaluation questions, and a post-test. The video incorporated concepts from the PDF interpretation guide and addressed misconceptions identified in the score report interpretation interviews with teachers.

The pre- and post-test questions included in the tutorial were written by the researchers and by DLM item writers who are familiar with the DLM score reports. The evaluation questions were written by the researchers and included four Likert scale items and two open-ended items.

Participants

Interview participants included 12 teachers from two states and two parent advocates from one state. In the first state, eight teachers taught in a school that exclusively served students with

intellectual and multiple disabilities from 6th grade through age 21. Teacher participants in the first state taught in secondary grades (grades 6-8, grades 9-10, or grades 11-12). Two of the teachers in the second state taught students with intellectual and multiple disabilities at a regional high school. The remaining two teachers taught student with disabilities at two elementary schools in the same district. The students of these teachers received instruction in mixed settings, with some instruction in inclusive classes with their general education peers, and some instruction in separate classrooms for students with disabilities. All of the teachers in both states taught two or more academic subjects. Their years of teaching experience ranged from 1 to 26 years. Eight teachers participated in individual interviews and four more participated in two paired interviews.

Teachers were recruited for participation in the score report tutorial by state partners participating in the DLM consortium. A completion certificate was made available to teachers who passed with post-test with a percentage of 80%, which could be used for professional development credit depending on individual state guidelines. A total of 93 teachers participated in the study. Table X summarizes the number of teachers participating in the study from each state. Of the participating teachers, 58 (62%) indicated they had seen a DLM score report before, while 33 (36%) had not and 2 (2%) did not respond.

Table X. Number and Percent of Teachers Participating in Tutorial by State

State	<i>n</i>	%
Alaska	1	1.1
Iowa	29	31.2
Kansas	20	21.5
Missouri	18	19.4
Oklahoma	23	24.7
Wisconsin	2	2.2

The video tutorial lasted approximately 20 minutes for participants from year-end model states (Alaska, Oklahoma, and Wisconsin) and approximately 30 minutes for participants from integrated

model states (Iowa, Kansas, Missouri). Only 64 participants (69%) completed the video in its entirety, indicating an attrition rate of about 31%.

Data Analysis

Interviews were coded using a two-step process. First, each researcher reviewed each transcript to mark responses related to the primary research questions (i.e., reading/interpretation, explanation to parents, resources to support interpretation, and uses of report contents). During the second step, each researcher added codes to identify the part of the report to which the participant was referring. Thematic codes were also used to identify processes or elements associated with the primary codes. At least two researchers coded each transcript. Final codes were then reviewed to describe overall thematic areas. For example, within responses coded as reading/interpretation, statements were also coded to indicate the types of behaviors (e.g., direct read, question about contents, misinterpretation). A tentative list of codes was developed prior to analysis, based on review of the literature. Codes were added and refined as new ideas emerged from the data. Paired interviews relied on the same codes as individual interviews, but the emphasis was primarily on uses of the contents rather than interpretation.

Results

In the conceptual framework for this study (Figure 1), accurate interpretation of the report is a prerequisite to communication with parents and to the use of results for instructional planning. One of those uses combines instructional considerations and communication with parents through the IEP development process.

Reading and Interpretation of Reports

Participants varied in the parts of the report that they tended to rely on for information. Results are described with numeric references back to the report component listed in Table 1.

Explanations to Parents

Use of Reports

Uses for instructional planning include individualized decision-making, such as in developing Individualized Education Plans (IEPs); planning for the content of instruction; instructional groupings; and identifying needs for further assessment. Participants described a range of uses of the report contents beyond sharing the results with parents. For this paper, uses are roughly grouped into planning for instruction and IEP development.

Impact of Tutorial

Prior to completing the tutorial, teachers reported a range of confidence in their ability to interpret and use DLM score reports. Most teachers reported being *somewhat confident* in their ability to interpret and use the reports, with the least number of teachers reporting being *not at all confident*.

Table X. Confidence in Ability to Interpret and Use DLM Score Reports

Level of Confidence	<i>n</i>	%
Very confident	11	12.0
Somewhat confident	33	35.9
Neither confident nor unconfident	25	27.2
Somewhat unconfident	13	14.1
Not at all confident	10	10.9

Following the training video, evaluation questions were presented to the participants. A total of 55 participants completed the evaluation questions. All respondents either strongly agreed (40%) or agreed (60%) that the tutorial covered important information. Most respondents strongly agreed (25%) or agreed (64%) that explanations provided in the tutorial were clear. Additionally, most respondents reported that they felt prepared to explain DLM score report information to parents [strongly agreed (11%) or agreed (76%)] and to use DLM score reports to inform instruction [strongly agreed (15%) or agreed (65%)].

The evaluation included two open-ended items. The first asked teachers to indicate if they had remaining questions about interpreting DLM score reports. The second asked teachers to indicate what additional resources would help with interpretation and use of DLM score reports. Most teachers

indicated they did not have remaining questions about the score reports. Additional feedback included requests for local training and additional materials that would support instructional planning and decision-making. One participant requested a repository of training videos on different aspects of DLM, which is already available, and indicates a need to make teachers aware of the resources that are available to them to support their use of score reports. Several participants also requested transcripts and print-outs of the sample reports used in the video, which will be made available online for subsequent use.

Post-test items were included following the evaluation section of the tutorial to prevent performance on the quiz from impacting participant evaluation of the tutorial. A total of 18 participants passed the post-test with at least 80% accuracy on their first try. In instances where the participant did not respond to 80% of the items correctly, the tutorial was presented again so that they could re-attempt. A total of 24 participants completed the post-test a second time, with two participants reaching the 80% threshold on their second attempt. Ten participants completed the tutorial a third time, but none achieved the 80% correct threshold.

Practical Implications

The DLM assessment system is designed to support several intended uses of assessment results. Accurate and appropriate interpretations are necessary to facilitate these intended uses. This presentation adds to the body of validity evidence related to the DLM assessment system and to the general body of research on score report interpretation. It also advances research on the use of large-scale assessment results to improve teaching and learning.

In general, teachers reported liking the reports, finding them helpful when talking to parents and useful for guiding instruction. There were relatively few signs of misconceptions when describing report contents. This was probably due in part to reliance on the exact text in the report, preference for mastery statements over aggregated information, and a tendency to not rephrase key meanings. The

fact that we did not see broad evidence of misinterpretations may mean that teachers' misconceptions are more hidden. For instance, they may correctly identify that the conceptual area results are percentages rather than percentiles, but may have a misunderstanding of what "mastery" means that was not evident in these interviews. Different probes and more structured techniques (e.g., paraphrase testing) will be needed to assess this deeper understanding in future interviews.

Results from both the interviews and tutorial were based on limited sample sizes and participants in a subset of tested states. However, the feedback received can Due to the high attrition rate following the video portion of the tutorial, it is recommended to limit the length of tutorial videos while balancing breadth of content coverage to ensure its utility to teachers.

The interpretations and uses of score report information in this study also highlight potential risks and unintended consequences we will want to evaluate in the future. For instance, when the learning profile is viewed as a fixed map for the student's future instruction, a teacher may not teach concepts that come between the linkage levels and support student growth. A choice to focus only on the next linkage level, without considering the grade-level target, may lead to viewing the linkage levels as a sequential learning progression – unintentionally lowering expectations and limited progress. Finally, the linkage levels on a summative report are associated with one grade, and these reports are used to plan for the next grade – one which has different Essential Elements and associated linkage levels. Helping teachers understand the learning profile as a source of information about the student's academic skills rather than a concrete road map for the next grade's instruction will prevent misaligned instruction and limited opportunity for growth.

References

Authors (2016, April). *Alternate assessment score report interpretation and use: Implications for instructional planning*. Presentation at the 2016 annual meeting of the National Council on Measurement in Education, Washington, DC.

Authors. (2015, April). *Designing alternate assessment score reports that maximize instructional impact*. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, Illinois.

Karvonen, M., Wakeman, S., Moody, S., & Flowers, C. (2013). The relationship of teachers' instructional decisions and beliefs about alternate assessments to student achievement. *Exceptionality, 21*, 238-252. doi: 10.1080/09362835.2012.747184

Nitsch, C. (2013). *Dynamic Learning Maps: The Arc parent focus groups*. Unpublished manuscript. Washington, DC: The Arc.

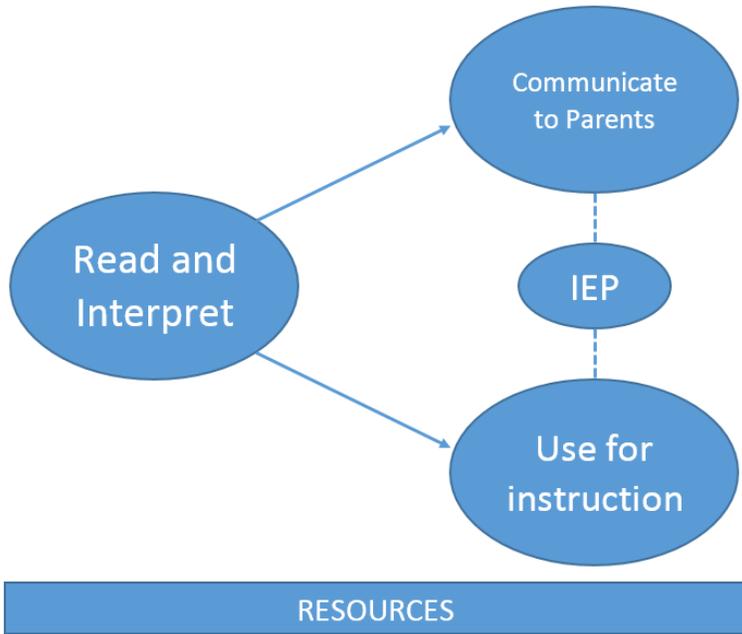


Figure 1. Conceptual framework for the study