Facilitating the Use of Data From Multiple Sources for Formative Learning in the Context of Digital Assessments: Informing the Design and Development of Learning Analytic Dashboards

Priya Kannan and Diego Zapata-Rivera*

Educational Testing Service, Learning and Assessment Foundations and Innovations Research Center, Princeton, NJ, United States

Learning analytic dashboards (LADs) are data visualization systems that use dynamic data in digital learning environments to provide students, teachers, and administrators with a wealth of information about student's engagement, experiences, and performance on tasks. LADs have become increasingly popular, particularly in formative learning contexts, and help teachers make data-informed decisions about a student's developing skills on a topic. LADs afford the possibility for teachers to obtain real-time data on student performance, response processes, and progress on academic learning tasks. However, data presented on LADs are often not based on an evaluation of stakeholder needs, and have been found to not be clearly interpretable and actionable for teachers to readily adapt their pedagogical actions based on these insights. We elaborate on how insights from research focused on interpretation and use of Score Reporting systems and research on open learner models (OLMs) can be used to inform a research agenda aimed at exploring the design and evaluation of LADs.

Keywords: dashboards, learning analytics, Score Reporting, open learner models, data visualization, user-oriented research

INTRODUCTION

With COVID-19 and the consequent radical shift to online and hybrid learning environments, there has been a lot of interest in exploring approaches to better support student learning and assessment in formative teaching and learning contexts. These instructional contexts are considered formative since they are intended to provide teachers and students with information about learning as it develops—not after the fact, such as after a unit or term. Formative tasks are woven into instruction, and are intended to provide teachers with on-going, and in many instances real-time,

feedback about their students' current level of understanding in relation to a specific learning goal (Black and Wiliam, 1998; Shepard, 2005; Shute, 2008; Bennett, 2011, 2019). Therefore, in these formative, everyday teaching and learning contexts, feedback should be presented to help teachers identify what students know and can do and guide teachers in making instructional decisions and plan lessons both at the individual and classroom level. With effective feedback, teachers should know how to modify their teaching practices by diagnosing gaps in their students' current learning.

Online and digital learning environments, adaptive instructional technologies, and game-based learning and assessment environments have seen a rise in recent years (Heffernan and Heffernan, 2014; Feng et al., 2018; Sinatra et al., 2020; Rahimi and Shute, 2021). Large and varied types of data about students' overall learning experiences (including process and log data) are now available within these digital environments, and it would be most helpful if these data are used to provide interpretable, useful, and actionable feedback to teachers in the classroom context. Learning analytic dashboards (LADs) have become increasingly popular for providing feedback in these digital contexts (Papamitsiou and Economides, 2014; Sahin and Ifenthaler, 2021). The data visualizations and reports used within LADs are intended to help students understand their progress toward goals and help teachers make data-informed decisions in formative learning contexts (Bayrak et al., 2021; Dickler, 2021; Keskin and Yurdugül, 2021; Rahimi and Shute, 2021). However, research has found that the data presented within LADs are often not based on an evaluation of stakeholder needs (Sahin and Ifenthaler, 2021), and are often not clearly interpretable and actionable (Molenaar and Knoop-van Campen, 2018; Sahin and Ifenthaler, 2021; Valle et al., 2021) in a way that lends effective pedagogical support to teachers; therefore, more research on LADs is warranted (Rahimi and Shute, 2021).

We elaborate on how insights from research on Score Reporting systems (Hambleton and Zenisky, 2013; Kannan et al., 2018a; Zapata-Rivera, 2019) and open learner modeling (Bull, 2020; Zapata-Rivera, 2020) can be used to inform a research agenda aimed at exploring design and evaluation of user-centric LADs that provide interpretable and actionable feedback for teachers. Through formative feedback, LADs can help teachers and students make better teaching and learning decisions. Supporting users (teachers and students) in understanding the data and making appropriate decisions should be at the forefront of research in the area of digital and AI-based systems, since at the end of the day humans (and not the AI systems) are the ones using this data and making decisions for instructional and learning purposes.

FEEDBACK AND REPORTING IN FORMATIVE CONTEXTS

In today's digital learning context, with the influx of computers in classrooms (e.g., tablets and laptops), students have access to various types of digital and online learning resources (e.g., intelligent tutoring systems; game-based learning systems).

Moreover, with the increase in digital and online assessments and learning tools, there is a lot of detailed background data (including log data and data about student response processes) that is available. Examples of such data include: *number of times a student accesses various features within the learning environment, where and when the student clicks, how the student navigates, the amount of time a student spends on the assigned task, the number of attempts a student takes to answer an item correctly, number of hints and scaffolds used, etc.* Such data could be used to analyze student behaviors and interactions with a digital environment and could be used to inform instructional planning and decision-making (Bennett, 2019). The addition of process and log data in feedback not only provides teachers with a richer context of the student's learning and provide complementary information about a student's current state of understanding, but such data can also provide some opportunities to support more effective and personalized learning experiences for each student (Zapata-Rivera et al., 2016; Hao and Mislevy, 2018; Andrews-Todd et al., 2021; Sahin and Ifenthaler, 2021; Zapata-Rivera and Arslan, 2021).

With the large amount of data available to teachers, it is important to scaffold this information and present it to teachers in a way that is interpretable and useful in informing instruction (Kuosa et al., 2016; Bennett, 2019). In addition and particularly in formative assessment and learning contexts, teachers require timely and actionable feedback that can inform their immediate instructional next steps (Kulik and Kulik, 1988; Black and Wiliam, 1998; Nicol and MacFarlane-Dick, 2006); this type of ongoing need for high-quality actionable information has been referred to as "who needs to be taught what next" (Brown et al., 2019, p. 109) guidance for teachers. In other words, feedback provided to teachers in the formative context should be immediate and designed to inform instruction and student groupings such that teachers can tailor their instructional next steps specifically to support gaps in student's conceptual understanding (Zapata-Rivera et al., 2007; Shute, 2008).

Learning analytics-based systems, as a set of tools for measuring and reporting data about learners in digital learning environments, have become popular in the last decade or more since the "Big Data" revolution (Papamitsiou and Economides, 2014). With the increasing availability of various data types, fields such as learning analytics and education data mining have emerged. These large amounts of data from digital environments can be made available to users through a Learning Analytics Dashboard (LAD) wherein algorithmic analyses and information visualizations could be used to synthesize and present data to users in meaningful ways. For example, these systems can support personalized learning pathways (through learner models), and provide adaptive feedback through sequencing of activities and tasks with multiple opportunities for gathering student responses and underlying process data (Leonardou et al., 2019; Bull, 2020). Such data can then be presented to students and teachers on interactive dashboards to support self-reflection and instructional decision-making. In the next section, we will briefly describe Learning Analytics Dashboards (LADs) as one type of feedback mechanism in digital learning contexts, provide a couple of examples of LAD implementations, and discuss some problems

with data representations within LADs particularly with respect to providing useful and actionable feedback for teachers.

LEARNING ANALYTICS DASHBOARDS

Learning Analytics Dashboards (LADs) are information visualization dashboards that are intended to provide students and teachers with a wealth of feedback about students' current and historical learning status to inform instructional decision-making. Development of LADs have been informed by research in information visualization and educational data mining, wherein the latent learning patterns of students in digital learning environments are discovered through educational data mining algorithms, and these patterns are then presented to learners using visualization techniques and dashboards through learning analytics (Yoo et al., 2015; Schwendimann et al., 2017; Sahin and Ifenthaler, 2021).

LADs have been described as a specific type of "personal informatics" applications (Verbert et al., 2013). There has been an increasing number of "personal informatics" systems across domains ranging from medicine to sports and fitness (e.g., Fitbit). These "personal informatics" systems are typically built to enable users to collect and review personally relevant information and receive actionable feedback for the purposes of self-awareness, self-monitoring, and self-reflection (Verbert et al., 2013; Kersten-van Dijk et al., 2017). Personal Informatics systems have been touted as allowing their users to receive actionable data-driven feedback and extract meaningful insights that would result in positive behavioral changes (Verbert et al., 2013; Kersten-van Dijk et al., 2017).

LADs are used to translate a large amount of usage data into interpretable formats to assist users, who are primarily teachers and students (Liu et al., 2021; Sahin and Ifenthaler, 2021). Student-facing LADs can be used to automate a lot of feedback that teachers normally provide to students in formative contexts (Rahimi and Shute, 2021), and can be helpful to students in setting personal goals and seeing their progress toward those goals and also obtain immediate feedback about their learning and what to do next (Bodily et al., 2018; Sedrakyan et al., 2020; Rahimi and Shute, 2021). Student dashboards can also help by providing the appropriate frame-of-reference (norm or criterion referenced) in helping evaluate their progress toward goals (Aljohani and Davis, 2013). For example, providing norm-referenced comparisons enable a student to compare their progress toward goals with their peers, while criterion-referenced comparisons are aimed at providing feedback on progress toward designated levels of mastery (Bloom, 1956; Angoff, 1974; Betebenner, 2009). Research has found that norm-referenced comparisons may not be ideal and lead to unhealthy competition, while providing criterion-referenced comparisons toward one's own mastery goals has been consistently shown to have a positive impact on student motivation and learning (Rahimi and Shute, 2021).

Teacher-facing LADs often include data visualizations that help teachers understand students' current state and could be used to reflect on student understanding and act upon it (Rahimi and Shute, 2021). Therefore, teacher-facing LADs could either be information-oriented or action-oriented; however, it is the action-oriented LADs (which provide insights about possible next steps) that are likely most beneficial for teachers in formative contexts by providing them with real-time information about their students' time on task, progress toward goals, their overall level of conceptual understanding, and their strengths and needs relative to ongoing formative goals (Molenaar and Knoop-van Campen, 2018; Michaeli et al., 2020; Sahin and Ifenthaler, 2021; Valle et al., 2021).

LADs were originally developed in the context of higher education, specifically student interactions within popular learning management systems (LMSs such as Blackboard or Moodle), and used to translate large amounts of system usage data (e.g., clickstream data, course content summaries, time spent on content, and forum participation) into interpretable visualizations to assist college professors (Khosravi et al., 2021). One early example of a LAD system developed in the higher education context is Course Signals (Arnold and Pistilli, 2012). Course Signals used a traffic light (signal) visual representation to provide students in collegiate courses (at Purdue) with real-time feedback based on their interactions with Blackboard and other supplementary information such as past academic performance. Another example of an early LAD system is Student Activity Meter (SAM; Govaerts et al., 2012). SAM visualizes student actions (such as time spent and resource use) using easy to understand box plots for students to be able to compare themselves with their peers. In the context of higher education, these early studies on dashboard visualizations have been followed by years of research on the effectiveness of various visualization techniques (e.g., bar charts, line graphs, tables, network graphs) and the ability of these systems to support informed decision-making for both learners and instructors (Sahin and Ifenthaler, 2021).

In the K-12 context, the use of dashboards has been explored within Intelligent Tutoring Systems (ITS; Sinatra et al., 2020) such as ASSISTments (Heffernan and Heffernan, 2014; Feng et al., 2018) and MATHia (Ritter et al., 2016; Fancsali et al., 2018) that support student learning based on models of how students learn. ASSISTments is a web-based platform intended to support students as they solve mathematics problems, and is designed to provide detailed student-level and class-level data to teachers in informing their instructional planning and pacing in the formative context (Heffernan and Heffernan, 2014). MATHia, part of the Carnegie Learning Math Series (CLMS), is an ITS developed to support mathematics instruction for students in grades 6–8. With built-in formative assessments, MATHia is designed to provide teachers with *real-time feedback* about what students know thereby helping support their instructional decision-making based on student needs. LADs have now become increasingly popular in K-12, particularly in formative learning contexts (Mazza and Dimitrova, 2007; Aljohani and Davis, 2013; Xhakaj et al., 2017; Bayrak et al., 2021; Dickler, 2021; Keskin and Yurdugül, 2021; Rahimi and Shute, 2021), and have been found to be particularly useful in the reporting of data through scaffolds and visualizations (Valle et al., 2021) to help teachers make data-informed decisions about students'

developing skills on a topic. LADs often provide feedback about a student's learning using interactive graphical representations, and such feedback may either be provided in real-time (e.g., as students are engaged in a reading activity) or periodically at the end of various intervals of learning or after the learning activity has been completed (Bodily et al., 2018).

LADs are especially useful for providing real-time feedback about learning processes that cannot be easily captured via conventional classroom monitoring strategies (Liu et al., 2021). For example, if the teacher had assigned all students in a class to independently read aloud from a digital reading tool for 20 min, it would be hard for the teacher to know which students are "on-task" and reading, which students are not engaged or completing the read-aloud task, and which students may need support. A teacher-facing dashboard which provides them with real-time information on students' engagement in their reading activity would be useful to teachers in determining which of their students may need immediate attention and when and whom to provide with additional scaffolding and support. One example mockup of a teacher-facing dashboard (from Kannan et al., 2019) where such real-time feedback can be provided to teachers is presented in **Figure 1**. These mock dashboards were developed iteratively by first engaging the intended stakeholders (in this case teachers) in an audience-centric needs assessment, and will be further described in the last section of this paper.

Challenges in the Area of Learning Analytics Dashboards

With large amounts of data about students' overall learning experiences (including process and log data) available within LADs (Sahin and Ifenthaler, 2021), it is important to ensure that this information is appropriately scaffolded and presented in an interpretable and actionable format to teachers (Kuosa et al., 2016; Bennett, 2019). Moreover, research has also indicated that the data provided in LADs are often not actionable—teachers struggle with selecting the appropriate feedback from the plethora available, and also appropriately using the data to support pedagogical actions and allocating instructional time across students of different abilities (Knoop-van Campen et al., 2021). A number of issues and challenges have been identified with the ways in which data is currently presented within LADs. We discuss some of these challenges here, particularly with regard to ensuring that LADs are designed with the intended stakeholder's needs in mind, and ensuring that the data within LADs are presented to ensure appropriate interpretation and use.

Choosing Appropriate Data

First, the lack of consistent quality in the types of data collected within LADs may pose a major challenge to the ways in which these data are appropriately understood by stakeholders and are then effectively utilized (Kuosa et al., 2016). Some data may not be supported by enough evidence to support claims or warrant action. Moreover, we cannot just assume that the system captures the right information, and automatically present data based on all collected information. One solution that has been proposed in terms of identifying appropriate data is "feature selection" (Sahin and Ifenthaler, 2021, p. 590), wherein educational data

mining is used to define metrics and identify appropriate types of data for stakeholders. However, automated data selection based on algorithms may not be a sufficient solution, and the data presented to users should also be informed by their context-specific needs. Therefore, there is a need to identify appropriate slices of data that are supported by evidence, and a need to evaluate which of these slices of data would be considered informative, useful, and actionable by stakeholders. Without taking a user-centered design approach into account, it is possible that information in LADs may not be useful in supporting decision-making and instead be distracting, confusing, or mislead users to make inappropriate interpretations.

An Overwhelming Amount of Data

When it comes to teachers as stakeholders, it is important to remember that they are often overwhelmed with vast amounts of data, and often feel inundated with information that they are unable to process. This phenomenon is referred to as "data rich—information poor" or DRIP which was first proposed in the field of healthcare (Goodwin, 1996), and later extended to refer to the overwhelming amounts of data available to educators (Charman, 2009) in today's context of ever-increasing assessments. Manual drill-downs of large volumes of data can be overwhelming to users like teachers who are already strapped for time. This might result in an unwanted increase in the cognitive processing required to understand and effectively use the data (Kuosa et al., 2016) and might result in "*curiosity-driven explorations*" (Wise and Jung, 2019; Khosravi et al., 2021, p. 3) of irrelevant questions that are not directly informative to their instructional needs. Feedback presented to teachers in LADs should be based on teachers' needs, and should appropriately consolidate various pieces of information in a way that supports formative hypotheses about their students' understanding and inform their next instructional steps. Therefore, there is a need for improving the alignment of design and evaluation aspects of LADs in order to support the appropriate interpretation and use for teachers (Valle et al., 2021).

Interpretation and Use of Data and Visualizations

Another important issue in LADs is that the visualizations are often presented in a way that make them difficult for the stakeholders to understand (Sahin and Ifenthaler, 2021). The design process is often ignored in dashboard design and development (Bodily et al., 2018), and stakeholders are not typically involved in the design process. Therefore, in designing dashboards, it is critically important to take into account stakeholders' information needs and abilities to understand various visualizations (Zapata-Rivera and Katz, 2014; Yoo et al., 2015; Sedrakyan et al., 2019; Sahin and Ifenthaler, 2021). It is also important to ensure that the information presented in LADs is based on what stakeholders would consider most useful (Yoo et al., 2015). Finally, LAD design may also benefit from being directly linked to learning theories (Yoo et al., 2015; Bodily et al., 2018; Sahin and Ifenthaler, 2021), which, in addition to needs, also considers the underlying principles of how students learn and developmental trajectories of student
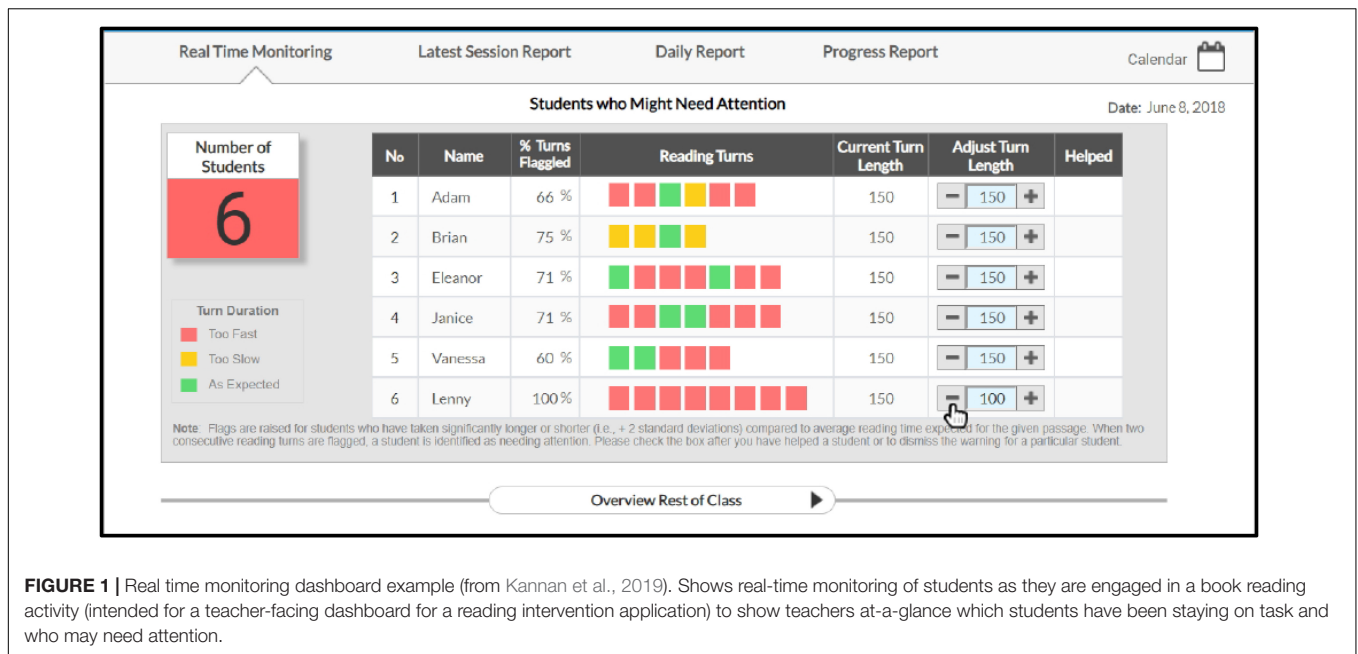
FIGURE 1 | Real time monitoring dashboard example (from Kannan et al., 2019). Shows real-time monitoring of students as they are engaged in a book reading activity (intended for a teacher-facing dashboard for a reading intervention application) to show teachers at-a-glance which students have been staying on task and who may need attention.

conceptual understanding in appropriately presenting feedback to teachers and students (Kannan et al., 2021a).

LESSONS FROM SCORE REPORTING AND OPEN LEARNER MODELS

As pointed out in the previous section, LADs may contain volumes of data that may not be designed and presented in a way that is most easily interpretable and usable by the intended stakeholders. Moreover, the feedback provided may also not be actionable and clearly targeted toward appropriate instructional next steps. We feel that the literature and research on Score Reporting and Open Learner Models (OLMs) can be extremely useful in informing a research agenda for LADs. Particularly research in these areas suggests that dashboards should be designed appropriately for various stakeholders with their specific needs at the forefront and evaluated for accurate interpretation and appropriate use with intended audiences. Lessons from these areas of research could help inform a research agenda for LAD research and ensure that LADs provide interpretable, useful, and actionable feedback to the intended score users. So, in this section, we will provide a broad overview of how Score Reporting and OLM research can inform the development and evaluation of LADs.

Score Reporting

In the context of large-scale assessments, results—particularly insights into the underlying knowledge and skills of the test taker—are communicated to various stakeholders (e.g., teachers, administrators, and parents) through some form of a score report that uses graphical representations and data tables to communicate results for individual students or groups of test takers (Zapata-Rivera et al., 2012; Hambleton and Zenisky,

2013). However, Score Reporting, as a field, goes beyond just communicating the scores obtained on a test (Zapata-Rivera, 2019). Score Reporting research is grounded in validation (Kane, 2006) and focuses primarily on the accuracy of inferences drawn from score reports by critical stakeholders (Tannenbaum, 2019); in fact, the validity of the assessment is dependent upon the interpretation and use of scores as communicated in score reports. Therefore, Score Reporting research has been grounded in contextualizing the results to the needs of the intended stakeholders in a way that is meaningful and actionable. In addition, Score Reporting research has followed a recommended iterative multistage approach (see Zapata-Rivera et al., 2012; Hambleton and Zenisky, 2013) to the design and evaluation of prospective score reports before they are operational for any assessment. In the last decade, research on issues surrounding Score Reporting has substantially increased with a focus on audience specificity (Zapata-Rivera and Katz, 2014) in the design and development and on stakeholder interpretation and use (Kane, 2006) in the evaluation of score reports.

Audience Specificity in Score Reporting

Each stakeholder group such as parents, teachers, administrators, and students are likely to have different needs for information, have different levels of pre-existing knowledge about the assessment and its context, and have different attitudes, feelings, or biases that might color their interpretations of the information shown in the reports (Zapata-Rivera and Katz, 2014). Results from Score Reporting research (e.g., Underwood et al., 2010; Kannan et al., 2021a) focused on specific stakeholder groups (e.g., parents, teachers, administrators) have highlighted the diverse needs, pre-existing knowledge, and attitudes for these groups.

For example, research shows that while parents mainly want to know how their child has performed in an assessment, what these scores mean, and how they can help their child improve

(Kannan et al., 2018a), teachers are interested in information that can directly guide instruction (Brown et al., 2019), and administrators value results that can help them appropriately allocate resources and evaluate interventions based on average performance of their school or district population (Zapata-Rivera and Katz, 2014). So, in Score Reporting research, best-practice suggests that an in-depth audience analysis be conducted prior to designing score reports so that it caters to audience needs, thereby ensuring that users can understand and use the information appropriately given their context and needs.

Evaluating Score Reports for Interpretation and Use

As noted previously, needs, pre-existing knowledge, and attitudes may vary across stakeholder groups. In addition, cognitive aspects (Hegarty, 2019) such as perception, attention, and working memory, which varies across individuals, may also vary largely between various stakeholder groups. All of these factors tend to play a critical role in the extent to which stakeholders can comprehend the information presented in score reports. Therefore, using varied methodologies such as cognitive interviews, focus groups, and surveys, Score Reporting research and practice has focused on ensuring that the intended stakeholders understand the information presented and know how to use these results appropriately.

Score Reporting research has found that each stakeholder group has their own time, resource, and contextual constraints that hinder their ability to spend sufficient time to understand the information presented in score reports (e.g., Marshall and Drummond, 2006; Underwood et al., 2010; Kannan et al., 2018a). For example, parents, who are a particularly diverse and heterogeneous group, and tend to have different levels of education and English language proficiency, in general struggle to understand technical terms such as standard error of measurement (SEM) presented in score reports (Kannan et al., 2018a). Teachers have also been found to struggle to parse out some of the technical information presented in score reports (e.g., Impara et al., 1991; Zapata-Rivera et al., 2012). And, administrators and policy makers, who are often strapped for time, have been shown to become overwhelmed with large volumes of data (e.g., Underwood et al., 2010) and tend to draw unwarranted conclusions from the information presented in score reports.

Several methods have been used in Score Reporting research to ensure that stakeholders are able to understand and use the information presented appropriately. Wainer et al. (1999) used within subject design where various alternative visual displays were presented to policymakers—they found that the simplified visual displays led to better comprehension for the policymakers. Other studies (e.g., Kannan et al., 2018b) have used a hybrid cognitive interview style which combines retrospective verbal probing where participants respond to directed questions with concurrent think-aloud methods where participants verbalize their thoughts as they are interacting with the report or reporting system. These cognitive interviews are intended to identify the elements in the score report that are most salient to the stakeholders and if they are able to access and use all

the information as intended in addition to evaluating if the information presented in these reports is interpreted accurately.

In other studies (e.g., Kannan et al., 2021b), specific comprehension questions pertinent to the range of information provided in the reports were embedded in online surveys. These survey studies have included several questions that are quick to answer (such as multiple choice, true/false), where the questions have focused on aspects of the representations that may likely be confusing given the specific prior knowledge and other constraints of the intended stakeholder group. Participant responses to these comprehension questions then enable us to evaluate the extent to which the data visualizations and other information presented in the reports are being understood correctly and identify areas where additional clarity may be needed. These survey-based methods (e.g., Kannan et al., 2021b) have combined the within-subject design methodology proposed by Wainer et al. (1999) by using alternative visual displays to evaluate comprehension based on each display. These methods help in weeding out displays and technical details that are not being correctly interpreted, and help identify the visual displays and report formats that aid stakeholder interpretation.

Finally, various stakeholders, particularly administrators and policy makers, are often strapped for time, and have been shown to become overwhelmed with large volumes of data. To help stakeholders grapple with large volumes of data (e.g., large-scale assessment results for a district), Underwood et al. (2010) proposed an evidence-based framework for designing administrator and policy-maker reports that link student data to focal questions that are informed by stakeholder needs and the types of decisions made by these stakeholders. Such reports, that use a "question-based scaffolding" methodology have been shown to result in better comprehension and foster appropriate use among administrators and policymakers (VanWinkle et al., 2011).

Open Learner Models

Open Learner Models (OLMs) are a special case of learner awareness tools where the system's representation of the learner (i.e., learner model) is made available/open to students, teachers, and other users (Bodily et al., 2018; Sergis and Sampson, 2019; Bull, 2020). These learner models can include information about a learner's knowledge, skills, and other attributes (KSAs). In other words, learner models can hold information about a learner's current knowledge and skill level (e.g., competencies, understandings, misconceptions, and progress toward mastery), and hold information about other learner attributes (e.g., motivation, engagement, effort, and affective state). Since this information is automatically inferred and dynamically updated based on student responses to questions and other process data (e.g., time taken to view material and complete tasks, navigation routes), learner models enable systems to adapt to a learner's educational needs (Bull, 2020). In many cases, it also allows for additional input/evidence from the user (learner) as an additional source of evidence (Zapata-Rivera et al., 2007; Bull, 2020).

Learner models are key components of adaptive instructional systems. A variety of open learner modeling approaches have been implemented and evaluated including guided exploration,

negotiation with a human or an agent, and collaboration with a human or a virtual peer (Zapata-Rivera and Greer, 2002; Shute and Zapata-Rivera, 2012; Bull and Kay, 2016; Dimitrova and Brna, 2016; Bull, 2020). Evidence-based approaches to interacting with OLMs have been designed and evaluated with teachers and students (e.g., Van Labeke et al., 2007; Zapata-Rivera et al., 2007). These approaches are designed following human-computer design principles to create graphical interfaces that allow users to explore and use the information maintained by the system in support of their learning and teaching goals. OLM interfaces are evaluated with the target audiences through usability studies and large-scale studies aimed at evaluating their effectiveness in supporting learning and other goals.

Bull and Kay (2016) describe various approaches used to evaluate OLMs. These approaches include studies in authentic contexts, laboratory and field evaluations (Kay, 1995; Zapata-Rivera and Greer, 2004; Czarkowski and Kay, 2006), small-scale and large-scale studies using qualitative and quantitative analysis. In addition, various techniques have been recommended for evaluating OLMs such as think-aloud protocols, evaluating the comprehension and usability of the interface by learners, evaluating affect and emotions, and evaluating the effectiveness of the approach for the intended purpose (e.g., improving the accuracy of the learner model, facilitating control over the model, and supporting learning and reflection). For example, Mitrovic and Martin (2002) reported on positive effects on learning outcomes associated for those who interacted with the learner model.

In addition, other studies have evaluated the effectiveness of the OLMs for teacher use. For example, Zapata-Rivera et al. (2007) used focus groups to evaluate the types of supports teachers would need to interact with an OLM. In this paper, they offer an evidence-based approach to evaluating the interaction of teachers with Open Student/Learner Models. Results of their study indicated that teachers found the information provided by the system useful in deciding their next instructional actions for individual students or small groups of students. However, teachers expressed the need for additional support to help them focus on the most relevant/high priority cases due to time limitations. Teachers suggested the use of automated messages to inform teachers about particular high-priority cases and involving teacher assistants in the process. In another study, Mazza and Dimitrova (2007) used surveys, focus groups, and interviews to evaluate teacher understanding of social, behavioral, and cognitive aspects of learners using graphical representations created from log data generated by course management systems in an online distance learning context. Results showed that teachers were able to use these graphical representations successfully to identify main trends at the group level as well as individuals that may need special attention. Kay et al. (2022) describe an OLM-driven learning data design approach for teachers. This approach is used to enhance learning analytics platforms used by teachers and students.

Overall, OLMs have been designed and developed within various contexts to support student self-regulation, self-reflection, knowledge awareness, group formation, student model accuracy, and learning (Brna et al., 1999; Hartley and

Mitrovic, 2002; Dimitrova, 2003; Zapata-Rivera and Greer, 2004; Mazza and Dimitrova, 2007; Bull, 2020; Hooshyar et al., 2020). Various useful approaches and methods have been offered in these contexts to evaluate the graphical interfaces and guidance mechanisms aimed at supporting learning and teaching goals. Similar to the recommendations offered within Score Reporting research, these OLM interfaces have also been developed taking into account the needs of various stakeholders such as learners, teachers, and parents (Lee and Bull, 2008; Bull and Kay, 2016; Ginon et al., 2016; Bull, 2020). Therefore, we think that the methods and approaches offered within OLM research can also inform the proposed research agenda for the design and evaluation of LADs.

Dealing With Identified Challenges

In this section, we offer some suggestions for dealing with the challenges mentioned in section "Challenges in the Area of Learning Analytics Dashboards." In addition, we will offer some illustrative examples of dashboard designs that follow these methods, and hope that these methods and approaches could be useful in informing a research agenda in the design and evaluation of LADs.

Strategies for Identifying Appropriate Data

One of the first issues raised in LAD design and development was the lack of consistent quality of data, and the need for appropriate selection of data to present to stakeholders (Kuosa et al., 2016; Sahin and Ifenthaler, 2021). Even though "feature selection" through educational data mining is offered as a solution to presenting appropriate data, automated data selection based on algorithms may not be a sufficient solution. It is critical that the data presented to users are evidence-based and are informed by users' context-specific needs. In other words, based on best practices recommended in the Score Reporting and OLM literatures, we recommend that in-depth audience analyses (Zapata-Rivera and Katz, 2014) and stakeholders-specific needs assessments should be conducted in determining what pieces of information would be considered most useful by the intended users.

The iterative multistep approach (Zapata-Rivera et al., 2012; Hambleton and Zenisky, 2013) used in the design and evaluation of score reports always starts with an audience-focused needs assessment. This iterative approach was also applied to the design and development of a few teacher-facing dashboards in formative contexts such as the dashboard for supporting teachers in monitoring students as they engage in a reading intervention as presented in **Figure 1**. In developing a teacher-facing dashboard for classroom implementations of a reading intervention tool (see Kannan et al., 2019), we applied the iterative multistep approach recommended in Score Reporting literature and started with an audience focused needs assessment to elicit teacher needs for feedback as they monitor students' progress on the reading intervention tool.

In this study, we first allowed teachers to interact with the reading app, and in a series of whole- and small-group discussions elicited some of their needs for feedback if such a reading intervention tool were to be implemented in their

classroom. Though a number of different types of feedback could be provided based on the log data and process data collected in this app about students' reading activity, teachers' elicited needs were very helpful in prioritizing the dashboard screens for the next stage of the iterative design and evaluation cycle. For example, results from the needs assessment indicated that in addition to the ability to monitor students' real-time engagement with the reading activity (see **Figure 1**), teachers were also interested in feedback about students' reading fluency and their ability to comprehend the materials they read (see **Figure 2**) after each reading session (Kannan et al., 2019).

Similarly, in the context of LAD development, stakeholder-specific needs thus generated should then be examined against data that can be collected within the system and substantiated with evidence, and then be used in designing additional mockups which are iteratively evaluated for interpretability and usefulness before deployment. For example, Zapata-Rivera et al. (2020) provide a list of assessment information needs for various types of users of adaptive instructional systems. Therefore, LAD development can be informed by using the iterative multistep approach recommended in Score Reporting literature (Zapata-Rivera et al., 2012; Hambleton and Zenisky, 2013).

In starting with a needs analysis focused on the intended stakeholder group (whether it be teachers or students), LADs can be designed so that the data presented is based on evidence and directly actionable based on the needs of the user.

Dealing With an Overwhelming Amount of Data

Another issue when it comes to the large volumes of data that can be available and presented within LADs relates to the DRIP issue for teachers referred to earlier. Manual drill-downs of large volumes of data can be overwhelming and result in an unwanted increase in the cognitive processing (Kuosa et al., 2016) for users like teachers who are already strapped for time. Therefore, feedback presented to teachers in LADs should appropriately consolidate various pieces of information in a way that supports formative hypotheses for users like teachers.

One way to alleviate the DRIP issue for teachers is to use question-based drill-downs (VanWinkle et al., 2011) that may help teachers in informed explorations of the data. Appropriate and interpretable visualizations, which are designed to respond to specific need-based questions can help teachers process this information better (Kuosa et al., 2016). It is anticipated that using a guided exploration method (such as question-based drill-down) can increase cognitive resources and reduce distracting information (Hegarty, 2019), thereby guiding the user through insightful drill-downs (Khosravi et al., 2021) that use a set of pre-determined and audience-specific probe questions.

For example, we developed score reports for administrators to provide feedback on district performance where administrators can easily drill-down into data by using a question-based method (Zapata-Rivera, 2020; see **Figure 3**). So, instead of puzzling over the overwhelming amounts of data and tables based on student performance in the district, administrators would use directed questions and drill-down to arrive at pre-canned views of data that is more directly suited to their needs. Similarly, question-based drill-downs that are informed by audience-specific needs

analysis can be implemented in the design of LADs and evaluated with the intended stakeholders to see if this results in insightful drill-downs and targeted explorations of the data to support formative hypotheses about students and inform instructional decision-making.

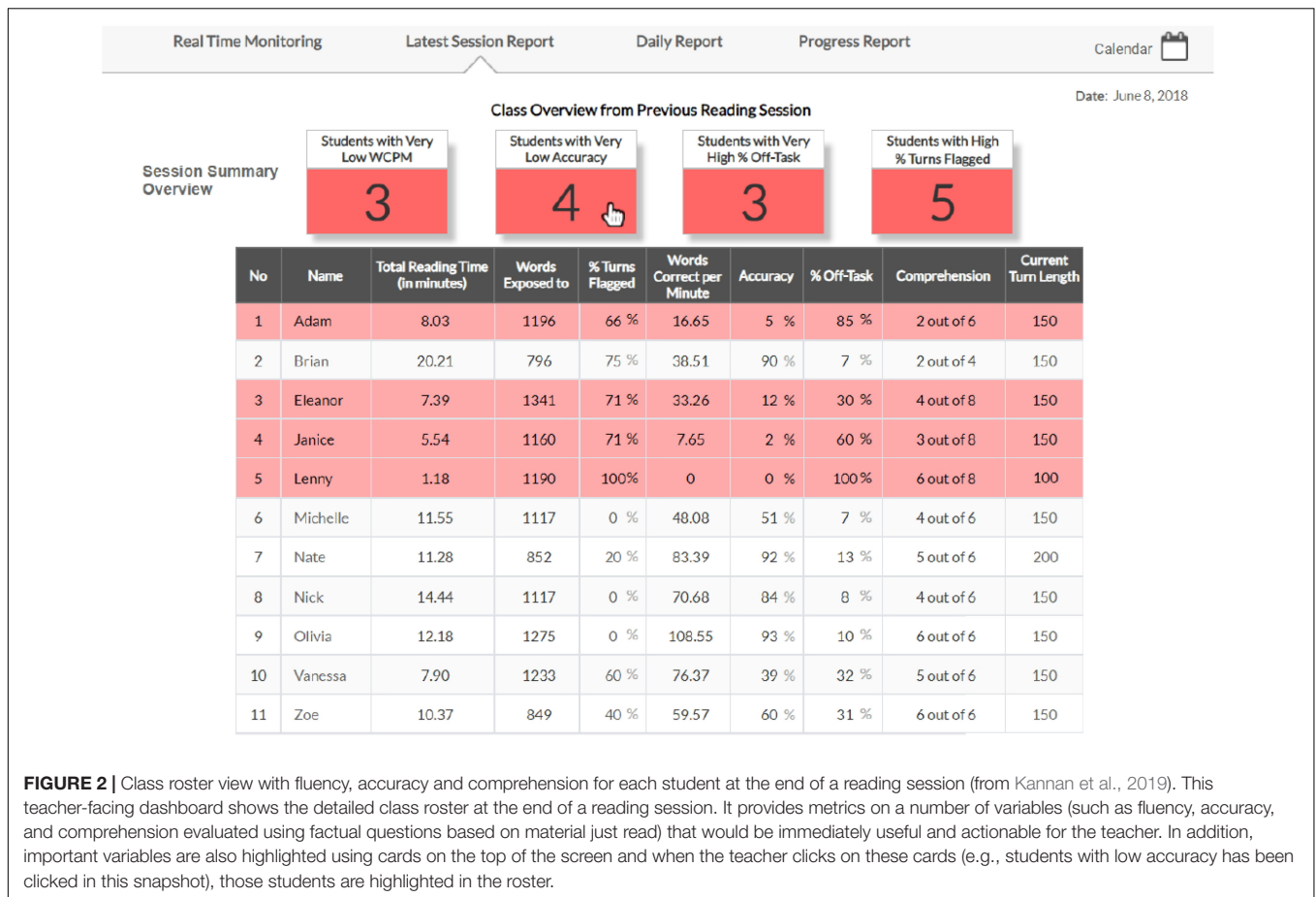### Improving Interpretation and Use of Data and Visualizations

Another important issue in LADs is that the visualizations are often presented in a way that make them difficult for the stakeholder to understand (Sahin and Ifenthaler, 2021). In designing dashboards, not only is it critically important to take the intended stakeholder's needs into consideration, but it is equally important to consider their ability to understand various visualizations given their background (Zapata-Rivera and Katz, 2014). As previously noted, Score Reporting research has found that each stakeholder group has various time, resource, and contextual constraints that hinder in their ability to spend sufficient time to understand the information presented (e.g., Marshall and Drummond, 2006; Underwood et al., 2010; Kannan et al., 2018a). For example, in one previous study (Kannan et al., 2019) we found that even though teachers really wanted a measure of their students' Oral Reading Fluency, communicating this information to teachers meaningfully based on normative distributions was challenging. Teachers expected a numeric score for oral reading fluency, while in this context, fluency measurements for every student resulted in a wide distribution of scores. In the first couple of design iterations, we found that this information was hard for teachers to correctly comprehend. Therefore, we used a color gradient based visual representation with appropriate legends in subsequent iterations and found that teachers were more successful in making appropriate inferences from this type of a visual representation.

Therefore, it is important to evaluate the visualizations with the intended stakeholder groups to ensure that they are able to understand the information presented and use it appropriately. Several methods such as cognitive laboratories, usability studies, focus groups, and surveys are suggested within Score Reporting and OLM research to evaluate stakeholder interpretation and use (see Bull and Kay, 2016; Kannan et al., 2018a, 2021b; Demmans Epp et al., 2019; Zapata-Rivera, 2020). In addition to recommending that LADs are evaluated using similar methods for stakeholder interpretation and use, we offer the following recommendations for LAD design as informed by Score Reporting research.

SUGGESTIONS FOR FUTURE WORK

This section includes suggestions to inform future work in the area of LADs:

Ensuring that teachers and students understand the results presented and use them appropriately is critical to developing actionable dashboards. Therefore, we recommend that LA dashboards should be iteratively evaluated for interpretation and use by the intended stakeholder groups using recommended methods such as cognitive laboratories and large-scale surveys.

FIGURE 2 | Class roster view with fluency, accuracy and comprehension for each student at the end of a reading session (from Kannan et al., 2019). This teacher-facing dashboard shows the detailed class roster at the end of a reading session. It provides metrics on a number of variables (such as fluency, accuracy, and comprehension evaluated using factual questions based on material just read) that would be immediately useful and actionable for the teacher. In addition, important variables are also highlighted using cards on the top of the screen and when the teacher clicks on these cards (e.g., students with low accuracy has been clicked in this snapshot), those students are highlighted in the roster.

Results from such cognitive labs and surveys should reveal aspects, features, and data elements presented in the dashboard that stakeholders are not able to clearly understand. These revelations should be used to inform redesign of the dashboards, and these redesigned dashboards should again be iteratively evaluated to ensure appropriate stakeholder interpretation and use of LADs data.

Once such information is identified in the first iterative evaluation with stakeholders, necessary steps should be taken to ensure that visualizations are appropriately redesigned, and that any complex technical information (e.g., reliability, measurement error) is clearly scaffolded using footnotes and explanatory text. In addition, it would be important to provide any necessary guidelines to interpret the data, and clearly articulate all of the explanatory metadata (e.g., what topics does this cover/what content does it not cover, what do these data mean). The additional supplementary information provided should then be evaluated through cognitive labs and focus groups to ensure that stakeholders recognize the intended relationships among the data presented and the explanatory metadata.

As the use of dashboards continue increase in digital learning and assessment environments, more attention should be placed in the design of evaluation of their interactive features. Results from OLM research on designing and evaluating interactive graphical interfaces and guidance mechanisms can inform the development of interactive components of dashboards. For example, OLM evaluation approaches to support particular uses (e.g., student learning and student reflection) are relevant to the use of dashboards in supporting teaching and learning.

Insights from OLM research in how to support teachers and students in their use of interactive graphical components can inform the design of dashboards. For example, research results about teacher use of OLMs to support instruction can facilitate the development dashboards (e.g., by providing alerts or notifications to teachers as a mechanism for reducing the cognitive load associated to monitoring dashboard indicators).

Potential inappropriate uses of data should be identified through focus groups and surveys. Then, clear recommendations for appropriate use should be provided, while intentionally steering stakeholders away from inappropriate use. Clear guidelines should be laid out describing what the data is intended for and how it be used. All of this supplementary information should be available at the click of a button and usability studies should be used to evaluate if stakeholders are able to access and interpret the supplementary information appropriately. Evidence from these usability studies should indicate that stakeholders are attending to the salient features and guidelines and interpreting this information as intended.

The stakeholder needs assessments conducted at the outset should inform how the dashboards are designed and appropriate
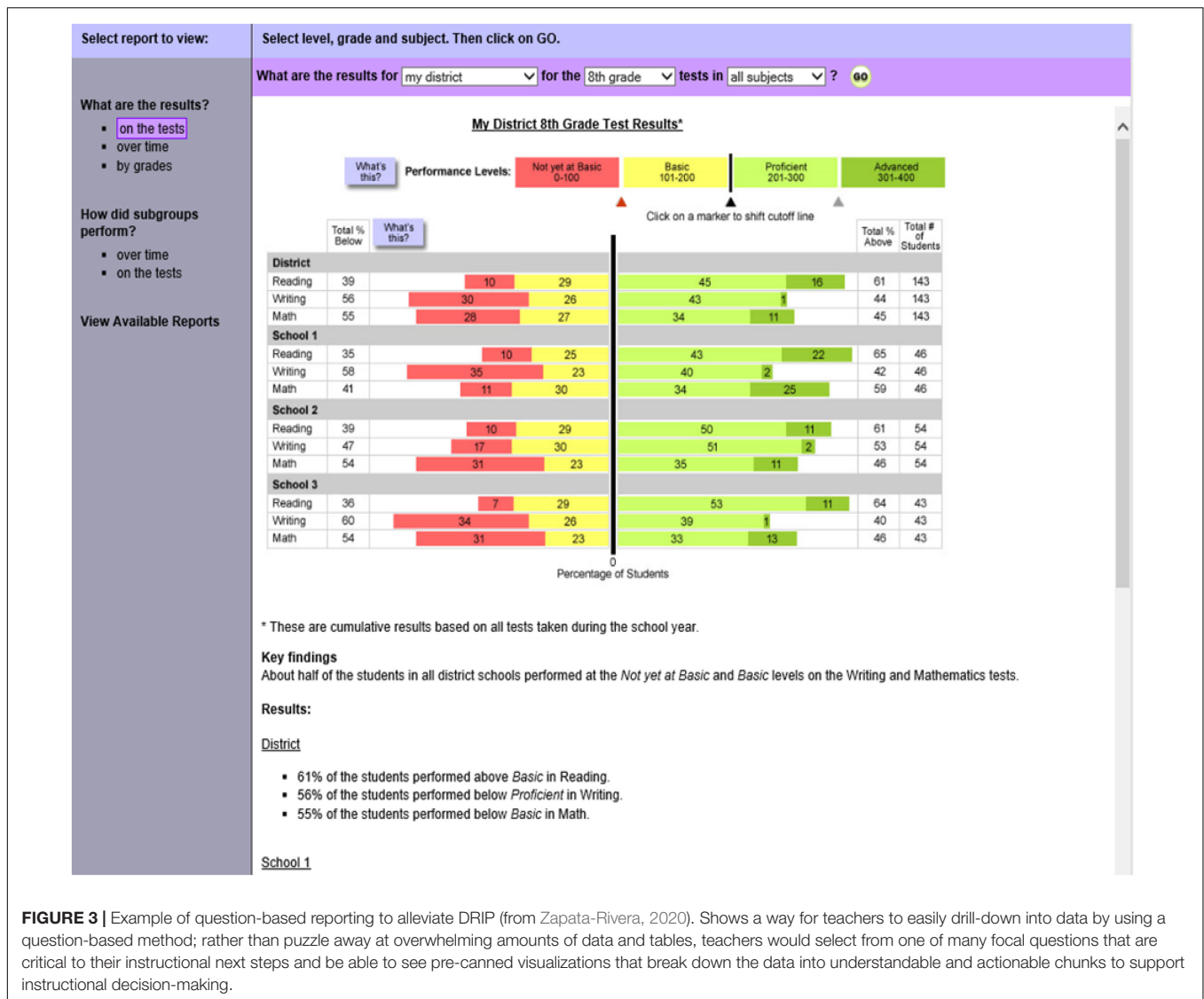
FIGURE 3 | Example of question-based reporting to alleviate DRIP (from Zapata-Rivera, 2020). Shows a way for teachers to easily drill-down into data by using a question-based method; rather than puzzle away at overwhelming amounts of data and tables, teachers would select from one of many focal questions that are critical to their instructional next steps and be able to see pre-canned visualizations that break down the data into understandable and actionable chunks to support instructional decision-making.

actionable next steps should be provided to stakeholders to directly cater to their needs. Guidelines should be provided for appropriate use (e.g., which pieces of data are supported by evidence, which data needs to be used with caution or has contradictory evidence and may need further evaluation). Guidelines should also be clearly provided for the types of decisions that these data support. For example, data from ongoing formative assessments should not be used to support high-stakes placement decisions. And, finally, again ensuring that stakeholders understand these caveats and know how to use this information should be evaluated through focus groups and surveys, and additional changes should be made, if warranted.

CONCLUSION

We presented insights from research on Score Reporting systems (Hambleton and Zenisky, 2013; Kannan et al., 2018a; Zapata-Rivera, 2019) and open learner modeling (Bull, 2020;

Zapata-Rivera, 2020) to inform a research agenda for the design and evaluation of user-centric LADs. Based on lessons learnt in these other bodies of research, we provided some methodological recommendations to ensure that LADs are designed with the intended users' needs at the forefront and are evaluated for stakeholder interpretation and use. The goal would be to develop actionable LAD systems that can consolidate various disparate sources of information and facilitate appropriate interpretation and use of data that is useful and actionable to the intended stakeholders. We hope that the various suggestions and recommendations laid out in this paper provide methodological guidelines in the design and evaluation of user-centric, interpretable, and actionable LADs.

AUTHOR CONTRIBUTIONS

Both authors contributed to the article and approved the submitted version.

REFERENCES

Aljohani, N. R., and Davis, H. C. (2013). "Learning analytics and formative assessment to provide immediate detailed feedback using a student-centered mobile dashboard," in *Paper presented at the Seventh International Conference on Next Generation Mobile Apps, Services and Technologies*, (Prague), 262–267. doi: 10.1109/NGMAST.2013.54

Andrews-Todd, J., Mislevy, R. J., LaMar, M., and Klerk, S. D. (2021). "Virtual performance-based assessments," in *Computational Psychometrics: New Methodologies for a New Generation of Digital Learning and Assessment*, eds A. A. von Davier, R. J. Mislevy, and J. Hao (Cham: Springer), 45–60. doi: 10.1007/978-3-030-74394-9_4

Angoff, W. H. (1974). *Criterion-Referencing, Norm-Referencing and the SAT. Research Memorandum*. Princeton, NJ: Educational Testing Service.

Arnold, K. E., and Pistilli, M. D. (2012). "Course signals at Purdue: using learning analytics to increase student success," in *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, (Vancouver, BC), 267–270. doi: 10.1145/2330601.2330666

Bayrak, F., Nuhoǧlu Kibar, P., and Kocadere, S. A. (2021). "Powerful student-facing dashboard design through effective feedback, visualization, and gamification," in *Visualizations and Dashboards for Learning Analytics. Advances in Analytics for Learning and Teaching*, eds M. Sahin and D. Ifenthaler (Cham: Springer), doi: 10.1007/978-3-030-81222-5_7

Bennett, R. E. (2011). Formative assessment: a critical review. *Assess. Educ.* 18, 5–25. doi: 10.1080/0969594X.2010.513678

Bennett, R. E. (2019). "Integrating measurement principles into formative assessment," in *Handbook of Formative Assessment in the Disciplines*, 1st Edn, eds H. L. Andrade, R. E. Bennett, and G. J. Cizek (London: Routledge), doi: 10.4324/9781315166933

Betebenner, D. W. (2009). Norm- and criterion-referenced student growth. *Educ. Meas.* 28, 42–51. doi: 10.1111/j.1745-3992.2009.00161.x

Black, P., and Wiliam, D. (1998). Assessment and classroom learning. *Assess. Educ.* 5, 7–75. doi: 10.1080/0969595980050102

Bloom, B. S. (ed.) (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals: Handbook I, Cognitive Domain*. New York, NY: Longman.

Bodily, R., Kay, J., Aleven, V., Jivet, I., Davis, D., Xhakaj, F., et al. (2018). "Open learner models and learning analytics dashboards: a systematic review," in *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, (Sydney, NSW: Association for Computing Machinery), 41–50. doi: 10.1145/3170358.3170409

Brna, P., Self, J., Bull, S., and Pain, H. (1999). "Negotiated collaborative assessment through collaborative student modelling," in *Proceedings of the workshop Open, Interactive, and other Overt Approaches to Learner Modelling at AIED99*, (Le Mans), 35–44.

Brown, G. T. L., O'Leary, T. M., and Hattie, J. A. C. (2019). "Effective reporting for formative assessment: the asTTle case example," in *Score Reporting Research and Applications*, ed. D. Zapata-Rivera (London: Routledge), 107–125. doi: 10.4324/9781351136501-9

Bull, S. (2020). There are open learner models about! *IEEE Trans. Learn. Technol.* 13, 425–448. doi: 10.1109/TLT.2020.2978473

Bull, S., and Kay, J. (2016). SMILI: a framework for interfaces to learning data in open learner models, learning analytics and related fields. *Int. J. Artif. Intell. Educ.* 26, 293–331.

Charman, P. (2009). "Data rich, information poor: creative and innovative approaches to results analysis to support teaching and learning," in *Paper Presented at the 35th Annual Conference of the International Association for Educational Assessment*, (Brisbane, QLD).

Czarkowski, M., and Kay, J. (2006). "Giving learners a real sense of control over adaptivity, even if they are not quite ready for it yet," in *Advances in Web-based Education: Personalized Learning Environments*, eds S. Chen and G. Magoulas (London: IDEA), 93–125. doi: 10.4018/978-1-59140-690-7.ch005

Demmans Epp, C., Perez, R., Phirangee, K., Hewitt, J., and Toope, K. (2019). "User-centered dashboard design: iterative design to support teacher informational needs in online learning contexts," in *Paper Presented at the American Educational Research Association (AERA) Annual Meeting*, (Toronto, ON).

Dickler, R. (2021). *An Intelligent Tutoring System and Teacher Dashboard to Support Students on Mathematics in Science Inquiry* Doctoral dissertation. New Brunswick, NJ: Rutgers The State University of New Jersey, School of Graduate Studies.

Dimitrova, V. (2003). StyLE-: interactive open learner modelling. *Int. J. Artif. Intell. Educ.* 13, 35–78.

Dimitrova, V. G., and Brna, P. (2016). From interactive open learner modelling to intelligent mentoring: STyLE-OLM and beyond. *Int. J. Artif. Intell. Educ.* 26, 332–349.

Fancsali, S. E., Zheng, G., Tan, Y., Ritter, S., Berman, S. R., and Galyardt, A. (2018). "Using embedded formative assessment to predict state summative test scores," in *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, (New York, NY: Association for Computing Machinery), 161–170. doi: 10.1145/3170358.3170392

Feng, M., Krumm, A., and Grover, S. (2018). "Applying learning analytics to support instruction," in *Score reporting research and applications*, ed. D. Zapata-Rivera (New York, NY: Routledge), 145–159. doi: 10.4324/9781351136501-10

Ginon, B., Johnson, M. D., Turker, A., and Kickmeier-Rust, M. (2016). "Helping teachers to help students by using an open learner model," in *European Conference on Technology Enhanced Learning*, eds K. Verbert, M. Sharples, and T. Klobuèar (Cham: Springer), 587–590. doi: 10.1007/978-3-319-45153-4_69

Goodwin, S. (1996). Data rich, information poor (DRIP) syndrome: is there a treatment? *Radiol. Manag.* 18, 45–49.

Govaerts, S., Verbert, K., Duval, E., and Pardo, A. (2012). "The student activity meter for awareness and self-reflection," in *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, (Austin, TX), 869–884. doi: 10.1145/2212776.2212860

Hambleton, R., and Zenisky, A. (2013). *Reporting Test Scores in More Meaningful Ways: A Research-Based Approach to Score Report Design. APA Handbook of Testing and Assessment in Psychology*. Washington, DC: American Psychological Association, 479–494.

Hao, J., and Mislevy, R. J. (2018). The evidence trace file: a data structure for virtual performance assessments informed by data analytics and evidence-centered design. *ETS Res. Rep. Ser.* 2018, 1–16. doi: 10.1002/ets2.12215

Hartley, D., and Mitrovic, A. (2002). "Supporting learning by opening the student model," in *Proceedings of ITS 2002*, (Berlin: Springer), 453–462. doi: 10.1007/3-540-47987-2_48

Heffernan, N., and Heffernan, C. (2014). The ASSISTments ecosystem: building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *Int. J. Art. Intell. Educ.* 24, 470–497. doi: 10.1007/s40593-014-0024-x

Hegarty, M. (2019). "Advances in cognitive science and information visualization," in *Score Reporting Research and Applications*, ed. D. Zapata-Rivera (New York, NY: Routledge), 19–34. doi: 10.4324/9781351136501-3

Hooshyar, D., Pedaste, M., Saks, K., Leijen, Ä, Bardone, E., and Wang, M. (2020). Open learner models in supporting self-regulated learning in higher education: a systematic literature review. *Comput. Educ.* 154:103878. doi: 10.1016/j.compedu.2020.103878

Impara, J. C., Divine, K. P., Bruce, F. A., Liverman, M. R., and Gay, A. (1991). Does interpretive test score information help teachers? *Educ. Meas. Issues Pract.* 10, 16–18. doi: 10.1111/j.1745-3992.1991.tb00212.x

Kane, M. T. (2006). "Validation," in *Educational measurement*, 4 Edn, ed. R. L. Brennan (Westport, CT: American Council on Education), 17–64.

Kannan, P., Beigman-Klebanov, B., Shao, S., and Long, R. (2019). "Evaluating teachers' needs for on-going feedback from a technology-based book reading intervention," in *Paper presented at the 2019 Annual Meeting of the National Council for Measurement in Education*, (Toronto, ON).

Kannan, P., Bryant, A. D., Shao, S., and Wylie, E. C. (2021a). *Identifying Teachers' Needs for Results From Interim unit Assessments. (Research Report No. RR-21-08)*. Princeton, NJ: ETS. doi: 10.1002/ets2.12320

Kannan, P., Zapata-Rivera, D., and Bryant, A. D. (2021b). Evaluating parent comprehension of measurement error information presented in score reports. *Pract. Assess. Res. Eval.* 26:12. doi: 10.7275/rgwg-t355

Kannan, P., Zapata-Rivera, D., and Leibowitz, E. A. (2018a). Interpretation of score reports by diverse subgroups of parents. *Educ. Assess.* 23, 173–194. doi: 10.1080/10627197.2018.1477584

Kannan, P., Zapata-Rivera, D., Mikeska, J., Bryant, A. D., Long, R., and Howell, H. (2018b). "Providing formative feedback to Pre-service teachers as they practice facilitation of high-quality discussions in simulated mathematics and science methods classrooms," in *Proceedings of Society for Information Technology & Teacher Education International Conference*, eds E. Langran and J. Borup (Washington, D.C), 1570–1575.

Kay, J. (1995). The um toolkit for cooperative user modeling. *User Model. User Adapted Interact.* 4, 149–196. doi: 10.1007/BF01100243

Kay, J., Bartimote, K., Kitto, K., Kummerfeld, B., Liu, D., and Reimann, P. (2022). Enhancing learning by Open Learner Model (OLM) driven data design. *Comput. Educ. Art. Intell.* 3:100069. doi: 10.1016/j.caeai.2022.100069

Kersten-van Dijk, E. T., Westerink, J. H. D. M., Beute, F., and IJsselsteijn, W. A. (2017). Personal informatics, self-insight, and behavior change: a critical review of current literature. *Hum. Comput. Interact.* 32, 268–296. doi: 10.1080/07370024.2016.1276456

Keskin, S., and Yurdugül, H. (2021). "Linking assessment results and feedback representations in e-assessment: evidence-centered assessment analytics process model," in *Visualizations and Dashboards for Learning Analytics. Advances in Analytics for Learning and Teaching*, eds M. Sahin and D. Ifenthaler (Cham: Springer), 565–584. doi: 10.1007/978-3-030-81222-5_26

Knoop-van Campen, C. A. N., Wise, A., and Molenaar, I. (2021). The equalizing effect of teacher dashboards on feedback in K-12 classrooms. *Interact. Learn. Environ.* 1–17. doi: 10.1080/10494820.2021.1931346

Khosravi, H., Shabaninejad, S., Bakharia, A., Sadiq, S., Indulska, M., and Gašević, D. (2021). Intelligent learning analytics dashboards: automated drill-down recommendations to support teacher data exploration.åJ. *Learn. Analyt.* 8, 133–154. doi: 10.18608/jla.2021.7279

Kulik, J. A., and Kulik, C. C. (1988). Timing of feedback and verbal learning. *Rev. Educ. Res.* 58, 79–97. doi: 10.3102/00346543058001079

Kuosa, K., Distante, D., Tervakari, A., Cerulo, L., Fernández, A., Koro, J., et al. (2016). Interactive visualization tools to improve learning and teaching in online learning environments. IJDET *Int. J. Dist. Educ. Technol.* 14, 1–21. doi: 10.4018/IJDET.2016010101

Lee, S. J., and Bull, S. (2008). An open learner model to help parents help their children. *Technol. Instr. Cogn. Learn.* 6:29.

Leonardou, A., Rigou, M., and Garofalakis, J. D. (2019). "Open learner models in smart learning environments," in *Cases on Smart Learning Environments*, eds A. Darshan Singh, S. Raghunathan, E. Robeck, and B. Sharma (Hershey, PA: IGI Global), 346–368. doi: 10.3390/pharmacy8040197

Liu, M., Han, S., Shao, P., Cai, Y., and Pan, Z. (2021). "The current landscape of research and practice on visualizations and dashboards for learning analytics," in *Visualizations and Dashboards for Learning Analytics. Advances in Analytics for Learning and Teaching*, eds M. Sahin and D. Ifenthaler (Cham: Springer). doi: 10.1007/978-3-030-81222-5_2

Marshall, B., and Drummond, M. J. (2006). How teachers engage with assessment for learning: lessons from the classroom. *Res. Pap. Educ.* 21, 133–149. doi: 10.1080/02671520600615638

Mazza, R., and Dimitrova, V. (2007). CourseVis: a graphical student monitoring tool for supporting instructors in web-based distance courses. *Int. J. Hum. Comput. Stud.* 65, 125–139. doi: 10.1016/j.ijhcs.2006.08.008

Michaeli, S., Kroparo, D., and Hershkovitz, A. (2020). Teachers' use of education dashboards and professional growth. *Int. Rev. Res. Open Distrib. Learn.* 21, 61–78. doi: 10.19173/irrodl.v21i4.4663

Mitrovic, A., and Martin, B. (2002). "Evaluating the effects of open student models on learning," in *Proceedings of Second International Conference: Adaptive Hypermedia and Adaptive Web-Based Systems*, eds P. De Bra, P. Brusilovsky, and R. Conejo (Berlin: Springer-Verlag), 296–305. doi: 10.1007/3-540-47952-X_31

Molenaar, I., and Knoop-van Campen, C. A. N. (2018). How teachers make dashboard information actionable. *IEEE Trans. Learn. Technol.* 12, 347–355.

Nicol, D. J., and MacFarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Stud. High. Educ.* 31, 199–218. doi: 10.1080/03075070600572090

Papamitsiou, Z., and Economides, A. (2014). Learning analytics and educational data mining in practice: a systematic literature review of empirical evidence. *Educ. Technol. Soc.* 17, 49–64.

Rahimi, S., and Shute, V. (2021). "Learning analytics dashboards in educational games," in *Visualizations and Dashboards for Learning Analytics. Advances in Analytics for Learning and Teaching*, eds M. Sahin and D. Ifenthaler (Cham: Springer), doi: 10.1007/978-3-030-81222-5_24

Ritter, S., Yudelson, M., Fancsali, S. E., and Berman, S. R. (2016). "How mastery learning works at scale," in *Proceedings of the Third (2016) ACM Conference on Learning @ Scale*, (New York, NY: Association for Computing Machinery), 71–79. doi: 10.1145/2876034.2876039

Sahin, M., and Ifenthaler, D. (2021). "Visualization and dashboards: challenges and future directions," in *Visualizations and Dashboards for Learning Analytics. Advances in Analytics for Learning and Teaching*, eds M. Sahin and D. Ifenthaler (Cham: Springer),

Schwendimann, B. A., Rodríguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., et al. (2017). Perceiving learning at a glance: a systematic literature review of learning dashboard research. *IEEE Trans. Learn. Technol.* 10, 30–41. doi: 10.1109/TLT.2016.2599522

Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., and Kirschner, P. A. (2020). Linking learning behavior analytics and learning science concepts: designing a learning analytics dashboard for feedback to support learning regulation. *Comput. Human Behav.* 107:105512. doi: 10.1016/j.chb.2018.05.004

Sedrakyan, G., Mannens, E., and Verbert, K. (2019). Guiding the choice of learning dashboard visualizations: linking dashboard design and data visualization concepts. *J. Comput. Lang.* 50, 19–38. doi: 10.1016/j.jvlc.2018.11.002

Sergis, S., and Sampson, D. (2019). "An analysis of open learner models for supporting learning analytics," in *Learning Technologies for Transforming Large-Scale Teaching, Learning, and Assessment*, eds D. Sampson, J. M. Spector, D. Ifenthaler, P. Isaías, and S. Sergis (Cham: Springer), 155–190. doi: 10.1007/978-3-030-15130-0_9

Shepard, L. A. (2005). "Formative assessment: caveat emptor," in *Paper presented at the 2005 ETS Invitational Conference on The Future of Assessment: Shaping Teaching and Learning*, (New York, NY).

Shute, V. J. (2008). Focus on formative feedback. *Rev. Educ. Res.* 78, 153–189. doi: 10.3102/0034654307313795

Shute, V., and Zapata-Rivera, D. (2012). "Adaptive educational systems," in *Adaptive Technologies for Training and Education*, eds P. Durlach and A. Lesgold (Cambridge: Cambridge University Press), 7–27. doi: 10.1017/CBO9781139049580.004

Sinatra, A., Graesser, A. C., Hu, X., Goldberg, B., and Hampton, A. J. (Eds) (2020). *Design Recommendations for Intelligent Tutoring Systems: Volume 8-Data Visualization*. Orlando, FL: US Army Combat Capabilities Development Command–Soldier Center.

Tannenbaum, R. J. (2019). "Validity aspects of score reporting," in *Score Reporting Research and Applications*, ed. D. Zapata-Rivera (New York, NY: Routledge), 9–18. doi: 10.4324/9781351136501-2

Underwood, J. S., Zapata-Rivera, D., and VanWinkle, W. (2010). *An Evidence-Centered Approach to Using Assessment Data for Policymakers. (Research Report No. RR-10-03)*. Princeton, NJ: Educational Testing Service, doi: 10.1002/J.2333-8504.2010.TB02210.X

Valle, N., Antonenko, P., Dawson, K., and Huggins-Manley, A. C. (2021). Staying on target: a systematic literature review on learner-facing learning analytics dashboards. *Br. J. Educ. Technol.* 52, 1724–1748. doi: 10.1111/bjet.13089

Van Labeke, N., Brna, P., and Morales, R. (2007). Opening up the interpretation process in an open learner model. *Int. J. Art. Intell. Educ.* 17, 305–338.

VanWinkle, W., Vezzu, M., and Zapata-Rivera, D. (2011). *Question-Based Reports For Policymakers (Research Memorandum No. RM-11–16)*. Princeton, NJ: Educational Testing Service.

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., and Santos, J. L. (2013). Learning analytics dashboard applications. *Am. Behav. Sci.* 57, 1500–1509. doi: 10.1177/0002764213479363

Wainer, H., Hambleton, R. K., and Meara, K. (1999). Alternative displays for communicating NAEP results: a redesign and validity study. *J. Educ. Meas.* 36, 301–335. doi: 10.1111/j.1745-3984.1999.tb00559.x

Wise, A. F., and Jung, Y. (2019). Teaching with analytics: towards a situated model of instructional decision-making. *J. Learn. Anal.* 6, 53–69. doi: 10.18608/jla.2019.62.4

Xhakaj, F., Aleven, V., and McLaren, B. M. (2017). "Effects of a teacher dashboard for an intelligent tutoring system on teacher knowledge, lesson planning, lessons and student learning," in *Proceedings of the European conference on technology enhanced learning*, (Cham: Springer), 315–329. doi: 10.1007/978-3-319-66610-5_23

Yoo, Y., Lee, H., Jo, I. H., and Park, Y. (2015). "Educational dashboards for smart learning: review of case studies," in *Emerging Issues in Smart Learning. Lecture*

Notes in Educational Technology, eds G. Chen, V. Kumar, Kinshuk, R. Huang, and S. Kong (Berlin: Springer), 145–155.

Zapata-Rivera, D. (2020). Open student modeling research and its connections to educational assessment. *Int. J. Art. Intell. Educ.* 31, 380–396. doi: 10.1007/s40593-020-00206-2

Zapata-Rivera, D. (ed.) (2019). *Score Reporting Research and Applications.* New York, NY: Routledge.

Zapata-Rivera, D., and Arslan, B. (2021). "Enhancing personalization by integrating top-down and bottom-up approaches to learner modeling," in *Adaptive Instructional Systems. Adaptation Strategies and Methods. HCII 2021. Lecture Notes in Computer Science*, Vol. 12793, eds R. Sottilare and J. Schwarz (Cham: Springer), 234–246. doi: 10.1007/978-3-030-7787 3-6_17

Zapata-Rivera, D., Graesser, A., Kay, J., Hu, X., and Ososky, S. (2020). "Visualization implications for the validity of ITS," in *Design Recommendations for Intelligent Tutoring Systems: Volume 8 – Data Visualization*, eds A. Sinatra, A. C. Graesser, X. Hu, B. Goldberg, and A. J. Hampton (Orlando, FL: U.S. Army CCDC - Soldier Center), 61–68.

Zapata-Rivera, D., Hansen, E. G., Shute, V. J., Underwood, J. S., and Bauer, M. I. (2007). Evidence-based approach to interacting with open student models. *Int. J. Art. Intell. Educ.* 17, 273–303.

Zapata-Rivera, D., and Katz, I. (2014). Keeping your audience in mind: applying audience analysis to the design of score reports. *Assess. Educ.* 21, 442–463. doi: 10.1080/0969594X.2014.936357

Zapata-Rivera, D., Liu, L., Chen, L., Hao, J., and von Davier, A. (2016). "Assessing science inquiry skills in immersive, conversation-based systems," in *Big Data and Learning Analytics in Higher Education*, ed. B. K. Daniel (New York, NY: Springer International Publishing), 237–252, doi: 10.1007/978-3-319-06 520-5_14

Zapata-Rivera, D., VanWinkle, W., and Zwick, R. (2012). *Applying Score Design Principles in the Design of Score Reports for CBAL*[TM] *Teachers. ETS Research Memorandum RM-12-20.* Princeton, NJ: ETS.

Zapata-Rivera, J. D., and Greer, J. (2002). "Exploring various guidance mechanisms to support interaction with inspectable learner models," in *Proceedings of the Intelligent Tutoring Systems. ITS 2002. Lecture Notes in Computer Science*, Vol. 2363, eds S. A. Cerri, G. Gouardères, and F. Paraguaçu (Berlin: Springer), 442–452. doi: 10.1007/3-540-4798 7-2_47

Zapata-Rivera, J. D., and Greer, J. (2004). Interacting with Bayesian student models. *Int. J. Art. Intell. Educ.* 14, 127–163.