Review of Research on Student-Facing Learning Analytics Dashboards and Educational Recommender Systems — Source link

Robert Bodily, Katrien Verbert

Institutions: Brigham Young University, University of Copenhagen Faculty of Science

Related papers:

- Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research

- Learning dashboards: an overview and future research opportunities

- License to evaluate: preparing learning analytics dashboards for educational practice

- Learning analytics dashboard applications

- Awareness Is Not Enough: Pitfalls of Learning Analytics Dashboards in the Educational Practice

This article can currently be accessed on IEEE Xplore at the following web address: http://ieeexplore.ieee.org/document/8010828/

Review of research on student-facing learning analytics dashboards and educational recommender systems

Robert Bodily and Katrien Verbert, *Member, IEEE*

Abstract—This article is a comprehensive literature review of student-facing learning analytics reporting systems that track learning analytics data and report it directly to students. This literature review builds on four previously conducted literature reviews in similar domains. Out of the 945 articles retrieved from databases and journals, 93 articles were included in the analysis. Articles were coded based on the following five categories: functionality, data sources, design analysis, student perceptions, and measured effects. Based on this review, we need research on learning analytics reporting systems that targets the design and development process of reporting systems, not only the final products. This design and development process includes needs analyses, visual design analyses, information selection justifications, and student perception surveys. In addition, experiments to determine the effect of these systems on student behavior, achievement, and skills are needed to add to the small existing body of evidence. Furthermore, experimental studies should include usability tests and methodologies to examine student use of these systems, as these factors may affect experimental findings. Finally, observational study methods, such as propensity score matching, should be used to increase student access to these systems but still rigorously measure experimental effects.

Index Terms—Data mining, Data and knowledge visualization, Self-assessment technologies, Homework support systems, Adaptive and intelligent educational systems, Literature review

———————————— ◆ ————————————

1 INTRODUCTION

ONLINE learning continues to grow, in part, due to reduced costs, increased flexibility regarding class schedules, and improved mobility when taking classes (Allen & Seaman, 2014). As online learning becomes more widespread, it becomes increasingly important to understand how to help learners succeed in online environments. The focus of the emerging field of learning analytics is to achieve this goal. Learning analytics is defined as "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" (Siemens, 2010, para. 6). This definition is used because it was provided during the first conference on learning analytics and has since been adopted by the Society of Learning Analytics Research (SoLAR). The learning analytics process includes selecting, capturing, analyzing, and reporting data, and then refining this process based on what has been learned (Clow, 2012; Greller, & Drachsler, 2012). The majority of learning analytics systems report student interaction data to instructors or administrators (Schwendimann et al., 2016). However, these systems restrict student autonomy as administrators and instructors make decisions affecting student learning without direct student involvement. Student autonomy is defined within the self-determination theory framework as

the level of control students are given in their learning. Students with high levels of autonomy are likely to be intrinsically motivated to succeed (Ryan & Deci, 2000). Student-facing learning analytics systems can enable student autonomy, giving students more control over their learning and helping them feel more intrinsically motivated to succeed. For this reason, the focus of this review is on student-facing learning analytics reporting systems.

In this paper, we first discuss previous literature reviews related to this topic and how our review builds upon their work. We then discuss the methodology for identifying and including articles in our review. Then, we report on the coding and analysis methodology. Finally, we discuss our findings related to current research trends surrounding learning analytics reporting systems, give recommendations for future research, and provide implications for practice to improve online teaching and learning.

2 LITERATURE REVIEW

The scope of student-reporting systems would encompass all assessment and feedback systems in the literature and would be far too large for a single review. To narrow the focus, this literature review will focus exclusively on learning analytics systems that collect *click-level student data* and *report this data directly to students*. This data reporting may take the form of text feedback, recommendations, visualizations, or dashboards. These systems are found in a variety of educational technology fields such as intelligent tutoring systems, educational recommender systems, educational data mining systems, and learning analytics dashboard systems.

———————————————
- *R. Bodily is with the Instructional Psychology and Technology department, at 150 MCKB Brigham Young University, Provo, UT 84604. E-mail: bodilyrobert@gmail.com*
- *K. Verbert is with the Human and Computer Interaction group in the department of Computer Science, KU Leuven, Celestijnenlaan 200A, B-3001 Heverlee, Belgium Office: 04.49. E-mail: katrien.verbert@cs.kuleuven.be*

Intelligent tutoring systems are electronic systems which seek to improve learning that "must possess: (a) knowledge of the domain (expert model), (b) knowledge of the learner (student model), and (c) knowledge of teaching strategies (tutor)" (Hartley & Sleeman, 1973, p. 808). Educational recommender systems are defined as "any system that produces individualized recommendations as output or has the effect of guiding the user in a personalized way to interesting or useful objects in a large space of possible options" (Burke, 2002, p. 1). Educational data mining systems "[seek] to use…data repositories to better understand learners and learning, and to develop computational approaches that combine data and theory to transform practice to benefit learners" (Romero & Ventura, 2010, p. 1). Learning analytics dashboards "support users in collecting personal information about various aspects of their life, behavior, habits, thoughts, and interest. [They also] help users to improve self-knowledge by providing tools for the review and analysis of their personal history" (Verbert, Duval, Klerkx, Govaerts, & Santos, 2013, p. 2). A diagram illustrating the focus of this literature review is indicated with the student-facing systems gray oval seen below (Figure 1).
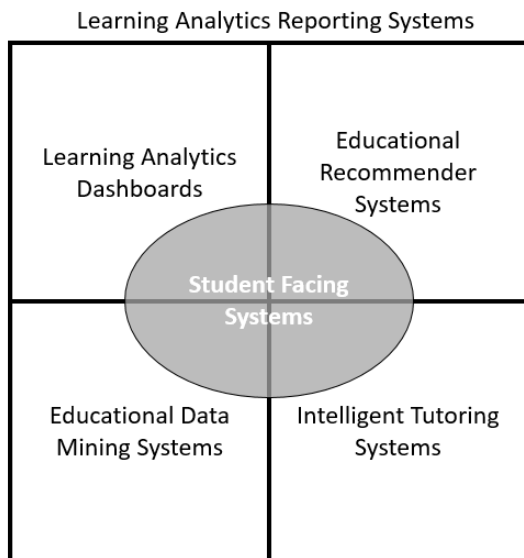

Learning Analytics Reporting Systems

Fig. 1. A diagram illustrating the focus of this literature review situated between various educational technology sub-fields.

2.1 Previous Reviews

Four previously published literature reviews are relevant to this review. Verbert et al. (2013) reviewed 15 learning analytics dashboards (LAD). These LAD were selected for review in order to provide an illustration of their conceptual framework and provide interesting examples for the reader. Verbert et al. (2013) coded these articles based on the target user (e.g., instructor, student), what data was tracked (e.g., resources use, time spent), and what type of evaluation was conducted (e.g., usability, effectiveness). This categorization of LAD based on data type, target user, and evaluation conducted was the first published review of LAD, so instead of including a large number of articles, it provided an example for future research.

Verbert et al. (2014) built upon the work done in Verbert et al. (2013) by expanding the categorization of LAD and including additional LAD in the analysis. Their article is still not an exhaustive search of the literature, but instead seeks to provide a variety of interesting articles that will benefit the reader. Additional categories added to this analysis beyond the Verbert et al. (2013) article include devices used (e.g., laptop, cell phone, tabletop), some extra types of evaluation conducted (e.g., efficiency), and data tracking technology used (e.g., microphone, depth sensor, manual reporting). Additional systems were included in this analysis when compared with the previous study, but it still only included a small number of articles.

Yoo, Lee, Jo, and Park (2015) took 10 articles from the previous two literature reviews—Verbert et al. (2013) and Verbert et al. (2014)—and extended their framework by adding a more extensive evaluation criteria. They found that the current research on LAD is lacking in evaluation, so they created an evaluation framework of 11 sub-categories for dashboard evaluation: goal-orientation, information usefulness, visual effectiveness, appropriation of visual representation, user friendliness, understanding, reflection, learning motivation, behavioral change, performance improvement, and competency development. The sub-categories in the evaluation framework were excellent and were instrumental in the development of pieces of the categorization framework for this literature review.

Schwendimann, Boroujeni, Holzer, Gillet, and Dillenbourg (2016) conducted the first exhaustive search of the literature on LAD. Their methodology included searching for the phrases "dashboard AND ("learning analytics" OR "educational data mining" OR "educational datamining" in the databases ACM Digital Library, IEEE Xplore, Springer Link, Science Direct, and Wiley (Schwendimann et al., 2016, p. 1). Their search included all learning dashboards regardless of the stakeholder the system was intended for. They found that the majority of systems are instructor facing (74%) in a higher education context, and researchers do not conduct much research on the impact of these systems on teaching and learning.

We are interested in examining what types of systems exist in the student-facing learning analytics reporting system literature regarding their purpose, functionality, and types of data collected. Schwendimann et al. (2016) addressed this question at a broader level discussing the purpose, data sources, platforms, indicator types, visualization types, and technologies used. However, this analysis looked at each level separately, or at most, two levels at a time. We are interested more specifically in the mechanisms by which student-facing systems attempt to improve teaching and learning, which would require an analysis across categories. This has not previously been done, and would require a more comprehensive search of the literature beyond learning analytics dashboards (Schwendimann et al., 2016), along with a method and categorization scheme that allowed for comparison across codes. Schwendimann et al. (2016) suggested this was a gap in the current research and said, "The field still lacks comparative studies among different dashboards or dashboard design options" (p. 9). One of the first topics to address in

order to compare dashboards or dashboard design options is to understand what types of systems exist in student-facing reporting literature based on their purpose, functionality, and data types collected.

We are also interested in examining which methods are being used to increase the rigor of student-facing reporting systems research. Verbert et al. (2013, 2014), Yoo et al. (2015), and Schwendimann et al. (2016) all partially addressed this question. However, these four previous reviews focused on summative evaluation of systems that had already been created. We are interested in both summative and formative evaluation, specifically looking for evaluation or research methods being used to increase the rigor of design and development in student-facing learning analytics reporting systems research.

Lastly, we are interested in examining, across the field, the effect of student-facing learning analytics reporting systems on student achievement, student behavior, and student skills. This has not been previously addressed in a literature review and would provide a synthesis of the effect of these systems on student behavior, achievement, and skills.

In summary, we will address the following questions in this review:

1. What types of systems exist within the student-facing learning analytics reporting system literature based on their purpose, functionality, and the types of data they collect?
2. Which methods are being used to increase the rigor of research in student-facing learning analytics reporting system literature?
3. What is the effect of having access to a student-facing learning analytics reporting system on student behavior, student achievement, and student skills?

3 ARTICLE SEARCH METHODS

Learning analytics reporting systems research is a multi-disciplinary research area that is a combination of education and computer science. Because of this, the following education and computer science journal databases have been included in our search: ERIC to capture education articles, IEEE Xplore to capture computer science conference proceedings, Computers and Applied Sciences to capture computer science journal articles, and ACM to capture additional computer science articles. We also conducted targeted searches in Google Scholar, reviewed the entire educational data mining (EDM) and learning analytics and knowledge (LAK) conference proceedings, and found relevant literature reviews for additional citations to ensure articles were not missed because they were not indexed in the previously mentioned databases. The searches conducted are explained in Table 1.

We chose to only include journal articles that were peer reviewed and published between January 2005 and June 2016. The year 2005 was the start year because no articles were found before that time. The only exception to journal articles is conference proceedings from IEEE

TABLE 1
DATABASES, JOURNALS, AND ARTICLES SEARCHED WITH THEIR CORRESPONDING TOPIC OR SEARCH TERM

Source	Search Term or Topic	Count
ERIC	(student OR students) AND ("data driven decision making" OR "resource use" OR analytics OR "student interaction" OR clickstream OR "online activity" OR "data mining") AND (dashboard OR visualization OR visual OR recommendation OR recommendations OR recommender)	193
LAK & EDM Proceedings	dashboard OR visualization OR visual OR recommendation OR recommender OR feedback	24
IEEE Xplore	(student OR students) AND (.QT.data driven decision making.QT. OR .QT.resource use.QT. OR analytics OR .QT.student interaction.QT. OR clickstream OR .QT.online activity.QT. OR .QT.data mining.QT.) AND (dashboard OR visualization OR visual OR recommendation OR recommendations OR recommender)	260
Computers and Applied Sciences	(student OR students) AND ("data driven decision making" OR "resource use" OR analytics OR "student interaction" OR clickstream OR "online activity" OR "data mining") AND (dashboard OR visualization OR visual OR recommendation OR recommendations OR recommender)	102
ACM database	(student OR students) AND ("data driven decision making" OR "resource use" OR analytics OR "student interaction" OR clickstream OR "online activity" OR "data mining") AND (dashboard OR visualization OR visual OR recommendation OR recommendations OR recommender)	172
Google Scholar: search 1	intitle:"feedback system" AND intitle:"learning"	66
Google Scholar: search 2	intitle:"learning analytics" AND intitle:"feedback"	9
Google Scholar: search 3	intitle:"learning dashboard" OR intitle:"learning analytics dashboard"	14
Google Scholar: search 4	intitle:"dashboard" AND intitle:"feedback"	8
Google Scholar: search 5	intitle:"learning analytics" AND (intitle:"reflection" OR intitle:"reflect")	7
Google Scholar: search 6	intitle:"learning analytics" AND intitle:"awareness"	6
Google Scholar: search 7	intitle:"data mining" AND (intitle:"recommendations" OR intitle:"recommendation" OR intitle:"recommend") AND intitle:"learning"	17
Drachsler, Verbert, Santos, & Manouselis, (2015)	Literature review on educational recommender systems	37
Romero & Ventura, 2010	Literature review on educational data mining	30
Verbert et al., 2013, 2014; Schwendimann et al., 2016	Literature reviews on learning analytics dashboards	20

Xplore, the Learning Analytics and Knowledge conference, and the Educational Data Mining conference. IEEE Xplore is a database for computer science conference proceedings, so conference presentations within this database were included in our search. The learning analytics and educational data mining conference proceedings are the two conferences most closely related to learning analytics reporting systems, so they were included in this review as well.

To increase the rigor of our search criteria, literature review articles in similar domains were found and reviewed to identify relevant articles to this literature review. From this search, the following literature reviews were identified: an educational recommender system review article (Drachsler et al., 2015), an educational data mining review article (Romero & Ventura, 2010), and three learning analytics dashboards review articles (Verbert et al., 2013; Verbert et al., 2014; Schwendimann et al., 2016). We were not able to find any relevant articles from intelligent tutoring system review articles.

Finally, to ensure important articles were not missed, the titles of all of the previously found articles were examined for keywords. Keywords included learning dashboard, feedback system, recommendations, dashboard, learning analytics, feedback, reflection, and awareness. Once these words were identified, they were entered in Google Scholar and relevant articles were either kept as part of the review or rejected based on our inclusion criteria (see Table 1 for the exact searches). Once all articles had been identified, duplicates were removed because some

articles showed up in multiple databases. There were 945 articles remaining for further analysis.

3.1 Inclusion Criteria

There were two main inclusion criteria used to narrow the pool of articles for this literature review. First, the article must have discussed a learning analytics system. This means the system had to automatically track student inter-action data. For example, this data could be resource use, time spent data, or social interaction data. Furthermore, assessment data alone did not count. Second, the system must automatically report data directly to students. For example, this could be in the form of visualizations, text-based feedback, dashboards, or recommendations.

Articles that did not meet both of these two inclusion criteria were eliminated from the analysis. This narrowed the scope of this literature review to 93 articles. The list of articles included in this analysis can be viewed at the following web address www.bobbodily.com/article_list.

4   CATEGORY AND SUB-CATEGORY DEFINITIONS

The 93 articles included in this analysis were coded based on the following five categories: functionality, data sources, design analysis, student perceptions, and measured effects. The functionality and data sources categories were used to determine the type of each system for our first research question, the design analysis and student perceptions categories were used to examine what kinds of methods are being used to increase the rigor of the design and development process of student-facing reporting systems for research question two, and the measured effects category will review the effect of having access to a learning analytics reporting system on student behavior, student achievement, and student skills for research question three. Each of these five categories was composed of subcategories. The categories and subcategories are defined below. These categories and subcategories were determined using both an open and closed coding approach. Some categories and subcategories were based on the coding categories used in previous literature reviews and some categories emerged as common themes from the articles in this review. We have included two learning analytics dashboards (see Figure 2 and 3) that provide multiple data views for students in order to provide a visual context for the categories and subcategories in this review (Santos et al., 2012; Grann & Bushway, 2014).

4.1 Functionality

The purpose of the functionality category is to determine what affordances the learning analytics reporting system offered to students and is broken down into the following subcategories: intended goal of the system, data mining, visuals, visual technique, recommendations, feedback, class comparison, and interactivity. These subcategories are defined in Table 2.

4.2 Data Sources

The data sources category examines the inputs to the learning analytics reporting systems to determine what types of
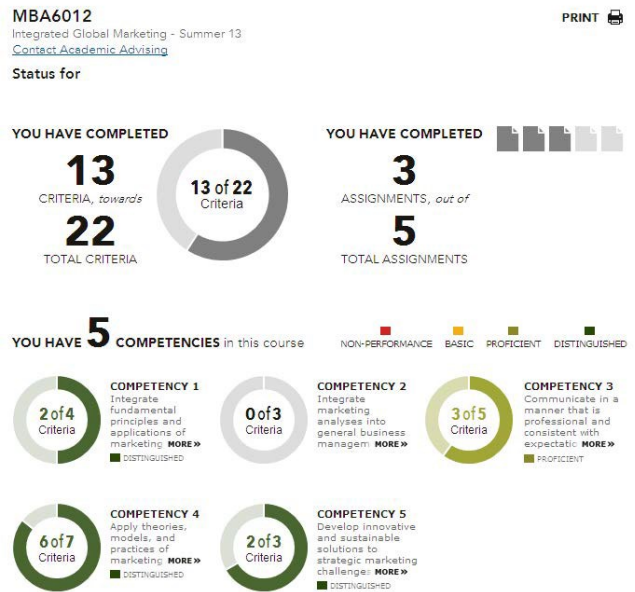
Fig. 2. J. Grann and D. Bushway, "Competency Map: Visualizing Student Learning to Promote Student Success," *Proc. Fourth Int'l Conf. Learning Analytics and Knowledge (LAK14)*, 2014; http://doi.org/10.1145/2567574.2567622. Figure 3.

Fig. 3. J.L. Santos et al., "Goal-oriented visualizations of activity tracking: A case study with engineering students," *Proc. Second Int'l Conf. Learning Analytics and Knowledge (LAK12)*, 2012; http://doi.org/10.1145/2330601.2330639. Figure 5.

TABLE 2
SUB-CATEGORY DEFINITIONS FOR THE FUNCTIONALITY CATEGORY

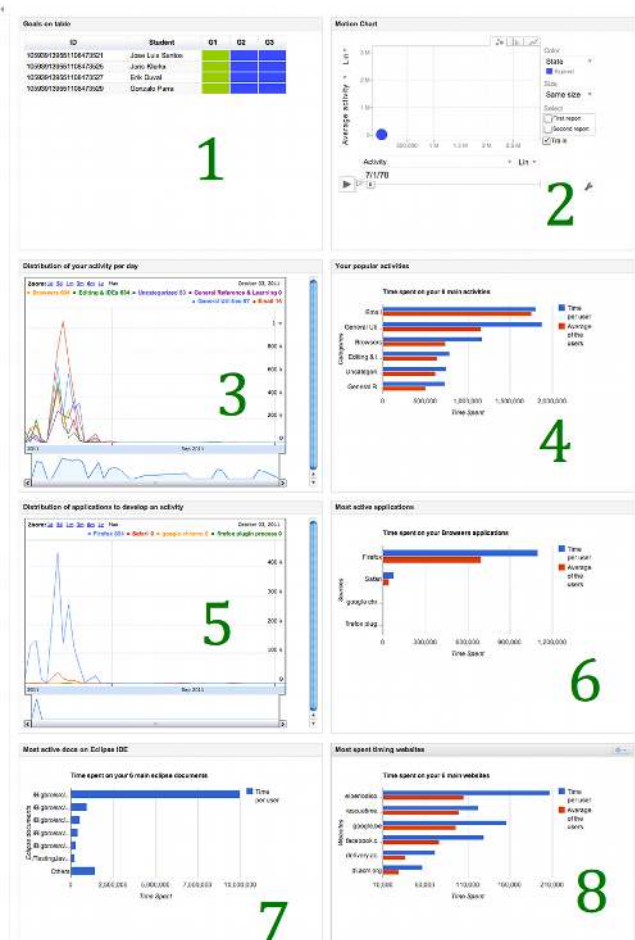| Sub-category	Definition
Intended goal of the system	The goal that the system was attempting to achieve
Data mining	Any automated statistical process before data reporting
Visual	Reported data to students in a visual or graphical form
Visual technique	The visualization method used to display data
Recommendations	Told students to do something based on the data
Feedback	Provided text feedback to students
Class comparison	Allowed students to compare their data with other students data
Interactive	Allowed students to interact with the reporting system

data are being collected, analyzed, and reported to students. This category is broken down into the following subcategories: resource use, assessment, social interaction, time spent, other sensors, and manually reported data. These subcategories are defined in Table 3.

TABLE 3
SUB-CATEGORY DEFINITIONS FOR THE DATA SOURCES CATEGORY

Sub-category	Definition
Resource use	The number of times a resource was accessed
Assessment	Student mastery data as measured by assessment
Social interaction	Any activity performed interacting with others
Time spent	The amount of time spent accessing resources
Other sensors	Other sensors include mouse tracking, heartrate, steps, etc.
Manually reported	Data manually reported by student, instructor, or other party.

4.3 Design Analysis

The design analysis category examines the design considerations that should be made before testing or implementing a reporting system with students in an actual class. The design analysis framework we use in this paper includes four sub-categories: needs assessment, information selection, visual design, and usability test. These sub-categories are defined in Table 4.

TABLE 4
SUB-CATEGORY DEFINITIONS FOR THE DESIGN ANALYSIS CATEGORY

Sub-category	Definition
Needs analysis	Analysis conducted to determine the needs of an end-user
Information selection	Justification provided for the types of data collected
Visual design	Justification provided for the types of visualizations used
Usability testing	Conducted a system usability test beyond student perceptions

4.4 Student Perceptions

The student perceptions category groups a variety of student perceptions on learning analytics reporting systems into the following subcategories: usability, satisfaction/usefulness, perceived behavior change, perceived achievement change, and perceived skills change. These sub-categories are defined in Table 5.

4.5 Measured Effects

The measured effects category deals with articles that conducted a research experiment to determine what effect the learning analytics reporting system had on students. The

TABLE 5
SUB-CATEGORY DEFINITIONS FOR THE STUDENT PERCEPTIONS CATEGORY

Sub-category	Definition
Usability	Asked students about the usability of the system
Usefulness	Asked students about the usefulness of the system
Perceived behavior change	Asked students if they perceived any behavior changes
Perceived achievement change	Asked students if they perceived any achievement changes
Perceived skills change	Asked students if they perceived any skills changes (e.g. meta-cognition, self-regulation, etc.)

measured effects category is broken down into three sub-categories: behavior, achievement, and skills. Each of these sub-categories is defined in Table 6.

TABLE 6
SUB-CATEGORY DEFINITIONS FOR THE MEASURED EFFECTS CATEGORY

Sub-category	Definition
Behavior	Conducted an experiment to determine if the system had an effect on student behavior
Achievement	Conducted an experiment to determine if the system had an effect on student achievement
Skills	Conducted an experiment to determine if the system had an effect on student skills (e.g. meta-cognition, self-regulation, etc.)

4.6 Student Use

The student use category deals with articles that track student data, report this data back to students, and then track how students interact with or use the reporting system. This interaction could be in the form of clicks, online sessions in the system, or page views in the system.

5 RESEARCH METHODS

A researcher examined the methods, results, discussion, and conclusion section of each article to determine whether the article would receive a one or zero in each of the categories (except visualization type and intended goal of the system which received a text description). Every article received either a one (indicating the article included the subcategory topic) or a zero (indicating the article did not include the subcategory topic). None of the subcategories within these five categories were mutually exclusive, which means an article could receive a one on every sub-category within every category. In order to ensure an objective coding approach, 20% of the articles were coded by a second reviewer. The agreement between the two coders was 86%.

In order to determine the types of systems that were discussed in the final set of articles, the functionality and data sources categories were grouped together to identify patterns across categories. Then, the number of unique article types was counted based on the data sources and functionality sub-category codes.

To determine what methods were being used to increase the rigor of student-facing learning analytics reporting systems, we used an open coding approach. Some category ideas were taken from previous literature reviews but the final categories emerged throughout the process as we read and coded articles.

To report on the effects of giving students access to these systems on student behavior, achievement, and

skills, each article that was coded as having done experimental research was analyzed again in more detail. The sample size, research methodology, and results for each article were extracted and summaries were provided for each sub-category.

6 RESULTS AND DISCUSSION

The results from analyzing each coding category will be discussed in the following sections, organized based on our research questions. First, we provide an overview of the frequency counts and percentages of total for each sub-category (see Table 7). Then, we discuss the results for each research question. Please note that the sub-category names have been extended or slightly modified to provide a better description for the context of the table.

TABLE 7
FREQUENCY COUNTS AND PERCENTAGES OF TOTAL FOR EACH SUB-CATEGORY IN OUR ANALYSIS

Category	Sub-category	# of Articles	% of Articles
Data Sources	Resource use data	70	75%
Functionality	Visualizations	62	67%
Functionality	Data mining	46	49%
Functionality	Recommendations	43	46%
Functionality	Class comparison	35	38%
Data Sources	Assessment data	34	37%
Student Perceptions	Student perceptions on usefulness	34	37%
Data Sources	Social interaction data	32	34%
Student Perceptions	Student perception on usability	32	34%
Functionality	Interactive	29	31%
Data Sources	Time spent data	28	30%
Functionality	Text feedback	17	18%
Actual Effects	Actual effect on student behavior	17	18%
Student Perceptions	Behavior change	16	17%
Actual Effects	Actual effect on student achievement	16	17%
Student Perceptions	Skills change	15	16%
Design Analysis	Justification for data selection	14	15%
Data Sources	Manual reporting data	12	13%
Design Analysis	Conducted a visual design analysis	12	13%
Student Use	Tracked student use of system	12	13%
Design Analysis	Conducted a usability test	10	11%
Data Sources	Other sensor data	7	8%
Design Analysis	Conducted a needs assessment	6	6%
Student Perceptions	Achievement change	2	2%
Actual Effects	Actual effect on student skills	2	2%

The most prevalent system characteristics were tracking resource use data, reporting data in visualizations, using data mining to process data, and providing recommendations to students. The least prevalent system characteristics were tracking other sensor data, conducting a needs assessment to identify the needs of the system end-user, asking students if they perceived an achievement change based on their system use, and examining the effect of these systems on student skills (e.g., awareness, meta-cognition, motivation, etc.).

6.1 Types of Learning Analytics Reporting System

The results presented in this section specifically address our first research question: What types of systems exist within student-facing learning analytics reporting system literature based on their purpose, functionality, and the

data types they collect? We aggregated the co-occurrence of various functionality and data sources categories in order to identify patterns in the types of student-facing learning analytics reporting systems discussed in the articles in this review. The groupings for the functionality category are reported in Table 8 below. A *visualization* type means data was displayed visually in a graphic or dashboard (Few, 2013). If it is an *enhanced visualization* type it means the visualization included a class comparison feature or an interactivity feature. If it is a recommender system or includes recommendations it means it is a *recommendations* or *recommender system* type. If data mining was conducted on the data before it was reported to students it is included as a *data mining* type.

TABLE 8
TYPES OF STUDENT-FACING LEARNING ANALYTICS REPORTING SYSTEMS AS CATEGORIZED BY FUNCTIONALITY

Type of reporting system	# of articles	% of articles
Enhanced visualization	27	29%
Data mining recommender system	25	27%
Enhanced visualization with recommendations	8	9%
Visualization	6	6%
Visualization with recommendations and data mining	6	6%
Other	22	23%

The most prevalent systems were the enhanced visualizations and the data mining recommender systems. This makes sense because enhanced visualizations would come from the learning analytics dashboard literature and the data mining recommender systems would come from the educational recommender systems literature.

One example of an enhanced visualization tool is a learning analytics dashboard that provides students with their mastery level on each concept in the class. The dashboard also provides class comparison functionality and interactivity. This means users can compare themselves to their class by looking at a visual representation of data generated in the class. One example of a data mining recommender system is a resource recommender that uses collaborative filtering techniques to recommend resources to a student based on their similarity to other students.

The groupings for the data sources category are reported in Table 9. The most common data source combinations are included, and the least common data source combinations are grouped into the "other" category. The most common data source collected was social interaction and resource use (17% of articles). All of the data sources within the top six categories collected some combination of social interaction, resource use, time spent, and assessment data.

TABLE 9
TYPES OF STUDENT-FACING LEARNING ANALYTICS REPORTING SYSTEMS AS CATEGORIZED BY DATA SOURCES

Type of reporting system	# of articles	% of articles
Social interaction and resource use	16	17%
Resource use, time spent, and assessment data	9	10%
Resource use and assessment data	8	9%
Social interaction data	8	9%
Social interaction, time spent, and resource use data	6	6%
Resource use and time spent	6	6%
Other	40	43%

To look for more detailed patterns, the two categories were combined and we searched for trends across all of the sub-categories within these two categories. The top three types of reporting systems as categorized by data sources and functionality are displayed in Table 10. The rest of the system types are not displayed because there were two or less occurrences.

TABLE 10
TYPES OF REPORTING SYSTEMS AS CATEGORIZED BY DATA
SOURCES AND FUNCTIONALITY

Types of reporting systems	# of articles	% of articles
Social interaction, resource use, data mining, recommendations	9	10%
Social interaction, resource use, enhanced visualization	6	6%
Social interaction, resource use, assessment data, data mining, recommendations	4	4%
Other	74	80%

There were two examples of learning analytics reporting systems that merit additional discussion from the Other category. The first is called The NTU Student Dashboard (Ferguson, Sharkey, & Mirriahi, 2016). The NTU Dashboard was implemented at Nottingham Trent University. This system integrates tutor comments, student biographical information, door swipes (tracked by student ID card), library loans, virtual learning environment use, Dropbox submissions, and attendance in classes. This comprehensive data collection contains much more information than Learning Management System data and could potentially increase the predictive power of current early-alert warning system prediction algorithms. However, there are not published results yet from this dashboard on the effect of the dashboard on student behavior, achievement, or skills.

The other example of a system in the Other Sensors category was an educational resources recommendation system (Holanda et al., 2012). This system provided students with blog article recommendations based on a blog web crawler, student post behavior, and student post content on instructor blogs. This article included a text mining component not common among other learning analytics reporting systems. In a small experimental study, they found that interacting with the resource recommender increased the percentage of interaction by 83.3% (N=12). More research is needed on learning analytics reporting systems that incorporate text mining as additional data sources to see the effect on student achievement, behavior, and skills.

There are no major trends across reporting systems. When combining the sub-categories for the functionality and data sources categories, there were 68 unique types of systems. One reason for this could be that student-facing learning analytics reporting systems are tools that are context dependent. Each circumstance has unique instructors, students, needs, and goals, which means each system needs to track unique data sources and report it in unique ways to strive to achieve a unique goal. Another reason for this could be that researchers do not know what is best to track and report to students, so there are a wide variety of approaches in use. In summary, more research should investigate which types of data and functionality elements lead to increased student success to help guide the student reporting system field.

6.2 Methods for Rigorous Research

The results presented in this section specifically address our second research question: Which methods are being used to increase the rigor of research in student-facing learning analytics reporting system literature? The methods identified using an open coding approach in the article analysis stage include the following methods: needs assessment, information selection analysis, visual design analysis, student usability perceptions, conducting usability tests, and tracking student use of the reporting system. Each of these sub-categories along with examples extracted from the literature will be discussed in the following sections.

6.2.1 Needs Assessment

A needs assessment is common in instructional design. The purpose is to understand what the stakeholder or end user needs. It answers the question, "What problem needs to be solved?" Out of the 93 total articles in this analysis, only 6% of articles (N=6) included a description of their needs analysis. It is likely that an informal needs assessment is still happening for some of the other 87 articles included in our analysis; however, it is important to be more explicit about the kinds of needs analyses we are conducting. Santos, Verbert, Govaerts, & Duval (2013) conducted a needs assessment on their system called StepUp!. They conducted three brainstorm sessions with different groups of students to identify problems students faced in their courses. Next, each student group rated the previously identified problems to determine which were most important to them. The problems that could be addressed by a learning dashboard were then selected and sorted based on student ranking of importance. Solutions to the final list of problems were then implemented into the learning dashboard. We need more research on learning analytics reporting systems that conduct rigorous needs assessments.

6.2.2 Information Selection

The learning analytics process definition commonly includes a data selection stage (Campbell & Oblinger, 2007). The data selection stage is determining what data should be collected. In our analysis, we call this the information selection sub-category. Only 15% of articles (N=14) provided information about why they were collecting certain types of data. It is likely that there are good reasons the rest of the articles are collecting the types of data they are collecting; however, it would be beneficial if researchers started examining why they are collecting some types of data but not other types of data.

From the articles in this literature review, we have identified three articles that conducted a meaningful information selection process. In order to identify performance indicators, Ott, Robins, Haden, & Shephard (2015) reviewed the literature that examined predictors of student success in programming courses. Then, they created indicators that had been previously shown to predict student success to use in their models. These indicators include pre-course grades, number of submitted laboratory tasks, time of submission, and mid-semester exam result. These

indicators were then visualized in the infographic they created for their course (Ott et al., 2015).

Feild (2015) conducted exploratory data analysis in order to determine which indicators were worth reporting to students as feedback. The author analyzed various levels of data including days of the semester, days of the week, hours of the day, and start and submit times of student assignments. Based on their findings in the exploratory data analysis, Feild identified four messages they could include in their feedback engine (Feild, 2015).

Iandoli, Quinto, De Liddo, and Buckingham Shum (2014) used a theoretical framework in order to determine what feedback to give to students. The three categories of feedback they identified were community (who), interaction process (how), and knowledge absorption (what, where). Based on the purposes identified for each of these categories, they were better able to frame what, when, and where to represent information to students (Iandoli et al., 2014). Based on this analysis, more learning analytics reporting systems should report on reviewing previous literature, conducting exploratory data analysis, and using a theoretical framework to guide the information selection process.

6.2.3 Visual Design

For learning analytics dashboards, this would be deciding which visualization is the best representation of the data. For educational recommender systems, this would be deciding when and where is best to provide a recommendation. Olmos & Corrin (2012) provide an excellent example of the benefits of an iterative design process when designing the visual component of a learning analytics reporting system. The first visualization they tried was a table. Then, after reflecting on the advantages and disadvantages of their table design, they tried a Gantt chart. After additional reflection, their next iteration used a line chart which was more vertically compact and showed additional information not found in the Gantt chart. After a final round of reflection, their final design expanded on the line chart by adding additional symbols as markers along the lines as well as color coding the lines based on student. While this process was devoid of any kind of user-testing, this iterative visual design process allowed them to create a much cleaner, succinct, and informative visualization than they would have been able to create otherwise. More learning analytics reporting systems need to take into account and report on the visual design process in order to improve the visualizations in these systems.

6.2.4 Student Perceptions of Usability

The majority of these articles administered surveys to students to assess student perceptions of usability. This can apply to learning analytics dashboards as well as educational recommender systems. In a recommendation system, the questions might include if the recommendations were presented at an appropriate time or if the recommendations were easy to understand. In a dashboard system, the questions might include if the visualizations were easy to understand or whether they were easy to access. To better analyze student perceptions beyond administering a

survey, Wise, Zhao, & Hausknecht (2014) conducted interviews with seven students and the instructor to evaluate the usability of their system. These interviews focused on student understanding of and reactions to the system analytics. One of the benefits to using interviews instead of surveys to assess usability is participants can be led through a think-aloud process to give feedback as they interact with the system. This interview process can provide additional insights into the usability of a system than a single response on a survey (Wise et al., 2014). More learning analytics reporting systems should include interviews and think-aloud protocols in their usability testing in addition to survey work.

6.2.5 Usability Testing

This is separate from the student perceptions usability category (discussed in the student perceptions category section above) because usability testing has to be more rigorous than simply asking students if they thought the system was user-friendly or easy to use. The usability test subcategory included usability tests such as: (1) an assessment on how easily students could find information in the system, (2) an assignment to see whether students could accomplish tasks within the system, or (3) a validated system usability survey (Brooke, 1996). Only 11% of articles (N=10) included a report on a usability test.

Two methods of conducting a usability test were selected that merit further discussion. Santos, Verbert, & Duval (2012) and Santos, Govaerts, Verbert, & Duval (2012) both used the System Usability Scale (SUS) to assess the usability of their system. One of the benefits of using this scale is it has been previously used by hundreds of other research papers evaluating online systems, so it allows systems to be compared on an equal scale. Santos, Boticario, and Perez-Marin (2014) conducted the most rigorous usability assessment and brought in a usability and accessibility expert to evaluate their system. The usability expert interviewed faculty to determine how they were using the system as well as students to see how they were using the system. The expert also evaluated the learning system environment as well as student interactions within the learning environment. With help from the usability expert, the authors were able to (1) enhance the learning management system for "adaptive navigation support", (2) semantically model course recommendations, (3) create recommendations and configure services in the learning space, (4) prepare data collection methods, and (5) assess the learning experience based on the data collected. In future reporting system research, a system usability scale, evaluation expert, or other appropriate methods should be used to improve system evaluation.

6.2.6 Student Use

This category was not discussed in previous literature reviews (Verbert et al., 2013; Verbert et al., 2014; Schwendimann et al., 2016; Yoo et al., 2015), but has important implications for research on the effect of reporting systems on student behavior, achievement, and skills. If students are not using the system, the results about the effects of the system on students are not meaningful. Furthermore, the way in which students use these systems can provide valuable information to guide future research and development of reporting systems. Research on student-facing reporting systems should address this by tracking the frequency and duration of student use as students interact with the visualizations, recommendations, or feedback provided in reporting systems

6.3 Measured Effects

The results presented in this section specifically address our third research question: What is the effect of having access to a student-facing learning analytics reporting system on student behavior, student achievement, and student skills? Out of 93 articles in this analysis, 16% of the articles (N=15) examined the effect of their system on student behavior; 15% of the articles (N=14) examined the effect of their system on student achievement; and 3% of the articles (N=2) examined the effect of their system on student skills. The effects found by these articles are summarized in Table 11, Table 12, and Table 13.

Table 11 summarizes the articles that examined the effect of a reporting system on student behavior. Three articles did not include sample sizes, seven articles had sample sizes less than 75 students, and five articles had sample sizes greater than 75. Randomized control trials and descriptive statistics were the predominant methods used to identify if students' behavior had changed. Results on the effectiveness of these systems is mixed. Future research should consider quasi-experimental methods to provide all students with the reporting tool and still evaluate effectiveness; use larger sample sizes; and continue to examine the effect of reporting systems on student behavior.

TABLE 11
ARTICLE SUMMARIES INCLUDED IN THE MEASURED
EFFECTS CATEGORY FOR BEHAVIOR CHANGE

Citation	Sample Size	Context	Result
(Hsu, 2008)	Not listed	Used association rule mining in an ESL student nursing course	21% of students accepted the system recommendation to view additional content
(Grann & Bushway, 2014)	Not listed	Used logistic regression to predict if use of competency map increased reregistration behaviors (controlling for student engagement)	Competency map behavior accounts for 7% additional variance beyond student engagement variables in predicting reregistration behavior
(Arnold, Hall, Street, Lafayette, & Pistilli, 2012)	about 8,000	Used descriptive statistics and chi-square comparisons between students that take and do not take course signals courses.	Students that participated in course signals courses were more likely to continue taking classes than those that did not enroll in course signals courses
(Xu & Makos, 2015)	73 students	Correlation and ANOVA analyses conducted to determine effect of student activation of a notification tool on collaboration behaviors	Students that enabled notifications (on 2 out of 3 systems) showed increased contributions in the social network space
(Nakahara, Yaegashi, Hisamatsu, & Yamauchi, 2005)	9 treatment, 53 control	Non-parametric mean difference testing was used to determine the effect of i-tree on student online social activity	Students in the i-tree treatment group visited the discussion space more frequently than those that did not have i-tree, however, they did not post more frequently.
(Wise et al., 2014)	9 students	Descriptive statistics were used to examine changes in student behavior over time.	After the analytics reporting was introduced, the percentage of posts viewed increased for all students (except 2 that were already at 100%). Some students had week by week changes based on personal goals, but there were few sustained changes due to the analytics reporting in this study
(Chen, Chang, & Wang, 2008)	27 treatment, 27 control	Descriptive group comparisons as well as t-tests were used to determine what effect having mobile access to learning awareness modules has on student learning behavior	The number of students completing assignments increased and LMS use increased when comparing treatment and control groups
(Lee, 2005)	15 control, 10 experiment	One-way ANOVA was used to determine what group differences exist between treatment and control in terms of web search activity.	Students that used their system, VisSearch, we able to better search the web when compared with students using traditional search engines; this is defined as reading more unique web pages, creating more bookmarks, extending more bookmarks, having an increased length of average search query, and revisiting web pages.
(Beheshitha, Hatala, Gašević, & Joksimović, 2016)	169 students	Hierarchical linear mixed models were used to determine the effect of access to data visualizations on the quantity and quality of discussion board posts taking into account their goal orientation	On two of the three visualizations, students post quantity increased; on the third student post quantity decreased. This was also seen with post quality as measured by discourse features.
(Huang, Huang, Wang, & Hwang, 2009)	57 treatment, 56 control	A Markov chain model and an entropy-based approach were used to see if the recommender system could provide helpful learning paths to students	Almost 50% of students accepted recommendations from the recommender system.
(Vesin, Klašnja-Milićević, Ivanović, & Budimac, 2013)	35 treatment, 35 control	T-test were used for mean difference testing to determine whether Protus, an adaptive and personalized recommendation engine, had an effect on student learning.	Students in the treatment group were able to complete assignments more quickly and were able to complete the entire course more quickly than students in the control group.
(Holanda et al., 2012)	12 students	Descriptive statistics were used before and after initial discussion posting to see what effect recommendations had on posting behavior	There was an 83.3% student interaction increase after recommendations were given.
(Santos et al., 2014)	173 students	T-tests were used to compare treatment and control groups to determine the effect of recommendations on student resource use	Students that received recommendations logged in more frequently, completed their coursework more quickly, completed more questions, and answered more questions correctly on assignments than students in the control group.
(Janssen et al., 2007)	410 treatment, 398 control	T-tests were used to compare how quickly learners were able to finish the course	There were no significant differences between the treatment and control groups in terms of learning efficiency
(Kuosa et al., 2016)	Not listed	Descriptive statistics were used to compare treatment and control groups to determine the effect of their visualizations on student behavior.	Students in the treatment group were able to find correct answers more quickly and were able to more quickly pick keywords to search on.

TABLE 12
ARTICLE SUMMARIES INCLUDED IN THE MEASURED EFFECTS ACHIEVEMENT CHANGE CATEGORY

Citation	Sample Size	Context	Result
(Grann & Bushway, 2014)	Not listed	Mean difference testing was used to determine whether students that used the competency map had higher levels of performance than students that did not	Students that used the competency map had slightly higher achievement rates, however, this was not statistically significant
(Arnold et al., 2012)	about 8,000	By comparing student achievement before and after course signals, descriptive statistics were used to determine the effect on student achievement.	Classes with course signals (compared with the same course before course signals) saw increased A's and B's and decreased C's, D's, and E's.
(Park & Jo, 2015)	36 treatment, 37 control	A randomized control trial research design was used to determine the effect of the LAPA dashboard on student achievement. Mean difference testing was used to determine if there was a significant difference between groups.	Although the treatment group had slightly higher achievement rates than the control group, there were no significant differences between the treatment and control group regarding their achievement rates.
(Kim, Jo, & Park, 2015)	72 treatment, 79 control	A randomized control trial research design was used to determine the effect of the learning dashboard on student achievement. Mean difference testing was used to determine if there was a significant difference between groups.	The treatment group (received access to dashboard) had significant higher achievement rates on the final exam than the control group.
(Denley, 2014)	about 50,000 students	Descriptive statistics were used to compare students in Degree Compass courses to those not enrolled in Degree Compass courses.	Compared with previous students that did not use Degree Compass, students that used Degree Compass received more passing grades (A, B, or C), especially if the student belonged to an at risk population. The prediction algorithm accuracy was 90%.
(Ott et al., 2015)	512 students	T-tests were used to determine if there was a significant achievement difference between previous semesters without the infographic and later semesters with the infographic. Assessments did not change between years and course curriculum stayed the same.	There was no significant difference after the introduction of the class infographic on student achievement.
(Dodge, Whitmer, & Frazee, 2015)	442 treatment, 440 control	T-tests to compare treatment and control groups of a randomized control trial were used to determine the effect of trigger events (recommendation emails) on student achievement.	There was no significant difference between treatment and control groups in terms of achievement. However, in one course there was a significant treatment effect on pell eligible students. This effect was not seen in the other course included in this study.
(Chen et al., 2008)	27 treatment, 27 control	T-tests were used to compare treatment and control groups to determine the effect of the ubiquitous learning website as well as the device used (cell phone, laptop, PDA) on student achievement and learning goal achievement.	Use of the ubiquitous learning website had significant effects on "testing results, task-accomplished rate, and learning-goal-achieved rate" (Chen et al., 2008, p. 90).
(Saul & Wuttke, 2014)	about 80 students	Comparisons were made between students that used the askMe! system and the students that did not use the system.	The average grade of students that used the system was higher than those that did not. In addition, the failure rate was four times lower for those that used the system when compared with those that did not.
(Beheshitha et al., 2016)	about 100 students	Controlling for achievement goal orientation, what effect do learning analytics visualizations have on the quality of student social media posts? A linear mixed-effects analysis was conducted.	The frequency and quality of student posts were affected positively and negatively, depending on the visualization.
(Huang et al., 2009)	57 treatment, 56 control	A Markov chain model and an entropy-based approach were used to see if the recommender system could provide helpful learning paths to students.	Learners in the treatment group performed significantly better than the control group on the evaluation system task.
(Vesin et al., 2013)	35 treatment, 35 control	T-test were used for mean difference testing to determine whether Protus, an adaptive and personalized recommendation engine, had an effect on student learning.	Student learning efficiency was improved, but no analyses were conducted to determine change in grade based on treatment effect.
(Santos et al., 2014)	173 students	T-tests were used to compare treatment and control groups to determine the effect of recommendations on student achievement	There were no significant differences between the treatment and control groups in terms of learning achievement
(Wang, 2008)	40 treatment, 40 control	A t-test was used to determine the effect of content recommendations on student exam score.	The treatment group performed equivalently to the control group on the pre-test, and then the treatment group had significantly higher scores than the control group on the post-test.

Table 12 summarizes the articles that reported on assessing the effect of reporting systems on student achievement. One article did not report sample size, three articles had sample sizes below 75, and ten articles had sample sizes greater than 75. These sample sizes were larger, on average, than those in the behavior change category. In addition, more articles used randomized control trials to determine student achievement differences when compared with behavior differences articles. Despite larger sample sizes and more rigorous methods, the results are mixed. Some studies showed benefits, some studies showed detrimental effects, and some studies showed both benefits and detrimental effects. Future research should use large sample sizes, continue to use randomized control trials or preferably quasi-experimental methods, and continue to examine the effect of reporting systems on student achievement.

TABLE 13
ARTICLE SUMMARIES INCLUDED IN THE MEASURED EFFECTS CHANGE IN SKILLS CATEGORY

Citation	Sample Size	Context	Result
(Kerly, Ellis, & Bull, 2008)	30 students	A randomized control trial and t-test analysis were used to determine whether a chatbot with the CALM system had an effect on student learning	All participants became more aware of their own knowledge, but the treatment group (with chatbot) had a significant increase in self-awareness accuracy above that of the control group.
(Muldner et al., 2015)	209 students	Excitement and interest surveys were used before and after to establish a pre and post baseline. In addition, a one question excitement and interest question was used about every 10 minutes to gauge in the moment student affect. T-tests were used with groups split randomly to determine the effect of the student progress page on student affect.	Female students reported higher interest when they had the choice to use the student progress page, whereas male students reported higher interest when student progress page usage was enforced with a notification.

Table 13 summarizes the two articles that examined the effect of a reporting system on student skills. Both articles found differences in student skills, the first in self-awareness and the second in interest. Due to the lack of research in this area, more research is needed on how reporting systems affect student motivation, interest, self-regulation, awareness, or self-efficacy.

7 LIMITATIONS

One of the major limitations to this analysis is there is not a common vocabulary for learning analytics reporting systems (Van Barneveld, Arnold, & Campbell, 2012). As evidence, articles from educational recommender system literature, intelligent tutoring system literature, educational data mining system literature, and learning analytics dashboard literature were all included in this review. Because there are so many different keywords associated with these systems, there may be articles that were not included in our analysis that should have been. However, to address this limitation, we made our methodology especially rigorous in an effort to include as many relevant articles as possible. For example, we included education and computer science journals, we used various broad keywords in our initial search to catch as many articles as possible, we conducted targeted Google Scholar searches based on keywords we saw from our initial search, and we found related literature reviews to try to include as many relevant articles as possible.

Another limitation to this analysis is we limited our search to research articles, conference proceedings, or book chapters that discussed learning analytics reporting systems. There are undoubtedly many learning analytics systems that have not been researched or written about. These systems are not included in this analysis. However, we feel that the most effective learning analytics reporting systems will be empirically tested for their effectiveness, so we are satisfied with the inclusion criteria for this article.

The final limitation we address is the potential for subjectivity in the coding process because all of the articles in this analysis were coded by human researchers on a number of categories and subcategories. To mitigate this, 20% of the articles were randomly chosen and double coded by a second reviewer. The two reviewers had an 86% agreement.

8 RECOMMENDATIONS FOR PRACTICE AND FUTURE RESEARCH

The results discussed previously have direct implications in practice and for future research. We first discuss considerations for those developing learning analytics reporting systems and then discuss future research topics.

8.1 Considerations for Practice

When starting to create a student-facing learning analytics reporting system, it is important to consider the questions listed below to guide the development process (Table 14). These questions correspond to categories discussed in previous literature reviews and form the outline for the categories discussed in this review. The importance of these questions and categories has been discussed in the results and discussion section.

TABLE 14
QUESTIONS TO GUIDE THE PROCESS OF CREATING A
STUDENT-FACING LEARNING ANALYTICS REPORTING
SYSTEM

Question	Category	% of Articles
What is the intended goal of the system?	Intended Goal	100
What visual techniques will best represent your data?	Visualizations	13
What types of data support your goal?	Information Selection	15
What do students need? Does it align with your goal?	Needs Assessment	6
Is the system easy and intuitive to use?	Usability Test	11
Why use the visual techniques you have chosen?	Visual Design	13
How do students perceive the reporting system?	Student Perceptions	17
What is the effect on student behavior/achievement?	Actual Effects	18
How are students using the system? How often? Why?	Student Use	13

In addition to considering these questions in the development of a learning analytics reporting system, it is also important to include justification for the questions found in Table 14 in the reporting of results. The number of articles that included answers to the questions above was less than 20% in all cases except intended goal of the system, which could be inferred from the article regardless if it was explicitly stated. The creators of these systems were likely thinking about and answering these questions, but the majority failed to report the results in their written work. The field of student-facing learning analytics reporting systems will be greatly improved by addressing and reporting on the questions listed above.

Another important consideration for practice is while many educational technology products have student-facing reporting systems, such as learning management system analytics tools, online homework system dashboards, or cognitive tutor reports, many of these systems do not conduct any research on their system. This means that their system might look well-presented, but that does not mean it has been empirically proven to help students. As instructors or administrators, you should question the claims of these systems unless they have evidence from research to support their claims.

The final consideration for practice deals with student use of reporting tools. From the student use category, 13% of articles reported on tracking student use of their system. In general, the articles reported low student use, around 30% of students access systems on average. However, systems that sent notifications to students through email or text had higher use than static systems students had to visit themselves. As an instructor or administrator, you should consider how to increase student use of these reporting tools. Factors to consider include student familiarity with the system, use of notifications and reminders, student perceptions of usefulness, and effectiveness of the system.

8.2 Recommendations for Future Research

Because student-facing learning analytics reporting systems is an emerging research field, there are many areas of future research. These topics will be addressed corresponding to categories evaluated in this review.

8.2.1 Student Use

The articles that reported on student use, on average, reported low use of their systems. Because of this, more research should be conducted to examine how to increase student use of these tools, specifically in supporting students to act on the feedback they receive in these reporting systems. In addition, more articles should track and report on the way in which students are using reporting systems. Additional research should be conducted to understand student help seeking behavior in online environments in order to support student motivation in engaging with learning analytics reporting systems.

8.2.2 Actual Effects

Based on the low number of articles evaluating the actual effects of these systems on students, more research is needed examining the effect of these systems on student behavior, achievement, or skills. In the actual effects tables (Table 11, 12, & 13), the results are mixed, and therefore not sufficient to make a conclusion about the effect certain types of systems have on student behavior, achievement or skills. In order to add additional rigor to this area of research, (1) larger sample sizes should be used for greater statistical power and the ability to make generalizations beyond the current sample, (2) more detail should be provided (see Table 14) about the reporting system to understand what features are causing the changes to students, and (3) random controlled trials or quasi experimental studies should be used to identify true effects in the place of correlation analyses or simple descriptive statistics com-

parisons. While these results may seem intuitive, many researchers are not using these methods. For example, none of the articles included in this analysis used an observational study to measure impact, such as propensity score matching.

8.2.3 Intended Purpose

It is interesting to note that while the most common purpose of these systems was to increase student awareness and reflection (N=35), only 2% of the articles (N=2) conducted an experiment to determine the effectiveness of the system on student skills (e.g., student reflection, awareness). It is also interesting that the majority of these student-facing systems are not trying to directly increase student retention or improve student engagement. Instead, they focus on student reflection or awareness. In future research, authors should be explicit about the purpose of their system, and should make sure their research questions and analyses align with that purpose.

However, it is possible that many of these systems were used by instructors as a part of mandatory activities in classes with the primary purpose of increasing student awareness and reflection. Student retention would then be less important as higher-level thinking and learning become the focus. If this is the case, researchers and practitioners should consider conducting research on what effect these tools have on student reflection and awareness.

One reason conducting rigorous research on learning technologies in the classroom is a challenge is because it requires a multi-disciplinary effort. Technically savvy team members must come together with researchers and teachers in order to create an appropriate research design, collect the data, analyze the data, and write up the results.

8.2.4 Student Perceptions

Based on the low article count in the student perceptions category, more research is needed to examine student perceptions of these systems and on the perceived effects of these systems on student behavior, student achievement, and student skills. This is important because a perceived effect on student behavior, achievement, or skills could lead to an actual effect on student behavior, achievement, or skills, similar to a pygmalian effect (Rosenthal & Jacobson, 1968). Student perceptions are also important to how students use these systems because as student perceptions improve about a system, they are more likely to use it.

8.2.5 Recommendations

There are two important pieces to a learning analytics reporting system: (1) helping students understand what has happened (through feedback or visualizations) and (2) helping students know what to do because of what they know (through recommendations). To see how many systems are currently doing both of these things, we examined the number of articles that had a recommendations component and a feedback or visualization component. Only 17% of articles (N=16) met these requirements. Future systems should address both what to tell the students to do in recommendations and why students should act on the information in text feedback or visualizations.

8.2.6 Usability

Many of the systems in this review failed to conduct a usability test. This is detrimental to the research field of learning analytics reporting systems because a lack of usability could be the reason why students do not like or use a system. More learning analytics reporting systems need rigorous usability tests conducted, either by administering a standard system usability survey, conducting think-aloud interviews with students, or bringing in a usability expert to evaluate the system. Once a system has been sufficiently evaluated from a usability perspective, additional questions such as what effect systems have on students can then be addressed.

8.2.7 Interactive/Exploration

We hypothesize that interactive or exploratory features in a learning analytics reporting system will lead to increased student use. Only a few articles included an interactive visualization component. Ji, Michel, Lavoue, and George (2014) created an excellent example of an interactive student dashboard called DDART. While they discuss other dashboards that allow dashboard customization, DDART is the first customizable dashboard that does not require computer programming experience in the visualization creation. The authors' dashboard, DDART, allowed students to select parameters, create new indicators, and choose their own visualization method. They provided a graphical interface for students to use to remove the need for computer programming experience. Allowing students to select their own parameters, create their own indicators, and choose their own visualizations may increase student motivation to use the dashboard as they would be more invested in the experience. This level of customization might also increase student awareness or self-reflection because students would have to decide which indicators and visualizations to create. Additional research should examine the effect of various types of dashboard interactivity on student behavior, achievement, and skills.

9 CONCLUSION

This article is a comprehensive literature review on learning analytics reporting systems that track student click-level data and report that data directly to students. In this analysis, we have discussed the types of student-facing learning analytics reporting systems based on system functionality and data sources collected, the methods used to increase the rigor of reporting systems, and the current findings of the effect of these systems on student behavior, achievement, and skills. Future research should focus not only on evaluating the final product of a reporting system, but also on evaluating the design and development process. This process includes administering a needs assessment, providing justification for information selection, justifying the visual design used, and conducting a usability test. More research is also needed with large sample sizes and rigorous experimental methods to examine the perceived and actual effects of learning analytics reporting systems on student behavior, student achievement, and student skills. There were not any articles in this review

that used observational studies. Quasi-experimental methods, such as propensity score matching, should be used in observational studies to allow all students to have access to these systems and still conduct rigorous impact studies. If the goal of a system is to improve student awareness or reflection, the focus of the experimental study should be on student skills, giving a validated pre- and post-survey to determine differences. Student use of reporting systems is not well studied nor understood. Practitioners and researchers should track student use of these systems to understand how to support student motivation to improve the effectiveness of these student-facing learning analytics reporting systems.

ACKNOWLEDGMENTS

REFERENCES

[1] I.E. Allen and J. Seaman, "Grade Change: Tracking Online Education in the United States," 2014; http://www.onlinelearning-survey.com/reports/gradechange.pdf.

[2] K.E. Arnold et al., "Course Signals at Purdue: Using Learning Analytics to Increase Student Success," *Proc. Second Int'l Conf. Learning Analytics and Knowledge* (LAK12), 2012, pp. 2–5.

[3] S.S. Beheshitha et al., "The Role of Achievement Goal Orientations When Studying Effect of Learning Analytics Visualizations," *Proc. Sixth Int'l Conf. Learning Analytics and Knowledge* (LAK16), April 2016; http://doi.org/10.1145/2883851.2883904.

[4] J. Brooke, "SUS—A quick and dirty usability scale," *Usability evaluation in industry*, vol. 189, no. 194, 1996, pp. 4–7.

[5] R. Burke, "Hybrid recommender systems: Survey and experiments," *User modeling and user-adapted interaction,* vol. 12, no. 4, 2002, pp. 331–370.

[6] J.P. Campbell and D.G. Oblinger, "Academic analytics," *Educause Review*, vol. 42, no. 4, July/August 2007, pp. 40–57.

[7] G.D. Chen, C.K. Chang, and C.Y. Wang, "Ubiquitous learning website: Scaffold learners by mobile devices with information-aware techniques," *Computers and Education*, vol. 50 no. 1, 2008; http://doi.org/10.1016/j.compedu.2006.03.004.

[8] D. Clow, "The learning analytics cycle: Closing the loop effectively." *Proc. Second Int'l Conf. Learning Analytics and Knowledge* (LAK12), 2012, pp. 134-138.

[9] T. Denley, "How predictive analytics and choice achitecture can improve student success," *Research & Practice in Assessment*, vol. 9, no. 2, 2014, pp. 61–69.

[10] B. Dodge, J. Whitmer, and J.P. Frazee, "Improving undergraduate student achievement in large blended courses through data-driven interventions," *Proc. Fifth Int'l Conf. Learning Analytics and Knowledge* (LAK15), 2015; http://doi.org/10.1145/2723576.2723657.

[11] H. Drachsler et al., "Panorama of Recommender Systems to Support Learning," *Recommender Systems Handbook*, 2015; http://doi.org/10.1007/978-1-4899-7637-6.

[12] J. Feild, "Improving Student Performance Using Nudge Analytics," *Proc. Eighth Int'l Conf. Educational Data Mining* (EDM15), 2015, pp. 464–467.

[13] R. Ferguson, M. Sharkey, and N. Mirriahi, "The NTU student dashboard: Implementing a whole-iInstitution learning analytics platform to improve student engagement," *Proc. Sixth Int'l Conf. Learning Analytics And Knowledge* (LAK16), 2016, pp. 14–15.

[14] S. Few, *Information Dashboard Design: Displaying data for at-a-glance monitoring*, Analytics Press, 2013.

[15] J. Grann and D. Bushway, "Competency Map: Visualizing Student Learning to Promote Student Success," *Proc. Fourth Int'l Conf. Learning Analytics and Knowledge (LAK14)*, 2014; http://doi.org/10.1145/2567574.2567622.

[16] W. Greller and H. Drachsler, "Translating Learning into Numbers: A Generic Framework for Learning Analytics," *Educational Technology & Society*, vol. 15, no. 3, 2012, pp. 42–57.

[17] J.R. Hartley and D.H. Sleeman, "Towards more intelligent teaching systems," *International Journal of Man-Machine Studies,* vol. 5, no. 2, 1973, pp. 215–236.

[18] O. Holanda et al., "Educational Resources Recommendation System Based on Agents and Semantic Web for Helping Students in a Virtual Learning Environment," *International Journal of Web Based Communities*, vol. 8, no. 3, 2012; http://doi.org/10.1207/s15327752jpa8502.

[19] M.H. Hsu, "A personalized English learning recommender system for ESL students," *Expert Systems with Applications*, vol. 34, no. 1, 2008; http://doi.org/10.1016/j.eswa.2006.10.004.

[20] Y.M. Huang et al., "A Markov-based recommendation model for exploring the transfer of learning on the Web," *Educational Technology and Society*, vol. 12, no. 2, 2009; http://doi.org/10.1207/S15328015TLM1504.

[21] L. Iandoli et al., "Socially augmented argumentation tools: Rationale, design and evaluation of a debate dashboard," *International Journal of Human Computer Studies*, vol. 72, no. 3, 2014; http://doi.org/10.1016/j.ijhcs.2013.08.006.

[22] J. Janssen et al., "Self-organising navigational support in lifelong learning: How predecessors can lead the way," *Computers and Education*, vol. 49, no. 3, 2007; http://doi.org/10.1016/j.compedu.2005.11.022.

[23] M. Ji et al., "DDART, a dynamic dashboard for collection, analysis and visualization of activity and reporting traces," *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, LNCS 8719, 2014; http://doi.org/10.1007/978-3-319-11200-8_39.

[24] A. Kerly, R. Ellis, and S. Bull, "CALMsystem: A Conversational Agent for Learner Modelling," *Knowledge-Based Systems*, vol. 21, no. 3, 2008; http://doi.org/10.1016/j.knosys.2007.11.015.

[25] J. Kim, I.-H. Jo, and Y. Park, "Effect of learning analytics dashboard: analyzing the relations among dashboard utilization, satisfaction, and learning achievement," *Asia Pacific Education Review*, vol. 17, no. 1, 2016, pp. 13–24.

[26] K. Kuosa et al., "Interactive Visualization Tools to Improve Learning and Teaching in Online Learning Environments," *International Journal of Distance Education Technologies (IJDET)*, vol. 14, no. 1, 2016; http://doi.org/10.4018/IJDET.2016010101.

[27] Y.J. Lee, "VisSearch: A collaborative Web searching environment," *Computers and Education*, vol. 44, no. 4, 2005; http://doi.org/10.1016/j.compedu.2004.04.009.

[28] K. Muldner et al., "Exploring the impact of a learning dashboard on student affect," *Artificial Intelligence in Education Seventeenth Int'l Conf.* (AIED15), 2015; http://doi.org/10.1007/978-3-319-19773-9.

[29] J. Nakahara et al., "ITree: Does the mobile phone encourage learners to be more involved in collaborative learning?" *Proc. Int'l Conf. Computer Supported Collaborative Learning* (CSCL2005), 2005, pp. 470–478.

[30] M. Olmos and L. Corrin, "Learning analytics: A case study of the process of design visualizations," *Journal of Asynchronous Learning Networks*, vol. 16, no. 3, 2012, pp. 39–49.

[31] C. Ott et al., "Illustrating performance indicators and course characteristics to support students' self-regulated learning in CS1," *Computer Science Education*, vol. 25, no. 2, 2015; http://doi.org/10.1080/08993408.2015.1033129.

[32] Y. Park and I. Jo, "Development of the Learning Analytics Dashboard to Support Students' Learning Performance," *Journal of Universal Computer Science*, vol. 21, no. 1, 2015, pp. 110–133.

[33] R.M. Ryan and E.L. Deci, "Intrinsic and extrinsic motivations: Classic definitions and new directions," *Contemporary educational*

psychology, vol. 25, no. 1, 2000, pp. 54–67.

[34] C. Romero and S. Ventura, "Educational Data Mining: A Review of the State of the Art," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 6, 2012; http://doi.org/10.1109/TSMCC.2010.2053532.

[35] R. Rosenthal and L. Jacobson, "Pygmalion in the classroom," *The Urban Review*, vol. 3, no. 1, 1968, pp. 16–20.

[36] J.L. Santos et al., "Goal-oriented visualizations of activity tracking: A case study with engineering students," *Proc. Second Int'l Conf. Learning Analytics and Knowledge* (LAK12), 2012; http://doi.org/10.1145/2330601.2330639.

[37] J.L. Santos, K. Verbert, and E. Duval, (2012). "Empowering students to reflect on their activity with StepUp!: Two case studies with engineering students," *CEUR Workshop Proceedings*, vol. 931, 2012, pp. 73–86.

[38] J.L. Santos et al., "Addressing Learner Issues with StepUp!: An Evaluation," *Proc. Third Int'l Conf. Learning Analytics and Knowledge* (LAK13), 2013; http://doi.org/10.1145/2460296.2460301.

[39] O.C. Santos, J.G. Boticario, and D. Perez-Marin, "Extending web-based educational systems with personalised support through User Centred Designed recommendations along the e-learning life cycle," *Science of Computer Programming*, vol. 88, 2014; http://doi.org/10.1016/j.scico.2013.12.004.

[40] C. Saul and H.D. Wuttke, "Turning learners into effective better learners: The use of the askMe! System for learning analytics," *CEUR Workshop Proceedings*, vol. 1181, 2014, pp. 57–60.

[41] B. Schwendimann et al., "Perceiving learning at a glance: A systematic literature review of learning dashboard research," *IEEE Transactions on Learning Technologies*, 2016; doi: 10.1109/TLT.2016.2599522.

[42] G. Siemens, *1st International Conference on Learning Analytics and Knowledge 2011*, July 22, 2010, retrieved March 30, 2016 from https://tekri.athabascau.ca/analytics/.

[43] A. Van Barneveld, K.E. Arnold, and J.P. Campbell, "Analytics in higher education: Establishing a common language," *Educause Learning Initiative*, vol. 1, 2012, pp. 1–11.

[44] K. Verbert et al., "Learning Analytics Dashboard Applications," *American Behavioral Scientist*, 2013; http://doi.org/10.1177/0002764213479363.

[45] K. Verbert et al., "Learning dashboards: An overview and future research opportunities," *Personal and Ubiquitous Computing*, vol. 18, no. 6, 2014; http://doi.org/10.1007/s00779-013-0751-2.

[46] B. Vesin et al., "Applying recommender systems and adaptive hypermedia for e-learning personalization," *Computing and Informatics*, vol. 32, no. 3, 2013, pp. 629–659.

[47] F.-H. Wang, "Content Recommendation Based on Education-Contextualized Browsing Events for Web-based Personalized Learning," *Educational Technology & Society*, vol. 11, no. 4, 2008, pp. 94–112.

[48] A. Wise, Y. Zhao, and S. Hausknecht, "Learning Analytics for Online Discussions: Embedded and Extracted Approaches," *Journal of Learning Analytics*, vol. 1, no. 2, 2014, pp. 48–71.

[49] Z. Xu and A. Makos, "Investigating the impact of a notification system on student behaviors in a discourse-intensive hybrid course," *Proc Fifth Int'l Conf. Learning Analytics and Knowledge* (LAK15), 2015; http://doi.org/10.1145/2723576.2723663.

[50] Y. Yoo et al., "Educational Dashboards for Smart Learning: Review of Case Studies," *Emerging Issues in Smart Learning*, 2015; http://doi.org/10.1007/978-3-662-44188-6.

Robert Bodily Robert Bodily is a graduate student in the Instructional Psychology and Technology program at Brigham Young University (BYU). He received his B.S. in Neuroscience from BYU in 2014. He works as a graduate researcher at BYU and as a data scientist for Lumen Learning. As a graduate student, he has published 8 papers and has 5 more under review. He has also presented 14 conference presentations as a graduate student. His research interests include student-facing learning analytics reporting systems and educational data mining. He is a Society of Learning Analytics Research (SoLAR) member and participates regularly in the Learning Analytics and Knowledge conference.

Katrien Verbert Katrien Verbert obtained a master's degree in Computer Science in 2003 and a doctoral degree in engineering in 2008 at KU Leuven, Belgium. She was a post-doctoral researcher of the Research Foundation – Flanders (FWO) at KU Leuven and held Assistant Professor positions at TU Eindhoven, the Netherlands (2013-2014) and VUB, Belgium (2014-2015). She is currently an Assistant Professor at the HCI research group of KU Leuven. She is an Associate Editor of IEEE Transactions on Learning Technologies. She obtained two post-doc fellowships of the FWO and several European as well as Flemish research grants. She was also involved in the organization of several conferences and workshops (general chair EC-TEL 2017, program co-chair EC-TEL 2016, workshop co-chair EDM 2015, program co-chair LAK 2013 and program co-chair of the RecSysTEL workshop series). She has published over 75 papers. Her research interests include learning analytics, visualization techniques, recommender systems for learning, and digital humanities.