This is an author-prepared version of accepted peer-reviewed work. The definitive Version of Record for this publication can be found in the ACM Digital Library here: http://dl.acm.org/citation.cfm?id=3027403.

DOI: 10.1145/3027385.3027403

This article can be cited as:
Bodily, R., & Verbert, K. (2017, March). Trends and issues in student-facing learning analytics reporting systems research. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 309-318). ACM.

Trends and Issues in Student-Facing Learning Analytics Reporting Systems Research

Robert Bodily
Brigham Young University
bodilyrobert@gmail.com

Katrien Verbert
University of Leuven
katrien.verbert@cs.kuleuven.be

ABSTRACT

We conducted a literature review on systems that track learning analytics data (e.g., resource use, time spent, assessment data, etc.) and provide a report back to students in the form of visualizations, feedback, or recommendations. This review included a rigorous article search process; 945 articles were identified in the initial search. After filtering out articles that did not meet the inclusion criteria, 94 articles were included in the final analysis. Articles were coded on five categories chosen based on previous work done in this area: functionality, data sources, design analysis, perceived effects, and actual effects. The purpose of this review is to identify trends in the current student-facing learning analytics reporting system literature and provide recommendations for learning analytics researchers and practitioners for future work.

CCS Concepts

• **Information systems ~ Decision Support Systems**
• **Human centered computing ~ Visualization**
• *Information systems ~ Data Mining*
• *Information systems ~ Web Mining*

Keywords

Learning analytics; learning analytics dashboards; educational recommender systems; student-facing systems; literature review

1. INTRODUCTION

As online learning continues to grow, it becomes increasingly important to identify design and teaching strategies to improve student success in online and technology mediated environments [1]. Learning analytics (LA) is commonly 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", and could be used to help improve student success in online environments [2]. Within the LA process, there are a number of stages that have been identified: select, capture, predict, use, refine, and report [3]. This article focuses on the reporting stage of the LA process. Learning analytics dashboards, educational recommender systems, intelligent tutoring systems, and automated feedback systems are commonly used in the reporting stage to close the feedback loop and provide information to stakeholders that can be easily understood in a short period of time.

There have been previous literature reviews conducted in this area (see [4], [5], [6], and [7]) which focus on learning analytics dashboards for all stakeholders (e.g. administrators, instructors, students). In order to enable student autonomy and compare student-facing reporting systems across disciplines, we focus exclusively on student-facing systems (collecting student data and reporting the data back to students) that report data back in a learning analytics dashboard, educational recommender system, educational data mining system, intelligent tutoring system, or automated feedback system.

This review has implications in the learning analytics community because student-facing reporting systems close the feedback loop and in best case scenarios, give students real-time access to their data to increase student awareness, reflection, and achievement.

This review identifies research trends and issues related to designing, developing, and evaluating student-facing reporting systems. Based on the analysis from this review, we provide recommendations to (1) aid researchers in conducting more rigorous research in this area, and (2) enable practitioners to increase the impact of their systems on student success.

2. PREVIOUS LITERATURE REVIEWS

This review builds on four literature reviews conducted within the past four years.

Verbert, et al. [2013] selected interesting dashboard articles and provided a framework for coding various types of systems [4]. Their framework included what types of data were tracked, which stakeholder the dashboard was intended for, and whether the system was evaluated or not. This article did not have a systematic search of the literature so it is hard to make comprehensive statements about learning dashboards from this article alone. However, this article is an excellent example as the first review article on learning analytics dashboards.

Verbert, et al. [2014] built on the previous review by including a few additional systems not included in the previous review. The authors also expanded the article categorization framework discussed in Verbert et al. [2013]. The expanded framework included what kind of technology was used to track the data, additional evaluation categories, and the presentation medium (tablet, cell phone, computer, etc). This study was a good follow-up to Verbert et al. [2013], but in order to generalize across learning dashboards, a comprehensive literature review is still needed [5].

Yoo et al. [2015] used the Verbert et al. [2014] review to find articles about learning analytics dashboards that conducted system evaluations. They excluded articles if they did not conduct an evaluation and ended up with 10 articles in their final analysis. The purpose of this article was to find learning analytics dashboard articles that conducted evaluations in order to develop an evaluation framework. Yoo et al. [2015] provided an evaluation framework at the end of their article to guide future

evaluations of dashboard systems. Our literature review categories have included elements from this evaluation framework [6].

Schwendimann, et al. [2016] conducted the first systematic search in the literature for learning analytics dashboard articles. Their final analysis included 53 articles. They analyzed all types of learning analytics dashboards, including administrator-, counselor-, instructor-, and student-facing systems. Some of their findings include that most dashboard systems are developed predominantly for instructors, and mainly exist in higher education [7]. In addition, most articles do not report on research experiments to determine effects on students.

The majority of analytics systems focus on providing teacher- or administrator-facing views in their systems [7]. These can be beneficial in helping teachers or administrators accomplish their goals, however, these approaches generally increase teacher control and decrease student autonomy. Ryan and Deci [2000] suggest that from a self-determination theory perspective, students will have greater intrinsic motivation to succeed in their coursework when they have greater autonomy [8]. Student-facing reporting systems enable, rather than inhibit, student autonomy, and could increase student motivation in ways that teacher or administrator systems could not.

Additionally, many articles use different terminology and are presented in different venues (e.g., automated feedback systems, educational recommender systems, intelligent tutoring systems, or educational data mining systems). However, the goal of these student-facing systems is the same: to provide some kind of feedback to learners to improve teaching and learning. Because each of these systems has a common purpose, we wanted to review all systems trying to accomplish the same goal to better compare the strengths and weaknesses of each type of system.

In order to enable student autonomy and compare student-facing reporting systems across disciplines, we build on the research that has been previously conducted by reviewing student-facing learning analytics reporting systems.

Our review builds on the previous reviews in the following ways: (1) we use the evaluation framework proposed by Yoo et al. [2015] in the creation of the categories for this literature review, (2) we use the categorization frameworks from Verbert et al. [2013] and Verbert et al. [2014] as part of our literature review categories, (3) we build on the work of Schwendimann et al. [2016] by enlarging the search criteria from learning analytics dashboards to all learning analytics reporting systems, and (4) we narrow our scope by focusing exclusively on student-facing reporting systems. Instead of focusing on the tool (learning analytics dashboards) we focus on the stakeholder (students) in order to provide practical suggestions for all student-facing learning analytics reporting systems (see Figure 1).

The research questions that will be addressed in this review include the following:

1. What types of features do student-facing learning analytics reporting systems have?
2. What are the different kinds of data collected in these systems?
3. How are the designs of these systems analyzed and reported on?
4. What are the perceptions of students about these systems?
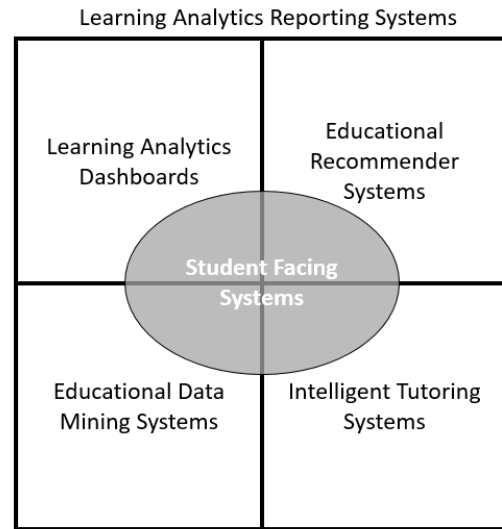5. What is the effect of these systems on student behavior, student skills, and student achievement?

Figure 1. An illustration of the focus of this literature review.

3. METHODS

Learning analytics is a multidisciplinary field situated between education and computer science. Because of this, searches were conducted in both education databases and computer science databases. Specifically, searches were conducted in the following databases or conference proceedings: ERIC (main education database), Learning Analytics and Knowledge conference proceedings (main conference in this research field), Educational Data Mining conference proceedings (main conference in this research field), IEEE Xplore (main computer science database), Computers and Applied Sciences (main computer science database), and ACM database (main computer science database). The exact search keywords can be found in Appendix A.

In order to ensure that we did not miss important articles due to missed keywords, we searched for literature reviews in similar areas (educational data mining, educational recommender systems, learning analytics dashboards) and included articles found in the following literature reviews: Drachsler et al. [2015], Romero and Ventura [2010], Verbert et al. [2013], Verbert et al. [2014], and Schwendimann et al. [2016].

As another check to make sure we did not miss important articles, we took the keywords from the titles of the articles that we had already found (e.g., awareness, dashboard, feedback, etc.) and conducted targeted Google Scholar searches. Articles that met our inclusion criteria that we found in these Google Scholar searches were included in our analysis. These searches can be seen in Appendix A.

After removing duplicates, the initial search yielded 945 articles.

3.1 Inclusion Criteria

For an article to be included in our final analysis, the system described in the article had to (1) collect learning analytics data (e.g., resource use, time spent) and (2) provide a report of this data to students. Because we wanted to focus our review on student-facing *learning analytics* reporting systems, learning analytics data was defined as resource use, time spent data, social media activity, or additional unobtrusive data collected. Notice the system could not simply report assessment data. To be included, the system had to provide some kind of feedback, reporting, recommendation, or visualization directly to students. A student is defined as someone attending a course in a higher education context.

One researcher read through the titles and abstracts of the articles to determine if it was possible to exclude articles after only reading those sections of the articles. If it was certain the article would not be included, it was excluded from the analysis. If the researcher was unsure, the article was skimmed, specifically focusing on the methods and results to understand whether the system described in the article was collecting the right kind of data and reporting it directly to students. After this process of removing articles that did not meet our inclusion criteria, the final analysis included 94 articles.

3.2 Coding Categories

Each of the articles in our final analysis was coded based on the categories *functionality*, *data sources*, *design analysis*, *perceived effects*, and *actual effects*. These categories were synthesized from the categorization and evaluation frameworks identified in previous literature reviews. Each of these categories was then separated into sub-categories which will be discussed in the results and discussion sections below. None of the sub-categories were mutually exclusive, meaning an article could be coded as having included every sub-category of every category. One researcher coded all of the articles. A second researcher double-coded 20% of the articles to ensure a rigorous and consistent article coding process. The two coders achieved an 86% agreement rate.

4. RESULTS AND DISCUSSION

4.1 Functionality

The functionality category describes the features of the system reported on in each article of our analysis. This category includes the following sub-categories: purpose of the system, data mining, visualizations, class comparison, recommendation, feedback, and interactivity.

4.1.1 Purpose of the system

The *purpose of the system* sub-category describes what the authors indicated as the purpose of the system in their article. The purpose of each article was extracted and then coded using an open coding approach to identify common themes across articles. Table 1 summarizes the results of the analysis.

Table 1. Summary of purposes of articles in this review.

Category Name	# of articles	% of articles
Awareness or reflection	35	37
Recommend resources	27	29
Improve retention or engagement	18	19
Increase online social behavior	7	7
Recommend courses	3	3
Other	4	4

Most instructor systems focus on improving engagement or retention by helping instructors identify struggling students so they can intervene (e.g., OLI Dashboard [9], Moodle Dashboard [10], Student Inspector [11]). It is interesting that awareness or reflection is the primary purpose of student-facing systems. Only 19% of the articles in this analysis had the purpose of improving retention or engagement. Why are there not more student-facing systems with the purpose of increasing student engagement or student achievement? Other common purposes for student-facing reporting systems included recommending resources, increasing online social behavior, or recommending courses.

There are various purposes of student-facing reporting systems, as shown in Table 1. These purposes largely depend on the problems the system is trying to solve. We advise researchers and practitioners to be explicit in identifying the purpose of their system so research findings and implications for practice can be generalizable within a concrete problem domain.

4.1.2 Data Mining

Articles were coded in the data mining category if there was some kind of statistical analysis (beyond descriptive statistics) that happened between data collection and data reporting. We acknowledge that not all of the methods in these articles coded with this category can be called data mining, but we use the name data mining for our category name for simplicity. Just under half (N=46) of the articles in this analysis were coded in the data mining category. The other half of the articles used simple text feedback, descriptive statistics reporting, or simple visualizations of what happened.

Data mining as a methodology was more common in the educational recommender system and educational data mining literature while visualizations and dashboards were more common in the learning analytics literature. This is interesting because visualizations/dashboards provide information on what has happened or provides context for why something happened, and data mining or recommender systems provide information on what to do because of what has happened. There were only a few systems (N=16) that included both a visualization and recommendation component. More systems should consider bridging the gap between these fields by including both what has happened as well as what to do because of what has happened.

4.1.3 Visualizations

The visualizations sub-category is defined as any type of visual used to display data. For example, showing a picture of a smiley face if students are doing well and showing a picture of a frowny face if students are doing poorly would count in this category. Another example would be showing a complicated visual dashboard website that students could visit to see their activity compared to the class. Table 2 shows how often common visualization types were used.

Table 2. Common visualization types

Visualization Type	# of Articles
Bar chart	25
Line chart	19
Table	15
Network graph	10
Scatterplot	10
Donut graph	5
Radar chart	4
Pie chart	3
Timeline	3
Word cloud	3
Stoplight	2
Other	21

Most visualizations in dashboard articles were basic visualization types, such as bar charts, line charts, or tables. While these can be

helpful, more research is needed on additional visualization types and how they compare to bar charts, lines charts, and tables. In the visualization type analysis, the *Other* category included visualizations that were only mentioned in one article and provide examples of additional visualizations that merit further research. These include the following: learning path visualization (with squares and arrows), box and whisker plot, tree map, explanatory decision tree, parallel coordinates graph, editable planning and reflection tool, plant metaphor visual, and tree metaphor.

4.1.4 Class Comparison
An article was coded in the class comparison category if it included a system that allowed students to compare their data with another student's data. This could be comparing student grades to the "A" students in the class or comparing students based on social media posting frequency. There were 35 articles that included some type of class comparison functionality. An issue that still needs to be addressed is the effect of a class comparison tool on student motivation. If a student's achievement is above the class average, do they become complacent in their coursework? If a student's achievement is below the average, do they become discouraged in their coursework? This is an issue that has yet to be addressed in the literature and merits additional research.

Another issue that can be addressed in future research is the effect of different class comparisons on students. For example, a student may want to compare their activity to the "A" students in the class, the "B" students in the class, the top 10% of the class, students (anonymized) that are most similar to them in activity, or historical students that are similar to them in activity.

4.1.5 Recommendation
The recommendations category is defined as any article that included a system that provided a recommendation to a student. A recommendation is defined as telling or suggesting the user to do something based on what has happened. Just under half (N=43) of the articles included a recommendations component. The recommendations category was most similar to the data mining category. Out of the 45 articles with a data mining component, 35 of them also had a recommendations component. This shows that the systems using data mining are the ones providing recommendations or suggestions to let the student know what to do based on what has happened.

Many of these recommender systems are not transparent in the recommendations that they provide. In other words, the system does not tell the student why they are receiving a specific recommendation. Additional work should examine the effect between transparent recommendations and more traditional black-box recommendations on student motivation to use the system or follow recommendations. This is important because if students know why they receive a particular recommendation, it could increase their trust in the system along with the likelihood of them following feedback provided by the system.

4.1.6 Feedback
Feedback in this context is defined strictly as text feedback because we have another category for visualization feedback. There were only 17 articles with a feedback component (18%). Text feedback was descriptive in nature, telling students what happened in the past or how they were doing in the course up to that point. Text feedback is used frequently for just-in-time feedback, but is not used as frequently for unit-level or concept-level data reports.

4.1.7 Interactivity
An article was coded in the interactivity section if the reporting system gave the student the opportunity to click around to explore their activity data. Examples of interactivity include providing additional content as links, allowing the user to filter their data based on type of activity or grade, or giving students a simple and advanced view based on their preferences. There were 29 articles that discussed systems that were interactive in some way. An interesting further line of research should investigate whether interactive visualizations or recommendations change student behavior with the reporting system. How do students use the interactive features? Do these interactive features increase student achievement more than systems without those features?

4.2 Data Sources
The data sources category has sub-categories to describe the types of data collected in reporting systems. The sub-categories include resource use, time spent, social interaction data, other sensor data, assessment data, and manually reported data. Resource use is counting the number of times students accessed materials in the course or performed course actions. Time spent data is tracking how long students accessed materials or performed actions in the course. Social interaction data is tracking student use or posts in blog, wiki, discussion board, or messaging systems. Other sensor data was collected from sensors such as face recognition, mouse tracking, or biometric sensors. Manually reported data asked the students to provide answers to surveys or track their own time and input it into the system. The number of articles that tracked each data type is presented below in Table 3.

Table 3. Article counts for each data source

Subcategory Name	# of Articles
Resource use	71
Assessment data	34
Social interaction	33
Time spent	29
Other sensor data	7
Manually reported data	5

The majority of systems in our analysis collected resource use data. Then assessment data, social interaction data, and time spent were all collected about one-third of the time. Future research should investigate what additional information could be useful to include in a student-facing reporting system. Potential data sources include biometric sensor data (heart rate, EEG, skin conductance), mouse tracking, GPS location, university access card swipes, library use, sports facilities use, fit bit tracker steps, social media use not related to school, or internet use not related to school. There are very few systems integrating multiple data sources together into a student-facing reporting system, so research should focus on the impact of adding additional data sources to these systems.

4.3 Design Analysis
The design analysis category describes the effort that went into the reporting system design. What goal is the system trying to achieve? How did the authors identify this goal? How did the authors attempt to achieve the goal? Did they evaluate how well they achieved the goal? To address these questions, the sub-categories of the design analysis category include needs assessment, information selection, visual design, and usability testing.

4.3.1 Needs Assessment

A needs assessment is a common step in many design models. The purpose of a needs assessment is to determine the needs of the stakeholder for which you are designing something. In this case, the stakeholder is the student, and the design is the reporting system. This was not common in the articles we analyzed, and only six articles included a report of a needs assessment.

The solutions to the student needs identified by these six articles included alerting a student if something goes wrong in the course, showing students how they use their time, facilitating group communication for group projects, supporting student motivation in engaging with their course, providing relevant learning material when it is needed, showing students what is important to study, and increasing awareness of tools and resources available to students.

Needs assessments are critical to make sure a designed system is fulfilling stakeholder goals [12]. More researchers and practitioners should adopt this approach when designing a student-facing reporting system.

4.3.2 Information Selection

Select is one of the stages in the learning analytics process [3], however it is not discussed very much in the learning analytics reporting system literature. Information selection is defined as including justification for why data was included in system reports. There were three good examples of articles that had justification for the information selection stage. Ott et al. [2015] conducted a literature review to provide justification for the variables included in their reporting system [13]. Feild [2015] used exploratory analysis to identify which variables to include in their reporting system [14]. Iandoli et al. [2014] used a theoretical framework to guide their information selection process [15]. The majority of the articles in our analysis did not include any justification for the data included in their reporting system. It seems most research is including the data that is easily accessible and not many people are going out of their way to include additional data sources. A justification for why data was included in reporting systems is key for other researchers and practitioners in the early stages of designing a student-facing reporting system to help them determine what data sources they will include in their system.

Future research should also investigate the benefits or drawbacks to using certain kinds of data in a reporting system. Do students respond better to certain kinds of data over other kinds? If so, why? The collection of certain types of data requires expensive system architecture. Are there data sources that are not useful to students and are not worth collecting?

4.3.3 Visual Design

An article was coded in the visual design category if the article had justification for why the data was visualized or reported in the way it was reported. There were 12 articles that included a visual design component. Most of the authors of reporting system articles have likely considered why they are visualizing or reporting data in the ways they have chosen, however, many of them are not reporting it in their reporting system articles. We advise learning analytics researchers to include justifications of design choices in the reporting of their work, as they are key to guide the selection of good visualizations.

Additional research should try to identify the affordances and constraints of each type of visualization to better illustrate when certain types of visualizations should be used and when certain visualizations should be avoided.

4.3.4 Usability Testing

An article is included in usability testing if the authors conducted and reported on a usability test of their reporting system. This usability test is more in depth than simply asking students if the system was user-friendly. If the system only asked about student perceptions of the system, it will be included in the student perceptions usability category, discussed below. There were 10 articles that included some sort of usability test. A few examples of effective usability tests that were conducted include (1) asking students to answer questions about a demo view of the system to see if they can navigate and understand the system, (2) conducting a think-aloud-protocol with the students to understand how students are thinking about the system as they interact with it, (3) using the validated System Usability Scale (SUS) [16] to get one number describing the usability of the system (that can be compared to other systems using that scale), or (4) bringing in a system usability expert to professionally evaluate the usability of the reporting system.

There were more articles that conducted a randomized control trial (RCT) to determine the effect of the reporting system on student achievement (discussed in 4.5 Actual Effects) than articles that conducted a usability test on their system. This is problematic because it is hard to trust the results of an RCT if the authors did not control for system usability by making sure it would not affect students as they interacted with the reporting system. As research in this field becomes more mature, authors should be sure to include usability test reports on their system so we can make generalizable conclusions about RCT results.

4.4 Student Perceptions

The student perceptions category included sub-categories for how students perceived the learning analytics reporting system. Sub-categories include usability, usefulness, and perceptions on the effect the system had on the student.

4.4.1 Student Perceptions of Usability

An article was included in the student perceptions of usability category if the authors asked the students about the usability of the system and reported it in their article. This section is different from section 4.3.4 in that this section deals with student perceptions of usability while section 4.3.4 deals with other forms of system usability tests. There were 32 articles that were coded in the student perceptions of usability category. There were three times as many articles that asked students about the usability of the system instead of conducting a more rigorous usability test. System usability can affect how students perceive and use a reporting system, so in order to better understand how students use these systems, more rigorous usability tests should be conducted. In future research, authors should consider conducting a more rigorous usability test instead of simply asking students if their system was easy to use.

4.4.2 Usefulness

The sub-category usefulness is defined as asking students if they thought the system was useful or if students were satisfied with the system. There were 34 articles coded in this sub-category, which is about the same as the usability perception category. There were 25 articles that included both perceived usability and perceived usefulness, so these questions were usually asked together. Student perceptions of usefulness were generally positive regardless of the system they were using, however this information is not very helpful in helping us understand anything about student use with the system or the effect the system has on students. Instead of asking about system usefulness, we

recommend to consider asking about perceived effect on student behavior, student skills, or student achievement.

4.4.3 Student Perceptions of System Effects
This category is concerned with identifying articles that included a discussion of perceived system effects on student behavior or student achievement. There were 16 articles that asked about perceived behavior change, 2 articles that asked about perceived achievement change, and 15 articles that asked about perceived student skills. Student skills are defined as metacognitive strategies or self-regulated learning strategies. Because of the low number of articles in this sub-category, future research should ask students what effect they believe the reporting system had on them and what feature of the reporting system led to that effect.

Research should also consider why there are so many more perceived behavior and skills change articles when compared with perceived achievement. Is it difficult to trust student perceptions of achievement changes? Are RCTs preferred to investigate the effect of a reporting system on student achievement more than for student behavior or skills? Additional research topics related to these questions are discussed in the actual effects section.

4.5 Actual Effects
The actual effects category is different than the student perceptions of system effects category because the actual effects had to include some sort of research experiment to try to determine the effect of the reporting system on student behavior, skills, or achievement. There were 15 articles that looked at student behavior, 14 articles that looked at student achievement, and 2 articles that looked at student skills. The articles reviewing student achievement have been summarized and are included in Appendix B. The articles coded in the actual effects category, on average, used small sample sizes, descriptive statistics, and did not have very many significant results. There were a few articles with large sample-sizes that conducted randomized control-trials [17, 18], but these were rare. More research is needed on the actual effects of these reporting systems on student behavior, student achievement and student skills.

Very few articles are using RCTs to investigate student skill change in this context. This may be because it is easier to ask students about their perceived awareness, motivation, or self-regulation change than to find a validated scale to use in a pre- and post-survey research design. Future research should not only use RCTs to investigate student achievement changes, but should also give validated scales to students before and after the class to see if student skills are changing as well.

Most of the methodologies examining experimental effects used RCTs or descriptive statistics. Because reporting systems are helpful to students and we want all of our students to have access to these systems, it is hard to argue that RCTs are the best research methodology because not all students will have access to the tool. Quasi-experimental methods should be considered in the future to give all students access to the reporting systems during the course instead of a random selection of students.

5. LIMITATIONS
One challenge we faced in conducting this literature review was the lack of a common vocabulary across fields talking about learning analytics reporting systems. For example, an intelligent tutoring system that tracks resource use is similar to an educational recommender system that also tracks resource use and provides recommendations in real-time to the student. However, these systems are from two different but related research fields. This also applies to learning analytics dashboards and automated feedback systems. Then, there were also systems that did not use any of these keywords in their manuscripts and called their system a visualization system or gave their system a specific name (e.g., ECoach, StepUp!, itree, etc). To mitigate this issue, we added in Google Scholar searches based on popular title words and we conducted related literature review searches to find articles we might not have found using keyword search criteria.

Another limitation we faced in conducting this literature review is the potential for bias in the article coding process because we used human coders. To address this issue, two coders coded 20% of the articles and the codes were compared to find their percent agreement. The coders had an 86% agreement.

This review presents the state of the art in the student-facing learning analytics reporting literature, however because we restricted our search to research articles and conference presentations, many systems that are not reported on in research will not be included in our review. We feel justified in only selecting conference presentations and research articles because, in general, the best student-facing reporting systems will have research conducted on them to determine their efficacy.

6. RECOMMENDATIONS FOR PRACTICE
The recommendations for practice included in this section are based on the extensive analysis of articles included in this review. For practitioners and researchers thinking about or starting to implement a student-facing learning analytics reporting system, use the questions in Table 4 as guiding points to make sure you are addressing the items needed to make a tool that will benefit students the most.

Table 4. Questions to guide in implementing reporting systems

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 this need align with your goal?	Needs Assessment	6
Is the system easy and intuitive to use?	Usability Test	11
Why are you using the visual techniques or recommendations 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

Most of the articles did not report on the categories in Table 4, however, the authors of these articles were probably thinking about these questions informally. It is important to document the answers to these questions in final research manuscripts and conference presentations to increase the rigor of the learning analytics reporting systems field. Eventually, there are going to be enough articles published on the effects of these systems on student achievement and behavior to start to make inferences about the types of design, data, visualization, or functionality that best help students succeed. However, these generalizations cannot

be made if the research articles written about these systems were not explicit in addressing the questions in Table 4.

If you are an administrator or instructor thinking about adopting an educational technology system that includes learning analytics tracking and student-facing reporting systems, you should consider the questions in Table 4 during the selection process. The creators of many systems have not conducted rigorous research on their student-facing systems, so they may over-promise on the results of these systems.

7. CONCLUSION

Student-facing learning analytics reporting systems is an emerging area of research and practice. In this review, we conducted a systematic search of the literature in education databases, computer science databases, Google Scholar, and related literature review articles. We only included articles that tracked student learning analytics data and then reported that data directly back to students. Our final analysis consisted of 94 articles. We coded the articles using a closed coding approach into categories synthesized from Verbert et al. [2013], Verbert et al. [2014], and Yoo et al. [2015]. The categories were functionality, data sources, design analysis, perceived effects, and actual effects.

There were two types of systems that emerged from this analysis. First, there were systems that used data mining to analyze the data and then provided recommendations to students. Second, there were systems that used descriptive statistics and then provided data visualizations in the form of a dashboard for students. Only a few systems conducted a data mining analysis and then used visualization to report the results. This may be because of the differences between the educational recommender and learning analytics dashboard fields. Similarly, intelligent tutoring systems and automated feedback systems use their own methods to try to achieve a similar purpose. Researchers and practitioners should consider interdisciplinary efforts across these fields to bring expertise together in order to accomplish their goals.

This systematic literature review on student-facing learning analytics reporting systems was the first to examine student use of reporting systems across multiple articles. Student use is important in experimental research because the way students use a tool will determine the effect it has on them. None of the studies included in the student use category broke down student use by demographic, learner characteristics, or student achievement levels. In order to better personalize recommendations and dashboards to students, we need to put more emphasis on understanding student use of these systems.

The previous literature reviews identified in this review ([3], [4], [5], & [6]) provided categorization frameworks for dashboards after they had already been designed and developed. However, there is a lot of work that goes into designing and developing a dashboard that is rarely discussed in the literature. This review has enumerated a number of practices to increase the rigor of designing and developing a student-facing reporting system: needs assessment, information selection, visual design, and usability testing. A needs assessment ensures that the system being developed will actually accomplish the goal, the information selection process determines the information needed to accomplish the goal, the visual design stage establishes how the information will be provided to students, whether in a dashboard, feedback system, recommender system, or text feedback, and the usability test phase assesses the user-friendliness and usefulness of the system. These practices will greatly enhance the rigor of the design and development process in student-facing learning analytics reporting systems research.

9. REFERENCES

[1] Allen, I. E., and Seaman, J. 2014. *Tracking Online Education in the United States*, 1–45. Retrieved from http://www.onlinelearningsurvey.com/reports/gradechange.pdf

[2] Siemens, G. 2010. In *Proceedings of 1st International Conference on Learning Analytics and Knowledge 2011*. Retrieved March 30, 2016 from https://tekri.athabascau.ca/analytics/

[3] Elias, T. 2011. Learning Analytics: The Definitions, the Processes, and the Potential. DOI=10.1.1.456.7092.

[4] Verbert, K., Duval, E., Klerkx, J., Govaerts, S., and Santos, J. L. 2013. Learning Analytics Dashboard Applications. *American Behavioral Scientist*, 1–10. DOI=http://doi.org/10.1177/0002764213479363

[5] Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., and Klerkx, J. 2014. Learning dashboards: An overview and future research opportunities. *Personal and Ubiquitous Computing*, *18*(6), 1499–1514. DOI=http://doi.org/10.1007/s00779-013-0751-2

[6] Yoo, Y., Lee, H., Jo, I., & Park, Y. 2015. Educational Dashboards for Smart Learning: Review of Case Studies. *Emerging Issues in Smart Learning*, 145–155. DOI=http://doi.org/10.1007/978-3-662-44188-6

[7] Schwendimann, B. A., Boroujeni, M. S., Holzer, A., Gillet, D., and Dillenbourg, P. 2016. Understanding learning at a glance : An overview of learning dashboard studies. In *Proceedings of 6th International Conference on Learning Analytics and Knowledge.* (pp. 3–4).

[8] Ryan, R. M., and Deci, E. L. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist, 55*(1), 68.

[9] Dollar A., and Steif P. S. 2012. Web-based statics course with learning dashboard for instructors. In: *Uskov V (ed) Proceedings of computers and advanced technology in education* (June 25–27, 2012, Napoli, Italy). CATE 2012.

[10] Podgorelec V., and Kuhar S. (2011) Taking advantage of education data: advanced data analysis and reporting in virtual learning environments. *Electron Electr Eng 114*(8):111–116.

[11] Scheuer O., and Zinn C. 2007. How did the e-learning session go? The student inspector. In: *Luckin R et al. (eds) Proceedings of the 2007 conference on artificial intelligence in education: building technology rich learning contexts that work.* IOS Press, Amster- dam, pp 487–494.

[12] Altschuld, J. W., and Kumar, D. D. 2009. *Needs assessment: An overview*. Sage Publications.

[13] Ott, C., Robins, A., Haden, P., and Shephard, K. 2015. Illustrating performance indicators and course characteristics to support students' self-regulated learning in CS1. *Computer Science Education*, *25*(2), 174–198. DOI=http://doi.org/10.1080/08993408.2015.1033129

[14] Feild, J. (2015). Improving Student Performance Using Nudge Analytics. In *Proceeding of the 8th International Conference on Educational Data Mining* (pp. 464–467).

[15] Iandoli, L., Quinto, I., De Liddo, A., and Buckingham Shum, S. (2014). Socially augmented argumentation tools: Rationale, design and evaluation of a debate dashboard. *International Journal of Human Computer Studies*, *72*(3), 298–319. DOI=http://doi.org/10.1016/j.ijhcs.2013.08.006

[16] Brooke, J. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4-7.

[17] Dodge, B., Whitmer, J., and Frazee, J. P. 2015. Improving undergraduate student achievement in large blended courses through data-driven interventions. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15* (pp. 412–413). DOI=http://doi.org/10.1145/2723576.2723657.

[18] Janssen, J., Tattersall, C., Waterink, W., van den Berg, B., van Es, R., Bolman, C., and Koper, R. 2007. Self-organising navigational support in lifelong learning: How predecessors can lead the way. *Computers and Education*, *49*(3), 781–793. DOI=http://doi.org/10.1016/j.compedu.2005.11.022

[19] Drachsler, H., Verbert, K., Santos, O. C., and Manouselis, N. (2015). Panorama of Recommender Systems to Support Learning. In *Recommender Systems Handbook* (pp. 421–451). DOI=http://doi.org/10.1007/978-1-4899-7637-6.

[20] Romero, C., and Ventura, S. 2010. Educational Data Mining: A Review of the State of the Art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40*(6), 601–618. DOI=http://doi.org/10.1109/TSMCC.2010.2053532

[21] J. Grann and D. Bushway, Competency Map: Visualizing Student Learning to Promote Student Success, in *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge - LAK '14*, 2014, pp. 168–172.

[22] K. E. Arnold, Y. Hall, S. G. Street, W. Lafayette, and M. D. Pistilli. 2012. Course Signals at Purdue: Using Learning Analytics to Increase Student Success," in *LAK '12, 2012*, no. May, pp. 2–5.

[23] Y. Park and I. Jo. 2015. Development of the Learning Analytics Dashboard to Support Students' Learning Performance. *J. Univers. Comput. Sci.*, vol. 21, no. 1, pp. 110–133, 2015.

[24] J. Kim, I.-H. Jo, and Y. Park. 2015. Effect of learning analytics dashboard: analyzing the relations among dashboard utilization, satisfaction, and learning achievement. *Asia Pacific Educ. Rev.*, 2015.

[25] T. Denley. 2014. How predictive analytics and choice achitecture can improve student success. *Res. Pract. Assess., vol. 9*, no. 2, pp. 61–69, 2014.

[26] C. Ott, A. Robins, P. Haden, and K. Shephard. 2015. Illustrating performance indicators and course characteristics to support students' self-regulated learning in CS1. *Comput. Sci. Educ., 25*(2) pp. 174–198, 2015.

[27] B. Dodge, J. Whitmer, and J. P. Frazee. 2015. Improving undergraduate student achievement in large blended courses through data-driven interventions *in Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15*, 2015, pp. 412–413.

[28] G. D. Chen, C. K. Chang, and C. Y. Wang. 2008. Ubiquitous learning website: Scaffold learners by mobile devices with information-aware techniques. *Comput. Educ., 50*(1), pp. 77–90, 2008.

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

[30] S. S. Beheshitha, M. Hatala, D. Gašević, and S. Joksimović. 2016. The Role of Achievement Goal Orientations When Studying Effect of Learning Analytics Visualizations. *Learn. Anal. Knowl.*, no. April, 2016.

[31] Y. M. Huang, T. C. Huang, K. Te Wang, and W. Y. Hwang. 2009. A Markov-based recommendation model for exploring the transfer of learning on the Web. *Educ. Technol. Soc.,* 12(2), pp. 144–162, 2009.

[32] B. Vesin, A. Klašnja-Milićević, M. Ivanović, and Z. Budimac. 2013. Applying recommender systems and adaptive hypermedia for e-learning personalization. *Comput. Informatics, 32*(3), pp. 629–659.

[33] O. C. Santos, J. G. Boticario, and D. Perez-Marin. 2014. Extending web-based educational systems with personalised support through User Centred Designed recommendations along the e-learning life cycle. *Sci. Comput. Program*., *88*, pp. 92–109.

[34] F.-H. Wang. 2008. Content Recommendation Based on Education-Contextualized Browsing Events for Web-based Personalized Learning. *Educ. Technol. Soc. 11*(4), pp. 94–112.

10. APPENDIX A

Table 5. The search criteria used in this literature review.

Source	Search Term or Topic	Article 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
[19]	Literature review on educational recommender systems	37
[20]	Literature review on educational data mining	30
[4], [5], and [7]	Literature reviews on learning analytics dashboards	20

11. Appendix B.

Table 6. Article summaries included in the experimental effects achievement change category.

Citation	Sample Size	Context	Result
[21]	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
[22]	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.
[23]	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.
[24]	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	The treatment group (received access to dashboard) had significant higher achievement rates on the final exam than the control group.

		was a significant difference between groups.	
[25]	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%.
[26]	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.
[27]	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.
[28]	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).
[29]	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.
[30]	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.
[31]	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.
[32]	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.
[33]	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
[34]	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.