Development of the Learning Analytics Dashboard to Support Students’ Learning Performance

Yeonjeong Park
(Ewha Womans University, Seoul, South Korea
ypark78@ewha.ac.kr)

Il-Hyun Jo
(Ewha Womans University, Seoul, South Korea
ijo@ewha.ac.kr)

Abstract: The Learning Analytics Dashboard (LAD) is an application to show students’ online behavior patterns in a virtual learning environment. This supporting tool works by tracking students’ log-files, mining massive amounts of data to find meaning, and visualizing the results so they can be comprehended at a glance. This paper reviews previously developed applications to analyze their features. Based on the implications from the review of previous studies as well as a preliminary investigation on the need for such tools, an early version of the LAD was designed and developed. Also, in order to improve the LAD, a usability test incorporating a stimulus recall interview was conducted with 38 college students in two blended learning classes. Evaluation of this tool was performed in an experimental research setting with a control group and additional surveys were conducted asking students’ about perceived usefulness, conformity, level of understanding of graphs, and their behavioral changes. The results indicated that this newly developed learning analytics tool did not significantly impact on their learning achievement. However, lessons learned from the usability and pilot tests support that visualized information impacts on students’ understanding level; and the overall satisfaction with dashboard plays as a covariant that impacts on both the degree of understanding and students’ perceived change of behavior. Taking in the results of the tests and students’ open-ended responses, a scaffolding strategy to help them understand the meaning of the information displayed was included in each sub section of the dashboard. Finally, this paper discusses future directions in regard to improving LAD so that it better supports students’ learning performance, which might be helpful for those who develop learning analytics applications for students.

Keywords: Learning analytics, Dashboard, Visualization, Learning Management System, Usability Test, Pilot Test, Perceived Usefulness

Categories: L.2.1, L.3.5, L.3.6

1 Introduction

Learning Analytics is an emerging area that explores the measurement, collection, analysis and reporting of data which is associated with students’ learning and their environment [Brown, 2011; Chatti, Dyckhoff, Schroeder, & Thüs, 2012]. Among the diverse approaches to such learning analytics, the use of log-data in Learning Management Systems (LMS) is one of the most popular research orientations due to its ubiquity in many educational institutions.
It has been well recognized that a LMS is a core of the cyber campus (a.k.a. virtual campus or virtual learning environment) where instructors provide various kinds of learning materials and students can easily access them; instructors and students communicate both synchronously and asynchronously; and students conduct their team projects and discuss each other. Above all, since the large amount of students’ behavioral data is left and accumulated as log files in the LMS, researchers in diverse areas have collaborated to extract them, conduct data-mining, and utilize them to improve students’ learning performance.

Recent studies, using such log-files, focused on predicting drop-outs or academic successes [Bayer, Bydzovská, Géryk, Obšťavc, & Popelinský, 2012; Lauria & Baron, 2011; Romero, Espejo, Zafra, Romero, & Ventura, 2013]. Further, remarkable applications were implemented beyond the prediction of learning achievement. For example, Course Signal in Purdue University provided real-time academic status in order for students to be informed of their current scores as well as give alarm signals. It has been shown that these signals helped prevent students dropping-out and enhances their success [Pistilli & Arnold, 2010]. Another example that presents the power of visualization is the SNAPP tool, which incorporates the technology of social network analysis in learning analytics [Bakharia & Dawson, 2011]. This tool is an open-source application that displays students’ interactions in LMS that are visualized in the format of sociogram that can be viewed in a web-browser. The networking information identifies central or isolated students and patterns of discussion activities effectively [Bakharia & Dawson, 2011; Dawson, Bakharia, & Heathcote, 2010].

Including the two examples above, Verbert, Duval, Klerks, Govaerts, and Santos [2013] compared fifteen dashboard applications. In their brief introduction, LAD applications were reviewed and compared in terms of the target users (teachers vs. students), tracked data (time spent, social interaction, document and tool use, artifacts produced, exercise results, and quizzes), and evaluation focus. Dashboards developed for diverse purposes supporting teachers, students or both throughout a synergetic approach that combines design principles and technologies including data-extraction, data-mining and visualization. However, as Verbert et al [2013] indicated, the previous studies have limitations in assessing users’ behavior changes by using such a dashboard treatment and also in proving the overall effectiveness of the tool since it requires consistent and iterative refinements as well as longitudinal studies.

The purpose of this study is to present the process and outcomes of developing a LAD that supports students’ learning performance, from the review of previous case studies to the evaluation of newly developed LAD. Multiple data-collections and analysis methods were employed with the interviews in need assessment and usability tests, and surveys and experimental approach in a real-world setting. Implications from each stage influenced the refinement of the initial version of LAD.

2 Learning Analytics Dashboards (LADs)

2.1 Meaning

Originally, a “dashboard” is a board or panel placed on the front of a horse-drawn carriage to protect mud or dirt from being splashed into the interior. The word has evolved to mean a control panel in front of the driver in automobiles, showing
information to help driving [Wikipedia, n.d.]. In the business community, a dashboard was recognized as an emerging performance management system, for example, to monitor productivity, analyze cost-effectiveness and improve customer satisfaction [Eckerson, 2010]. Key performance indicators (KPIs) are displayed at a glance in a dashboard such that decision-makers can receive alerts as to whether the performance has deviated from predefined targets [Podgorelec & Kuhar, 2011].

For the past decades, with the exponential growth in data volume and the applications of big data, enterprise dashboards, similar to a dashboard in an aircraft which is even more complicated than that of automobiles and requires a large number of indicators, came to be SMART (synergetic, monitor KPIs, accurate, responsive, and timely) and IMPACT (interactive, more data history, personalized, analytical, collaborative, and traceability) [Malik, 2005]. Moreover, due to the influence of information technology and digital devices, dashboards received more attentions from both the professional world and in our daily lives. Finally, the dashboard was defined as a “visual display of the most important information needed to achieve one or more objectives that has been consolidated on a single computer screen so it can be monitored at a glance” [Few, 2013, p. 26].

After applying the aforementioned meanings of a dashboard to LAD, the result was an interactive, historical, personalized, and analytical monitoring display that reflects students’ learning patterns, status, performance, and interactions. The outlook of LAD includes visual elements such as charts, graphs, indicators and alert mechanisms [Podgorelec & Kuhar, 2011]. While visualizing information is an important consideration externally, a series of data-mining processes are at the heart of the system, these find the unknown and implicit information from the patterns of relationships in large, noisy and messy datasets [Nisbet, Elder, & Miner, 2009]. Ultimately these external visualization and internal mechanisms of data-mining help amplify human cognition and support learning performance [Card, Mackinlay, & Shneiderman, 1999].

### 2.2 Previous Studies

LADs have been developed by a number of researchers. We reviewed related studies published or presented in different journals and conferences between 2005 and 2013. To select sample case studies, we filtered what Verbert, et al. [2013] reviewed through our criteria for analysis. The final nine dashboards were thoroughly reviewed and compared against the groups of criteria to get an overview of the important features of state-of-the-art LADs. These criteria include 1) intended goals and target users, 2) data-extraction and mining, 3) visualization, and 4) evaluation.

#### 2.2.1 Intended Goals and Target Users

The intended goal implies what is changed throughout the visualization tool. That is, it is assumed that a visualization tool is only meaningful when it causes the intended behavior. For example Duval [2011] states, “visualization of eating habits can help to lead a healthier life … a visualization of mobility patterns can help to explore alternative modes of transport” (p.12). Similarly, visualized information in LAD tries to influence the users’ psychologies and actions to drive effective teaching and learning.
As shown in Table 1, previously developed dashboards can be generally divided into three types: dashboards for 1) teachers only, 2) both teacher and students, and 3) students only. The target users determine the distinct intended goals. Generally a dashboard for teachers informs them of a student’s learning status, in real-time, in a scalable way. The system aims to monitor multiple students’ learning progress and helps teachers perform their roles effectively in areas including class management, learning facilitation, provision of feedbacks, and evaluation and grading. For example, LOCO-Analysis [Ali, Hatala, Gašević, & Jovanović, 2012] focused on feedback generations; Students Success System [Essa & Ayad, 2012] identified at-risk students so that teachers could provide appropriate treatment to them. SNAPP [Bakharia & Dawson, 2011] is a powerful tool to visualize participants’ relationships in a discussion forum and to identify the key or isolated discussers.

<table>
<thead>
<tr>
<th>Target users</th>
<th>Names of Tools</th>
<th>Intended Goals</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teachers</td>
<td>LOCO-Analyst</td>
<td>• to provide feedbacks on students’ learning activities and performance</td>
<td>• Formative (2 iterations)</td>
</tr>
<tr>
<td></td>
<td>Student Success System</td>
<td>• to identify and treat at-risk students</td>
<td>• NA</td>
</tr>
<tr>
<td></td>
<td>SNAPP</td>
<td>• to visualize the evolution of participant relationships within discussion forums</td>
<td>• NA</td>
</tr>
<tr>
<td>Teachers and Students</td>
<td>Student Inspector</td>
<td>• to keep track of learners’ interaction in e-learning systems</td>
<td>• Summative</td>
</tr>
<tr>
<td></td>
<td>SAM</td>
<td>• to enable students’ self-reflection and awareness of what and how they are doing</td>
<td>• Formative</td>
</tr>
<tr>
<td></td>
<td>StepUp!</td>
<td>• to promote reflection and awareness of their activity</td>
<td>• Summative</td>
</tr>
<tr>
<td>Students</td>
<td>Course Signal</td>
<td>• to improve retention and performance outcomes</td>
<td>• Summative</td>
</tr>
<tr>
<td></td>
<td>Narcissus</td>
<td>• to help students see how well they are contributing to the group</td>
<td>• Summative</td>
</tr>
</tbody>
</table>

Table 1: The lists of Learning Analytics Dashboards
On the other hand, a dashboard for students presents learning patterns to students themselves, helping them modify their learning strategies and motivating learning. The system aims at stimulating students’ psychological changes through the improvement of self-knowledge [Verbert et al., 2013] and social awareness [Lambropoulos, Faulkner, & Culwin, 2012]. SAM [Govaerts, Verbert, Duval, & Pardo, 2012] and StepUp! [Santos, Verbert, Govaerts, & Duval, 2012] indicated that the major purpose of LAD is students’ self-reflection, awareness, and self-assessment. Narcissus [Upton & Kay, 2009] and SNAPP [Bakharia & Dawson, 2011] supported their group work and collaboration. Also, it can be a means to provide real-time feedback in regard to their learning performance like in Course Signal. The ultimate goal of LADs is to motivate students’ learning and, consequently, to improve their retention and performance outcomes.

2.2.2 Data-extraction and Mining

One of the most important and challenging processes in developing a LAD is to collect and extract the accumulated data from the system and to mine them in a meaningful way. There are several approaches to data-extraction and mining. Figure 1 shows a framework to categorize the different approaches and roughly positions the case studies into four types. The Y-axis represents the methods of data-tracking where two distinct approaches are revealed. While some cases such as Student Inspector, SAM, and GLASS used a LMS such as Moodle, WebCT, and Blackboard to track students’ activities automatically, another group of studies either developed a tracking tool such as SNAPP [Dawson et al., 2010] or utilized tracking software in the case of StepUp! [Santos et al. 2012] and collaboration tools in Narccissus [Upton & Kay, 2009]. Santos et al. [2012] explained that the former is more secure and scalable way and the latter is accomplished by the concept of “quantified self” where users themselves carry sensors, apps on mobile devices, or specific devices. In addition, Web 2.0 tools help track the students’ log data in social media such as blogs, Twitter, and wikis.

The X-axis indicates the different approaches to data-mining. In reviewing tracked data in previous studies, we found that one group of case studies focused on the representation of raw data such as login information (e.g., time spent-on, login frequency), performance results (e.g., exercise, quiz, assignment grades), content usage (e.g., download of documents, learning materials), and indication of popular contents (e.g., documents, postings). The characteristic of this group is rather descriptive. The counting part on the X-axis is rather prescriptive and includes exemplar studies such as Course Signal [Arnold & Pistilli, 2012] or Student Success System (S3) [Essa & Ayad, 2012]. While the representation of raw data helps users detect the patterns easily, application of sophisticated data mining algorithms allowed them to make decisions effectively [Ali et al., 2012]. Both CS and S3 utilized a prediction algorithm and pursued an early warning system so that at-risk students get an alarm and proper feedback. CS combined the data of students’ assignments grade, attendance behavior, and past academic performance in a course management system (CMS) to predict their success and provide early-warnings. S3 developed predictive models, using a success index that is decomposed into five indices including preparation, attendance, participation, completion, and social learning to identify and treat at-risk students. However, Essa & Ayad [2012] explained that in using a
different methodology, S3’s approach is similar to when a patient sees a physician in terms of the workflow: understand the problem, reach a diagnosis, prescribe a course of treatment, and track the success.

2.2.3 Visualization Techniques

Visualization converts the abstract and complex to the concrete and simple by amplifying human cognition. Few [2013] explained that well-designed dashboards lead to effective communication as well as correct decision making. The several design principles are summarized here: First, the most important information should stand out from the rest in the dashboard, which usually has limited space; second, the information in the dashboard should support one’s situated awareness and help rapid perception using diverse visualization technologies, third, the information should be deployed in a way that makes sense, and elements of the information should support viewers’ immediate goal and end goal for decision making.

In reviewing previous case studies, Table 2 presents the diverse visualization techniques that were employed. Visualization techniques are related to the characteristics of the information in LADs. For example, rich information regarding students’ activities is effectively delivered with a histogram (bar graph). A pattern such as weekly login trends goes well with a line chart. Some dashboards (e.g., LOCO-Analyst, GLASS) [Ali et al., 2012; Leony, Pardo, de la Fuente Valentín, de Castro, & Kloos, 2012] that included message analysis incorporated a tag cloud. When positioning individual students compared with other students, a scatter plot was an effective visualization strategy and an example is found in S3 [Essa & Ayad, 2012]. Also, to present a summary of activity including both numbers and text, and a number of units, a table matrix was incorporated. Performance results and content usage were delivered through bar or pie chart to display the portion of each part in the whole, for example the most frequent misconceptions in Student Inspector [Scheuer
Finally, as a SNAPP tool [Bakharia & Dawson, 2011] present, online communication and networking is effectively illustrated in a sociogram which consists of nodes (students) and links (communication, message exchanges, replies of forum postings).

<table>
<thead>
<tr>
<th>Names of Tools</th>
<th>Information in dashboard</th>
<th>Visualization Techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOCO- Analyst</td>
<td>Login trends, performance results, content usage, message analysis</td>
<td>Bar graph, pie chart, table matrix, tag cloud</td>
</tr>
<tr>
<td>Student Success System</td>
<td>Performance results, social network, at-risk student prediction</td>
<td>Risk quadrant, scatterplot, win-lose chart, sociogram</td>
</tr>
<tr>
<td>SNAPP</td>
<td>Content usage, social network, message analysis</td>
<td>Sociogram</td>
</tr>
<tr>
<td>Student Inspector</td>
<td>Performance results, content usage</td>
<td>Bar graph, pie chart</td>
</tr>
<tr>
<td>GLASS</td>
<td>Login trends, performance results, content usage, message analysis</td>
<td>Timeline, bar graph</td>
</tr>
<tr>
<td>SAM</td>
<td>Login trends, performance results, content usage, message analysis</td>
<td>Line chart, bar graph, tag cloud</td>
</tr>
<tr>
<td>Course Signal</td>
<td>Login trends, performance results, content usage, message analysis</td>
<td>Signal lights</td>
</tr>
<tr>
<td>Narcissus</td>
<td>Content usage, social network</td>
<td>Wattle tree</td>
</tr>
</tbody>
</table>

*Table 2: Tracked data and visualization techniques in previous LADs*

2.2.4 Evaluation

Six case studies out of the nine studies we have reviewed conducted an evaluation of the tools as marked in the right column of Table 1. In general, there are two types of evaluation based on the purpose of the LAD: *formative* which aims at revising the quality of tools or programs and *summative* which is conducted to determine the effectiveness of them [Dick, Carey, & Carey, 2005]. In our review of previous studies, we found cases studies like LOCO Analyst [Ali et al., 2012] and SAM [Govaerts et al., 2012] conducted formative evaluation by implementing the dashboard to students and investigating the perceived usefulness and usability of the system with multiple iterations. Other cases such as Students Inspector [Scheuer & Zinn, 2007], StepUp! [Santos et al., 2012; Santos, Verbert, Govaerts, & Duval, 2013], Course Signal [Arnold & Pistilli, 2012], and Narcissus [Upton & Kay, 2009] attempted to prove the usefulness of their tools, thus the evaluations are summative. Course Signal presented improved retention rates and performance outcomes; Narcissus experimented with a control group. However, as Verbert et al. [2013] already mentioned, previous studies rather focused on the presentation of functions and usability in LADs, highlighting
potential impact on learning, and neglected to prove the effects of LADs as a pedagogical treatment.

2.3 Findings

This section introduced previous case studies that developed and evaluated a LAD application. Since we depended on literature and did not use the real tools, there is a limitation for deep analysis. However, we found several points that give insight into the design and development of a LAD.

First, the LADs had different intended goals that are determined by who are the target users. Since a LAD displays student activities, it is useful to both teachers and students. Therefore, although the LADs have different purposes: monitoring students in the case of teacher LADs and self-monitoring in the case of student LADs, a tool that works for both in the case of Student/Teacher LADs. Consequently, several cases were developed for both situations. We also adopted this approach in the initial stages of developing our LAD.

Second, LADs displayed diverse information. They ranged from the simple representation of raw data to output based on sophisticated algorithms. We found a spectrum from descriptive to prescriptive in nature in terms of the level of data-mining and tracking in formal, secure, and scalable LMSs to tracking in informal, open, and social environments. Developing a descriptive LAD that utilizes LMS is a good starting point; however, gradually, it is expected to move to other dimensions such as applying prediction models or including logs of activities in informal and open learning environments.

Third, social networks, at-risk student prediction and message analysis were attempted in a few cases. Social networks involve the targets’ social interaction, including discussion behavior and content or message exchange. At-risk student prediction is a function to alert students who are at risk of failure in the course semester. Message analysis involves text data analysis, often summarized in a tag cloud form. Since such functions that require careful analytics need to be verified through a series of pilot tests, we decided not to include such functions in the early version of our LAD; however we consistently tested its accuracy and usability.

Fourth, the dashboards incorporate visualization techniques in order to best present the information; however, only a few case studies considered designing such visualizations based on the dashboard design principles. Further, no case attempted to review the relationship between the visualization information and users’ reactions. As further research, it is necessary to investigate how users react to and understand such visualized information. Therefore, in our study, we attempted to observe students’ reactions to a LAD in the usability test by recording their facial expression and conducting stimulated recall interviews.

3 Design and Development Process

Based on the review of previous studies, we designed a LAD that is intended to be implemented with students in a large private university in South Korea. This section covers the design and development process involving 1) needs assessment, 2) rapid prototyping, 3) usability test, and 4) findings.
3.1 Needs assessment

Before developing the actual LAD, we investigated how students perceive needs in regard to the LAD and what should be considered in implementing the dashboard for students in higher education. Since the comparison of the different LADs provided insights to us, we explained the concepts and features of these LADs to students and investigated what they think about the idea of providing such a LAD. The reason for conducting this investigation is that users’ demands and fitness are clearly prioritized against richness of features in tools [Haintz, Pichler, & Ebner, 2014].

At this stage a needs assessment was conducted with eight college students at the university that we intended to implement. The selection of subjects was conducted to investigate intended target subjects and interpret specific phenomenon qualitatively. For this purpose, we searched for college students who had used a cyber campus based on the Moodle LMS. However, since we conceived different phenomenon and perceptions depending on their academic years, we selected four freshmen and four senior students to compare them more dramatically.

In-depth one-to-one interviews were employed using a pre-designed semi-structured interview guide approach. The results of the interviews include participant’s needs on 1) display of analysis of students’ learning pattern, 2) display of analysis about their learning performance.

3.1.1 Perceptions on analysis of learning pattern

Researchers asked a question regarding what extent a LAD that displays the analysis of their learning patterns would be useful for them to take courses and to achieve their learning outcomes. While three students revealed negative responses, five students expressed neural or positive opinions. Student 6 responded in the following way:

_I just log into the cyber campus to download learning materials and print them. I do not think my online learning behaviors such as log-ins would reflect my general efforts for learning and learning outcomes._

Student 3 responded that the information about their online learning patterns would be useful because they can plan their learning schedule, manage their learning process, and set their learning goals based on the learning analytics information. Also, student 1 mentioned that she would be able to trust the information because _dashboard data might be objective and accurate compared to what they themselves guess or perceive how much time and efforts they spend in the cyber campus_. An interesting response was made regarding the implementation of LAD. Student 1 revealed displeasure about the fact that her learning pattern would be disclosed to her instructor or other students. Student 3 also mentioned that she does not want such data impact on their final score and grade.

This has some implications for the design and implementation of the dashboard: the analysis of learning patterns should be _objective_ and _trustable_ and no subjective interpretations or relations with any kind of evaluation would be preferred.
3.1.2 Perceptions on analysis of learning scores

Learning Analytics Dashboard (LAD) provides not only online learning patterns but also learning performances such as midterm exam, quiz, and task reports that they submit. Regarding statistics about their scores, positive responses were revealed by all participants. While some instructors currently provide such information in the classroom, the idea of using a dashboard in LMS was revealed as a “warm welcome” to the students. Above all, students were able to recognize the meanings of the dashboard information, and further, they liked the idea of comparing themselves with other students throughout the statistical data including means and standard deviation of exams, quizzes and task scores.

Throughout the needs assessment, we found that students perceived the potential usefulness, especially in regard to the function that the information in LAD is objective and accurate; moreover, they can compare their position with peers. However, over-interpretation based on their login information and over-analogism should be careful since students are sensitive about their evaluation and grade.

3.2 Rapid prototyping: 1st version of LAD

When the instructional design output is required to be developed quickly and ensure the success of the final product throughout the iterative cycle of formative evaluation, rapid prototyping is a strategic and effective approach [Dick et al., 2005]. In our case, at the early stage of development, we conducted a design and development process in collaborations with researchers, designers and developers. Here we describe the process and outcomes of LAD with criteria that we reviewed in previous studies: 1) intended goals and target users, 2) data extraction and mining, and 3) visualization.

3.2.1 Intended goals and target users

The first version of LAD was designed and developed for both online learners and instructors. We named the LAD application LAPA (Learning Analytics for Prediction and Action), which was derived from a model that Jo [2012] developed. The model has three segments: learning, prediction, and action (intervention) as illustrated in Figure 2. This model presents several implications. First, learning is influenced by the learner’s psychological characteristics, their self-regulatory ability and instructional strategies. Second, learning is observed and measured via their learning behaviors; especially online learning behaviors can be observed through their log data and their learners’ characteristics, which are measured by survey or students’ records in the system. Third, such behavioral data can be utilized to predict their performances or cluster their learning styles. Fourth, based on their learning style and prediction model, precautionary actions (interventions) can be provided through the Service Oriented Architecture (SOA) dashboard for both learners and instructors.

The LAPA (from here in we use this term for LAD that we design and develop) dashboard intends to be a learning and teaching support tool that stimulates students’ learning related actions. From a narrow and short-term perspective, the goal of LAPA is to inform learners’ online learning behavior to learners themselves and the instructor; however, a broad and long-term goal is to guide their learning in a smart and personalized way. As we mentioned above, we categorized the function of LADs into descriptive and prescriptive in nature. Our current goal is to develop a descriptive
LAD, however, as we develop and elaborate the prediction models, in the long run, we aim at providing a prescriptive LAD so that the learners get more accurate feedbacks and instructors get more useful instructing guidelines.

3.2.2 Data-extraction and Mining

In our case, the educational institution, which is the context of this study, utilizes a cyber campus that works within the Moodle-based LMS as a supplementary environment to offline traditional lecture-based classes. Since students’ login data for the LMS is extracted automatically, we focused on mining the raw data and converting it to meaningful variables, especially those that are related to students’ self-regulatory ability. Our design work was based on a series of previous studies [Jo, Kim, & Yoon, 2014; Jo & J. Kim, 2013; Jo & Y. Kim, 2013] that explored the relations between students’ time management strategy and learning achievement. In addition, Jo and Yu [2014] reported that students’ total log-in time, log-in frequency and log-in regularity, and visits on the board and repository are meaningful variables that predict students’ learning achievement. Therefore, we decided to display such variables visually in LAD.

The first version of LAPA, shown in Figure 3, consists of 7 graphs. The graph chosen for the online activity summary is the scatterplot, where individual learners can choose the X-axis and Y-axis to locate their position in class. The other 6 graphs are provided with a trend line of their activity every week along with the average activity information of other learners in the class. All of the graphs in the LAPA are updated every week until the end of the semester.

Figure 2: LAPA (Learning Analytics for Prediction & Action) Model [Jo, 2012]
3.3 Usability Test: Use of Stimulated Recall Interview

After the first version of the dashboard was created, we conducted a usability test. 6 college students participated. A qualitative research method, including stimulated recall protocol, was employed to observe how students react to the LAPA and how they perceive its usability. Users’ experience on LAPA was recorded through the software ‘Morae’, and the interview was conducted by stimulating their memory as we showed the recorded video. Figure 4 presents one participant seeing our LAPA dashboard, her face as well as her mouse-movements on the dashboard were recorded. A program in ‘Morae’ allowed us to replay the recorded data and mark on each timeline when the user reacted to something. In the one-to-one interviews, we asked participants the degree of understanding of LAPA, and their opinions including suggestions to improve the quality of LAPA. In this process, we applied a stimulated recall interview that is useful to help students’ recall specific moment during the test [I. Jo, Heo, Lim, Choi, & Noh, 2013]. For example, by asking about the reason for the
reaction “why did you do this?” or “what did you think about at that moment?” the participant could easily recall their perceptions and answer more precisely based on their memory. Testing the usability of LAPA dashboard including the interviews with six participants was conducted within two days of the end of the trial in order to prevent further memory loss.

The recorded data using ‘Morae’ in the usability test

The results of the usability test provided useful insights for refining the dashboard. Here we summarize several points. First, students had difficulties understanding the meaning of the “log-in regularity” graph information, which shows how regularly students access the virtual class, they responded to a description that explains what the information implies and how the information can be applied to their learning and performance. Second, most participants reported difficulties understanding the first summary graph. We observed that they were trying to compare the summary graph with the other graphs. One student mentioned:

“The Summary graph was not comprehensible at first, so I looked down at the others…the summary graph presents too much information.”

In this regard, students pointed out that they needed a description on the meaning of graphs. While two students answered that they could not understand the graphs, another student mentioned that she could understand LAPA almost perfectly. She also explained how she could analyze the LAPA.

“I wondered why the log-in frequency does not match the total log-in time in the seventh week, so I inferred that the data is not complete. I think more data should be updated.”
Third, through participants’ diverse reactions to LAPA, we discovered that the level of students’ understanding, especially the literacy of graphic representation, varied and this might impact the perceived usefulness and the effectiveness of LAPA. This underlines the importance of providing enough instructions and guidelines to help them understand the meaning of graphs about their online learning patterns. Fourth, several students commented that the graphs illustrating their relative position in comparison with other students were useful. Moreover, they added that the function of graph comparing with other students would boost their motivation.

“I often wanted to know how other students are studying in the class. For example, I clicked all the board each time whenever I logged into the cyber campus. I really wondered how and what they did in there. Especially, in the forum board, students were interacting for their team project, so I looked for other teams’ performance as well. If they interacted more frequently than our team, I was motivated to do better.”

Finally, a few students answered that the correlation and interaction between students in the class is needed. When the class uses forum board, how students interact with other students can be an indicator for their participation in the group learning. Student ‘C’ said:

“Sometimes I would log in but don’t do anything in particular, so I think the number of replies could be helpful information for us. That information is about practical learning.”

3.4 Findings

This section presents the design and development process of the LAPA dashboard that is intended to support students’ learning performance. Although LAPA aims to provide not only descriptive displays representing students’ behavior in the cyber campus but also prescriptive and precautionary functions guiding their learning direction, the first version of LAPA attempted mainly to fulfill the former goal and to focus on supporting students’ time management and self-regulation. A usability test indicated participants perceived the LAPA in two ways. One is a reminder that they can reflect their previous behavioral patterns and the other is a comparison tool that they could learn their relative rank in the class. Lastly, the participants suggested a number of opinions for modifying LAPA.

4 Evaluation of LAPA

The first version of LAPA, after correcting critical issues that participants in the usability test pointed out, was implemented to two classes in a real-world setting. In this section, we share the results of the experiment as well as their responses in a series of surveys. We consider this implementation and experiment as a pilot test because we did not implement the system in all courses at the university, and we still plan to refine this tool.
4.1 Pilot test with a control group

During the second semester in July to December 2013, 73 college students were employed in a pilot test. Among them, we randomly chose 37 students in two classes as the treatment group. These students were invited to use LAPA by providing a link for them to access to the dashboard with brief instructions. Since the visualized information displaying students’ online learning pattern on the dashboard should only be drawn after a sufficient number of log files have been accumulated, we opened the dashboard 8 weeks after the semester began. Students were allowed to access and use the dashboard freely for the remaining 8 weeks of the semester. The remaining 36 students were in a control group, they were not invited to use LAPA.

As summarized in Table 3, we compared the mean value of students’ learning outcomes, measured by their final exam score in Class A and B, between the treatment and control group. Although students in each treatment group had slightly higher achievement (Class A: Mean=43.76, SD=11.12, B class Mean=80.02, SD=13.02) than the control group (Class A: Mean=43.01, SD=11.94, Class B: Mean=79.66, SD=5.34), the t-test results indicated that there were no statistical differences in the final scores between the treatment and control groups in both classes (Class A: t=0.22, \( p > .05 \), Class2: t=1.112, \( p > .05 \)). Through this we interpret that the first version of LAPA dashboard did not impact on students’ learning outcomes significantly.

<table>
<thead>
<tr>
<th>Groups</th>
<th>Class A</th>
<th></th>
<th>Class B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>SD</td>
<td>t-test</td>
</tr>
<tr>
<td>Treatment</td>
<td>21</td>
<td>43.76</td>
<td>11.12</td>
<td>0.22</td>
</tr>
<tr>
<td>Control</td>
<td>22</td>
<td>43.01</td>
<td>11.94</td>
<td>(( p &gt; .05 ))</td>
</tr>
</tbody>
</table>

Table 3: The comparison of treatment and control group using LAPA dashboard

4.2 Survey Results: 1st round

In addition to the comparison with a control group, we conducted two kinds of online surveys on the treatment group students in order to both verify its effect and to improve the quality of LAPA. The first survey asked about conformity, perceived usefulness, degree of understanding as well as opinions and suggestions. It was conducted right after the LAPA was opened to students. As shown in Table 4, the survey questionnaire consisted of 24 questions (21 Likert 5-scale questions and 3 open-ended questions).

22 out of 37 participants completed the survey. We analyzed and compared the mean value for each question. As shown in Table 5, in comparing the conformity, perceived usefulness, and degree of understanding for all graphs in the dashboard, the result indicate a higher degree of understanding (Mean=4.10, SD=.96), relatively lower degree of the perceived usefulness (Mean=3.22, SD=.91), and medium level of conformity (Mean=3.70, SD=.84). That is, students perceived the information in the dashboard reflected their online activities in the cyber campus. While they could understand the meaning of the graphs quite well, they didn’t perceive this information as a useful tool for their learning and performance.
1. Conformity Degree of conformity between learner’s perceived online activity and real data How much the total log-in time graph conforms to your perceived total log-in time? 7

2. Perceived Usefulness Degree of learner’s perceived usefulness of the information in LAPA How much do you think the total log-in information will help your learning process? 7

3. Degree of Understanding Degree of learner’s understanding of the graphs in LAPA How difficult it is to fully understand the total log-in time graph? 7

Opinion and Suggestion Participant’s opinion and suggestions on LAPA Please suggest any opinions you have on this dashboard. 3

Table 4: Summary of Survey Questionnaire: The 1st round

<table>
<thead>
<tr>
<th>Categories</th>
<th>Conformity</th>
<th>Perceived Usefulness</th>
<th>Degree of Understanding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Online Activity Summary</td>
<td>3.36</td>
<td>.79</td>
<td>3.36</td>
</tr>
<tr>
<td>Total Log-in Time</td>
<td>3.55</td>
<td>.96</td>
<td>2.68</td>
</tr>
<tr>
<td>Total Log-in Frequency</td>
<td>3.86</td>
<td>.77</td>
<td>3.09</td>
</tr>
<tr>
<td>Log-in Regularity</td>
<td>3.73</td>
<td>.83</td>
<td>2.59</td>
</tr>
<tr>
<td>Visits on Board</td>
<td>3.77</td>
<td>.92</td>
<td>3.68</td>
</tr>
<tr>
<td>Time Spent on Board</td>
<td>3.73</td>
<td>.94</td>
<td>3.27</td>
</tr>
<tr>
<td>Visits on Repository</td>
<td>3.91</td>
<td>.68</td>
<td>3.86</td>
</tr>
<tr>
<td>Total Mean</td>
<td>3.70</td>
<td>.84</td>
<td>3.22</td>
</tr>
</tbody>
</table>

Table 5: Descriptive statistics of conformity, perceived usefulness, and level of understanding (n=22)

4.2.1 Conformity

The answers to conformity questions were moderately high (Total Mean=3.70, SD=.84). Among them, the online activity summary (Mean=3.36, SD=.79) and the total log-in time (Mean=3.55, SD=.96) showed relatively low levels of conformity. There were relatively more students who answered the total log-in time and the log-in frequency information in LAPA was different from what they thought. In regard to this result, we suspect that students often stayed online without logging-out even after they achieved their goals. However, this result needs to be carefully observed since the result of repeated measures ANOVA indicated that no significant mean difference among the 7 variables was detected ($F_7 = 1.20, p > .05$).
4.2.2 Perceived Usefulness

Students generally had a medium perception on the usefulness of LAPA (Total Mean=3.22, SD=.91). Follow-up repeated measures ANOVA for perceived usefulness detected significant large mean difference among the 7 variables ($F_3 = 13.52, p < .001$). Participants perceived that the most useful item was the visits to repository, followed by the visit to board, the online activity summary, the time spent on board, the total log-in frequency, the total log-in time, and the log-in regularity, respectively. The total log-in time (Mean=2.68, SD=.84) and the log-in regularity (Mean=2.59, SD=1.01) showed the lowest levels of perceived usefulness. Also, perceived usefulness of the log-in regularity is the only question that received extremely negative answer, “I do not think this information will help my learning at all”.

4.2.3 Degree of Understanding

In terms of degree of understanding, students felt it was easy to understand most of the graphs in LAPA (Total Mean=4.10, SD=.96). Another repeated measures ANOVA for degree of understanding reports noticeably significant mean differences among the 7 variables ($F_3 = 3.17, p < .01$). Degree of understanding presented highest on the total log-in frequency and visits to board, followed by the total log-in time. The visits to board and the time spent on board have the same mean values, and the lowest degree of understand was shown in the log-in regularity and the online activity summary. There were many negative answers about the online activity summary (Mean=3.73, SD=1.08) and the log-in regularity (Mean=3.86, SD=1.13) suggesting that students were experiencing difficulties in fully understand those two graphs. Since the online activity summary is a scatterplot where users choose the X-axis and Y-axis by themselves, it can be assumed that it might be difficult to use it and understand the information without a detailed explanation or manual. In the case of the log-in regularity, the difficulty that students experienced was due to the misunderstanding of the concept itself. To solve this problem and help students’ understanding, it is necessary to provide in the help menu of LAPA the concepts of each item and how it is presented as graph.

4.3 Survey Results: 2nd round

The secondary survey measured the overall satisfaction with LAPA and perceived behavioral changes from the use of LAPA. It was conducted right after the semester was over. We developed different questionnaire from the first round of survey which focused on investigating which variables students perceive to be more reliable, useful, and understandable. The 2nd survey focused on “after students had utilized the LAPA dashboard for a significant time (8 weeks), were the users satisfied?” and was their behavior changed by using it?” Table 6 presents 12 items in the questionnaire.
Table 6: Summary of Survey Questionnaire: The 2nd round

<table>
<thead>
<tr>
<th>Part</th>
<th>Items</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Overall</td>
<td>1. LAPA consists of what I want to know.</td>
<td>3.18</td>
<td>.66</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>2. It consists of important information.</td>
<td>3.23</td>
<td>.69</td>
</tr>
<tr>
<td></td>
<td>3. Suggested information is useful for learning.</td>
<td>2.59</td>
<td>.96</td>
</tr>
<tr>
<td></td>
<td>4. The information in LAPA is constructed logically</td>
<td>3.36</td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td>5. The visualized information is attractive.</td>
<td>2.73</td>
<td>.99</td>
</tr>
<tr>
<td></td>
<td>6. The visualized information delivers messages effectively.</td>
<td>3.50</td>
<td>.96</td>
</tr>
<tr>
<td></td>
<td>Sub total</td>
<td>3.10</td>
<td>.83</td>
</tr>
<tr>
<td>2. Behavioral</td>
<td>7. LAPA stimulated learning motivation.</td>
<td>2.77</td>
<td>1.19</td>
</tr>
<tr>
<td>Changes</td>
<td>8. I monitored the information in LAPA regularly.</td>
<td>2.68</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>9. I reflected on my learning behavior based on the info in LAPA.</td>
<td>3.36</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>10. LAPA supported my time management strategy.</td>
<td>2.77</td>
<td>1.93</td>
</tr>
<tr>
<td></td>
<td>11. I utilized LAPA for my learning plan.</td>
<td>2.86</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>12. I felt confidence after reviewing my online behavior pattern in LAPA</td>
<td>2.50</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>Sub total</td>
<td>2.96</td>
<td>1.05</td>
</tr>
</tbody>
</table>

22 students out of 37 responded. The results indicated a moderate level of overall satisfaction (Mean=3.10, SD=.83) and a lower level of behavioral changes (Mean=2.82, SD=1.27). This implies that students could easily understand the information on the dashboard and had fair satisfaction; however, they were not able to connect the dashboard information to their behavioral changes. Consequently, we can say the LAPA dashboard did not influence their learning performance directly.

4.4 Findings

With the first version of the LAPA dashboard, unlike our expectation that participants would perceive the 1st version of LAPA useful and feel it was easy to understand all items, the LAD treatment did not significantly impact on their learning achievement directly. Throughout two surveys we learned that students did not perceive a high level of usefulness (1st Survey Mean=3.22) or satisfaction (2nd survey Mean=3.20). However, a correlation analysis provides further implications regarding the relations among students’ different perceptions. As shown in Table 7, while the correlation coefficients between degree of understanding and overall satisfaction indicated .679 ($p<.01$), such overall satisfaction was correlated with behavioral change ($r=.457$, $p<.05$). Through this result, we learned that overall satisfaction is a covariance that is related to both degree of understanding and behavioral changes caused. Therefore, in the future, when measuring the impact of the dashboard it is necessary to consider “to what extent students understand the visualized information”.

Table 6: Summary of Survey Questionnaire: The 2nd round
Table 7: Correlation coefficients among conformity, perceived usefulness, degree of understanding, and overall impression

<table>
<thead>
<tr>
<th>Categories</th>
<th>Conformity</th>
<th>Perceived Usefulness</th>
<th>Degree of Understanding</th>
<th>Overall Satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived Usefulness&lt;sup&gt;1&lt;/sup&gt;</td>
<td>.345</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree of Understanding&lt;sup&gt;1&lt;/sup&gt;</td>
<td>.293</td>
<td>.169</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Satisfaction&lt;sup&gt;2&lt;/sup&gt;</td>
<td>.109</td>
<td>.111</td>
<td>.679***</td>
<td></td>
</tr>
<tr>
<td>Behavioral Change&lt;sup&gt;2&lt;/sup&gt;</td>
<td>-.353</td>
<td>.133</td>
<td>.064</td>
<td>.457*</td>
</tr>
</tbody>
</table>

Note: 1) 1st survey 2) 2nd survey

(* p<.05, ** p<.01)

The first survey revealed that students had difficulties with understanding two graphs: log-in regularity and summary graph. The results of the repeated measure ANOVA shows the participants’ different perceptions on perceived usefulness and degree of understanding. Specifically, participants answered that the log-in regularity would not help their learning. This result is contradictory to the results of preceding research that the log-in regularity predicts higher learning achievement [Jo & Kim, 2013]. Considering the answers to open questions (i.e. “It is hard to understand log-in regularity”, “What is the unit or scale?” etc.) and low degree of understanding on the log-in regularity, it is possible that participants experienced difficulties of understanding this intuitively, so that they did not comprehend the concept of regular learning which was presented in LAPA.

5 Conclusions and Future Work

This study presents the design and development process of the first version of our LAPA dashboard, a supporting tool to facilitate students’ learning and performance through the display of their online learning patterns. A review of state-of-the-art LAD applications contributed to the design of this work; however, a challenge was to overcome the limitation of previous studies that rarely evaluated the tool in the terms of students’ learning performance. In our case study, a preliminary needs assessment, a usability test and a pilot test were conducted in a real world setting with a control group. Although, unlike what was expected, no significant difference in final score (which we regarded as the measure of students’ learning achievement) was found, a series of investigations suggested further steps to revise the early version of LAPA. We point out the meaning of this study and share several important conditions that should impact students’ learning performance more significantly.
Figure 5: Two revised versions of LAPA dashboard
First, in this study, a relatively lower level of perceived usefulness was found due to one extremely negative response. Moderate positivity on the usefulness of the LAPA dashboard indicated that students reacted to LAD positively in general, so that we still expect the quality and potential effects. We do think this study is meaningful in terms of suggesting implications for the development of more refined and effective dashboard treatments, and providing specific directions for future research.

Second, this study found that students’ overall satisfaction on LAD is correlated with both the degree of understanding and their behavior change. Thus, we made an effort to increase the users’ degree of understanding as we revised the first version of LAPA and developed the second version, as shown in Figure 5. Especially, in order to mitigate students’ difficulties in understanding graphs and to remind them of the goal of this dashboard, we decided to 1) include a ‘help’ menu as a guide on how to read the graphs, 2) consider an alternative of scatter graph to the one which was recognized as difficult to understand, and 3) provide scaffolding texts that allow students to reflect on the dashboard information and connect the information with their future learning and behavioral changes.

Third, this study suggests an option that users can choose in regard to referring to a LAD. In this study, because we learned that users have different levels of understanding, two kinds of dashboard (advanced and simple version) were made. The advanced version still keeps the scatter chart as a summary graph since some of students liked it due to its richness and the implications caused by exploring the relations between variables, and help menus on each sub section are optional so that they are opened only when clicked on. The simple version includes a bullet graph that shows a comparison of five variables at a glance. Also the help information and scaffolding texts are not optional, so the length of page is increased and the long scroll bar may distract from students’ quick perception, but there is an advantage that users don’t need to find the help. Also, since the usability test results revealed participants’ mixed reactions to overlapping an additional trend line on the bar graph, we made an option to view it in only in the advanced version. As the next phase of developing this LAPA dashboard, we are planning to verify students’ different preferences between two versions of output.

Fourth, comments from open-ended questions provided the direction for revision. For example, a student mentioned, “the usefulness is up to how each class utilizes LMS”. This implies the usefulness and effect of LAD would be very different between a full online class and a blended learning class. Therefore, as a further study, it is necessary to examine the effect of LAPA empirically in different contexts to prove its usefulness more precisely.

Finally, while the goal of the first version of LAPA was to develop a descriptive LAD that only displays students’ online behaviors, which are derived from their log-in information and visits to boards and repositories. However, as mentioned early, the long-term goal of LAPA is to provide prescriptive and precautionary interventions so that learners get more accurate feedback and instructors get more useful instructing guideline. We do think one of the biggest reasons for the low perceived usefulness is the descriptive nature of LAPA. If the dashboard not only described students’ online behavior but also provided a prediction of their final score and sent signals such as in the case of Course Signal [Arnold & Pistilli, 2012] or Student Success System [Essa & Ayad, 2012], the effects of LAD would be more powerful. However, as discussed
in the section of needs assessment, over-interpretation or over-analogism should be avoided since students are sensitive about being evaluated. Also, such interventions can stimulate negative behavior as well such as gaming (e.g., just logging to increase the number of login frequency or time spent). Moreover, since a wrong prediction can lead to students’ rage or demotivation, the ethical issues of this research need to be considered [Kruse & Pongsajapan, 2012]. Further research in the learning analytics area need to make efforts on increasing the accuracy of prediction and recommendation models as well as considering side-effect issues around LAD applications.

Acknowledgements
This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2013S1A5A2A0304410)

References


[Ha, K., Jo, I., & Lim, S., 2013] Ha, K., Jo, I., & Lim, S.: Usability study of visual dashboard as learning analytics interventions. Proc. International Conference of Educational Technology, Sejong University, 2013,


