

LADA: a Learning Analytics Dashboard for Academic Advising

Abstract

From the perspective of Learning and Educational Technologies, academic advising has been one of the most overlooked aspects of academic support systems, despite being critical for the learning process and final success of students. The majority of higher education institutions provides simple technical support to academic advisers with basic descriptive statistics. This article presents the general design and implementation of a Learning Analytics Dashboard for Advisers (LADA), to support the decision-making process of academic advisers through comparative and predictive analysis. Moreover, this work evaluates the use of this tool to support decision-making of actual advisers in two different higher education institutions (University A, University B), compared with more traditional procedures and tools. Results indicate that LADA enables expert advisers to evaluate significantly more scenarios ($Median = 2$), especially for high advising difficulty cases with students that failed many courses ($Median_A = 3$, $Median_B = 2.5$), in a not-significantly different amount of time. For inexperienced advisers, LADA is perceived as a valuable tool for more accurate and efficient decision-making, as they were able to make informed decisions in a similar amount of time compared to the experts. These results are encouraging for further developments in the field.

Keywords: Learning analytics, visualization, academic advising, academic adviser, data-driven decision-making

1. Introduction

Higher education students are confronted with a large number of academic choices during their studies. *Which courses should be taken next? How much study load can I successfully engage with? What should I do in case of failing a course? Can I still complete the program in the expected number of years?* All these decisions should usually be taken without complete knowledge of the intervening factors or their short, medium, and long-term consequences. To help students make better decisions, most universities offer an academic advising program.

Academic advising is a decision-making process that assists students in the clarification of their career/life goals and the development of an educational plan for the realization of these goals through communication and information exchanges with an adviser (Ender et al., 1982; Okewu and Daramola, 2017). In its most common instance, the academic advising occurs through one or more face-to-face meetings between the students and their advisers. During these meetings, the adviser answers questions that the student has about the program, recommends courses the student should take up, and offers advice on study strategies. The first formal academic advising programs were implemented in the late 1800s in American universities (Gordon et al., 2011) and currently the vast majority of Higher Education institutions offer them. The main reason for the ubiquity of academic advising is its proven positive impact on the academic performance and study completion of students (Young-Jones et al., 2013; Drake, 2011; Bahr, 2008; Sharkin, 2004).

Despite being highly beneficial to the success of students, academic advising has received very little attention in the technological overhaul of educational institutions that occurred during the past decades. Advising interviews and interventions are performed very similarly to how they were done during

the past century, as attested by the recent comprehensive reviews of the field (Barron and Powell, 2014). A specific analysis about the use of technology in academic advising (Steele, 2016) concludes that virtual meetings through e-mail or video conferences has been the leading innovation to deal with distance education students. While access to digital student information has been greatly improved, little has been done to automate the analysis and presentation of that information. Such automatic data analysis has gained increased interest (Stoneham, 2015). Of particular interest is recent work by the learning analytics community that focuses on producing suitable algorithms to predict study success, together with visualizations that empower users with actionable knowledge (Viberg et al., 2018). The overall objective is to represent the outcome of a prediction model, along with relevant data, to support different stakeholders (e.g. academic advisers, instructors, students and administrators) about the learning process in so-called learning analytics dashboards. Learning analytics dashboards are used mostly to facilitate blended or online learning, face-to-face learning, or group work (Verbert et al., 2014): increasing awareness and reflection and providing timely feedback are commonly supported goals. Little work has been done so far to use these dashboards to support academic advising. We did find some other interesting tools that enable staff to monitor student engagement and provide support to at-risk students (King, 2012; Choi et al., 2018), but a “one-size-fits-all” approach is often applied that is not adjustable to the requirements of different universities, faculties, and departments (Stoneham, 2015). In addition, the outcome of an academic risk prediction is often presented, but valuable information such as the quality of the prediction and uncertainty in the data because of missing values or program changes are not presented to the user.

In our work, we focus specifically on learning analytics tech-

niques to support academic advising scenarios: the overall objective is to support academic advisers with detailed insight into both relevant data and prediction models to improve the decision-making process. We present the design of LADA, a Learning Analytics Dashboard for Advisers, as well as results of elaborate user studies that assess the utility and usability of the dashboard. These user studies have been conducted at two different universities: University A (Europe) and University B (Latin America). The involved universities are part of a bilateral research project between Europe and Latin America and use different models of academic advising and different levels of technology adoption in the advising process. To strengthen the generalization of the results, elaborate user studies have been conducted in these two, very different, contexts.

2. Background

2.1. Academic advising process and models

Originally, the process of advising students has been centered on registration and record keeping (Winston, 1984). In this process, the advising session is seen as “prescriptive”, the adviser being the doctor and the student the patient. In this role of a doctor, the adviser prescribes a procedure so the student gets “better” in her/his academic life (Drake, 2011). In the 1970s, the importance of student-faculty interactions in the academic advising process was recognized (Grites, 2013), and academic advising evolved to a more comprehensive process performed by members of a university community (Grites, 1979). As the impact of academic advising on academic progress and retention was studied in social and behavioral sciences (Young-Jones et al., 2013; Drake, 2011; Bahr, 2008; Sharkin, 2004), the professionalization of academic advising in higher education institutes gained importance. As a result, academic advising became more student-centered, in which students’ needs and expectations for their personal and professional lives are part of the advising process (Chan, 2016).

Currently, higher education institutes typically provide each student with a personal academic adviser, who coaches the student in developing a pathway to the degree during face-to-face meetings (Phillips, 2013). The maturity of academic advising, underlying advising models, and their actual implementation in higher education institutes is, however, diverse. Universities promote and implement particular advising models such that there is an optimal fit with their organization, educational vision, national and regional context, faculty, staff, and students. These models have shifted in consistency with current trends in educational paradigms and, as elaborated on later, the advancement of technologies. Overall, the trend to move to models that support students in their development within the university is clear (Young and Valach, 2016; Wiseman and Messitt, 2010).

NACADA, the Global Community for Academic Advising, plays a key role in structuring, understanding, and appreciating the diversity in academic advising over institutes. The four dimensions of academic advising at university, recognized by NACADA in 2011, as stated by (Miller, 2012), are: who advises (faculty, staff, peers, etc.); is the advising service centralized (group of advisers) or decentralized (in departments); how

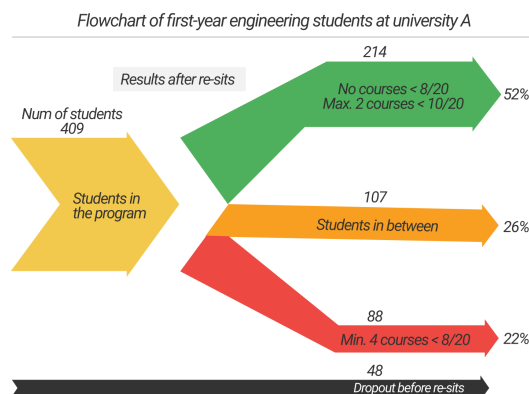


Figure 1: Current static visualization of University A: students are divided in three categories based on the number of mildly failed ($8/10 \leq \text{grade} < 10/20$) and strongly failed ($< 8/20$) courses. Students in the upper group (green flow) are likely to obtain their bachelor degree in nominal time, while students of the lower group (red flow) are likely to not obtain the bachelor degree.

the advising responsibilities are divided into the university, is the advising service divided according to topics or challenges (advisers per topic) or is the advising service divided according to levels of complexity from entry students to those that are ready to exiting a program? And; finally, where the advising service takes place (on-campus, off-campus).

2.1.1. Academic advising at University A

At the *Faculty of Engineering Sciences of University A*, there is both a central study advise centre with professionals focusing on generic study skills and questions transcending the program-level, and faculty-based services focusing on program-specific study skills and questions. The faculty-based service, the so-called Tutorial Services, is the first point of contact for a student. At the Tutorial Services, the academic advisers are professionals that are part of the educational support staff. They are either engineers or scientist with an additional training for academic advising and coaching. Their core tasks consist of academic advising and content-related support for first-year students. Therefore, academic advisers are both experts in both the content of the first-year courses, the program-specific and university-wide regulations, and academic advising in general.

Advising takes place on campus in a traditional setting with a face to face adviser-student meeting. Academic advisers obtain information on the student through multiple systems, including the students academic record (courses booked, grades obtained) and general personal information (e.g. gender and prior education). Academic advisers, however, lack information that allows them to make data-driven decisions, such as data that allows them to position students with respect to their peers and to understand the impact of the current students’ position on the future study career (a so-called predictive model). To accommodate the need for better data-driven support for academic advising, the tutorial services manually collected and visualized data from the university data-warehouse to understand the impact of a students first-year academic achievement on the future study career, presented in Figure 1.

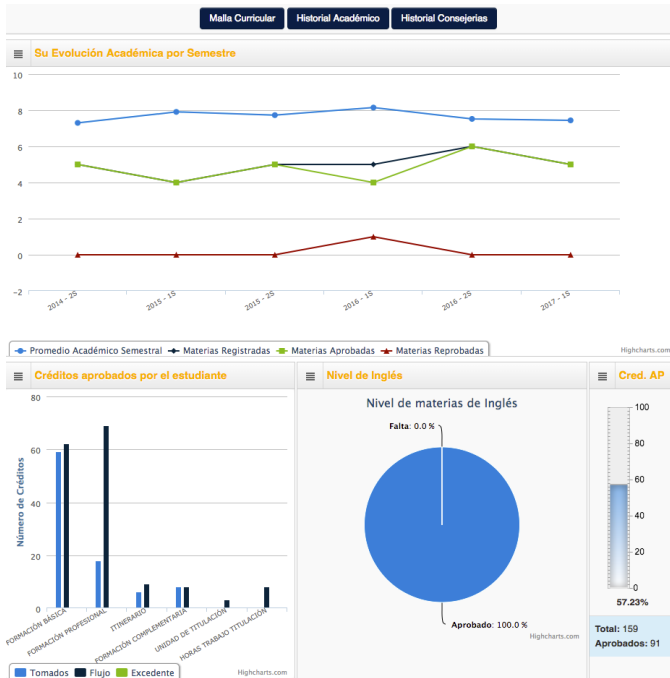


Figure 2: Academic advising system at University B, presenting information such as the number of courses taken and approved each semester.

2.1.2. Academic advising at University B

University B has a faculty-based advising model supported by an information system. Students are assigned by the program coordinator to a professor of the program. Each professor is responsible to advise between 20 and 30 students. Four times each year, the students have the opportunity to attend a 15-minutes academic advising session. In the case of students with good performance, only two of these sessions are mandatory (at the start of each semester). In case that the student presents a low GPA or has failed a course in the previous semester, their presence to all the advising sessions is mandatory. The professors advise students on courses to be taken in the next semester, as well of generic academic advising (study tips, practical information to graduate, optional courses, etc.). In case of situations that are not academic (mental illness, family problems, etc.), professors refer the student to the well-being service of the university. To facilitate their work, the university provides an information system to professors with an integrated view of the academic history of each student assigned. This system, as presented in Figure 2, presents static visualizations of this historic information (courses taken and approved during each semester, courses needed to finish the program, etc.) and is used before and during the academic advising session. As in the current approach of University A, little is done with the vast amount of data that the university has been collected to support more advanced data-driven decision-making.

2.2. Supporting data-driven advising with technology

Many research communities have been studying the potential of analyzing vast amounts of data that universities are collecting, including communities working on educational recom-

mender systems (Drachler et al., 2015), educational data mining (Papamitsiou and Economides, 2014), learning analytics (Siemens and Baker, 2012) and academic analytics (Campbell et al., 2007). Outcomes of a review on educational recommender systems (Drachler et al., 2015) indicate that the majority of recommender systems focus on fine-grained content suggestions or sequences of learning materials that are relevant for the student, although there are a few examples that suggest relevant courses to students. Research in the learning analytics and educational data mining communities is focused more often on the prediction of student performance and prediction of drop-out and retention (Papamitsiou and Economides, 2014). The latter is also at the core of research on academic analytics (Campbell et al., 2007). The focus of these research efforts is mainly on information discovery: educational data mining techniques are applied to student data to predict whether a student is likely to pass a course or is likely to be at risk. A key focus of research in the learning analytics community is to put this information in the hands of human experts to support decision-making (Lonn et al., 2012). The objective is to inform and to empower academic advisers, instructors and students of issues that are identified by data mining techniques and to leverage human judgement (Siemens and Baker, 2012).

While academic advisers are key-stakeholders in the educational development of students (Drake, 2011), little research has been done so far to use dashboards to support academic advising. A notable tool is LISSA, a dashboard to support the adviser-student dialogue Charleer et al. (2017). The LISSA dashboard merely visualizes global student progress to support the advising session and does not use predictive algorithms. *EAdvisor* is a combination of both a student and a staff-facing tool developed by the Arizona state university to support the choosing of a major, and choosing and scheduling particular courses (Phillips, 2013). It is claimed to better inform students about their majors and to help advisers to keep track of the changing requirements of majors and courses (Phillips, 2013), but no data is provided to support these claims.

Aguilar et al. (2014) designed *Bridge*, an adviser-facing tool intended to provide academic advisers with access to the achievement and engagement data of students. The only impact measured from *Bridge* was a non-significant increase in perceived usefulness. Fritz (2011) discussed the development and deployment of the *Check My Activity dashboard* that supports students' awareness of how their use of the learning management system and their current grades compares to that of their peers. They found that after promotion of the tool, students were using the tool during a longer amount of time, more intensively, and returned more often.

Some examples of the use of educational data mining for academic advising are mentioned by (Ranjan and Malik, 2007). The authors list among others the identification of the best program based on prediction of how students will perform in the selected courses as a target objective. To the best of our knowledge, the approach is described, but the usefulness and effectiveness of the approach has not been evaluated yet.

3. Design of LADA

In our work, we focus specifically on the use of learning analytics techniques to support academic advising scenarios: the overall objective is to support academic advisers with detailed insight into both relevant data and prediction models to improve the academic advising process.

Following a user-centered design approach, we designed and implemented the LADA dashboard¹, presented in Figure 3. LADA is a learning analytics dashboard that supports academic advisers in compiling a semester plan for students based on their academic history. LADA also includes a prediction of the academic risk of the student using a clustering technique, detailed in Section 3.2.1. UX/UI researchers and academic advisers worked together to design the dashboard through several brainstorm sessions. Feedback provided during all of the iterations resulted in different digital mockups. The final design was implemented and evaluated with academic advisers and students.

Close collaboration between the researchers and academic advisers was key to deploy the dashboard in real-life settings. Through collaboration with the advisers, historic data was also obtained that could be used to generate predictions. Data is restricted to internal access due to privacy regulations.

LADA was implemented following a client-server architecture using the Meteor Framework. On the client-side, we implemented LADA as a Web Application. On the server-side, a Python API was implemented to predict the academic risk of a student based on historical data.

3.1. Available data

LADA uses data that is available at any higher education institute: the grades of students, the list of courses in the program, the courses booked by a student, and the credits for each course. This data is collected both for students in previous cohorts and the current students. The data from previous cohorts is used in the predictive model that looks for similar past students in order to predict the academic risk of the student.

3.2. Client-side

The main interface of LADA consists of two sections: the top part, presented in Figure 3.a, depicts a prediction of the chance of success for the student and information regarding the quality of the prediction. The bottom part, presented in Figure 3.b, provides different components in the form of information cards. These information cards are attachable components that can be used to add or remove data sources to assist advisers in their decision-making process. First, we explain the chance of success component. Then, the different information cards are described.

3.2.1. Prediction algorithm

To predict the “chance of success”, LADA uses multilevel clustering, described by Clarke (2008) with adaptive specificity levels, following the framework defined by Ochoa (2016) for the prediction of academic risk. This technique was selected because of its capability to provide better predictions both for typical students (with a large number of similar students in the historical data) and fringe students (with few similar students in the historical data). The framework allows the automatic selection of different levels of features (from general to very specific) to create a clustering model to predict the academic risk for each student. This framework has been previously evaluated within an academic context showing a good performance compared to static models (Ochoa, 2016).

The algorithm clusters past students according to available features (their obtained grades and the number of courses that they took during the semester). The percentage of students in each cluster that completed the booked courses reflects the chance of success of this group (e.g., in cluster four, 8 out of 10 students passed the set of courses in the semester). The clustering is achieved through Fuzzy C-means algorithm (Bezdek et al., 1984).

When building the semester plan for the student, the algorithm calculates the most similar cluster for the student, according to the current grades and the courses selected. Also, the algorithm assigns the measured proportion of success as a prediction of the success of the student (e.g., 80% will be assigned to student x closest to cluster four for the chance of success).

In multilevel clustering, the features being used to classify the different students form hierarchical structures with increasing level of specificity. For example, in a *Mechanical Engineering* program, the feature “courses taken during the semester” can be analyzed at different levels with different levels of specificity. In a first level, only the number of courses taken are considered. In a second level, the number of courses can be counted according to topical grouping (e.g., “Mathematics”, “Design”, “Physical laws”, “General education”). In a third level, the individual course codes are used as features for the clustering. At the same time, level one (e.g., all students that have taken five courses in the semester) is expected to produce more massive clusters than level two (e.g., all students that have taken two Mathematical, one Design, and two General education courses in the same semester) and much more than level three (e.g., all students that have taken Math101, Math102, Des201, Com102 and Art103 in the same semester).

The adopted adaptation strategy allows adapting the specificity levels, based on the available data, thereby creating more accurate predictions. For example, “mainstream” students, that is, students that behave near the norm, are better served with particular models because they will be grouped with enough similar students to estimate their chance of success. On the other hand, the chance of success of “outlier” students, that is, students with a unique set of features, are better predicted with less specific models that allow clustering these students with a broader range of previous students, even if they are less similar than in the specific model.

¹<https://github.com/DoubleBlindReviewsHCI/LADA/>

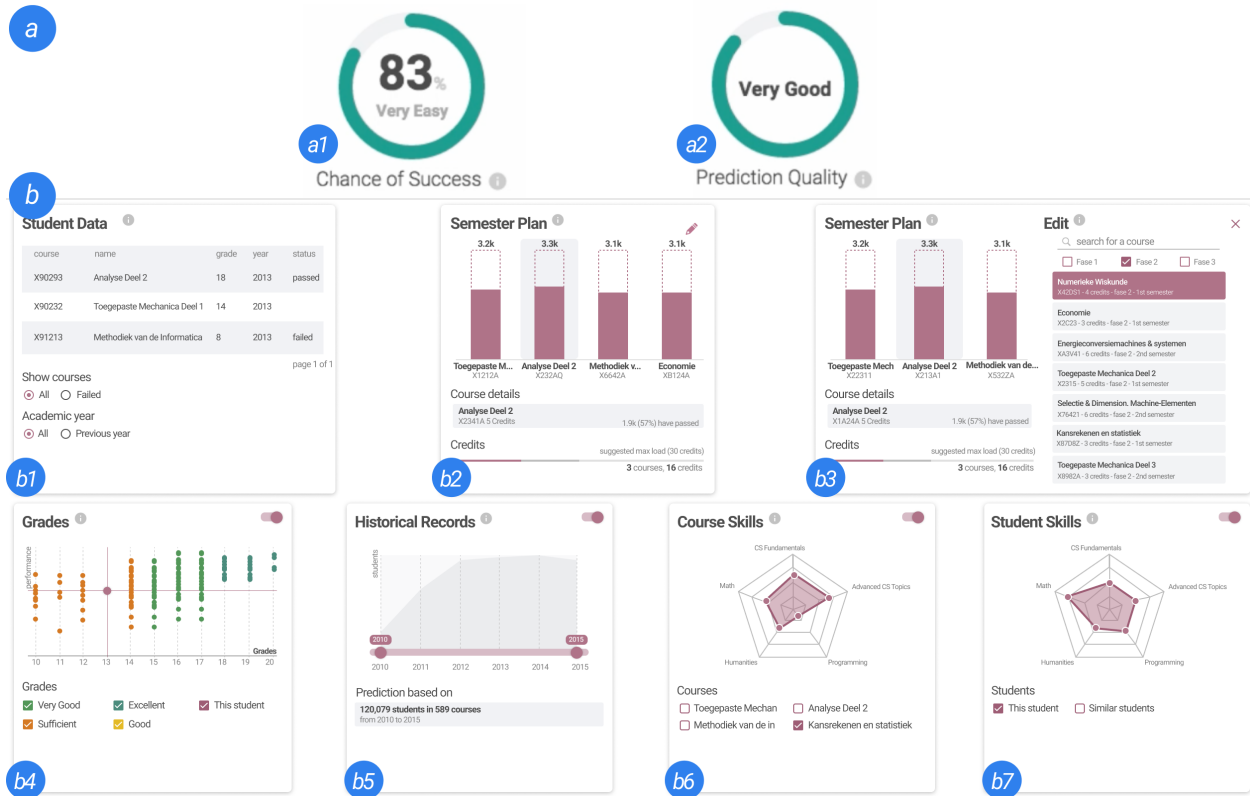


Figure 3: Overview of LADA: a learning analytics dashboard used for academic advising. a) The chance of success and prediction quality components b) The various information card components designed to support the adviser.

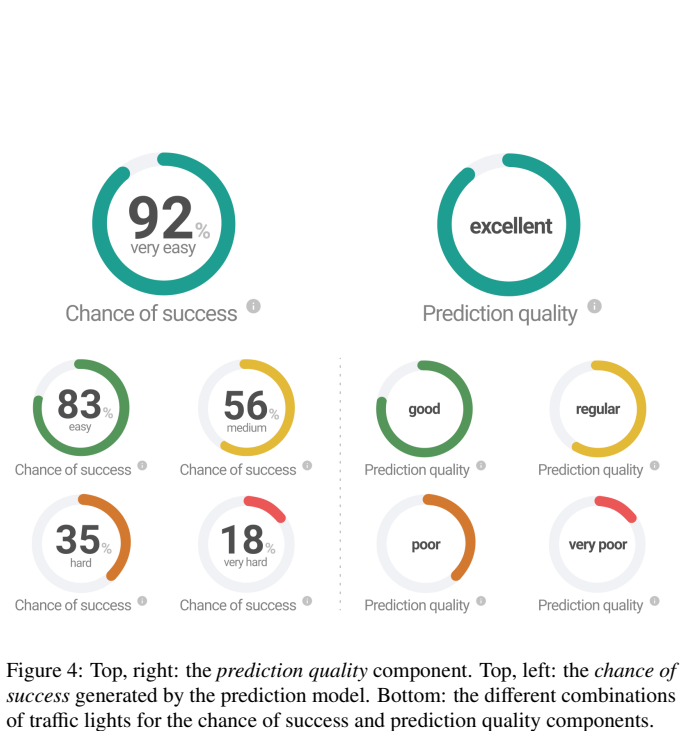


Figure 4: Top, right: the prediction quality component. Top, left: the chance of success generated by the prediction model. Bottom: the different combinations of traffic lights for the chance of success and prediction quality components.

3.2.2. Chance of success

The chance of success component, presented in Figure 4, uses the output from the prediction algorithm to indicate whether a student is likely to pass a course or the set of selected courses in the semester plan. The chance of success visualization updates every time the user adds or removes courses, or interacts with the historical data. To effectively communicate the prediction of the model, we have defined a set of traffic lights that correspond to different values of prediction results, as presented in Figure 4. Traffic lights correspond to the student's chance of success: between 100% and 80% (turquoise, *very easy*), between 80% and 60% (green, *easy*), between 60% and 40% (yellow, *medium*), between 40% and 20% (orange, *hard*), and between 20% and 0% (red, *very hard*). We use these color cues as well as descriptive words to present this predicted chance of success.

3.2.3. Prediction quality

Similar to the chance of success, we implemented a prediction quality component that uses the output from measuring the predictive power of the model (Gneiting and Raftery, 2007). We used the Brier score (Wilks, 2011; Brier, 1950; Van Oirbeek and Lesaffre, 2016) for representing the quality of the prediction, see equation 1, where f_t is the probability that was predicted, o_t the outcome of the event t and N is the number of predicted instances. The lower the Brier score, the more accurate the prediction is.

$$BS = \frac{1}{N} \sum_{i=1}^N (f_i - o_i)^2 \quad (1)$$

Every time the chance of success prediction is updated, the quality value updates accordingly. To keep consistency with the whole interface, the traffic light colors from the chance of success component were reused. Each color corresponds to the different levels of prediction quality: red (*very poor*), orange (*poor*), yellow (*regular*), green (*good*), and turquoise (*excellent*). Descriptive words were also included in the center of the component to explain the visualization.

3.2.4. Student Data Card

The Student Data card, presented in Figure 3.b1, shows the historical records of all the courses the student has previously taken. The listing of courses is sortable by course ID, course name, grade achieved, the year when the course was taken, and whether the course was failed or passed. In addition, users can filter the view by academic year, or by showing only failed courses to facilitate planning. The course list shows a maximum of five courses at a time, with a pagination allowing users to cycle through the list.

3.2.5. Semester Plan Card

The semester plan card provides two different views: a compact view, presented in Figure 3.b2, and an extended view, presented in Figure 3.b3. The compact view shows the selected courses in a bar chart. The height represents the number of students that have taken the course. The color indicates the amount of students that have succeeded in the course. The number at the top of each bar indicates the total amount of students from which the data was derived. Below the bar chart more detailed information is presented about the course: the total number of credits, the course name and the course ID.

By clicking on the pen symbol in the top right corner, the card doubles in size to change into the extended view, presented in Figure 3.b3, where the user can edit the course list. In the extended view, the user can navigate through a scrollable list of courses which can be filtered by year using check-boxes, or searched by course name using an input field at the top. By clicking on a course, the course is added to the list of selected courses on the left side of the card. When a course is added, the credit calculation updates simultaneously. A typical full-time program has 30 credits per semester. When this number is reached, a warning icon is shown on the card. Clicking the “close” icon in the extended views top right corner hides the course listing and shrinks the card back to the compact view. To remove a course from the semester plan, the user can hover over the course. When doing so, a remove button is shown. Any change made to the semester plan will result in an adjustment of the prediction model.

3.2.6. Grades Card

The main part of the Grades card shows the student position among peers based on the average grade achieved in the previous year, as presented in Figure 3.b4. By selecting a course

from the Semester Plan card, the user can inspect the grade of the student in relation to his/her peers of the respective course. While the distribution is course-specific, the position of the student is based on the average grade achieved in all courses and thereby only offers a rough impression of performance in comparison to other students.

The bottom part of the card allows the user to filter the scope of grades to be visualized in the distribution. Using a set of check-boxes, users can choose to show only students within a particular performance group, ranging from sufficient to excellent, or all of these in any combination.

3.2.7. Historical Records Card

The Historical Records card details the range of student data that are used for the prediction and is presented in Figure 3.b5. Using a slider, the user can define the range of student records to be used for the prediction. When courses have changed over time, the user can exclude certain parts of earlier data to improve the prediction’s validity and accuracy. Below the slider, the user can inspect the sample size of the current selection.

3.2.8. Course Skills Card

The Course Skills card is presented in Figure 3.b6 and details the skills of a particular course or a set of courses. Depending on the courses added to the Semester Plan card, the Course Skills card shows the skills required for completing the semester successfully. To do so, the card employs a radar chart showing the semester plan’s alignment with the skills Humanities, Math, CS Fundamental, Advanced CS topics, and Programming. The bottom part of the card provides a set of check-boxes that allow the user to add/remove a course from the visualization to gauge its impact on the skill distribution.

3.2.9. Student Skills Card

Similar to the Course Skills card, the Student Skills card employs a radar chart to represent the distribution of the students skills in the relevant categories (Humanities, Math, CS Fundamental, Advanced CS topics, and Programming), as presented in Figure 3.b7. In addition, the user can display a second web inside the radar chart that shows the average skill set of students with a similar performance in the previous year. Both can be toggled on or off using a set of checkboxes. This way, users can explore whether the current student differs from the average in his/her skill set and adjust the semester plan if necessary.

3.3. Server-side

To produce an estimation of the academic risk and the uncertainty of this prediction, the LADA dashboard uses a predictive model based on an Adaptive Multilevel Clustering technique (Ochoa, 2016). The model uses data gathered from historical academic records and the current course selection of the student to provide predictions. A pre-processing step was required to clean the data with courses that have changed names, split or joined over time. The model considers two features, the student performance and course load:

- *Student performance* is a good indicator of the probability to succeed in the courses that a student may take. The most common measurement of performance is obtained by averaging the grades acquired in previous courses. Depending on the institution, the average measure can be between 1 and 10, or between 1 and 20 and can be easily obtained for both current and past students from the academic records in the data.
- *Course load* is an indicator of the number of selected courses in relation with the required credits and the classification of the courses.

This model was implemented using the the Python library scikit-learn (Pedregosa et al., 2011). A REST API was implemented on top of this model to enable access by the client.

4. Study design

To strengthen the generalization of the results, we conducted the implementation and experimentation of LADA in two different universities: University A (Europe) and University B (Latin America)². The involved universities use different models of academic advising and different levels of technology adoption in the advising process, as explained in Section 2.1.1 and Section 2.1.2. As in evaluations of the LISSA adviser dashboard (Charleer et al., 2017; Millecamp et al., 2018), we recruited both experienced advisers (experts) and inexperienced advisers (laymen) to assess the potential benefits and challenges with different target user groups. The user groups are defined as follows:

- **Experts:** users who are trained to provide study advise. For example, at University A, experts are trained professionals who provide regular advise to undergraduate students. At University B, professors are trained to provide study advise to their students.
- **Laymen:** users who are familiar with the course program, but are not trained to provide study advise. For example, PhD and master/bachelor students work with students in teaching activities, but they are not experienced, nor trained, to provide study advise to undergraduate students.

In Figure 5, we illustrate the experimental settings used for both institutions. We asked participants to provide academic advise for three specific student cases. Cases were taken from the universities databases and fully anonymized. The chosen advising cases consisted of the following:

- A **low** advising difficulty case: the student has good scores and is following the program without observable delay or major difficulties.
- A **medium** advising difficulty case: the student has some good scores, but experiences difficulties in some other courses that might lead to a possible delay in the program.

- A **high** advising difficulty case: the student has failed many courses and has difficulties to continue with the program.

All of the selected students had begun their studies in 2013 and had completed one year. Using the information of the first study year, participants provided advise on the plan for the next semester. The advise had to be given for two conditions: using the traditional method of providing advise to the students (normal), and using the LADA dashboard. Both experienced (experts) and inexperienced (laymen) users participated in the studies to compare and contrast the utility of LADA for academic advisers and students in a counterbalanced way.

4.1. University A study

To evaluate LADA in University A, we conducted a user study with a total of 12 participants. We divided participants into two groups: academic advisers (experts) and PhD students (laymen). As experts (4F; 2M; mean age: 34.8 years; S.D.: 5.5 years), we selected six employees from the academic advising service of University A who regularly advised on semester planning for undergraduate students. As laymen, we recruited six PhD students (1F, 5M, mean age: 27.3 years, S.D.: 4.2) with no advising experience. Four out of these six PhD students participated in advising sessions during their undergraduate studies.

Normal condition: in the normal condition, we provided the participants with the traditional tools that are used at University A to provide advise to students. The tools include: a set of digital documents detailing the performance of the student in his/her courses during the first year, as well a link to an internal university website. The university website included the following information: a performance chart showing typical development streams of students in the first semester as explained in Section 2.1.1, a page showing the required courses and study points for each year and semester, a detailed page for each course that described requirements and course content. Also, we provided the participants with a blank text file for taking notes and recording their final decision regarding the recommendation of courses for the student in the upcoming semester.

LADA condition: in this condition, we presented the LADA dashboard to the participants in a Web browser. Three tabs were opened, each containing the tool with the data of the three selected student cases. The Web application was presented in a Google Chrome browser using the full-screen of a 27-inch display with a resolution of 2560 x 1440 pixels. Before the study started, an introduction video explained the different cards and two readouts of the application. After this introduction, participants could explore the application without restrictions for five minutes using a test case, together with the researcher answering any questions. The screen was recorded for a usability analysis in order to detect usability problems that were related to the tool operation. Similar as for the normal condition, participants had to provide advise for three student cases.

4.2. University B study

The experimental design for University B was similar to that of University A. Two groups participated in the study: eight

²The names of universities have been anonymized for the double-blind review process.

academic advisers (3F; 5M, mean age: 40 years. S.D.: 7.6 years) and six bachelor students (6M, mean age: 24 years, SD: 1.2 years). Both groups were shown the same introductory video as in the University A setting, and both groups could explore LADA for five minutes using an example case before commencing the study. We also introduced the current web-based system used for academic advising to students, who were inexperienced with such advising tools.

Normal condition: Participants had to use the current web-based information system available at University B which consisted of a set of websites, providing descriptive statistics about student progress.

LADA condition: the settings were identical as in the University A study. The dashboard was provided in a Web browser, with three opened tabs for each case. The user interface of LADA was translated to Spanish. The grading system was adjusted to a scale from 0 to 10. The historical records, course and student skills cards were adapted to the data provided by University B.

4.3. Task

Participants were asked to analyze the study results of the student and give a recommendation for the next semester. More specifically, participants had to complete the following steps with the traditional tools and LADA:

1. Explore the student data for each case.
2. Plan a set of courses for the next semester.
3. Explore potential different scenarios to recommend to the student.

The sequence of the steps is illustrated in Figure 5.

4.4. Recorded measurements

The think aloud protocol was used to collect data from participants while performing these steps. For each participant, we recorded the screen as well as the microphone for later evaluation. We measured the time from the first interaction with the tools until the participant had finished filling out a form with their final recommendation. At the end of the intervention, all the participants had to provide the following information:

- The list of recommended courses to take.
- The study load: the sum of the study points related to the chosen courses.
- An estimation about the advising difficulty of the recommended semester plan: very high, high, medium, low, very low.
- The number of alternative scenarios explored.
- Their confidence with the prediction: very low, low, regular, high, very high.

Once users had worked through all three cases, they were asked to fill out two questionnaires: the System Usability Scale (SUS) (Brooke et al., 1996) and a questionnaire about perceived usefulness on a five-point Likert scale (1: strongly disagree, 5: strongly agree), also giving them the opportunity to provide general feedback in a comments box. Participants were asked to provide feedback to the following statements:

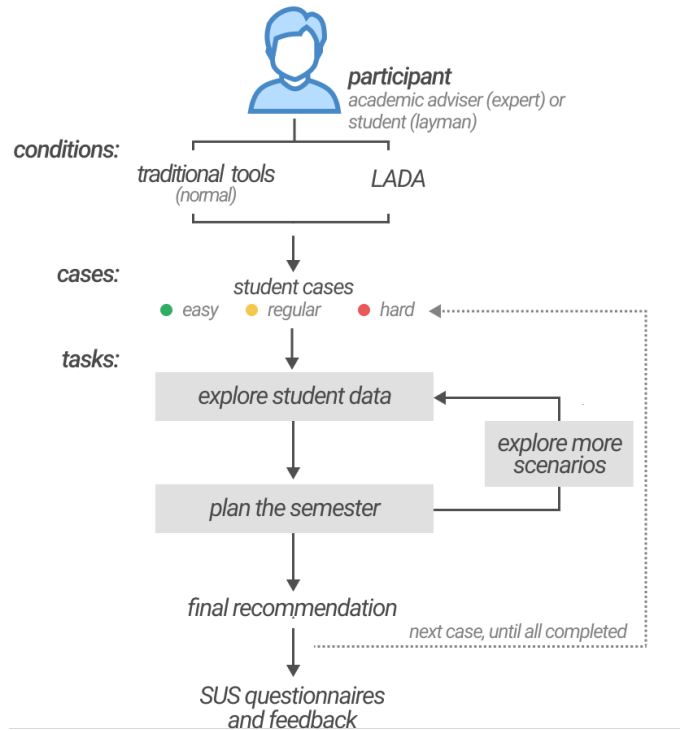


Figure 5: Experiment scenario showing the settings for both studies in University A and University B.

1. The tool helped me to make a more accurate decision in comparison to the traditional approach.
2. The tool helped me to make a decision quicker than I usually do.
3. I consider the tool as helpful.
4. The tool made a complex task easy.
5. The tool helped me understand the relations between the different parts of the data.
6. The tool helped me to better ascertain the performance of the student in relation to others.
7. I would use the tool as part of my analysis/advise session.
8. I would use the system as the only tool of my analysis/advise session.
9. The visualization of uncertainty (prediction quality) influenced my decision.
10. The visualization of uncertainty (prediction quality) is important.

Furthermore, the participants could provide free answers to the following questions:

1. What did you like about LADA?
2. What did you not like about LADA?
3. What are your general feedback and comments?

The study was counterbalanced by condition (normal or LADA) and student cases. It was conducted in a quiet office with one researcher present to conduct the study. The study took about one hour to complete.

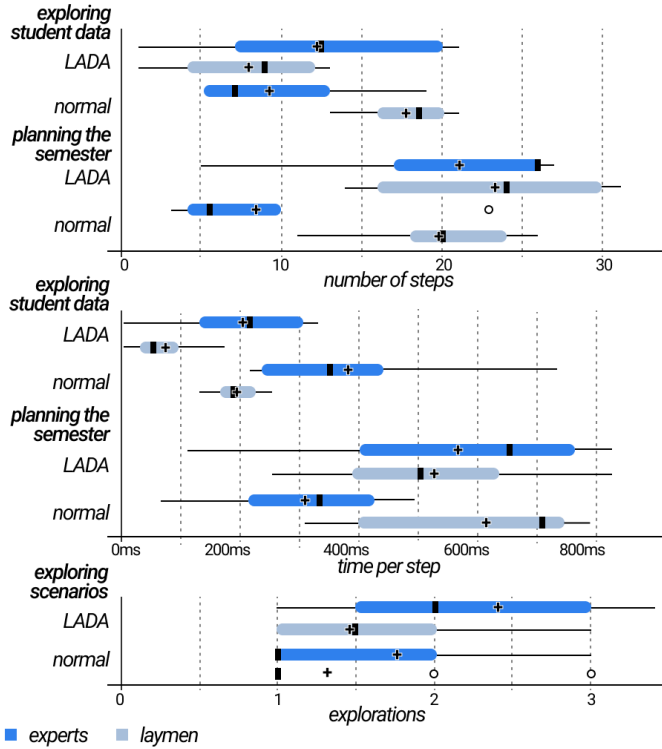


Figure 6: Confidence intervals of the number of steps, time per steps and explorations in the LADA and normal condition in the University A study.

5. Results

In this section, we report the results separately for University A and University B. We first analyze the interaction of the participants with LADA and compare it with the traditional method using quantitative measures. Afterwards, we analyze the qualitative data from the answers of the participants in the questionnaires, together with comments from the participants and observations derived from the think-aloud protocol. In the next section, these findings are discussed.

5.1. University A quantitative results

Due to the small sample size and the data not meeting parametric assumptions, we chose a Kruskal-Wallis test to obtain an indication of the possible effects of LADA on the decision making process. In particular, we investigated whether and to what extent the number of steps, time per step and amount of explorations may be affected by the tool when compared to the traditional method. While the results of quantitative methods applied to small sample sizes should be treated with caution, they nonetheless provide an objective measure with which to interpret the observations if we keep the likelihood of sampling-bias and skewness in mind. An overview of the interaction traces of the participants with LADA and the traditional method (normal) are presented in Figure 6. We describe the results for each step.

Exploring scenarios: experts explored significantly more scenarios ($p < 0.05$) while using LADA ($Mdn = 2$), compared

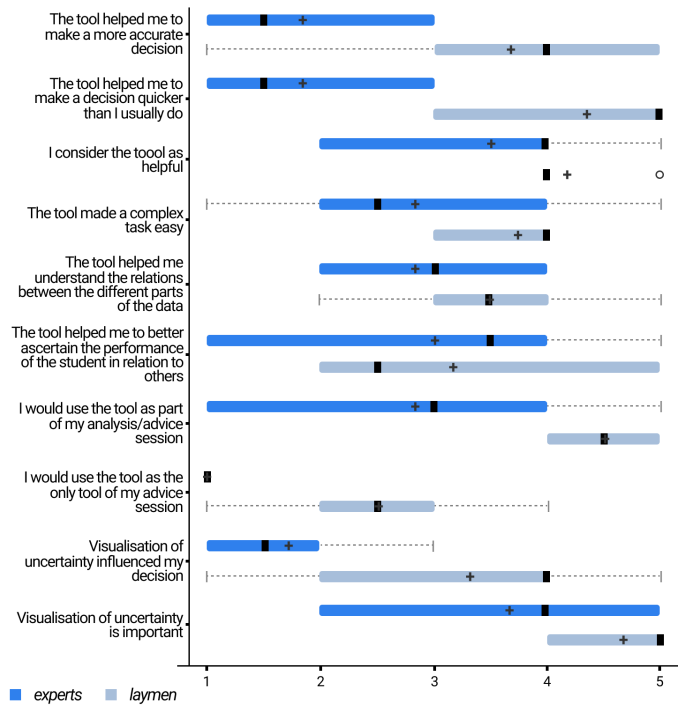


Figure 7: Qualitative feedback from the University A study: Boxplot indicating the median responses to the questionnaire given by Laymen and Experts. The dark line indicates the median, “+” the mean, and dots the outliers. (1: strongly disagree, 5: strongly agree)

to the normal condition ($Mdn = 1$). Laymen also tended to explore more scenarios while using LADA, however the difference with the traditional method was not significant. In Figure 10, we present the explorations sorted by the advising *difficulty* of the presented cases to the participants. We observed a tendency to explore more scenarios with LADA particularly during the high advising cases ($Mdn = 3$ for expert advisers).

Exploring student data: in the normal condition, laymen required significantly more time to complete the task than experts ($p < 0.05$), which can be explained by their limited experience. Also, laymen required significantly less steps ($p < 0.01$) when using LADA compared to the normal condition, indicating a potential time benefit.

Planning the semester: when using LADA, both experts and laymen tended to spend a similar amount of steps and time per step. The experts spend more steps with LADA compared to the normal condition ($p < 0.05$), which can be explained by the fact that they explored significantly more scenarios ($p < 0.05$). In the normal condition, laymen spend more time and steps compared to the experts ($p < 0.05$). In the LADA condition, laymen were equally efficient as experts to complete the tasks, indicating again a potential benefit of the tool for inexperienced users.

5.2. University A qualitative results

For the laymen, an evaluation of the SUS questionnaire resulted in an average score of 72.5 points. For the experts, the score was 70. These results suggest that the usability of LADA was considered as good.

As can be seen in Figure 11, experts indicated to feel more confident with their traditional tool ($Mdn = 4$), compared to LADA ($Mdn = 3$). However, laymen indicated to feel more confident with LADA ($Mdn = 4$) compared to the traditional method ($Mdn = 3$).

A Mann-Whitney test was used to compare the subjective responses from experts and laymen towards their experiences with LADA. Results are presented in Figure 7. Opinions of participants differed significantly ($U = 5.5, p = 0.04$) when asked whether *the tool helped them to make a more accurate decision*: experts tended to be more negative ($Mdn = 1.5$) compared to the laymen ($Mdn = 4$). When asked whether *the tool helped them to make a decision quicker than they usually do*, experts tended to be more negative ($Mdn = 1.5$) compared to the laymen ($Mdn = 5$) ($U = 2, p = 0.009$). Also, when asked whether *they would use the tool as the only tool of the advise session*, experts tended to be more negative ($Mdn = 1.5$) compared to the laymen ($Mdn = 5$).

No other significant differences were found. However, when participants were asked about the *helpfulness of the application*, both experts and laymen tended to rate the tool positively ($Mdn = 4$). When participants were asked whether *the tool made a complex task easy*, the median indicated that experts tended to rate the tool lower ($Mdn = 2.5$) compared to the laymen ($Mdn = 4$). When asked whether *the tool helps them to understand the relations between the different parts of the data*, the median indicated that experts ($Mdn = 3$) and laymen ($Mdn = 3.5$) tended to rate the tool as neutral. When participants were asked whether *the tool helps them to better ascertain the performance of the student in relation to others*, the median indicated that experts ($Mdn = 3.5$) tended to rate the tool higher than the laymen ($Mdn = 2.5$). When participants were asked whether *they would use the tool as part of their advise session*, the median indicated that experts ($Mdn = 3$) tended to rate the tool lower than the laymen ($Mdn = 4.5$). When participants were asked whether *the visualization of uncertainty influenced their decision*, the median indicated that experts tended to rate the tool lower ($Mdn = 1.5$) than the laymen ($Mdn = 4$). Finally, when participants were asked whether *the visualization of uncertainty was important*, the median indicated that experts ($Mdn = 4$) and laymen ($Mdn = 5$) tended to consider the visualization of uncertainty as important.

5.3. Observations and comments

We also collected observations from the participants by analyzing think-aloud data.

Experts tended to compile a possible semester plan verbally while reading the academic history of the student. Then, they quickly jumped to selecting courses for the upcoming semester without looking at the performance of the student in relation to other students or the prediction of LADA. They only consulted the prediction once they had completed the course selection.

Some of the experts tried to increase the chance of success by checking the skills of the students and selecting courses accordingly. One expert did not trust the tool: *“I never trust models. I never trust models other people make, I don’t understand them. For me this is magic. I am a scientist, so it is important for me*

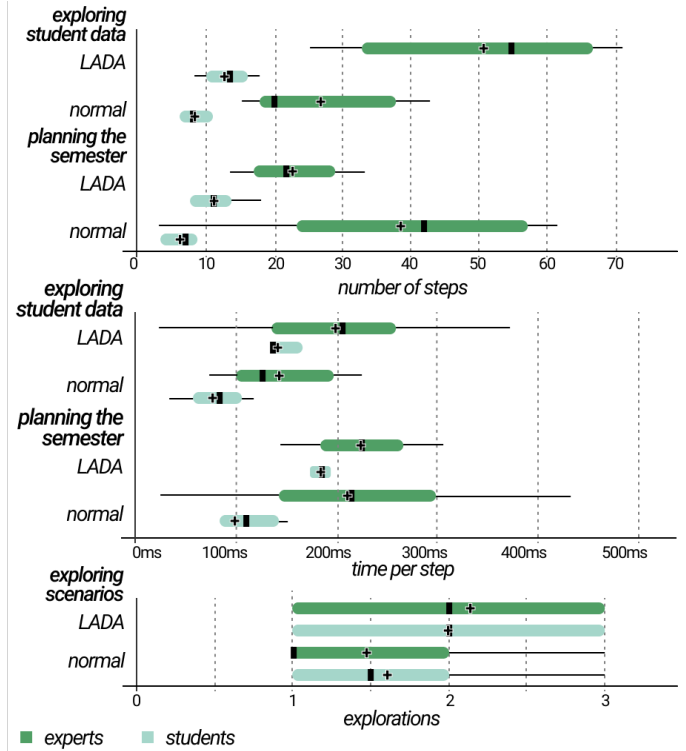


Figure 8: Confidence intervals of the number of steps, time per steps and explorations during LADA and normal conditions in the University B study.

to understand these to be able to use them”. Some users felt the need to investigate why the prediction was so low, particularly in the high advising case. Most of the expert users tried out different combinations of scenarios to gauge the impact of simulated course scenarios on the prediction. Some expert users wanted to see more background information about the student. They all tried to enable the student to achieve the Bachelor degree in 3 years by focusing on planning sufficient study credits to take up in the next semester.

Some laymen explored each card in detail, seemingly to ensure they had all the information they needed to compile a good semester plan. All of the laymen were observed to be strongly influenced in their course choices by the changing prediction. There was a clearly visible effort to achieve the highest possible prediction by selecting courses one-by-one and frequently checking the semester plan against the predicted performance. When the laymen made changes to the semester plan, especially those that had a negative impact on the students chance of success, they frequently consulted the skill cards to explore the reason for the change and to then adjust their recommendation. Users seemed to take great care to match the course list to the students skills.

5.4. University B quantitative results

To investigate the results, a Kruskal-Wallis test was conducted to compare the effects of explored scenarios (number of steps, time per step and explorations) between LADA and the traditional method. The overview of the interaction traces

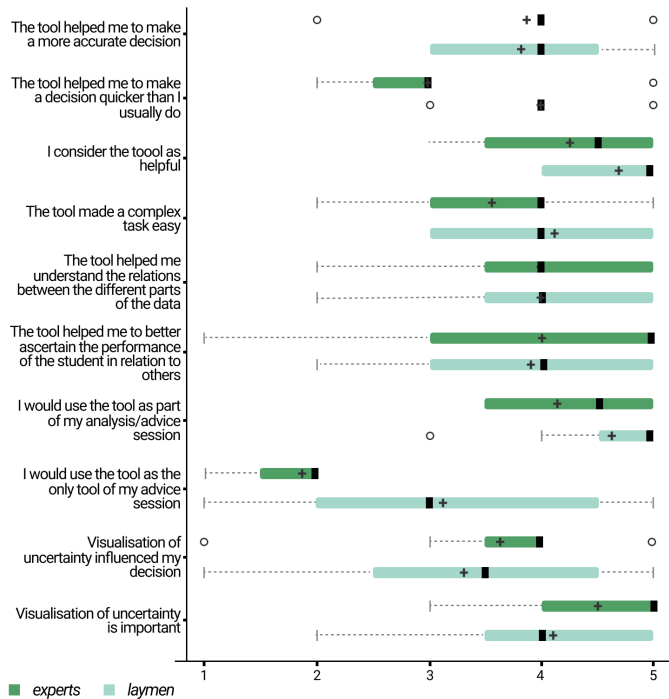


Figure 9: Qualitative feedback from the University B study: Boxplot indicating the median responses to the questionnaire given by Laymen and Experts. The dark line indicates the median, “+” the mean, and dots the outliers. (1: strongly disagree, 5: strongly agree)

of the participants with LADA and the traditional method (normal) are presented in Figure 8.

Exploring scenarios: experts explored significantly ($p < 0.05$) more scenarios while using LADA ($Mdn = 2$) compared to the normal condition ($Mdn = 1$). Also, laymen tended to explore more scenarios while using LADA ($Mdn = 2$) compared to the normal condition ($Mdn = 1.5$). In Figure 10, we present the explorations sorted by the difficulty of the presented cases.

Exploring student data: experts tended to spend more steps and time per step compared to the laymen in both conditions. Both experts and laymen spent more time per step and steps when using LADA compared to the normal condition, which can be explained by the fact that they explored significantly ($p < 0.05$) more scenarios in LADA than in the normal condition.

Planning the semester: when using LADA and the normal condition, experts tended to spend more steps and time per step compared to laymen. Notably, the experts required significantly more steps ($p < 0.05$) in the normal condition to complete the planning of the semester compared to laymen. Experts and laymen spend more time per step and steps while using LADA compared to the normal condition, as they explored significantly more scenarios.

5.5. University B qualitative results

For the laymen, the evaluation of the SUS questionnaire resulted in an average score of 71.25 points. For the experts, the score was 56.25. This suggests that laymen found the tool to

be reasonable usable, while the experts appeared to somewhat struggle with its operation.

As can be seen in Figure 11, both experts and laymen indicated to feel confident with the traditional tool and LADA ($Mdn = 4$).

A Mann-Whitney test was used to compare the subjective responses from experts and laymen towards their experiences with LADA. Results are presented in Figure 9. Opinions of participants differed significantly ($U = 16, p = 0.01$) when asked whether *the tool helped them to make a decision quicker than they usually do*: experts tended to have a more neutral attitude ($Mdn = 3$) compared to laymen ($Mdn = 4$). When asked whether *they would use the tool as the only tool of the advice session*, experts tended to be more negative ($Mdn = 2$) compared to laymen ($Mdn = 3$) ($U = 22.5, p = 0.04$).

No other significant differences were found. However, when asked whether *the tool helped them to make a more accurate decision*, the median indicated a positive attitude from experts and laymen ($Mdn = 4$). When asked about the *helpfulness of the application*, both experts ($Mdn = 4.5$) and laymen ($Mdn = 5$) tended to rate LADA more positively. When participants were asked whether *the tool made a complex task easy*, the median indicated that both experts and laymen tended to rate the tool positively ($Mdn = 4$). When participants were asked whether *the tool helps them to understand the relations between the different parts of the data*, the median indicated that both experts and laymen tended to rate the tool positively on this aspect ($Mdn = 4$). When participants were asked whether *the tool helps them to better ascertain the performance of the student in relation to others*, the median indicated that experts ($Mdn = 5$) tended to rate the tool higher than laymen ($Mdn = 4$). When participants were asked whether *they would use the tool as part of their advice session*, the median indicated that experts ($Mdn = 4.5$) and laymen ($Mdn = 5$) both rated this aspect positively. When participants were asked whether *the visualization of uncertainty influenced their decision*, the median indicated that experts tended to rate the tool slightly higher ($Mdn = 4$) than laymen ($Mdn = 3.5$). Finally, when participants were asked whether *the visualization of uncertainty was important*, the median indicated that both experts and laymen ($Mdn = 5$) tended to consider the visualization of uncertainty as important.

5.6. Observations and comments

We also collected observations from the participants by analyzing the think-aloud data.

Participants acknowledged that LADA is an interesting tool because of the possibility of making different combinations of courses: *“it helped me to make different combinations of courses to see what courses were better in the semester.”*

Participants indicated that they liked the prediction, and the possibility of comparing with other students based on similar experiences: *“It can help to choose courses, especially if one does not have references of how challenging the course is. Also, the visualization of skills is useful in specific areas that a course requires.”*

Participants also acknowledged the usefulness of the tool in the decision-making: *“The user interface was very fulfilling*

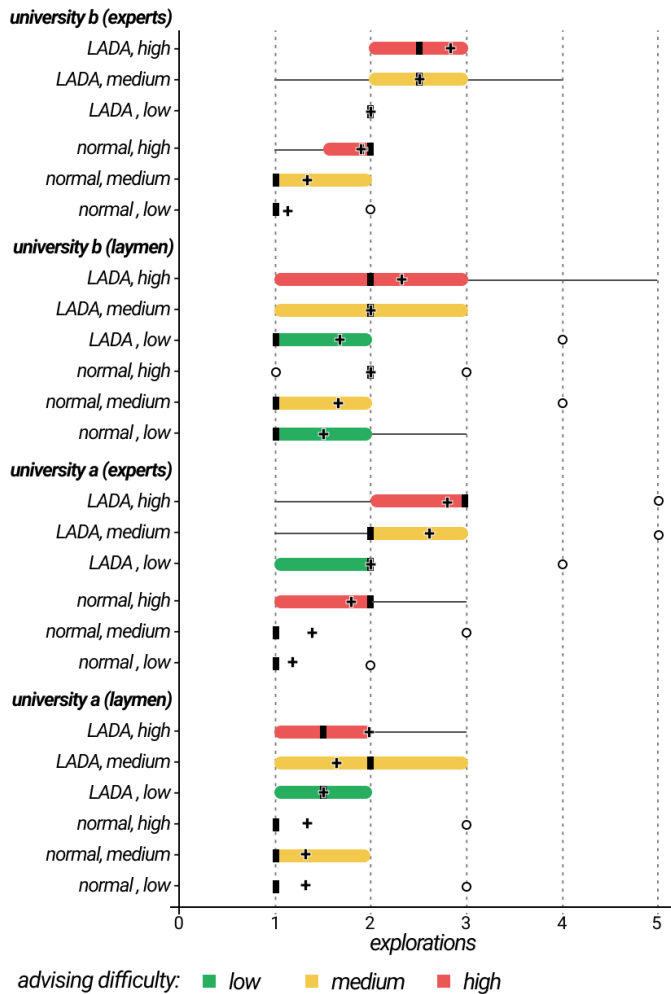


Figure 10: The mean with confidence intervals of the explorations made by the participants of the University A and University B studies, sorted by the advising difficulty of the presented cases (low, medium and high) and the condition: normal or LADA.

with the information presented, together with the synchronization in real time and that allowed me to compare with similar students. This helped me to make certain decisions.”

Participants indicated the usefulness of the tool towards improving their scores by planning their studies accordingly: “I think this was wanted by all the students since we entered here. I think that LADA will be very useful for us because it will reduce the failure rate.”

Participants indicated the ease of use of the system, and highlighted the short learning-curve: “The system at first is a little complicated to understand, after using it for a moment the system becomes easy to use and very comfortable.”

Participants also indicated that they paid attention to the quality of prediction component, indicating that it influenced their decisions: “When I observed that this was low, I changed the course for another and if I checked that it did not drop I removed a course that was not aligned with my abilities. I would not risk taking a course that doesn’t match my abilities.”

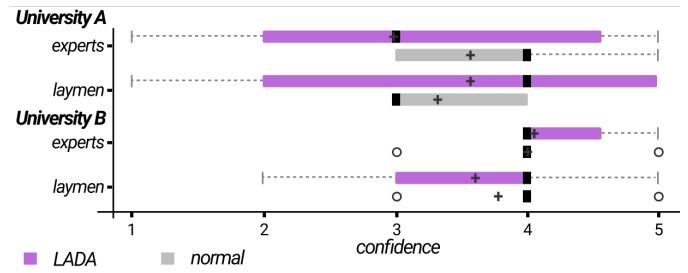


Figure 11: Confidence of the participants with their final predictions while using LADA and normal conditions for both University A and University B. From one: very low to five: very high.

6. Discussion

6.1. General results

Overall, participants tended to explore more scenarios with LADA than in the normal condition. What we observed during the sessions is that particularly the experts explored significantly more scenarios with LADA ($Mdn = 2$) than they would usually do ($Mdn = 1$), analyzing the history of a student and course combination possibilities, and then making a careful decision. The presence of the academic risk prediction information enables the experts to try different combinations of courses and select those that seem to generate a lower risk, both in their perception and in the computer-generated prediction. For laymen, LADA is perceived as a valuable tool for more accurate decision-making.

In general, these results are encouraging, as they indicate that LADA can offer support to make informed data-driven decisions: for experts by exploring more and different scenarios in the same amount of time and for laymen to increase confidence in the decision-making process. In the University A study, experts felt more motivated compared to the students to explore more scenarios and came to different conclusions for each of the cases. For laymen, LADA was efficient, as they were able to make informed decisions in a similar amount of time and steps compared to the experts. In this study, the overall experience of laymen was more positive than the experience of experts. Overall, experts felt that the representation of uncertainty was necessary, but they were not sure about the quality of the predictions. An important future line of research is to support transparency of the algorithm, so that end-users can understand how the algorithm works.

Usability results of the University A study indicate that the overall usability was perceived as good (SUS score ≥ 70 for both laymen and experts). In general, however, experts felt more confident with their traditional tools compared to LADA, whereas laymen indicated to feel more confident with LADA. The main reasons are evident from responses to our questionnaire: experts do not perceive the same time or accuracy benefits as laymen, which can be explained by the fact that experts also rely on a lot of expertise and knowledge. Laymen on the other hand often lack this knowledge and perceive LADA as a great help in making accurate and quick decisions. Both experts and laymen do find LADA helpful, and were in general positive about the overall experience as reflected in the SUS scores.

In the University B experiment, both groups of participants tended to be positive with their feedback towards their experiences with LADA. Similar to the University A study, the experts explored significantly more scenarios. They inspected the information of the students carefully and mainly explored more possible scenarios while using LADA compared to the traditional method, particularly for high advising cases with students that experienced most problems. They thought that LADA was a useful tool, providing compelling visual representations of the data. Notably, they indicated that the visualization of uncertainty is essential, and they felt confident with the quality representation. Students tended to spend fewer steps and time compared to the experts. They indicated that they felt motivated to explore the scores and how they performed compared to other students in the courses. Both experts and laymen of University B also asked for a more transparent representation of the prediction algorithm.

Usability results of the University B study indicate that the overall usability was perceived good for laymen (SUS score of 71.25), but only moderate for experts (SUS score of 56.25), hinting to potential usability issues. In this study, both experts and laymen indicated to feel equally confident with the traditional tool and LADA. In addition, they both perceived an accuracy benefit in decision making. Similar to the University A study, experts are interested in using LADA as an additional tool, but would not use LADA as the only tool during an advising session. They also do not perceive the same time benefit as laymen. This result indicates that additional functionality may be needed to fully cover the needs of experts during an advising session.

There are a few similarities and differences between the two studies. One of the main differences between the studies is the academic advising setting that is used. At University A, academic advising is done by professionals: i.e., trained academic advisers, whereas at University B advising is done by professors. A difference that can be observed in the study data is the lack of trust of professionals at University A in the prediction model. Professors of University B trusted the model, and the overall perceived usefulness was higher. Although professionals of University A also acknowledged the potential, they did not trust the prediction model and preferred to rely on their own expertise. A key issue that was identified was the lack of transparency of the underlying prediction model. We are currently working on visualizations that represent the clustering technique as a way to improve the confidence in the prediction outcome. Although this issue was also raised by both experts and laymen of University B, the overall confidence in the outcome was not influenced in this setting. The second difference is the use of tools in both universities: whereas at University B several tools are already in place for academic advising, University A still relies on manual effort to support advising sessions. The lack of technology use may be a second reason for the lower perceived usefulness scores in this setting.

Although the overall perceived usefulness was lower for experts than laymen in University A, we observed also many similarities in the two studies. Experts used LADA intensively for exploring multiple scenarios, such as different course combi-

nations. Particularly for high advising cases, such explorations were done very frequently. We also observed that laymen felt confident with LADA in both studies. In addition, they were able to make well informed decisions in a more efficient way. For experts, LADA was perceived as very useful to ascertain the performance of students in relation to peers: so while the prediction component still needs further work, the visualization and exploration of student data and comparison to peers seems to be a key component for academic advising support.

6.2. Design implications

Comparing to related studies that target study advisers, we can observe a few similarities and differences that can inform the design of dashboards for study advisers. Two types of dashboards have been elaborated to support study advisers: dashboards that are mostly descriptive in nature, presenting data about academic performance to advisers such as the LISSA dashboard (Charleer et al., 2017) and E-Advisor (Phillips, 2013), and dashboards that use predictive algorithms to try to infer additional information about study behavior that can inform study advisers, such as Bridge and Student Explorer (Aguilar et al., 2014). The latter example provides access to both achievement and engagement estimates of students based on data from the learning management systems, including login frequency (Lonn et al., 2013), to provide feedback to advisers as to whether students need to be “encouraged” to keep doing well, “explored” in more detail or immediately “engaged”.

The LISSA dashboard was evaluated with four study advisers (Charleer et al., 2017). In total, 15 sessions were observed. Results indicate that there is a difference in use between experienced study advisers and inexperienced study advisers. Experienced study advisers used the dashboard as a backup, glancing at it when needed. Inexperienced study advisers used the dashboard as a guide through the entire session. By recording the number and level of insights that were supported in these sessions, it was observed that LISSA was particularly helpful for inexperienced study advisers. In a follow-up study (Millecamp et al., 2018), the use of LISSA was observed with the same study advisers to try to gain insight into the utility of the dashboard for different student groups. Similar to results of our LADA study, results indicate that the dashboard was used most intensively to help students who struggle the most. More specifically, the number of insights obtained through the dashboard was higher for students doubting to continue the study program as opposed to easier decisions with regard to which particular course to retake. These results are inline with our results and indicate that learning analytics dashboards may be particularly useful for high advising cases.

Results of Bridge and Student explorer indicate that study advisers used the student explorer dashboard twice as often during a session with a student compared to before the session (Aguilar et al., 2014). In addition, use of the dashboard after the session was negatively associated with course grades (Lonn et al., 2015). Also in this study, the dashboard was used more frequently for the “high advising cases”, in this case students with lower grades.

Based on these results, we can infer a few recommendations for implementing dashboards for study advisers.

1. *Exploration support for high advising cases.* An interesting outcome that can be observed in different studies is that learning analytics dashboards for study advisers are particularly used more intensively for “high advising” cases, which include students with lower grades (Lonn et al., 2015) and students in doubt of continuing with the study program (Millecamp et al., 2018). Results of our study indicate that expert advisers like to explore more scenarios for such students to better understand which course selections would be a better fit. Multiple scenarios are explored with the dashboard by experts for such cases, as illustrated in Figure 10. Enabling advisers to explore the potential impact of different possibilities, thus, seems to be an interesting way forward for this particular class of learning analytics dashboards. Research in this line of “What If” scenarios has been put forward by the Intelligent User Interfaces community (Lim and Dey, 2010) and seems to offer a promising research direction for learning analytics dashboards.
2. *Actionable and immediate feedback.* To support “What If” scenarios, purely descriptive dashboards of course grades such as LISSA (Charleer et al., 2017) fall short of their potential. Dashboards such as the Purdue Course Signals (Arnold and Pistilli, 2012) and Student Explorer (Lonn et al., 2013) try to make use of other data that support timely interventions, by using additional data from learning management systems (as opposed to grades data only) and predictive algorithms. The objective is to support actionable and immediate feedback, instead of only being able to give feedback when course grades are final. In our LADA study, we used a predictive approach that uses grades, currently booked courses, and number of credits of the courses. The data from previous students is used to predict the chance of success or academic risk of the student using multi-level clustering. Compared to the descriptive LISSA dashboard that presents grades only (Charleer et al., 2017), it is interesting to observe that experienced study advisers used the LADA dashboard intensively, whereas they used LISSA mainly as a back-up (Charleer et al., 2017). To support experienced advisers, predictions thus seem an interesting way forward, as these predictions can complement their broad range of available expertise and provide additional insight. That said, there is a strong need for explaining the provenance of recommendations: similar to work of Lim and Dey (Lim and Dey, 2010), the rationale of predictions generated by different prediction algorithms needs to be explained to enable advisers to assess the overall quality of the prediction.
3. *Explainable predictions.* Data of our study demonstrates that data-driven decision support is valuable for both experts and laymen. Explanations of recommendations are needed to enable advisers to inspect the quality of recommendations. In our dashboard, we used a quality indicator

of the prediction models to better inform advisers about the quality of the prediction. Such indicators are insufficient to support expert advisers: such advisers want to understand the reasoning of a prediction. An interesting further line of research is to expand the current dashboard with visualizations of both the prediction model and the underlying data to increase confidence and trust in study advise scenarios.

6.3. Limitations

There are some limitations to this study that need to be articulated. First, the sample size of our study is comparably small due to the limited availability of expert users. For our work, study advisers and professors trained to provide study advice had to be recruited in two universities. These expert users are difficult to recruit due to their small numbers and tight time schedules. To generalize and validate the findings of our study, we replicated the study at two different universities that use different advising models. Our sample size was based on other learning analytics dashboard studies such as LISSA (Charleer et al., 2017), where five study advisers were recruited, and should provide a solid basis for deriving meaningful insights. The most common sample size Human-Computer Interaction studies is 12 (Caine, 2016): the validity of such sample sizes is justifiable by expert recommendations and data observes that are then re-evaluated. In total, 12 participants completed the study at University A, and 14 participants completed the study at University B: the validity is ensured by observations in previously collected data (both in our own study as in the LISSA study with five expert users), as well as the replication in two different settings. The explicit goal of deploying the dashboard in two different universities with different advising models was to enable generalizability of the results.

7. Conclusions

In this paper, we have introduced LADA, a learning analytics dashboard for supporting advisers in decision-making through comparative and predictive analysis. We evaluated the use of LADA in experimental settings that took place in two different universities with experts and non-expert participants. We have compared the learning analytics tool with more traditional procedures and tools. Results indicate that participants found the tool appealing, but we found that more transparency in the algorithm prediction is required in order to increase confidence. One aspect to point out, however, is the observation that LADA enabled advisers to evaluate a greater number of scenarios in a similar amount of time before making a final decision, particularly for difficult cases. Thus, the ability of the tool to increase the number of potential avenues that are evaluated with regards to the student’s future development, directly enables a better-informed decision-making process. Due to the increased amount of explorations in the same time window, decisions can be better-founded at no extra cost. For laymen, LADA is perceived as a valuable tool for more accurate decision-making. These results are no doubt encouraging for further developments in the field. In this regard, future work should consider

a longitudinal study that compares the academic performance of students that have been advised with the tool to those that have been advised with traditional methods. If the effect of the greater amount of explorations enabled by the tool could be determined, future development could be more targeted and the role of computers in predicting human behavior better understood.

References

- Aguilar, S., Lonn, S., Teasley, S. D., 2014. Perceptions and use of an early warning system during a higher education transition program. In: Proceedings of the Fourth International Conference on Learning Analytics And Knowledge - LAK '14. ACM Press, New York, New York, USA, pp. 113–117.
- Arnold, K. E., Pistilli, M. D., 2012. Course signals at purdue: Using learning analytics to increase student success. In: Proceedings of the 2nd international conference on learning analytics and knowledge. ACM, pp. 267–270.
- Bahr, P. R., 2008. Cooling out in the community college: What is the effect of academic advising on students chances of success? *Research in Higher Education* 49 (8), 704–732.
- Barron, K. E., Powell, D. N., 2014. Options on how to organize and structure advising. *Academic Advising*, 14.
- Bezdek, J. C., Ehrlich, R., Full, W., 1984. Fcm: The fuzzy c-means clustering algorithm. *Computers & Geosciences* 10 (2-3), 191–203.
- Brier, G. W., 1950. Verification of forecasts expressed in terms of probability. *Monthly Weather Review* 78 (1), 1–3.
- Brooke, J., et al., 1996. Sus-a quick and dirty usability scale. *Usability evaluation in industry* 189 (194), 4–7.
- Caine, K., 2016. Local standards for sample size at chi. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. CHI '16. ACM, New York, NY, USA, pp. 981–992.
URL <http://doi.acm.org/10.1145/2858036.2858498>
- Campbell, J. P., Oblinger, D. G., et al., 2007. Academic analytics. *Educause Quarterly*, 1–20.
- Chan, Z. C., 2016. A qualitative study of freshmen's and academic advisors' perspectives on academic advising in nursing. *Nurse education in practice* 18, 23–29.
- Charleer, S., Moere, A. V., Klerkx, J., Verbert, K., De Laet, T., 2017. Learning analytics dashboards to support adviser-student dialogue. *IEEE Transactions on Learning Technologies*.
- Choi, S. P. M., Lam, S., Li, K. C., Wong, B. T. M., 2018. Learning analytics at low cost: At-risk student prediction with clicker data and systematic proactive interventions. *Journal of Educational Technology Society* 21 (2), 273–290.
URL <http://www.jstor.org/stable/26388407>
- Clarke, P., 2008. When can group level clustering be ignored? multilevel models versus single-level models with sparse data. *Journal of Epidemiology & Community Health* 62 (8), 752–758.
- Drachler, H., Verbert, K., Santos, O. C., Manouselis, N., 2015. Panorama of recommender systems to support learning. In: *Recommender systems handbook*. Springer, pp. 421–451.
- Drake, J. K., 2011. The role of academic advising in student retention and persistence. *About Campus* 16 (3), 8–12.
- Ender, S. C., Winston, R. B., Miller, T. K., 1982. Academic advising as student development. *New directions for student services* 17 (17), 3–18.
- Fritz, J., mar 2011. Classroom walls that talk: Using online course activity data of successful students to raise self-awareness of underperforming peers. *The Internet and Higher Education* 14 (2), 89–97.
- Gneiting, T., Raftery, A. E., 2007. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association* 102 (477), 359–378.
- Gordon, V. N., Habley, W. R., Grites, T. J., 2011. *Academic advising: A comprehensive handbook*. John Wiley & Sons.
- Grites, T. J., 1979. *Academic advising: Getting us through the eighties*. aaehc/higher education research reports, no. 7, 1979.
- Grites, T. J., 2013. Developmental academic advising: A 40-year context. *NACADA Journal* 33 (1), 5–15.
- King, M., 2012. Co-tutor: a relationship management system to enable staff to monitor students' engagement and provide support to at risk students.
- Lim, B. Y., Dey, A. K., 2010. Toolkit to support intelligibility in context-aware applications. In: Proceedings of the 12th ACM international conference on Ubiquitous computing. ACM, pp. 13–22.
- Lonn, S., Aguilar, S., Teasley, S. D., 2013. Issues, challenges, and lessons learned when scaling up a learning analytics intervention. In: Proceedings of the third international conference on learning analytics and knowledge. ACM, pp. 235–239.
- Lonn, S., Aguilar, S. J., Teasley, S. D., 2015. Investigating student motivation in the context of a learning analytics intervention during a summer bridge program. *Computers in Human Behavior* 47, 90–97.
- Lonn, S., Krumm, A. E., Waddington, R. J., Teasley, S. D., 2012. Bridging the gap from knowledge to action: Putting analytics in the hands of academic advisors. In: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge. ACM, pp. 184–187.
- Millicamp, M., Gutiérrez, F., Charleer, S., Verbert, K., De Laet, T., 2018. A qualitative evaluation of a learning dashboard to support advisor-student dialogues. In: Proceedings of the 8th International Conference on Learning Analytics and Knowledge. ACM, pp. 56–60.
- Miller, M. A., 2012. Structuring the conversation: Shifting to four dimensional advising models. *Carlstrom*, A.
- Ochoa, X., 2016. Adaptive multilevel clustering model for the prediction of academic risk. In: 2016 XI Latin American Conference on Learning Objects and Technology (LACLO). pp. 1–8.
- Okewu, E., Daramola, O., Oct 2017. Design of a learning analytics system for academic advising in nigerian universities. In: 2017 International Conference on Computing Networking and Informatics (ICCNi). pp. 1–8.
- Papamitsiou, Z. K., Economides, A. A., 2014. Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Educational Technology & Society* 17 (4), 49–64.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12, 2825–2830.
- Phillips, E. D., 2013. Improving advising using technology and data analytics. *Change: The Magazine of Higher Learning* 45 (1), 48–55.
- Ranjan, J., Malik, K., 2007. Effective educational process: a data-mining approach. *Vine* 37 (4), 502–515.
- Sharkin, B. S., 2004. College counseling and student retention: Research findings and implications for counseling centers. *Journal of College Counseling* 7 (2), 99–108.
- Siemens, G., Baker, R., 2012. Learning analytics and educational data mining: towards communication and collaboration. In: Proceedings of the 2nd international conference on learning analytics and knowledge. ACM, pp. 252–254.
- Steele, G. E., 2016. Technology and academic advising. *Beyond Foundations: Developing as a Master Academic Advisor*, 305.
- Stoneham, R., 2015. Failing students need big data and learning analytics: Hype or reality? *Compass: Journal of Learning and Teaching* 7 (11).
- Van Oirbeek, R., Lesaffre, E., 2016. Exploring the clustering effect of the frailty survival model by means of the brier score. *Communications in Statistics-Simulation and Computation* 45 (9), 3294–3306.
- Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., Klerkx, J., 2014. Learning dashboards: an overview and future research opportunities. *Personal and Ubiquitous Computing* 18 (6), 1499–1514.
- Viberg, O., Hatakka, M., Blter, O., Mavroudi, A., 2018. The current landscape of learning analytics in higher education. *Computers in Human Behavior* 89, 98 – 110.
- Wilks, D. S., 2011. *Statistical methods in the atmospheric sciences*. Vol. 100. Academic press.
- Winston, R. B., 1984. *Developmental academic advising*. Jossey-Bass Incorporated Pub.
- Wiseman, C. S., Messitt, H., 2010. Identifying components of a successful faculty-advisor program. *NACADA Journal* 30 (2), 35–52.
- Young, R., Valach, L., 2016. Action and language: Contextual action theory in counselling. *Psychologie Française* 61 (1), 31–42.
- Young-Jones, A. D., Burt, T. D., Dixon, S., Hawthorne, M. J., 2013. Academic advising: does it really impact student success? *Quality Assurance in Education* 21 (1), 7–19.