

TELECI Approach for e-Learning User Behavior Data Visualization and Learning Support Algorithm

Atis KAPENIEKS, Iveta DAUGULE, Kristaps KAPENIEKS,
Viktors ZAGORSKIS, Janis KAPENIEKS Jr, Zanis TIMSANS,
Ieva VITOLINA

Distance Education Study Centre, Riga Technical University,
Kronvalda bulvaris 1, Riga, Latvia, LV1010

Atis.Kapenieks@rtu.lv, Iveta.Daugule_1@rtu.lv,
Kristaps.Kapenieks@gmail.com, Viktors.Zagorskis@rtu.lv,
jkapenieks@gmail.com, Zanis.Timsans@gmail.com,
Ieva.Vitolina@gmail.com

Abstract. The user behavior data generated in the TELECI learning environment with additional short, easy-to-use multiple-choice questions before and after each content subunit are used for visualization and correlation analysis. Three user behavior data clusters were identified in data landscape. The student behavior change among the TELECI-clusters was used for TELECI learning support algorithm design. The student performance data before and after learning the e-content were used for knowledge acquisition model design. This model is based on the assumption that knowledge acquisition of real e-content can be quantified by superposition of the impact of learning “perfect” content, too easy content, and too complicated content. The learner knowledge acquisition surface is calculated on this assumption. The data of real course learner knowledge acquisition are located on this surface as “telecides”. Telecides are the visualization of the appropriateness of an e-content unit for the needs of the specific learner or learners target group.1

Keywords: E-Learning, Learning Activity, Learning Data, Motivation, Telecide, User Behavior

1. Introduction

The main challenges in eLearning are content delivery quality, content quality, and the correspondence of content and delivery to learning needs and objectives.

The aim of the study is to create the complete landscape model of learner behavior and knowledge acquisition data and mapping the real learner performance data on it. These approaches are needed for building the TELECI learning forecasting and support algorithms.

Robinson and Cook point out that the metrics that include only evidence of the student's presence in the course “speak neither to the quality of the activity nor to its “stickiness”. Similarly, such indicators are poor predictors of repeat engagement (student motivation). Like the disengaged student in a physical classroom, logging in and out

does not prove much beyond the fact that a student entered and exited the learning environment” (Robinson and Cook, 2018).

To avoid the problem mentioned above, authors examined the student activity data in the context of a system of short e-content subunits and multiple-choice questions that are included in the course content. The goal of this system is to transform the student from a passive observer into an active thinker, thus facilitating his / her knowledge acquisition.

In our previous study (Daugule and Kapenieks, 2018) we found out that to determine the students’ initial motivation, there are several aspects that should be considered. The results of this study showed that students’ initial motivation aspects are complex. One aspect is confirmation that the evaluation given directly by students of their skills and the time required for their development without further processing is not applicable to the development of the course content and the learning material selection algorithm.

Also, the answers in the questionnaires should not be used as the only substantiation of the adaptation of the learning content to the students' needs.

The study (Daugule and Kapenieks, 2018) concluded that it is primarily necessary to evaluate student provided data in the context of student-generated data in an e-Learning environment rather than information presented by students using different questionnaires. Also, the information about their future intentions from this aspect is secondary and cannot be applied directly.

Other research in the field (Tsoni et al., 2019) also highlights the need to collect and analyze student data to keep track of their progress, emphasizing the need to create augmented reality environments that deliver it in real time. The authors of this study note that the availability of this type of data would reduce negative aspects such as poor grades and dropouts.

In our view, it is essential that the developed learning environment should be focused on the most important issues thus achieving possibly higher efficiency.

According to Koch it is possible and necessary to apply the 80/20 Principle in education. This Principle supplies the hypothesis that there are a few important reasons that explain superior educational performance, and that a few approaches or methods will prove to have exceptional results. To do this, it is necessary to distinguish between the approaches that bring the greatest benefits, and then multiply their incidence (Koch, 1998).

According to Szulanski's hypotheses, knowledge stickiness positively correlates with causal uncertainty, lack of motivation on the part of the knowledge source and recipient, lack of credibility of the source, lack of knowledge transfer and receiving capacity, poor organizational context and arduous relationship between the provider and the recipient (Szulanski, 2003).

The authors identified that motivation and its properties thereof is essential factor for the student's success.

Taking into account the problem highlighted in other studies (Robinson, Cook, 2018) to also consider the clicks made in the course in context with other parameters to determine the point where the student has lost focus and, instead of being an active learner, has become a presence imitator.

Reducing this circumstance was one of the tasks of the questions system developed by the authors - to keep the student's attention by regularly engaging in answering questions in the context of the subject being learned. The e-course used in the study was adapted to generate more user behavior data in each course unit. This was ensured by the placement of relevant questions at the beginning and end of each topic. Students were informed that the answers to these questions would not be considered in the final

assessment, however, answering the questions is part of the study process. The students accepted additional questions as motivating and helpful.

2. Course design for sufficient user behavior data generation

Considering that other studies (Gopalan et al., 2017) have found that intrinsic motivation and academic achievement share significant and positive bonding and such motivation is able to spread the positivity and ensure long-term sustainability of the gained knowledge, we focused our research on this type of motivation.

Although external evaluation may motivate action, it may not produce the desired result in the long run. Students can strive for rewards rather than their knowledge, that's why it is very important that students are motivated by intrinsic motivation (Gopalan et al., 2017).

Intrinsic motivation comes from the satisfaction of successfully completing a task. It does not directly affect academic success. However, it is strongly correlated to independent learning and cognitive strategy use. Nevertheless, Y Generation students tend to be extrinsic learners (Levy and Campbell, 2008).

Also, the positive impact of support on student motivation was highlighted. The findings of another study (Dennis et al., 2005) show that both personal / career motivation and lack of peer support are important predictors of college commitment, even when the impact of academic ability, as indicated by the high school Grade Point Average, is controlled. Although these college students support both individualistic and collectivist motivation, having a more individualized motivation is an important personal characteristic that at the end of the year provides a sense of satisfaction with the college environment and a commitment to graduate. In addition, the lack of contextual resources, such as peer support, results in lower grades and adjustments later that year. The authors of this study point to the need to create an ecosystem that includes the necessary support for students and helps them adapt to the study process in order to positively influence their learning success.

One of the factors used to predict student success is the students' initial motivation. However, as shown by previous research in this area, the issue is ambiguous. A study conducted in a computer science student group concluded that students are coming to introductory computer science courses with the positive motivational dispositions necessary to succeed, however, these entering motivations are not necessarily motivating course achievement. An exception is highly selective honors students. The conclusion of this study is the suggestion that the focus needs to be on within-course motivational and instructional strategies, and what instructors do affects students' motivation (Shell et al., 2016).

While investigating the initial motivation of students in an adult education program, the authors (Simpson, 1997) of the study found that adult education providers need to identify the motivational nature of their participants and to take this into account when managing their program. The study concluded that the interaction between each unique individual and environmental factors is important. Researchers point out that successful programs are unlikely to be carried out without the readiness to determine the characteristics of this interaction and the commitment to respond appropriately.

Therefore, our study focused on identifying and continuously monitoring the activity levels. The data obtained were compared with the success of the student's learning, thus

approaching identification of the most important parameters to follow along with the teaching process.

In order to provide the necessary environment for the study, the blended learning study course that was used in the previous study was updated. The e-course used in the study was designed in the Sakai e-learning environment. This study includes research on 61 student activities and learning outcomes in a Basic Business study course.

Each unit was split into 3 to 8 subunits. Each subunit was completed with multiple-choice questions with three answers before and after the subunit. The first introductory question of each subunit was mandatory. After answering this question, it was permitted to open the content and read it. It was strongly recommended to answer the second and final question of each subunit (See Fig. 1).

This concept allowed to record, track and visualize the student activity and student answers to questions after each subunit.

The used data set consists of data obtained from:

- student success in answering questions at the beginning and end of each study subunit;
- the student's success in the final test on the topics learned in the units.

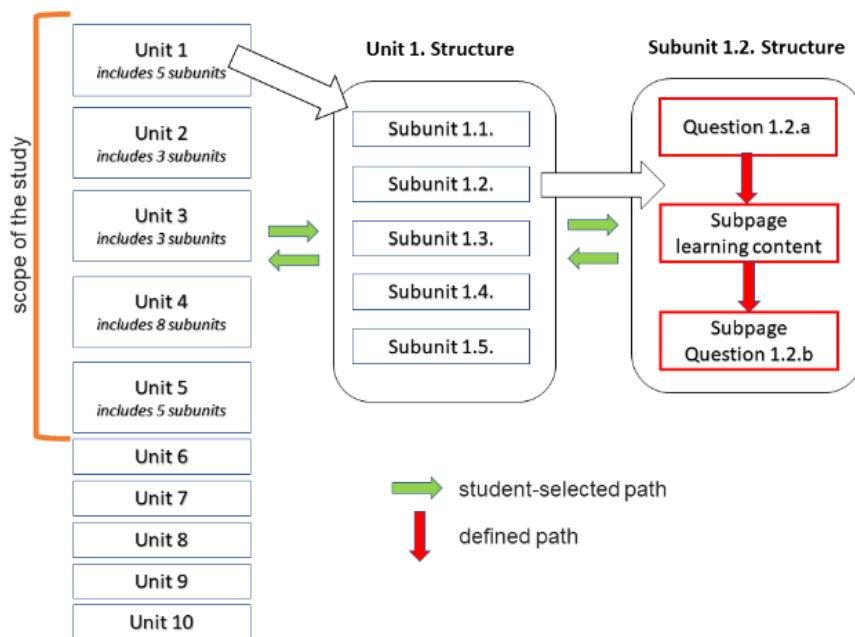


Fig. 1: Learning Content Management System with Multiple-Choice Questions Before and After Each Subunit

Each student's activity was measured, and their performance compared to that of the other students.

In order to improve the transparency of the results obtained, the course portion was divided into 5 two-week periods prior to the final test, during which the student activity measurements were compared.

The students' success in answering subunit questions in the study materials was recorded. A correct answer to each of these questions gives the student one point, so the total score of these points was determined by two factors - how well the student provided the answers (student knowledge) and how many questions he or she chose to answer (student activity and engagement).

The student's success and final grade on the first 5 chapters of this e-course were determined by a final knowledge test. This test included 30 multiple-choice questions. By answering the question correctly, the student received 1 point, the total possible amount of points in this test was 30. Among these questions were those with only one correct answer and those with multiple answers, so the student could also obtain 0.33, 0.5 or 0.66 points when answering the question. Conversely, if a student made the wrong choice, points were proportionally subtracted. The results of this test were used in this study as a reference point for student final success.

3. User behavior monitoring and support

One of the challenges of the research was to understand whether the data obtained during the course can determine the difference between the student's presence in the course and the student's involvement in the study process. In the authors' view the ability to differentiate presentism from engagement with the learning process is a step towards understanding students' current level of motivation.

During the study, we followed student activities and successes. This allowed us to notice several important nuances in student behavior and to highlight signs of several types of behavior.

Figures 3a, 3b, 3c, 3d and 3e show the learning activity trajectory of a single user. Our approach is to use this trajectory for learning support and encouraging decisions. A new snapshot of trajectory is presented after each two-week period (in our case) but it could be created after each significant time period or milestone in the course. The data are cumulative.

In order to design the next stage algorithm for monitoring learning support we studied data clustering.

We clustered our data using the k-means algorithm. Inputs: Variables – subunit test activity RANK, subunit test True/total answered (%); Level of Detail - each student; Scaling – Normalized.

After each time period we created 3 clusters with a total of 61 points.

To better evaluate our groups, as a third dimension we visualized a resulting characteristic of each student – the final test result of a mid-course test taken after these 5 time periods.

We found that student behavior in course content belongs to one of three clusters. Being in one or another cluster assigns the student a behavioral characteristic: 1. Active and successful; 2. Inactive and unsuccessful. 3. Inactive but successful; (Figure 2, Students' behavioral characteristic (week 43-44)).

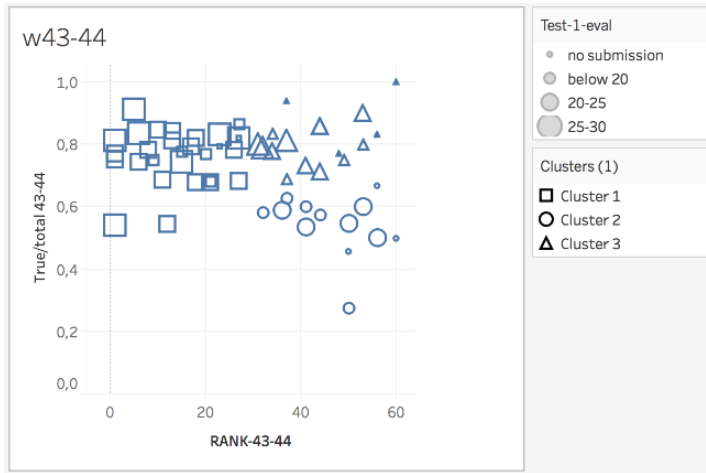


Fig. 2: Students’ behavioral characteristic (weeks 43-44)

We can follow each students path in time between axis and leaving and entering clusters. Students are being evaluated against their peers in the course.

The students learning activity and efficiency increase could be visualized by targeted movement from one cluster to another. Additionally, long time learner efficiency/activity localization in clusters with lower efficiency is indicating the need for learning encouragement.

The movement directions and speed can be used for automated encouragement action generation.

In the next Figures 3a, 3b, 3c, 3d and 3e each circle represents one student’s learning process and success (identifier s006). The True/Total means the relative amount of correct answers to the subunit questions. The RANK means the student activity rank between all 61 participating students. The circle diameter means the result in the course final assessment test. The bold circle shows the track of one selected student in e-user activity co-ordinates.

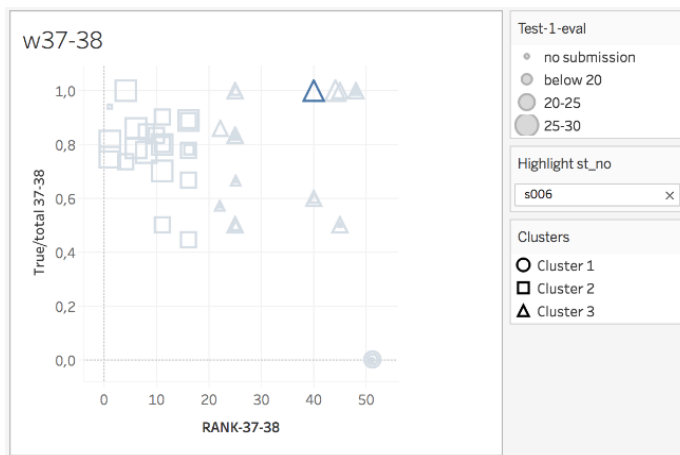


Fig. 3a: Student’s “s006” Performance (weeks 37-38).

As we can observe in the case of student s006, this student showed relatively low activity and good knowledge in answering the subunits questions at the start of the course (Figure 3a). This is evidenced by his position at the top right of the graph, in the Cluster 3 of inactive but knowledgeable students.

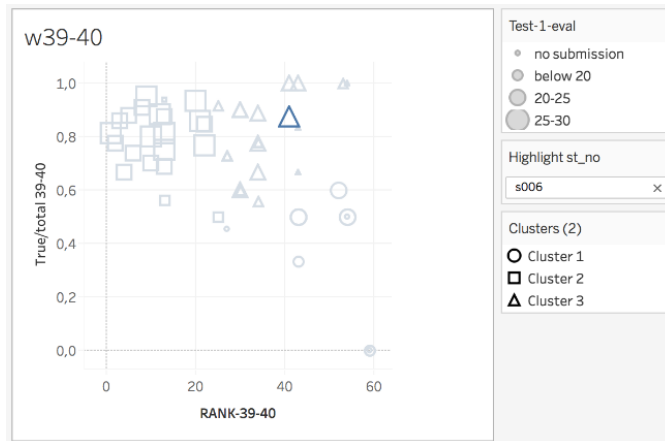


Fig. 3b: Student's s006 Performance (Weeks 39-40).

In the next period (Figure 3b) the student has shown more activity while still maintaining good success. This has helped him move closer to Cluster 1 while still retaining his place in Cluster 3.

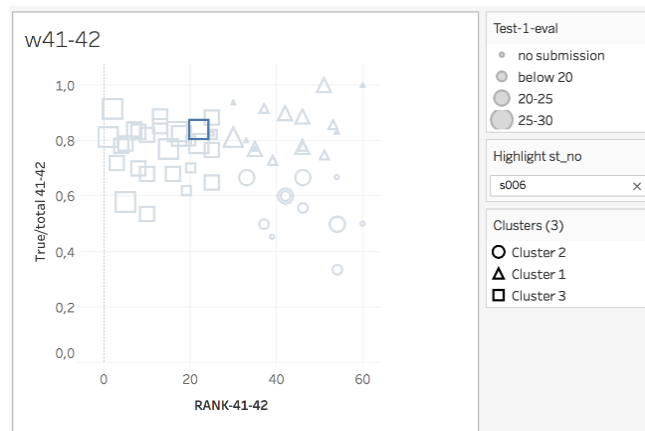


Fig. 3c: Student's s006 Performance (Weeks 41-42).

Further improving his activity in the 3rd period (weeks 41-42), the student's s006 position has moved even further to the left (Figure 3c). This has allowed him to take a place in top-left sided cluster, with knowledgeable and active students.

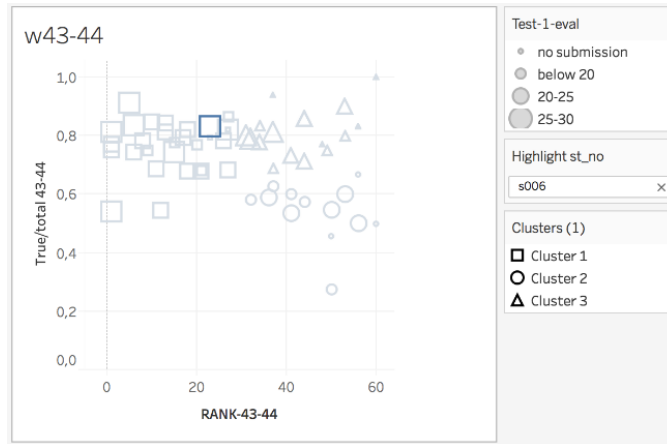


Fig. 3d: Student's s006 Performance (Weeks 43-44).

During the fourth study period (Figure 3d), student s006 maintained his position in the active and knowledgeable student group.

The student maintained his position in the 5th period. The size of the marking shape indicates that this student has also achieved good results in the final test (Figure 3e).

The data in Figures 3a, 3b, 3c, 3d and 3e confirm the possibility to track student activity and design the algorithms indicating the student's place in e-student activity coordinates and generate the supporting/encouraging messages accordingly.

As the student activity was recorded, ranked once every two weeks according to the activity displayed.

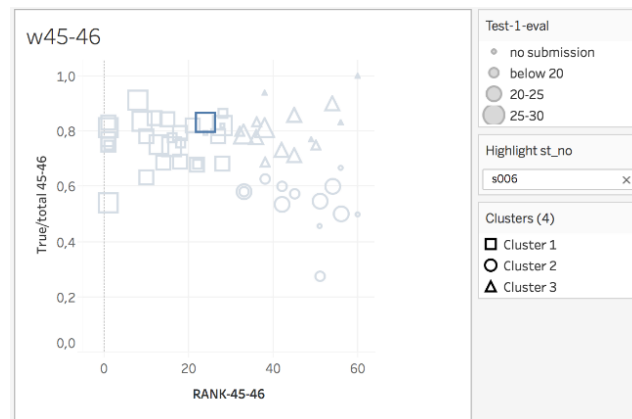


Fig. 3e: Student's s006 Performance (Weeks 45-46).

The data obtained were then compared with the final test success. When performing correlation calculations, the authors concluded that there is a moderate correlation between the student's rank in the activity rank and his / her final test performance (Fig. 4).

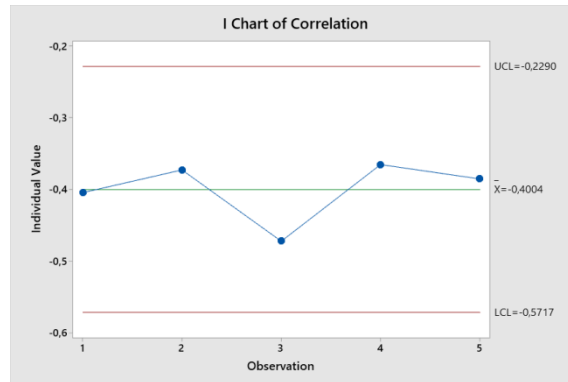


Fig. 4: Correlation Between Students Final Test Results and Activity.

Our obtained data on subunit question results were correlated with the final test success. When performing correlation calculations between the scores of individual questions in subunits and the student's final test performance, a moderate correlation of 0.45 was observed. It should be noted that students' success in learning is influenced by a variety of factors, many of which are not directly related to the learning process. Therefore, in our view, this relationship is worthy of further investigation by specifying the measurement points and the nature of the data to be obtained.

Even if student activity through reviewing study materials and answering questions, is not the only factor influencing student success, it can be used as part of an algorithm to predict other students' success in this or co-curricular course.

This allows authors to conclude that answering questions is both a useful way of reading a student's actual engagement in the learning process, and a positive influence on his or her learning success.

4. The correspondence of courses to user learning needs

The correspondence of courses to user learning needs is a critical issue in e-learning. If a course is too complicated, users lose interest. They do not benefit from a course if it is too simple. We will use the TELECI approach to model and get the quantitative data on the correspondence of courses to user learning needs based on the user behavior data in the e-learning environment.

Table 1: Calculated theoretical values of average relative probability for N-P, P-P, N-N, P-N for three types of e-content (too complicated content, too easy content, and ideally matching course content) if multiple-choice questions before and after the course unit have three answers.

	N-P	P-P	N-N	P-N
too complicated content	0,222	0,111	0,444	0,222
too easy content	0	1	0	0
ideally matching content	0,667	0,333	0	0

From user behavior data before and after subunit learning four records are possible:

- N-P record when the user selects an incorrect answer before learning and the correct answer after learning;
- P-P record when the user selects the correct answer before learning and the correct answer after learning;
- N-N record when the user selects an incorrect answer before learning and an incorrect answer after learning;
- P-N record when the user selects the correct answer before learning and an incorrect answer after learning.

The calculated data in Table 1 show:

- if the content is too complicated the N-P, P-P, N-N, P-N data have random probability;
- if content is too easy – users have all the knowledge before learning and PP=1, others are 0;
- if the content matches perfectly – users have no knowledge before learning (random distribution), and the right knowledge after learning.

The data calculated from real user behavior are between Table 1 presented extreme values. The real user data are quantitative measurement for matching e-content with user learning needs.

In the following Table 2 we present the integrated data of a group consisting of 61 students who were learning five course units.

Table 2: Experimental values of average relative probability for N-P, P-P, N-N, P-N for a student group learning five course units.

	N-P	P-P	N-N	P-N
Unit 1	0,26	0,60	0,07	0,07
Unit 2	0,14	0,75	0,04	0,07
Unit 3	0,37	0,48	0,06	0,08
Unit 4	0,26	0,55	0,04	0,14
Unit 5	0,56	0,35	0,08	0,02

The Table 2 data show the real user behavior data values between singular values presented in Table 1. The Unit 1, Unit 2 data show that the content is rather simple. Unit 5 is nearly ideal e-content with a small trend to be too complicated. Data show the increase of learning complexity with increasing of course unit number for our 61-student group.

In order to design the complete learning acquisition landscape, we generated a set of artificial data as superposition of Table 2 singular case data. The artificial data marked all possible learning acquisition space in coordinates N-P, P-P, X-N. We summarized (P-N)+(N-N) as X-N.

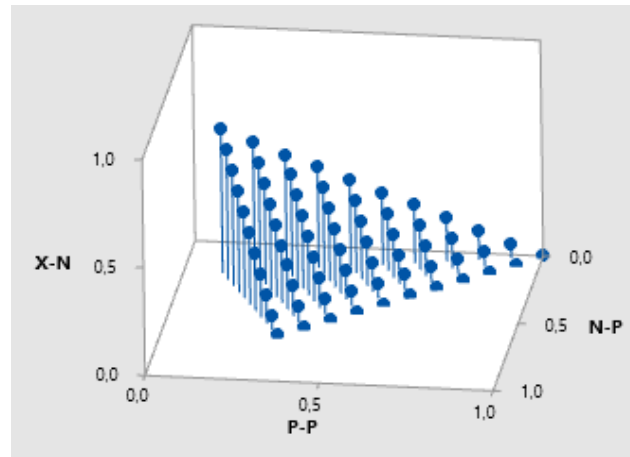


Fig. 5: Complete learning acquisition landscape designed from the generated set of artificial data.

All possible real data will be placed on the triangle Figure 5 between singular points in triangular peaks – too complicated content, too easy content, and perfectly matching content.

In the following Figure 6 we present the experimental data of Table 2 on the learning acquisition surface of Figure 5.

Each dark point corresponds learning acquisition of one of five course Units (Unit 1 – Unit 5).

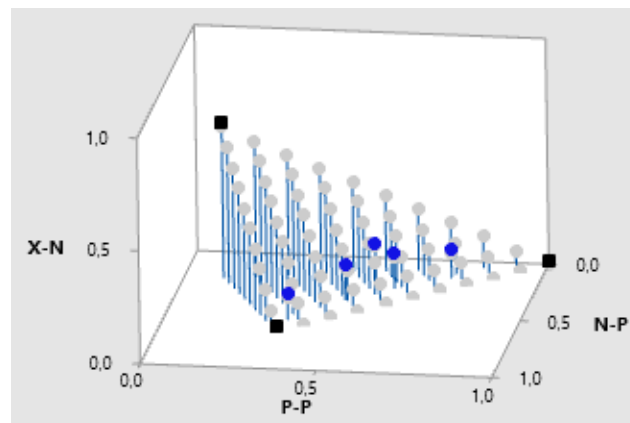


Fig. 6: Experimental data of learning acquisition on the learning acquisition landscape surface presented in Figure 5.

The experimental data in Figure 6 are positioned in curves called “telecides”. The telecide in Figure 6 describe the appropriateness of each course Unit to the learning needs of our 61 students’ group. Telecides have the potential to be used as student performance development trajectories on the learning acquisition landscape surface.

5. Discussion

The user behavior data visualization study provided the full landscape for user behavior activity and user knowledge acquisition.

In the framework of this research we have found a solution how to measure not only student's presence but also involvement in the study process. In our view, this approach will enable e-course designers to obtain data of equivalent quality on student motivation as observing student behavior in the classroom.

By structuring the e-course content and measuring points as used in the study, it is possible to get an idea of the students' current motivation as well as the suitability of the curriculum for the current group of students.

Telecides can be drawn for the study course, for each unit of the study course, for a specific group of students, or for an individual student. However, obtaining the necessary data requires appropriate restructuring of the e-content.

The points and their locations on the telecide can be used for the algorithm to design the content management system with additional motivation and support tools.

Regular retrieval and processing of TELECI type knowledge acquisition data sets the framework for course refinement in several directions. Based on the data obtained, it is possible to send relevant support messages to students and teachers. In this case, the messages sent to students would contain the necessary motivational content, while the messages sent to the lecturers would provide information on student success and potential output in the learning process. At the same time, information provided by telecide of a e-course unit or by a student can be used to make some adjustments to the content of the e-course, making it more complex or simplified, depending on the data provided by the system.

At the same time, the data obtained from the course question system can be used to design the algorithm for predicting student learning success. The question system developed by the authors also assess the students' initial knowledge of the course knowledge area. The experimental data on learning acquisition are reporting the appropriateness of learning e-content to the needs of the student.

The obtained results lead to the conclusion that the development of the question system has served as a motivating factor in attracting students to the course. Although the effort to answer the questions is not the only reason for the student's success, it is nevertheless important enough to be used in further algorithm development.

6. Conclusions and Future Work

The learning content in TELECI approach was divided into small parts with a multiple-choice test before the subunit and another test after it. This approach was better liked by students. It ensured obtaining a sufficient amount of user behavior data.

The method of structuring the curriculum developed within the framework of the study has provided important data, allowing a better understanding of student behavior in an electronic learning environment.

The e-course structure and data visualization enabled a quantitative description of the course's, or course unit relevance to the user group. It was done using multiple-choice test results before and after reading the content. The study demonstrates visually and

numerically the differences in the appropriateness of each e-course unit to the needs of the learner group.

Because of the research, several relevant indicators have been identified which allow to identify the student's involvement in the study process. The data are obtained by observing the student's behavior and its dynamics in the course. In our view, such an estimate is more accurate when compared to survey data.

The student's behavior in the course is a direct indication of his or her current motivation. By using our proposed course structuring and data acquisition points, you can obtain the information you need to dynamically motivate the student and / or improve the course content.

Data visualization and correlation studies showed that students can be divided according to three types of e-content users. The types differ with the final results of the course, activity in multiple-choice test decisions, and multiple-choice test results. This approach enabled the development of an algorithm for additional support and motivation of students if their e-content user model is less efficient and could be turned to more efficient one.

Being able to dynamically determine which of these groups a student belongs to allows you to successfully motivate and guide him or her for further learning

In the next phase of the study, the authors plan to develop an easy-to-use TELECI technology so that it can be applied in different courses and in different learners' groups.

Future research plans to develop techniques for obtaining this type of student behavior data. The goal is to find solutions that allow for the required engagement measurements without artificially complicating the course content. In our view, it is important that this type of measurement point is integrated into the curriculum so that it is perceived by the student as part of the curriculum and not as an unintended side activity.

Considering that answering short questions during the course of studies has facilitated student involvement, we plan to develop this technique by specifying the required number of measurement points and the scaling of the obtained results

To study the e-course appropriateness for the selected learner, we have to use larger volumes of content. The size of units, unit blocks or all course e-content could be efficient to get the usable data of e-course appropriateness to each student.

Acknowledgements

This research has been supported by a grant from the European Regional Development Fund (ERFD/ERAF) project "Technology Enhanced Learning E-ecosystem with Stochastic Interdependences - TELECI", Project No.1.1.1.1./16/A/154. Notarial Certificate 02.03.2020 LV02065412962.

References

- Daugule, I., Kapenieks, A. (2018). The Data of the Initial Motivation – a Valuable Source for the Development of the Course Content. A Case Study in the Group of Business Students. *International Journal of Engineering and Technology (UAE)*. 7. 89-94. 10.14419/ijet.v7i2.28.12886.

- Dennis, J. M., Phinney, J. S., Chuateco L. I. (2005) The Role of Motivation, Parental Support, and Peer Support in the Academic Success of Ethinc Minority First-Generation College Students. *Journal of College Student Development*, Vol 46 (3) pp.223-236, doi:10.1353/csd.2005.0023
- Gopalan, V., Abubakar, J., Zulkifli, A.N., Alwi, A., Che Mat, R. (2017). A review of the motivation theories in learning. *AIP Conference Proceedings*. 1891. 020043. 10.1063/1.5005376.
- Koch, R. (1998) *The 80/20 Principle. The Secret of Achieving More with Less*. Nicholas Brealey Publishing, pp. 6-7; 267-270.
- Levi, S., Campbell, H. (2008) Student Motivation: Premise, Effective Practice and Policy, *Australian Journal of Teachers Education*, Vol 33 (5) pp.14-26., <http://dx.dio.org/10.14221/ajte.2008v33n5.2>
- Robinson, A., Cook, D., (2018). "Stickiness": gauging students' attention to online learning activities, *Information and Learning Science*, Vol. 119 Issue: 7/8, pp.460-468, <https://doi.org/10.1108/ILS-03-2018-0014>
- Simpson, T. (1997) The Initial Motivation of Students Enrolling in an Adult and Workplace Education Programme, *Asia-Pacific Journal of Teacher Education*, 25:1, pp.67-77, doi: 10.1080/1359866970250106
- Shell, D., Soh, L., Flanigan, A., Peteranetz, M. (2016). Students' Initial Course Motivation and Their Achievement and Retention in College CS1 Courses. 639-644. 10.1145/2839509.2844606.
- Szulanski, G. (2003) *Sticky Knowledge: Barriers to Knowing in the Firm*. SAGE Publications, pp.25-30.
- Tsoni, R., Samaras, C., Paxinou, E., Panagiotakopoulos, C. and Verykois, V. From Analytics to Cognition: Expanding the Reach of Data in Learning. In *Proceedings of the 11th International Conference on Computer Supported Education (CSEDU 2019) – Volume 2*, pp. 458-465 ISBN: 978-989-758-367-4

Received February 18, 2020, accepted March 13, 2020