Going beyond the LMS logs. The complexity of analyzing learning evidences

Miguel Ángel Conde

1 University of León, Engineering School, Campus de Vegazana S/N, León, 24071, Spain

Abstract
The analysis of students' learning evidences is a key issue in the assessment process. This analysis has become more affordable because educational institutions use Learning Management Systems that track what students are doing. However, evaluation of students learning is much more complex than just presenting or summarizing some of their results; it requires to consider issues such as the context, the type of activities carried out, the learning objectives to achieve or the students’ interactions. This paper presents this situation and a sample of how to deal with Learning Analytics for the evaluation of teamwork competence. In order to do this, first we explore the issues related with complexity in LA approaches, later on explore how to assess teamwork competence and present a case study at the University of León, that describe how teamwork assessment was carried out along several academic years.

Keywords
Learning Analytics, Learning Evidences, Interactions analysis, Instant messaging tools, Complexity.

1. Introduction

Educational processes have never been simple and one of the most complex activities could be assessment. Assessment, whether summative or formative, requires evaluating evidence of what students have learned. For some time now, not only the assessment of learning outcomes has been considered, but also the assessment of process followed to achieve these outcomes. This needs attending to the learning process and all the activities that students develop on it. This, in many cases, is done through observation, but it is not always enough specially because not all the phases of the learning process are done in the institutional environments and because in many cases, especially with large students’ groups observation can be really complex [1].

With the emergence of online learning activities all the educational institutions have installed a Learning Management System (LMS). This type of platforms can be seen as an environment to manage course, contents, activities and tools for the teachers; and such as a fire camp for students, where they have their courses, contents, tests and so on [2]. Any of these stakeholders are going to leave evidence of what they have been doing in the platform. That is, if they have logged in, if they have accessed to a resource, the time they spend using it, the grade on a test, etc. The LMSs are going to provide us with reports with such kind of data; however, they are not enough to understand how to improve students learning, which is one of the key issues of Learning Analytics (LA).

LA can be defined as “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs” [3]. By applying LA, it is going to be possible to obtain more knowledge about what our students are doing and how they are learning. But at this point, it is important to be aware that not all the available data could be processed and analyzed, and this is a quite complex process.

Where is the complexity on LA analysis? It may depend on several issues such as:
The context. This a critical issue for the application of LA because the context can constrain the way in which the data exploration or the analysis is done. It is necessary to address the issues such: 1) the learning modality, quantity and type of data are different in contexts where learning is carried out only online vs those where classes are blended; 2) the technological systems employed, LMSs, videoconference tools, etc.; 3) the LA tools available for teachers and managers; and 4) the institutional LA policy, if any, and if the LA strategy is based on approaches closer to academic analytics, to learning analytics or both; the possibility to install or not new tools, etc.

The data access policy. Other important issue when considering applying an LA approach is to know how accessible data is. In this case we should consider the data protection law to be applied that can be defined at different levels: institutional, regional, national, European, etc. In addition, it is necessary to know if it is possible to access to all the data stored in the LMS or only to some parts and if it is possible to access to the logs of activities carried out in external tools employed by the institution.

Ethical policy. Sometimes, it is not considered ethical to use students’ information or track what they are doing. This should be defined at an institutional level. If some of the data could not be used because of ethical reasons this should be considered, specially at the beginning of the application of LA.

The application scope. At what level LA will be applied. The information required by the institution cannot be the same as that required by the teachers. It is not the same trying to gather information about students of several courses and exploring the results in a specific activity. Depending on this, metrics should be defined.

The metrics. What data do we need to explore and analyze to obtain certain knowledge? What would we like to assess? These are questions that could help us to evaluate the data, but the definition of the proper metrics is not so straightforward. It requires considering the previous mentioned factors and taking into account existing works in the literature.

Experts support. It is convenient to know if there is some technical team that can support the teacher in accessing the data, generating reports, or installing LA tools. In addition, it is desirable to have a team with a data science background that can help the teachers to understand which the best way is to explore the data.

This paper is going to present a case study related to the analysis of the individual acquisition of teamwork competence TWC. We are going to describe each of the previous defined factors and how the LA approach is developed at the University of León. In order to do so the rest of the paper is structured as follows: in section 2 we will present a brief description of the adopted LA approach. Later on, we will describe how it was applied taking into account each of the previously described items. After it, we will comment some of the results and tools employed and finally some conclusions are posed.

2. Teamwork competence assessment LA approach

One of the main aims of universities is preparing students to succeed in labor market. In order to do so, learners should achieve competences that later would need as professionals. Among them a relevant competence to acquire is Teamwork Competence. It is a highly demanded competence by companies [4], so it is very important to facilitate developing it in education institutions.

This type of competence can be developed by applying active methodologies such as Project Based Learning [5] or Challenge based learning [6] that in most cases require working in groups, very popular and especially useful in our actual learning contexts. However, although teamwork can be developed more or less easily the evaluation of the acquisition of such competence is quite hard. The common way to proceed in many cases is to do a summative assessment, that is, to evaluate the final outcome submitted by each of the groups. However, this is not enough to know if students are working properly in teams. It is necessary to check the final result, but also the intermediate results and how each of the students work together as a group while addressing the projects or challenges. This can be time-consuming, especially because it is necessary to track what each group member is doing [1]. One of the advantages of LMS and other institutional tools is that we can access to logs about what each student
has done and use them during the evaluation, but it is necessary to apply LA tools and techniques that make this evaluation affordable in time [7, 8].

At the University of León teamwork competence acquisition has been evaluating in several subjects. Since 2015 a methodology known as CTMTC (Comprehensive Training Model of the Teamwork Competence) [9] has been applied. This methodology has been adapted from IPMA method for project management [10]. These projects are addressed in groups by using different Information and Communication Technologies (ICT). CTMTC comprises several stages that produce different outcomes and require a formative and a summative evaluation. As the students use ICT to complete the activities and to interact between them, their activity is logged in the systems and can be later analyzed. To carry out the analysis several ad-hoc defined LA tools were developed. These tools extract data from Moodle and other contexts. Such data is analyzed and presented to the teachers to facilitate the assessment process. In the next sections we will present some of the applications of LA that have been used at the university in several academic courses of 5 subjects with a total of more than 600 students [1, 11-16].

3. Case study of the application of LA at the university of León

In this section we will present how the application of LA at the university of León has been, a description from the first years of application up to now, the methodology flexibility and the different tools used. First of all, it is important to introduce the case study focusing on the factors stated in the introduction:

- **Context.** The University of León is a Spanish public educational institution located at the northwest of Spain with 10,152 students in 2021/2022 academic year [17]. Most of the classes are face to face with a small percentage of online learning. Both for face-to-face classes and for online learning the LMS employed by the University of León is Moodle. It has a log system that tracks many of students’ activities. Some external tools are also provided by the institution, such as Google Suite, Office 365, AVIP Videoconference System, Turnitin, etc. Teachers can use other tools for educational activities. Regarding the LA tools, the teachers can use Moodle reports. However, although they are quite varied, they provide too much data and not so useful information. The university does not allow the installation of external plugins just by teacher request, they require a further study by the university ICT team. In 2021 the university acquired an Intelliboard Learning Analytic system, but it is not available for all the teachers. Although at the university different teachers use LA techniques and tools there is not a specific and public policy about this topic.

- **The data access policy.** The university has a data access policy [18] that follows the Spanish Regulation and the EU-GDPR. As in any educational institution, teachers can access to their students’ data and to the grades, but students’ data must be maintained in the institutional environments where the university is in charge of personal data security and confidentiality. In addition, when a research activity involves students’ data, the ethical commission of the university must approve it. With this consent data can be used anonymously and as aggregated numbers. In this case we have to follow the rules applied by the institution and also we pursue the ethical policy. We should point out that with external tools the data does not belong to the institution, so the companies or institutions in charge of such tools are which define the data access policy.

- **Ethical policy.** The university has an ethical committee that is in charge of all the ethical initiatives. There is not a specific protocol for ethics related to Learning Analytics, but the regional, national, European and international rules and recommendations are adhered to. In addition, each activity will explicitly publish what it is going to do with data, so the students can decline to participate. In the years of application of LA zero students declined to complete the activity in which we apply the tools and techniques.

- **The application scope.** The university has neither published LA initiatives for the whole institution, nor it has been applied at Faculty, Degree or Course level. The ones carried out were done in the context of a subject or an activity of a subject. For instance, the CTMTC case has been applied in several subjects of Computer Science Degree, in each of them differently. For
instance, in Operating Systems it is applied to a project that students should solve in teams and that has a weight of the 42.25% of the final grade. Other subjects, such as Computer Animation, applied the same methodology from the beginning of the subject to all the practical parts.

- The metrics. When it comes to the metrics, they are based on the methodology. It is possible to have metrics for different types and degrees of analysis. In the case of CTMTC application the most time-consuming part of the evaluation was to measure the interaction among students. It requires reading a lot of posts or messages to know which student was working more and how many of them less, check logs about reads and so on. To help measuring students’ interaction, a rubric was defined and employed [15]. It focuses on the issues such as responsibility and engagement (that can be measured considering the messages and if they are short or long), tracking what the team members are doing (visits to the threads or messages read), interaction (messages answering other or providing peers feedback about something), leadership (conversations started, problem solving, first and last messages about something, decisions made). The metrics depend on the data available, the type of data, the analysis techniques available and what you would like to explore. For instance, decision making is something difficult to measure based just on the number of messages, or the length of them. It requires, among other things, knowing what is written in the messages. To analyze this, it is possible to apply natural language processing and look for patterns [19, 20], sentiment analysis [21, 22] or dialog acts [23]. For the case study presented, the previous mentioned rubric was used; however, it has been evolving with the application of new LA tools.

- Experts support. For this case study the teachers are also experts on programming, so the technical part was not a problem, however, sometimes it was necessary the help of data scientists to know what data meant. For instance, to know if some kind of intervention can be associated to an improvement in students’ grades, something that is shown in Fidalgo et al. work [1].

4. Solutions adopted

After we have described the complexity of the case study, we are going to comment how it was carried out. As it was developed throughout several academic years and in different subjects it is important also to attend to how the application of LA has evolved in the context.

4.1. The subjects

The LA approach was carried out initially in Operating Systems, a second subject of Computer Science Degree Course. Within it, the CTMTC methodology and the LA tools were applied to a project development that has a weight of about 22% of the subject grade. After this first application, other subjects began to use the LA approach. More specifically six more subjects were involved, so it was possible to check the flexibility of the methodology and the tools as their contexts were very varied. For instance: the weight of the assessed activities goes from a 10% to a 60% of the final grade, students from 8 to 144, the way in which the students were divided into groups was is different, etc. In most of them the students interaction was carried out in Moodle forums and the intermediate results were described through a Moodle Wiki [11]. This application was maintained in the subjects during two academic years, after that only Operating Systems, Accessibility and Computer Animation maintained it. The rest continued with other kind of analytics approaches [24-27].

In the three subjects that continue with the methodology, some changes were also carried out. These changes are mostly based on experience. In Operating Systems the first change was the weight of the activity in which the LA approach was used, passing from a 22% to other activity that has a 42.25%. There were also changes in Accessibility subject regarding the group formation, at the beginning students were divided into groups by the teacher and later the students could decide on the team members. In Computer Animation, in which the methodology was applied twice, the one for parts of the project development in groups of about 6 persons and the other in the integration of such parts with groups of 20 participants, it finally changed to small groups and just one application.
4.2. The approach

As in most of the subjects, at least at the beginning, the approach was similar to what we are describing here. CTMTC in any of the subjects was applied to a practical work that usually implies carrying out a project in groups. The groups were formed by the students or the teachers; it depends on the subject. Once the groups have been formed, the students should conceive the project following the methodology stages. Stages such as group forming, planning, development are carried out using ICT tools, such as Moodle Forum, Moodle Wiki, Moodle Assignment Module, Google Drive, Control Version Systems, etc. When assessing the acquisition of teamwork competences, individual and group evidences were evaluated following the rubric previously commented [15]. The evaluation of the interaction between peers is an important issue in CTMTC, in order to facilitate this we used the LA tool defined [1]. It provides information about the group and about the team-members, as shown in Figure 1.

The application of CTMTC showed an improvement when it comes to the grades in those subject parts where it was used [15, 16]. However, one of the problems detected was that students do not use forums for their interaction - they use other tools and specially instant messaging tools [28]. If the interaction is carried out in these tools, it is necessary to take into account that the activity is developed beyond the institutional environment which requires changes in the methodology, the evaluation rubric and of course the tool [29].

The methodology was adapted to gather all students’ interaction in just one chat group instead of managing each stage in different threads. The rubric should reconsider how interaction is carried out, for instance, it is important to take into account when a message belongs to a conversation, what a reply is, what an emoji means, what can be considered long and short in this new context. The LA tool then requires an adaption. It needs to be adapted not only to show the information employed by the new rubric, but also to require facilitating students to upload their conversations. The teacher creates a
WhatsApp activity in the platform and the students can upload a text file with their conversations, although as each message is associated to a phone number it is necessary to preprocess the file to replace each phone number or contact name with an ID [30] (Figure 2).

CTMTC adaption to WhatsApp was successful, the number of messages increased, and grades were improved. However, not all students were happy using a proprietary tool and having to do the preprocessing. This meant to look for a new adaption, in this case to Telegram Instant Messaging Tool. The methodology remained the same, but the tool needed to change. In this case Telegram API provided much more information, such as the number of emojis, replies, etc; obtaining results quite useful for the review. A sample of the use of this new approach can be found at [31] and some screenshots of the new dashboard are seen in Figure 3.

Following all these experiments the teachers involved pointed out that it was very interesting to explore not only the number of messages but also the content of such messages, in this way it would be possible to explore issues such as teamwork behavior, leadership, sentimental analysis. In order to do so Natural Language Processing [33] approaches were used; the tool was improved including a complete new dashboard (Figure 4) and providing other metrics [34].
Figure 4: Group information in the PNL Telegram dashboard [34].

It was applied just one academic year so it is early to make conclusions about the performance of the new approach.

5. Conclusions

LA is a powerful discipline with the final aim to improve students’ learning. However, LA application requires a deep study Acknowledgment of the context where it is applied, the data nature, the institutional policies and, last but not least, to have a clear idea about what to study and how to assess it. The process is quite complex, but the results are really interesting.

In this paper, we have presented several issues that should be taken into account when applying LA and the case study that shows how a methodology was applied during several academic years and how this application and the associated LA tools have been evolving. From these experiments it is possible to see how important it is to learn from experience and refine the process, the success of the application could depend on several factors, but it is essential to consider opinions of all the involved stakeholders and explore the literature looking for the best solutions.

Acknowledgements

This work is partially supported by the Erasmus + project “Improving online and blended learning with educational data analytics” ILEDA - 2021-1-BG01-KA220-HED-000031121.
6. References


[18] ULE. Política de Privacidad. https://www.unileon.es/politica-de-privacidad


