Making understandable
Game Learning Analytics for teachers

Antonio Calvo-Morata[1][0000-0001-8701-7582], Cristina Alonso-Fernández[1][0000-0003-2965-3104], Manuel Freire[1][0000-0003-4596-3823], Iván Martínez-Ortiz[1][0000-0001-6595-5690] and Baltasar Fernández-Manjón[1][0000-0002-8200-6216]

[1] Facultad de Informática, Complutense University of Madrid, C/ Profesor José García Santesmases 9, 28040 Madrid, Spain
{antcal01, crisal03}@ucm.es
{manuel.freire, imartinez, balta}@fdi.ucm.es

Abstract. When used for education, games can increase students' motivation and engagement and provide a more authentic learning environment where they can apply knowledge, making them especially suited to schools. However, actual application of such (serious) games in schools is still limited. Teachers still consider that using games is a complex process that they do not fully master and that requires extra effort from them. We consider that simplifying teachers' tasks when deploying games is crucial to promote their use. In classroom scenarios, teachers can greatly benefit from knowing what is happening as a serious game is being played. Game learning analytics (GLA) is the process of collecting, analyzing and displaying student interaction data with the games to improve the educational experience. GLA can be used both at real-time, providing teachers with information while their students are still playing, and offline, inspecting already-finished game sessions. In both cases, analytics is only useful when it manages to bridge the gap between large collections of interaction data and pedagogically sound insight. Analytics dashboards should therefore provide not only complete but meaningful and easy-to-understand information, considering that teachers will most probably not know all the details of the analyses performed underneath. In this paper, we review our experiences on game learning analytics dashboards for teachers, and describe some of the steps we have taken to improve our dashboards.

Keywords: Learning Analytics, Serious Games, Dashboards, xAPI, Game-Based Learning.

1 Introduction

There are many characteristics that make games adequate for education, including their engaging and motivating nature [1, 2]. Despite these advantages, the adoption of serious games is still poor, partly due to a lack of standards for development, validation and deployment in schools [3]. Collecting and analyzing student gameplay is one of the keys to increase serious game adoption in schools, because collected data can
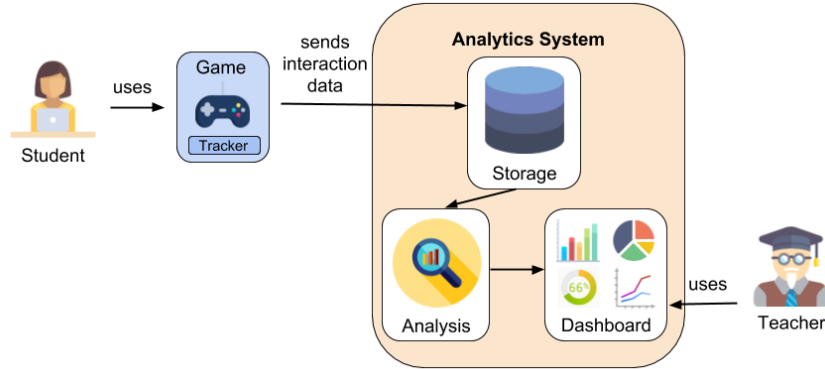
Fig. 1. Game Learning Analytics pipeline: a student plays a game that includes a tracker component. The tracker sends interaction data to the Analytics System for storage, analysis and visualization. Teachers, among other stakeholders, interact with dashboards to gain insights into student gameplay and educational outcomes.

provide insight and improve all steps in the process, and even become a major selling point in itself, thus driving adoption. In educational settings, Learning Analytics (LA) is used to provide insight into learners' actions to improve their learning process and contexts. When applied to games, Game Learning Analytics (GLA) [4] focuses on information gathered from players via in-game interactions.

Fig. 1 represents a generic Game Learning Analytics pipeline, focusing of two main stakeholders: students playing a serious game, whose information is tracked, stored and analyzed within the Analytics System; and teachers supervising the game session. Analyses and visualizations (embedded in a dashboard) provide information for teachers. Other stakeholders, such as students themselves, game developers, and academic officials can be presented with their own dashboards.

Dashboards are the usual way to communicate information to stakeholders [5]. They can display important metrics, and provide a visual overview of other information while allowing filtering and limited query capabilities to gather more in-depth data. Among the usual stakeholders of interest for game learning analytics (students, teachers, game developers), this paper focuses on teachers, as they oversee the actual educational environments and are, in our opinion, the first stakeholder that needs to be considered to improve analytics for games in education: student dashboards are also important, but teachers are who decide whether or not to use games. More narrowly, the last step of **Fig. 1** depicts the teacher using the dashboard, which requires teacher dashboards to be understandable, as described in [6].

Generating teacher dashboards starts with tracking interaction data, which must be done with care to guarantee privacy. Interaction data can then be displayed in different visualizations, for example displaying previously-identified KPIs *(Key Performance Indicators)*, the choice of which will be different from other stakeholders. Data collection and analysis should be entirely transparent to teachers. The final step, the visualization of the information, is where game learning analytics can provide value

to teachers. In this sense, dashboards need to be evaluated as pedagogical tools, taking into account their goals, affect and motivation, and usability [7].

The process of collecting, analyzing and displaying data from in-game interactions to yield useful teacher dashboards comprises several steps, each of them beset by possible issues:

a. Data collection: Collected interaction data cannot be easily shared unless a collection standard is being followed. Once standardized, privacy issues need to be addressed. Furthermore, what data should be captured depends on the games, and game developers are understandably more interested in designing games – rather than selecting what to send and then, on top, having to perform anonymization and sending it according to specific standards. While data collection is not, in itself, an issue that is specific to teacher dashboards, decisions made at this step (particularly what is collected and how it is anonymized) greatly influence dashboard outcomes.

b. Low teacher expectations: Teachers are often new to analytics dashboards, and do not really know what to expect. In our experience, when asked what they expect to see, teachers described only basic information, such as times of completion, difficulty, results in terms of counts of right and wrong answers, or number of attempts; possibly displayed using simple visualizations. Also, teachers assume that analytics will only be available after the intervention, and do not expect to receive any information while students are playing.

c. Dashboard design: The design space of possible dashboards is vast, and designing useful visualizations requires both pedagogical knowledge and game-specific information. Teachers are generally not experts in dashboard design (see point *b* above); and are unwilling to make significant investments in dashboard design upfront, before the game is even available.

d. Changing dashboard requirements: Teachers will often request additional visualizations for their dashboards after the game has been played (see points *b* and *c*). Fulfilling these may be costly or even impossible (for example, if the requisite data was not originally collected; see point *a*) – unless the whole system has been designed to allow the necessary flexibility.

e. Beyond stand-alone games: Teachers may want to use games as parts of larger courses, which may in turn be games as well. In these cases, dashboard granularity needs to be configurable, allowing the game to be analyzed not only by itself, but also as a part of a whole, and even as a whole with several parts.

As participants in two EU H2020 Projects, we have developed a complete architecture to track, collect, store, analyze and display the data collected from serious games in a systematized way [8, 9]. Section 2 of this paper describes how our architecture, and specifically our teacher dashboards, tackle the above issues. Section 3 describes how teacher dashboards could look like in subsequent iterations. Finally, Section 4 summarizes our conclusions and outlines future work.

2   Improving teacher dashboards

For data collection (issue *a* in the previous section), we use the Experience API for Serious Games Profile (xAPI-SG) as a standard collection and archival format. xAPI-SG defines a common set of interactions that are usual in serious games, as detailed in [10]. We provide an easy to use library that greatly simplifies adding analytics to serious games, isolating game developers from the details of the standard. To avoid privacy issues, we use pseudonymous tokens for students. Tokens are unique strings of 4 characters created at the server when the game is deployed, and provided by teachers to their students, who will then use them to access the game. We then rely exclusively on tokens to identify students across play sessions, using them also to display information in the visualizations. Only teachers can, if they choose, keep the correspondence between tokens and actual students.

Regarding low teacher expectations and dashboard design (issues *b* and *c* in the previous section), since we cannot expect teachers to provide detailed lists of what should be analyzed and how it should be displayed, we have developed a default dashboard, which does not require any setup and can display basic data for any game that sends valid xAPI-SG interaction data. For example, since xAPI-SG has a specific vocabulary to indicate that a student has made a choice, and whether the game considers the choice to be correct or not, the default dashboard can easily display counts of correct/incorrect student answers. The use of a default, generic dashboard immediately provides value to teachers, and provides a useful base to elicit requirements for more complex game-specific dashboards.

Additionally, since our dashboards are updated in near real-time, with delays of few seconds between receiving interactions and displaying updated visualizations, we also include a simple alerts and warnings mechanism that can be configured to notify teachers of possible issues as they arise. We consider alerts to be higher-priority than warnings, but the underlying mechanism is the same; and, asides from increasing the situational awareness of teachers, its existence reminds teachers that the use of analytics is not limited to presenting post-mortem information on playthroughs; and that their role during gameplay sessions need not be limited to proctoring.

To provide the necessary dashboard flexibility (issue *d* above), we are not limited to the default dashboard, and allow game developers (presumably with teacher feedback) to create customized game-specific analyses and visualizations for their games. Since our dashboards are built on top of the Kibana and ElasticSearch open-source projects, dashboard creation is developer-friendly, although not recommended for non-programmers. If need be, custom analyses and dashboards can re-evaluate old data, allowing dashboards to be updated to display existing information in new ways. This allows requirements to evolve as teachers and game designers refine their understanding of how students play and learn with a serious game, or when dashboard usability issues are identified. Subsections 2.1 and 2.2 contain two case-studies of such custom game-specific dashboards.

Finally, regarding issue *e*, games are sometimes part of more complex course structures. For example, a game may contain several mini-games, each of which can merit its own dashboards. However, it still makes sense to provide a global dashboard to

monitor progress across all the minigames. This functionality was required for the H2020 EU Project BEACONING and allows for multi-level analytics, where the results of, for example, games that launch other mini-games, can be combined and meaningfully analyzed and displayed to provide insights on overall progress.

2.1 A custom dashboard for a serious game on cyberbullying awareness

Conectado [11] is a serious game to raise awareness against cyberbullying by placing the students in the role of a transfer student that suffers bullying and cyberbullying after arriving at a new school. Players experience the life of this transfer student in first person, during each of 5 in-game days, while being exposed to feelings of impotence and increased (in-game) social isolation. The game keeps track of the level of friendship of the story's protagonist with each classmate. For instance, the variable called *friendship risk* indicates, based on the player's choices, the risk of being bullied in the game from 0 to 100, where a higher value corresponds to a worse social standing. There are indicators of risk for each character as well. The decisions that players make during the game, including whether the player decides to tell that is being bullied to the parents or the teacher or not, also determine the ending.

Some of the game-dependent visualizations developed for *Conectado* can be seen in **Fig. 2**. From left to right, and from top to bottom:

a. Average friendship risk: this general metric describes whether the average of the class has low, medium or high friendship risk (shown in green, yellow and red, respectively).
b. Number of players per game day: this bar chart provides a vision of the progress of players in the game. As there are five days, teachers can see in real-time how many students have played through each in-game day, and help students who are too far behind their classmates.
c. Number of players that have taken each possible action that determines the ending: this pie chart compares the number of player who have decided to complain about bullying to the in-game parents and teachers vs. those that have decided to remain silent.
d. Maximum friendship risk per student: this visualization provides more in-detail information that complements the general metric provided in the first visualization. This allows teachers to quickly identify which students are doing best and worse in the game. Since the visualization is sorted, it also provides an overview of the distribution of risk scores throughout the class.

These visualizations have been designed trying to cover some of the information usually required by teachers: progress (b), decisions taken (c) and specific metrics, both general (a) and per student (d).

By default, in our teacher dashboard, all visualizations were of the same shape and size. However, different sizes are possible as seen in **Fig. 2**. In some cases, this may even be required to fit the desired visualizations, as it is the case in the next example where a plug-in was created to provide the exact visualizations desired.
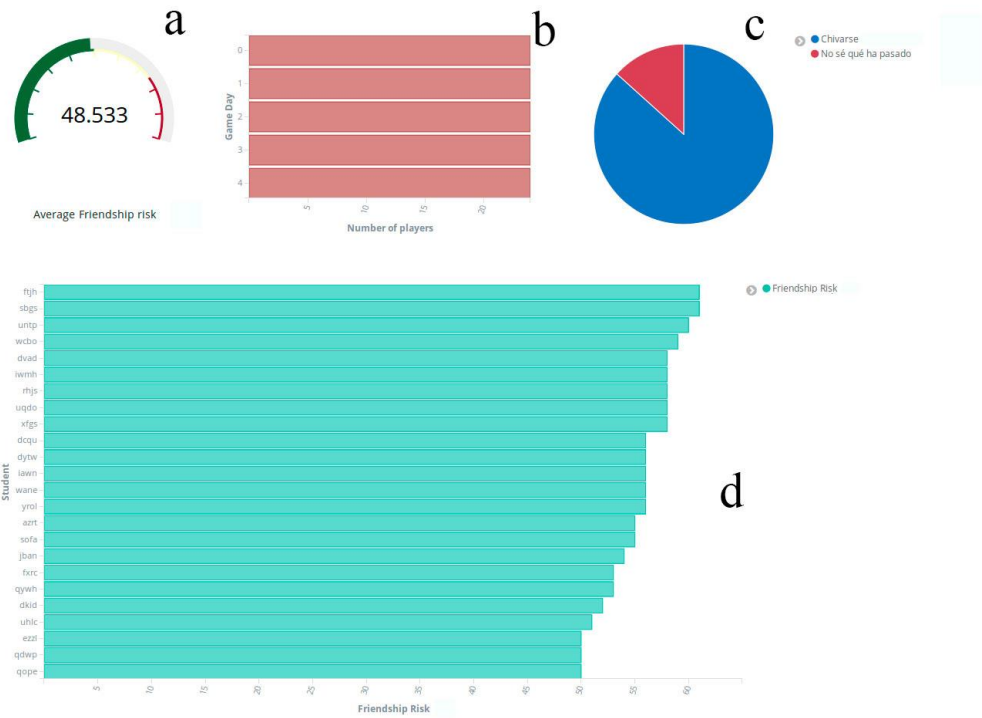
Fig. 2. Some of the game-dependent visualizations developed for the game *Conectado*.

2.2 A custom dashboard for a serious game on workplace interaction

Another scenario where both analysis and visualizations were developed ad-hoc was for a game centered on workplace interactions developed for the EU H2020 RAGE project. The game design included a requirement to use the Thomas–Kilmann Conflict Mode Instrument (TKI) to measure and display responses to the different conflict situations that the player is exposed to while working as a team leader in a simulated game development company. The TKI is based on two dimensions of behavior, assertiveness and cooperativeness; and defines five different approaches based on the balance between both dimensions: competing, accommodating, avoiding, collaborating and compromising. A specific analysis and visualization was developed to display the TKI categorization for each player. Additionally, certain situations allowed the player to exhibit, or avert, certain types of biases (for instance, based on gender, race, or fashion sense). Finally, the game allowed players to track office morale, productivity (in terms of shipped games), and awards for quality.
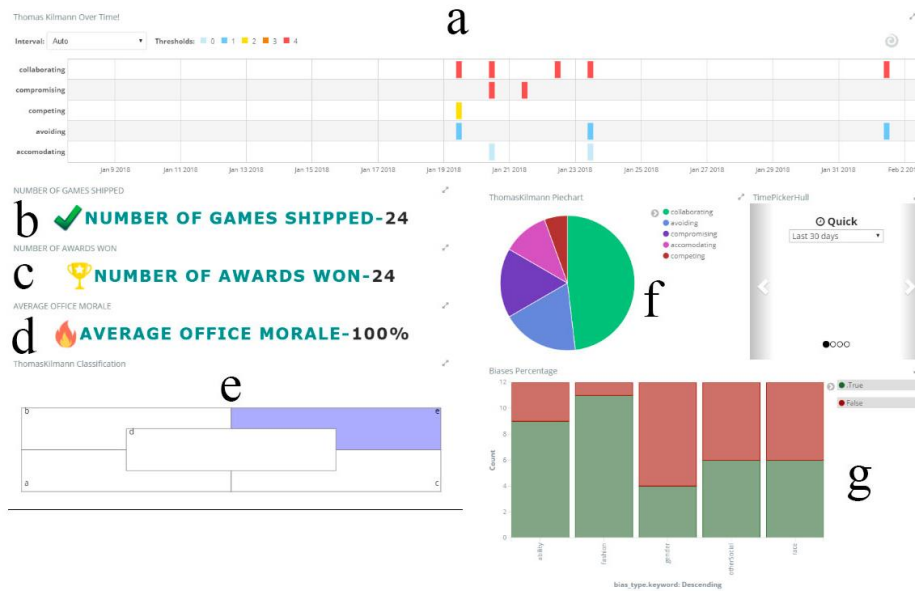
Fig. 3. Some of the game-dependent visualizations developed for the Thomas-Kilmann Conflict Mode Instrument.

Fig. 3 displays the seven visualizations developed for this dashboard:

a. Thomas-Kilmann classifications of a specific student's answers over time.
b. Games shipped, a measure of team productivity.
c. Awards won, a measure of team quality.
d. Office morale, a percentage indicating the degree to which the player's in-game co-workers are happy with the player's choices.
e. Overall Thomas-Kilmann classification for the player, displaying the category that has appeared more times from the five pre-set categories. This visualization is part of the standard TKI.
f. Pie chart displaying the distribution of answers according to the 5 TKI categories.
g. Bar chart displaying, for each bias, the ratio of responses where it was averted (green) or exhibited (red). Note that both counts are relevant, as most possible choices to the in-game conversations did not offer the opportunity to either avert or exhibit a bias.

This set of visualizations was a specific request made by the creators of the game in the project, sought to display all the information they considered to be of relevance in the game, and underwent several cycles of prototypes/changes until reaching its final iteration as displayed above.

3   Lessons learnt in teacher dashboard design

We have tested the default teacher dashboard, including an alerts & warnings visualization, in experiments with two games, played under experimental conditions by over 600 students and 150 teachers. Based on feedback from these experiments, we are currently redesigning the default dashboard to make it more understandable and actionable for teachers.

To make visualizations more understandable, we have completely reworked their names and labels, with a focus on clarity; and are including pop-up descriptions for the more complex visualizations. For instance, in **Fig. 3**, the visualization of biases exhibited/averted (g) requires, at first, a significant description to understand what is being displayed; but once understood, there should be no further need for bringing up the description, and it should therefore not receive permanent display space. Additionally, as can be seen from **Fig. 3**, we now allow dashboards to combine visualizations of different sizes, so that more visually complex visualizations can be rendered in larger areas. At first version all of visualizations were of the same size and at the second version, we had two possible sizes (**Fig. 2**). Also, while in the initial teacher dashboards the positions of individual visualizations were not fixed, we have now made positioning entirely predictable and controllable. This has considerable advantages when comparing the dashboards for different experiments based on the same game, for example.

Making a dashboard more understandable also makes it easier for teachers to reason on the underlying information, and to take actions based on these decisions, which is the goal of any analytics system: to induce new meaning or change behavior [12]. Therefore, usability is an important first step towards actionable feedback. We are also exploring other avenues to provide recommended actions for teachers. For example, we are considering the use of alerts to highlight statistical deviations from a baseline. This would first require sufficient baseline data to be gathered; for example, we can take all completion times from a validation run, and use these times to identify students who take significantly longer (say, one standard deviation) than their colleagues to finish. Since this analysis can be performed regardless of the game, it can be rolled into the default alerts system, benefitting all future users of the analytics system at essentially no increased cost for users. The view of alerts and warnings can also be improved, by making a better use of display area; for instance, showing triggered alerts and warnings directly in the general view if they are not too many, or showing only the most recent otherwise.

Fig. 4 depicts two versions of the general view of the alerts and warnings each student has triggered. In the original version, teachers must click on a student's name or access token to see the detailed alerts and warning that the student has triggered. In the updated version, teachers see details for each student directly on the main view.

As an upcoming approach, alerts could be used after validation of games to provide extra information: for instance, alerts could be deployed to identify students who are taking much longer than the expected time and show an alert in correspondence, helping to identify the outliers in terms of completion times.

Fig. 4. Previous (top) and updated (bottom) alerts & warnings view. The previous display only counts of alerts and warnings; teachers had to click on student names to view the actual alerts and warnings for those students. The updated version does not require such a context switch.

4   Conclusions

Teachers are key to increasing adoption of serious games by schools, and Game Learning Analytics should therefore focus on their specific needs. We consider that teachers' requirements should determine what information is to be collected and analyzed, to be later displayed on dashboards that are easy to understand for an average teacher. Dashboards should help teachers to make informed decisions not only after the games are played, but also while the game is ongoing and teacher interventions are still possible to help players make the most of their sessions.

In this paper, we have identified several issues with teacher dashboards, including privacy and data collection, low teacher expectations regarding the outputs of the dashboards, lack of initial input when creating initial dashboards vs. late dashboard design requirements, and the use of dashboards for non-standalone games; and we have described how we have met these challenges by using simple anonymization via tokens and the xAPI-SG standard, a default set of visualizations that provides teachers with quick and easy-to-understand information to act on the previous contexts, sup-

port for custom-built dashboards (we present two case-studies), a flexible alerts and warnings system, and hierarchical dashboards.

In our experiments using these dashboards, we have identified possible improvements (including alerts and warnings) to make them more understandable. These will be implemented on subsequent iterations, and tested and validated using the *Conectado* serious game and other games from RAGE and BEACONING H2020 projects.

References

1. Connolly, T.M., Boyle, E.A., MacArthur, E., Hainey, T., Boyle, J.M.: A systematic literature review of empirical evidence on computer games and serious games. Comput. Educ. 59, 661–686 (2012).
2. Denis, G., Jouvelot, P.: Motivation-driven educational game design: applying best practices to music education. In: Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology (2005).
3. Loh, C.S., Sheng, Y., Ifenthaler, D.: Serious Games Analytics. Springer International Publishing, Cham (2015).
4. Freire, M., Serrano-Laguna, Á., Iglesias, B.M., Martínez-Ortiz, I., Moreno-Ger, P., Fernández-Manjón, B.: Game Learning Analytics: Learning Analytics for Serious Games. In: Learning, Design, and Technology. pp. 1–29. Springer International Publishing, Cham (2016).
5. Shoukry, L., Göbel, S., Steinmetz, R.: Learning Analytics and Serious Games: Trends and Considerations. In: Proceedings of the 2014 ACM International Workshop on Serious Games (2014).
6. Schwendimann, B.A., Rodriguez-Triana, M.J., Vozniuk, A., Prieto, L.P., Boroujeni, M.S., Holzer, A., Gillet, D., Dillenbourg, P.: Perceiving learning at a glance: A systematic literature review of learning dashboard research, (2017).
7. Jivet, I., Scheffel, M., Drachsler, H., Specht, M.: License to evaluate: Preparing learning analytics dashboards for the educational practice. In: Learning Analytics & Knowledge Conference (2018).
8. Alonso-Fernandez, C., Calvo, A., Freire, M., Martinez-Ortiz, I., Fernandez-Manjon, B.: Systematizing game learning analytics for serious games. In: 2017 IEEE Global Engineering Education Conference (EDUCON). (2017).
9. Perez-Colado, I. J., Alonso-Fernández, C., Freire-Moran, M., Martinez-Ortiz, I., Fernández-Manjón, B.: Game Learning Analytics is not *informagic*!. In: 2018 IEEE Global Engineering Education Conference (EDUCON). (2018).
10. Serrano-Laguna, Á., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., Fernández-Manjón, B.: Applying standards to systematize learning analytics in serious games. Comput. Stand. Interfaces. 50, 116–123 (2017).
11. Calvo-Morata, A., Fernández-Manjón, B.: Conectado: concienciando contra el cyberbullying mediante un videojuego educativo. DIM Didáctica, Innovación y Multimedia.
12. Verbert, K., Duval, E., Klerkx, J., Govaerts, S., Santos, J.L.: Learning Analytics Dashboard Applications. Am. Behav. Sci. 57, 1500–1509 (2013).