

Improving serious games analyzing learning analytics data: lessons learned

Cristina Alonso-Fernández¹[0000-0003-2965-3104], Iván Pérez-Colado¹[0000-0002-1250-106X],
Manuel Freire¹[0000-0003-4596-3823], Iván Martínez-Ortiz¹[0000-0001-6595-5690] and
Baltasar Fernández-Manjón¹[0000-0002-8200-6216]

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

Abstract. Serious games adoption is increasing, although their penetration in formal education is still surprisingly low. To improve their outcomes and increase their adoption in this domain, we propose new ways in which serious games can leverage the information extracted from player interactions, beyond the usual post-activity analysis. We focus on the use of: (1) open data which can be shared for research purposes, (2) real-time feedback for teachers that apply games in schools, to maintain awareness and control of their classroom, and (3) once enough data is gathered, data mining to improve game design, evaluation and deployment; and allow teachers and students to benefit from enhanced feedback or stealth assessment. Having developed and tested a game learning analytics platform throughout multiple experiments, we describe the lessons that we have learnt when analyzing learning analytics data in the previous contexts to improve serious games.

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

1 Introduction

Serious games are being successfully applied in multiple fields (e.g. military, health); however, their uptake in formal education is still poor, and usually restricted to complementary content for motivation [1]. Several reasons can explain this, including the high development cost of new games, or the difficulty for teachers to assess the acquired learning, and therefore to effectively deploy and apply games in their classes.

Moreover, very few serious games have a full formal evaluation, and those that have been evaluated are usually tested with limited numbers of users [2]. This is hardly surprising, as large-scale formal evaluations can become as expensive as creating the game. Also, the feedback from formal evaluations is often obtained too late to improve the games or their educational experience. We consider that information from in-game users interactions can benefit all phases of a serious game's lifecycle, including game design, development, piloting, acceptance, evaluation and maintenance; and should be

used to improve the experience of all stakeholders involved (teachers, educators, and students), providing each with the specific information that they need for their purposes. But this process is still too game-dependent, complex and expensive.

Analysis of in-game user interaction data has been used to improve games development in the entertainment industry, in a discipline called Game Analytics (GA). This requires data to be obtained via telemetry, and then analyzed to extract metrics, such as performance or user habits. However, the usual focus of GA is increasing user retention, playing time and revenue [3]; while serious games, particularly in education, instead seek to maximize learning or improve the learning experience.

Learning Analytics (LA) is “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” [4]. LA seeks to lay the groundwork to go from theory-driven to evidence-based education, where data can be used to improve educational scenarios [5]. This approach can be extended to serious games, where in-game user interaction data can benefit their creation and applicability in real environments, in a discipline that we call Game Learning Analytics (GLA) [6].

In this paper, we wish to go beyond the usual post-game session analysis, and focus instead on three scenarios where GLA can be especially helpful. First, the data extracted, if done in a systematic and standardized way, can be not only used for improving the games but also openly shared for research purposes. Second, in the context of applying games in education, all stakeholders involved could benefit from information from the actions taken in the game, directly during the session. Finally, after sufficient data has been gathered, deeper analysis can be helpful to obtain richer information for all stakeholders, and inform improvements in several stages of the game’s lifecycle.

2 Obtaining in-game user interaction data

The first step to gain insights from in-game user interactions is to ensure that all data with the potential to yield such insights is adequately collected. Experimental design and deployment should comply with all the legal regulations (e.g. users’ consent). We consider three main pillars that data management must ensure:

- **Anonymization:** when possible, data must be adequately anonymized so no personal details are attached to the student data (e.g. using randomly generated codes). This will help to comply with regulations on data privacy [7].
- **Collection:** data collection must be non-intrusive and transparent, to avoid interrupting the students’ gameplay. Collection can be greatly improved using a standard tracking model that simplifies and standardizes this process.
- **Storage:** data received from games should be collected in a server that can efficiently manage large amounts of data in a secure way. If data is collected in a specific format, the storage system should also be prepared to validate and handle that format.

Our research group has developed a GLA System that is currently being improved and extended as part of two EU H2020 projects (RAGE and BEACONING). With this analytics platform, we have already conducted several experiments that follow the

above data-management guidelines, collecting data from thousands of game sessions. Some of the results and conclusions drawn from these experiences are detailed in the following sections, since we have used the resulting data for each of the three applications described in this paper: research, real-time reports, and deeper offline analysis.

2.1 Standardizing data collection: Experience API Serious Games Profile

To systematize and standardize data collection we propose the use of the Experience API Serious Games Profile (xAPI-SG for short), described in detail in [8].

As previously mentioned, it is mandatory to comply with all personal data privacy regulations, capturing only the relevant data and using anonymization whenever possible before storage, so no data can be traced back to specific students. For analytics, pseudo-anonymization techniques can be used, where the manager of a session assigns random tokens to players that use them to access the game. The tokens tie all data received from each player together, while providing no information of their identity. When required, teachers can retain the correspondence between anonymous tokens and students that use them; in such cases, this link must be managed outside the game and the analytics system. Additionally, best practices require informed consent forms disclosing both the intended experimental design and how collected data will be used.

Servers that can store xAPI data are usually called Learning Record Stores (LRS), and generally allow limited query capabilities. Every server and technology used in the tracking architecture should be ready to deal with large amounts of data (*big data*) as the number of traces generated by a single player may be large; and if the system is successful, large amounts of users generating many interactions per second can easily overwhelm low-capacity solutions. To ensure scalability, multiple servers that can share the load are an obvious choice; however, this also increases the chances of at least one of them failing or becoming unreachable, forcing truly scalable analytics implementations to be distributed, redundant, and fault-resistant.

Data tracked from serious games using an open format such as xAPI-SG, when suitably anonymized, can be easily shared with other researchers. Open sharing of research data and publications are among the tenets of the Open Science movement, with initiatives such as the European Commission's OpenAIRE [9] or CERN's Zenodo [10], which seek to ensure open access of research data and publications, respectively.

3 Uses of analytics data to improve serious games

Users of serious games can benefit from collected data at several stages: (1) at real time, to provide real-time feedback to teachers and students, (2) after the session is finished, to provide detailed feedback, and (3) after sufficient data has been collected, through enriched feedback based on data mining. In this paper we are mostly interested in (1) and (3), since (2) is generally known and explored in many other resources.

Fig. 1, has been adapted from the Learning Analytics Framework (LAF) described in [11], to add and highlight the use of open data, real-time feedback, and data-mining. The LAF did not envision serious games as sources for analytics data, and predates the

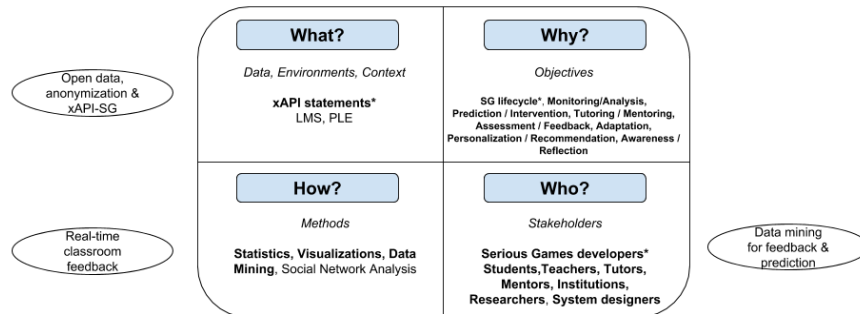


Fig. 1. Contributions of this paper (ovals), and an adapted version of the Learning Analytics Framework, as described in [11]. Elements with an asterisk were not present in the original framework. Bold-face text highlights framework elements affected by our contributions.

appearance of xAPI. As seen in the figure, the contributions described in this work have consequences for most if not all elements considered in the LAF. For each of the 4 questions considered in the LAF:

- **What** kind of data does the system gather, manage and use for analysis? While the LAF mentions e-learning systems as sources, we propose the use of xAPI anonymized statements from games.
- **Why** does the system analyze the collected data? We extend the goals of the LAF to improve serious games, from design to development, deployment, and maintenance. Real-time feedback is key for monitoring and classroom intervention, while data mining allows enhanced feedback once enough data has been gathered.
- **How** does the system perform the analysis of the collected data? We apply most of the methods envisioned by the LAF, even though our current focus is on single-player games.
- **Who** is targeted by the analysis? Stakeholders now include the actual game developers, in addition to the learners and teachers that gain, among others, feedback and assessment, or educational institutions that wish to know the outcomes of applying games in education. Researchers can also benefit from open research data.

4 Game learning analytics real-time applications

Different stakeholders can benefit from (near) real-time feedback. In this section, we adopt the common scenario of using serious games in education as part of a lesson in a classroom environment, focusing on real-time feedback for teachers and students.

Real-time feedback is available as soon as the game starts to be played; to allow teachers to monitor and perform timely interventions, easy-to-understand feedback must be quickly generated. For example, a student that stops playing can trigger an alert that allows the teacher to walk over to find out the cause; or a student that is advancing

much quicker than the rest of the class may benefit from the teacher suggesting additional tasks to attempt. Such simple scenarios illustrate real-time applications where the information collected from interaction data can help teachers and students. Certain types of visual analytics are particularly suited for real-time feedback, especially when combined in dashboards. The ideal content of these dashboards will depend on the game, delivery environment, and the metrics and KPIs that are most relevant to each stakeholder.

4.1 Real-time information for teachers

For teachers, visual analytics provide an easy way to explore the information gathered from their students' interactions. Analytics dashboards present aggregations of individual visualizations, each providing insight into specific aspects, such as progress, errors, or choices taken. Visualizations can also display actionable feedback to locate students that get stuck, or suggest additional work that may interest advanced students.

We have conducted multiple experiments with students to test our data gathering, real-time analytics and dashboards; the latest, as of this writing, with over 1000 students, seeking to validate a serious game that raises awareness on cyberbullying [12]. Previous experiments include games that teach first aid techniques [13], or that were geared towards cognitively impaired users (e.g. with Down Syndrome or Autism) [14], where dashboards were the only option to follow the progress of players.

Fig. 2 describes some of the visualizations included in the teacher dashboard used in the latest experiments to provide real-time feedback in classroom settings [15]. The dashboard uses xAPI-SG concepts such as *completables* (e.g. levels) and *alternatives* (e.g. multiple-choice questions). The visualizations depicted inform users on (a) *correct and incorrect alternatives selected*: the number of correct and incorrect answers selected as *alternatives* for each player, and therefore the general knowledge of players; (b) *total session players*: the number of students that have started the game; (c) *maximum progress of players per completable*: for each *completable*, the progress achieved by each player, and therefore whether students are finishing or struggling to continue; and finally (d) *games started and completed*: a pie-chart that displays the number of games that have been started and completed, providing an overview of the students that have started and finished; and indirectly, how many students are still playing.

These visualizations aim to provide general information from gameplays (e.g. progress, answers) to teachers, allowing them to understand it with minimal effort. Data can be used to trigger alerts or warnings in specific situations that require immediate action for teachers (e.g. a player has been inactive for too long). **Fig. 3** shows the general view of alerts and warnings; clicking on a specific student, teachers can see details of alerts and warnings triggered by that student's actions, and act accordingly.



Fig. 2. Some of the visualizations included on the default teacher dashboard.

Improvements are being considered for the visualizations in **Fig. 2**, based on the feedback collected from teachers. For example, simplifying the visualizations by adding clearer titles and legends, and showing general metrics that provide a quicker overview of the most critical information (e.g. questions failed most, critical areas in-game). We have also determined that providing additionally recommended actions is well-received by teachers (e.g. specific student needs help). These recommendations can help teachers to improve their classes, linking the information provided by LA with actions to support students learning [16]. For instance, teachers requested reports on the topics with the highest error ratios, to allow them to be reviewed before any others.

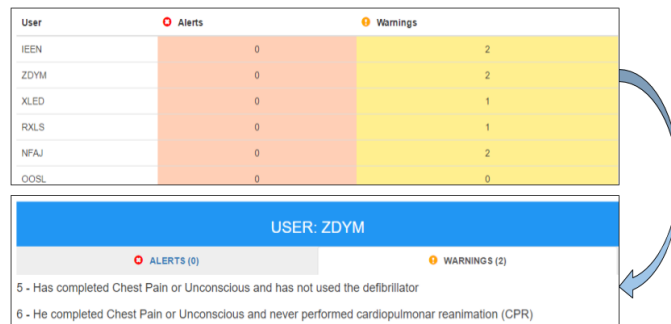


Fig. 3. Alerts and warnings general view (top part). Clicking on a specific student displays detailed information on the warnings and alerts triggered by that student (bottom part).

4.2 Real-time information for students

Student dashboards provide information on performance and in-game outcomes, allowing them to easily assess their strengths and weakness. Current solutions for learners' dashboards present several issues that should be considered. It is common to compare the results of students with their class or with average results (e.g. scores, times-to-finish) from their classmates; however, some researches have pointed out that this may demotivate those students who do not reach at least average rankings [17]. Authors of [18] concluded that most educational concepts used to design LA dashboards focus on self-regulated training by displaying their own data to players. These dashboards generally fail to use awareness and reflection to improve competencies (e.g. cognitive, behavioral or emotional) and usually promote competition instead of knowledge mastery.

Fig. 4 provides a sample dashboard with typical information shown to students: maximum score achieved (visualization labelled *a*); maximum, average and minimum scores of the player (*b*); a comparison with other players in a leaderboard (*c*), which uses traffic-light colors to display ranges of players scores; correct and incorrect answers (*d*); student progress over time (*e*); and levels completed (*f*).

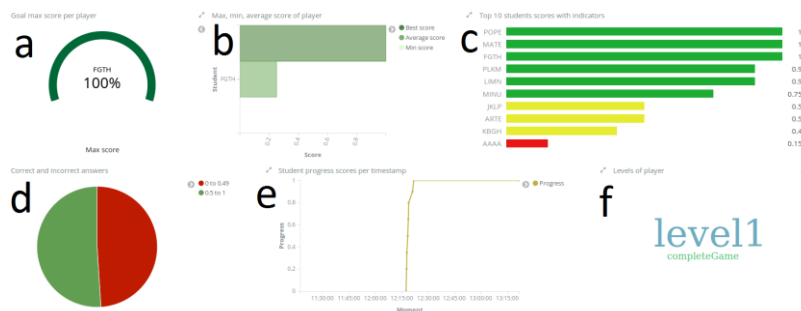


Fig. 4. Sample student dashboard showing information about scores, errors or progress.

Another area where data can be exploited at real-time to benefit students is that of adaptive learning experiences. Games can adapt their difficulty in real-time in response to players' in-game performance. The authors of [19] described the results of experiments comparing adaptive, non-adaptive games and other non-adaptive learning activities, concluding that, although all activities reached equal levels of motivation, the adaptive game resulted in significantly higher learning outcomes.

5 Offline data analysis

Apart from displaying real-time information, data collected during several gameplays can be further analyzed to yield additional insights. Data mining processes can be applied to extract patterns of use, which can be leveraged to improve the game for future deployments. Educational institutions and higher-education administrations can also benefit from aggregated data that can quantify the extent to which the use of games in

class benefits learning, allowing them to make evidence-based decisions on the value of using serious games in their classrooms.

Game developers and designers can obtain feedback from actual classroom gameplays to improve both the game and the learning design. For instance, they may find errors in the game, unreachable areas, levels that are too difficult or too easy for players, a more precise determination of average playing time, etc. In this sense, analysis can certainly help to improve the iterative design for subsequent versions of the game.

Data gathered can also be used to categorize players, creating different profiles for targeted feedback, a process which can be automated using data mining (considering some limits such as data complexity or the algorithms' efficiency). This feedback may include hints to help students or even changes in level difficulty to avoid a decrease in motivation [20]. Clusters of players that show different behaviors and characteristics may provide clues on how different learning elements affect each type of players [21]. In one experiment, we collected data for more than 200 students playing a serious game that teaches first aid techniques. Using data mining, we managed to classify players based on their actions and results to identify clusters of player profiles, and even the discovery of those in-game actions that had greater influence on player outcomes [13].

One of the latest steps we have carried out to improve the lifecycle of serious games using GLA data focuses on improving evaluation methods. So far, serious games are commonly evaluated using costly pre-post experiments [2]. We consider that their application in education would benefit from a quicker and cheaper evaluation process. To this end, we have proposed and tested the use of data mining to predict the results of the pre- and post- tests based on interaction data; while only possible once a sufficiently-large training set has been collected, the technique avoids the need of conducting tests – at least for players that are similar to those that the system was trained with.

This can be considered as another step in stealth assessment. Serious games can also become powerful assessment tools, even though the exact characteristics of games that best allow these assessments is still not entirely clear. Stealth assessment [22] is the practice of embedding assessment in a gaming environment in a non-intrusive way, analyzing gameplay actions to infer exactly what players know at each point in time; it is, in this sense, an extension of the evaluation without pre- and post-test described in the previous paragraph. For this discipline, it is still important to improve the serious games themselves, ensuring that their application is effective and making assessment more valid and reliable.

6 Conclusions

In-game users' interaction data from serious games can be exploited to provide a wide variety of insight on the educational process of different stakeholders. Developers can use it to improve the full lifecycle of games. Teachers can gain real-time insights of student behavior, allowing them to help students playing, or to summarize a session when discussing it with students once they have finished. Students can get feedback on their performance, including their strengths and weaknesses. Researchers can benefit

from open access to shared research data. Educators can obtain metrics on the efficacy of games application on their institutions.

When collecting data, many issues need to be addressed, including anonymization, which is especially relevant when working with minors or to allow collected data to be openly shared for research purposes.

In real-time scenarios, visualizations, alerts and warnings can help teachers to gain insights of the whole classroom, while students can track their performance and compare it to that of their peers; both uses of data provide information that allows teachers and students to make better decisions while the games are still in play.

After data has been collected, data mining techniques can provide further information to improve game design, deployment and evaluation. To improve evaluation, the latest line of work continues with experiments that follow the usual evaluation structure (pre-test, gameplay, post-test) and track interaction data to later predict previous and subsequent knowledge; and compare those predictions against actual data collected in the tests.

Adoption of serious games in schools could greatly improve with the feedback retrieved from interaction data, ideally with a general game learning analytics system that standardizes data tracking, collection, analysis and visualization, extracting useful information to be given back to the stakeholders involved. Whichever the approach, we consider that data-driven solutions that take advantage of the power of game learning analytics are essential to guide the future application of serious games.

Acknowledgments

This work has been partially funded by Regional Government of Madrid (eMadrid S2013/ICE-2715), by the Ministry of Education (TIN2017-89238-R) and by the European Commission (RAGE H2020-ICT-2014-1-644187, BEACONING H2020-ICT-2015-687676, Erasmus+ IMPRESS 2017-1-NL01-KA203-035259).

References

1. Popescu, M., Arnab, S., Berta, R., Earp, J., de Freitas, S., Romero, M., Stanescu, I., Usart, M.: Serious Games in Formal Education: Discussing Some Critical Aspects. Proc. 5Th Eur. Conf. Games Based Learn. 486–493 (2011).
2. Calderón, A., Ruiz, M.: A systematic literature review on serious games evaluation: An application to software project management. *Comput. Educ.* 87, 396–422 (2015).
3. El-Nasr, M., Drachen, A., Canossa, A.: Game analytics maximizing the value of player data. Springer, London ;;New York : (2013).
4. Long, P., & Siemens, G.: Penetrating the Fog: Analytics in Learning and Education. *Educ. Rev.* 31–40 (2011).
5. Bienkowski, M., Feng, M., Means, B.: Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. Washington, DC SRI Int. 1–57 (2012).
6. 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).
7. European Union: Regulation (EU) 2016/679, <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679> (last visited: Sep 2018).
 8. 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).
 9. European Commission: OpenAIRE, <https://www.openaire.eu/> (last visited: Sep 2018).
 10. CERN, OpenAIRE, EC: Zenodo, <https://zenodo.org/> (last visited: Sep 2018).
 11. Chatti, M.A., Lukarov, V., Thüs, H., Muslim, A., Yousef, A.M.F., Wahid, U., Greven, C., Chakrabarti, A., Schroeder, U.: *Learning Analytics: Challenges and Future Research Directions*. E-Learning Educ. (2015).
 12. Calvo Morata, A.: *Videojuegos Como Herramienta Educativa En La Escuela: Concienciando Sobre El Ciberbullying* (Master Thesis), (2017).
 13. Alonso-Fernández, C.: *Applying Data Mining Techniques to Game Learning Analytics* (Master Thesis), (2017).
 14. Cano, A.R., Fernández-Manjón, B., García-Tejedor, Á.J.: Using game learning analytics for validating the design of a learning game for adults with intellectual disabilities. *Br. J. Educ. Technol.* (2018).
 15. Alonso-Fernandez, C., Calvo, A., Freire, M., Martinez-Ortiz, I., Fernandez-Manjon, B.: Systematizing game learning analytics for serious games. In: *IEEE Global Engineering Education Conference (EDUCON)*. IEEE (2017).
 16. Bakharia, A., Corrin, L., Barba, P. De, Kennedy, G., Gasevic, D., Mulder, R., Williams, D., Dawson, S., Lockyer, L.: A Conceptual Framework linking Learning Design with Learning Analytics. *Proc. LAK16 6th Int. Conf. Anal. Knowl.* 2016. 329–338 (2016).
 17. Gašević, D., Dawson, S., Siemens, G.: Let's not forget: Learning analytics are about learning. *TechTrends*. 59, 64–71 (2015).
 18. Jivet, I., Scheffel, M., Drachsler, H., Specht, M.: Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice. In: *Lecture Notes in Computer Science*. pp. 82–96 (2017).
 19. Sampayo-Vargas, S., Cope, C.J., He, Z., Byrne, G.J.: The effectiveness of adaptive difficulty adjustments on students' motivation and learning in an educational computer game. *Comput. Educ.* 69, 452–462 (2013).
 20. Shute, V., Ke, F., Wang, L.: Assessment and Adaptation in Games. In: *Instructional Techniques to Facilitate Learning and Motivation of Serious Games*. pp. 59–78. Springer International Publishing, Cham (2017).
 21. Loh, C.S., Sheng, Y., Ifenthaler, D.: *Serious Games Analytics*. Springer International Publishing, Cham (2015).
 22. Shute, V.J., Moore, G.R.: Consistency and Validity in Game-Based Stealth Assessment. In: *Technology Enhanced Innovative Assessment: Development, Modeling, and Scoring From an Interdisciplinary Perspective* (2017).