Simva: Simplifying the scientific validation of serious games

Ivan J. Pérez-Colado, Antonio Calvo-Morata, Cristina Alonso-Fernández
Computer Science Faculty
Complutense University of Madrid
Madrid, Spain
{ivanjper, acmorata, calonsofernandez}@ucm.es

Manuel Freire, Iván Martínez-Ortiz, Baltasar Fernández-Manjón
Computer Science Faculty
Complutense University of Madrid
Madrid, Spain
{manuel.freire, imartinez, balta}@fdi.ucm.es

***Abstract*— Serious games validation is a highly complex and burdensome process. To ensure that games meet their intended educational goals, it is necessary to have a clear experimental design and the necessary tools to minimize the errors that may appear in the process. In this article, after describing the most common problems that we have found while validating our own games, we present *Simva*, a tool designed to simplify the process of validating serious games with formal questionnaires and relating them with learning analytics data, reducing time, cost, and error rates.**

Serious Games; experiments; validation; learning analytics; game-based learning;

I. INTRODUCTION

Video games are part of our current society as a form of art and leisure. Among this growth of the industry, videogames have also appeared for purposes other than entertainment, including awareness, training or, in general, education. These types of games are commonly referred to as *serious games*, and have had an increasing impact in different domains such as medicine, psychology or business, becoming more frequent every day. Moreover, research on serious games have found them to be an effective tool to teach very diverse subjects, from mathematics [1] to social problems [2], and even to enable scientific discoveries [3].

Currently, the most accepted method to validate serious games is the use of external questionnaires [4]. This method commonly comprises two questionnaires: a pre-test conducted before playing the game, to measure the initial characteristics (e.g. knowledge) of players; and a post-test, conducted after the game, which measures changes produced by playing. Comparing the results in both tests, and since the only intervention between them is the game, researchers can conclude whether the game produces a change in that players' characteristics. If a statistically relevant, positive change is identified, the formal validation is successful. As Learning Analytics (LA) [5] provides new opportunities to capture and analyze players' interaction data to make evidence-based decisions, use of analytics is becoming more common. However, for validation, use of analytics is generally still combined with questionnaires [6].

Validation through questionnaires entails several issues that can increase the complexity of the process, as well as its time and costs. Some of these problems are how to relate data of the same user, or how to collect and store the data. The validation stage of serious games is an imperative point when creating a serious game, since it is necessary to formally prove that these meet their objectives, especially in the case of educational video games that intend to be beneficial to the players, whether to practice, learn or change their perceptions. In this work, we propose a system that simplifies the creation and deployment of questionnaires and experimental data collection to facilitate formal validation of serious games. The goal of the system is to tackle the issues that appear when validating serious games.

When carrying out pre-post experiments, if using paper questionnaires, researchers must prepare and print them in advance, and later distribute them, wasting time in this process, especially for larger samples. The use of computer-based questionnaires also has drawbacks: if the questionnaire is to be completed online, Internet access is mandatory; and if it is deployed offline, there needs to be means and time to collect all responses. Combining questionnaires with LA data requires an extra effort during analysis to unify all responses for each participant. When anonymized data is desired, as it is commonly the case, and, if working with minors, mandatory, a unique identifier can be used, but it must be present in all data sources to link each users' responses together.

Storing the data collected if questionnaires are paper-based lead to huge accumulations of paper, which must then be transcribed into a computer for analysis, a burdensome and error-prone process. Conducting surveys on electronic devices also presents problems. If Internet connection is available, data can be sent directly to a server that stores the answers. However, it is necessary to have a backup plan in case of either the network or server failure. Otherwise data could be lost irrecoverably. If no Internet connection is available, data must be stored locally and manually collected for each computer at the end of the experiment.

The rest of this paper is structured as follows. Section II describes *Simva*, a tool that we have developed to facilitate the process of formally validating serious games using questionnaires combined with LA. Finally, Section III discusses the results and summarizes the main conclusions of our contribution.

II. SIMVA

To address the previously identified issues, we have developed a tool called *Simva*, which also allows monitoring of the different parts of the experiment, such as which questionnaires have been started and completed; and allows experimenters to verify that the data have been correctly received and stored. In the following subsections, we provide an overview of the main features of the tool, and describe a method of including and connecting *Simva* with the rest of our LA framework to solve the same-user-data relation issue.

A. *Features and architecture*

The main requirements of *Simva* directly address the issues that add complexity to the validation process. The tool must ensure that all data collected from a user is stored together; ensure that all questionnaire information is correctly received and stored; and that the questionnaire is correctly designed to be amenable to later analysis.

The tool must therefore be able to manage groups of students; and students need to be identified pseudonymously, so that their interactions remain together without revealing their actual identity. These pseudonymous identifiers must be easy to distribute to users. It is also required that the tool must notify other external tools about internal group management, in a way that allows, e.g. a Learning Management System (LMS) such as Moodle or a student-based Analytics Framework, to create and update the groups according to the groups in the tool, so the structure is shared between all the tools. The tool must also allow surveys to be created, modified, assigned to groups, and their state of completion to be followed, all linked to the pseudonymous identifiers.

For creating the questionnaires we have reused a free software tool called *LimeSurvey* [7]. This tool allows its users to create surveys, run them in a controlled environment, and export the responses in a wide variety of formats. The tool also included a *Remote Control API* that enables the possibility to manage LimeSurvey from an external tool. As this tool fits most of our identifier requirements, we decided to build our tool, called *Simva*, reusing the already available features of LimeSurvey. For this, the development of the *Simva* tool includes a controller capable of managing LimeSurvey remotely using an interface.

The use of LimeSurvey as the external survey management software allows users of *Simva* to easily create surveys, run them, and export the results. The rest of the requirements are directly available in *Simva*, including the management of surveys and groups of students, and the assignment of surveys to these groups. All students are issued random four-letter tokens to use as pseudonymous identifiers, provided automatically when the group of students is created in the tool. Also, as required for our use case as part of an Analytics System, when students are created in the tool, they are automatically created in the Analytics System.

Each survey is composed of three questionnaires: a pre-test, a post-test; and an additional questionnaire. Survey results are directly accessible from the tool (see Figure 1), as is the state of all questionnaires (one of "not found", "started", or "finished"). While raw survey data could have been exported using LimeSurvey, all answers would have been lumped together, instead of being separated by groups. *Simva* can request only the filtered answers, and download them as single file.

We have used *Simva* multiple times to validate serious games, together with an Analytics System that was actively collecting interaction data from these games. To simplify this process, which we consider to be a frequent use-case, *Simva* allows games to store any logs, traces, results, or data desired in the tool. In our case, games are connected to the Analytics System by means of a *tracker* module, which sends the interactions to the server in real time and saves a backup of these interactions to the local device. This trace backup is sent to *Simva* when the game is closed or all surveys are completed. On the right column of Figure 1, we can see the traces generated by the interactions of each student within the game (identified by their 4-letter pseudonymous tokens). Researchers can directly download these traces from this view, possibly after discarding users that have not completed their questionnaires.

B. *Additional features*

In our previous experience validating serious games, we have encountered several issues that were not easy to predict in advance. We have included additional features into *Simva* to address them:

1. **Use of identifiers**: the tool manages anonymous identifiers, and it must also provide a way to distribute them. To expedite this, the tool can generate, per group of students, printable PDF and MS Word files, participants can use the same identifier multiple times to keep their long-term information related. A screenshot is presented as Figure 2.
2. **Survey completion**: the tool controls the state of completion of the surveys so students do not start playing without completing the pre-game survey, and do not leave without completing the post survey. This helps to avoid inconsistent data, unlinked gameplays, and lost results.
3. **Additional metadata**: the tool can act as a reference location where games can retrieve metadata. This could be useful, for example, to enable the use of A/B

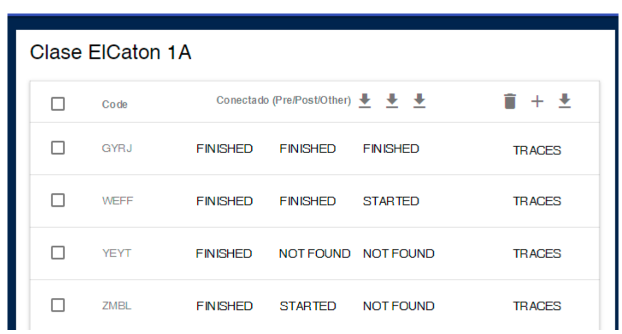

Figure 1. Fragment view of a class in Simva. View and access to the completed test and interactions collected from users in the game.

Clase ElCaton 1A:					
No.	Nombre	Código			
1		GYRJ	GYRJ	GYRJ	GYRJ
2		WEFF	WEFF	WEFF	WEFF
3		YEYT	YEYT	YEYT	YEYT
4		ZMBL	ZMBL	ZMBL	ZMBL
5		WSFJ	WSFJ	WSFJ	WSFJ

Figure 2. Printed list of anonymous tokens for students in a class.

scenarios, where a part of a group will be playing version A of the game, and the other part will be playing version B of the game. This metadata has been implemented to be extensible, so new types of metadata can be added to a survey easily.

III. DISCUSSION AND CONCLUSIONS

Carrying out experiments to validate serious games is a complex and costly process, which involves working with third parties and in facilities outside the control of researchers. For example, in the case of validating an educational game, problems may arise with the connection to the Internet or with the hardware available at the school labs. It is therefore necessary to create tools that simplify researchers' tasks to carry out these experiments and reduce the possibility of errors.

Simva has helped us in different validation experiments to simplify data collection and processing, saving us time and automating some of the processes that were most prone to errors. We have used *Simva* with more than 1000 students and 3 different serious games: in *Conectado* [2] we linked questionnaires with the collected LA data; in *First Aid Game* [8] we linked together data from two experiments over time; and in the *15 Objects test* [9] we used the metadata feature to include the information of which of the two versions of the game each student was using.

As presented in this article, *Simva* is a tool which effectively addresses these issues: it facilitates the assignment of questionnaires and the monitoring of their status (e.g. completed, in progress); it creates unique pseudonymous identifiers to anonymize players while linking all their data as collected from different sources (e.g. different questionnaires, interaction data); it allows to access to questionnaires and the application to be controlled, by querying whether a user has access to a questionnaire or exists in the corresponding class; and it simplifies repetition of experiments.

Although we consider that the developed tool greatly simplifies most scenarios, there are some limitations that *Simva* does not currently consider. For example, gathering data in cases where there is no Internet access during the experiment, neither before nor after, and where all the data must therefore be saved in the machine where the application itself is executed, remains a manual and burdensome task for researchers.

ACKNOWLEDGMENT

This work has been partially funded by Regional Government of Madrid (eMadrid P2018/TCS4307), by the Ministry of Education (TIN2017-89238-R), by the European Commission (BEACONING H2020-ICT-2015-687676, Erasmus+ IMPRESS 2017-1-NL01-KA203-035259) and by the Telefonica-Complutense Chair on Digital Education and Serious Games.

REFERENCES

[1] M. Bakker, M. van den Heuvel-Panhuizen, and A. Robitzsch, “Effects of playing mathematics computer games on primary school students’ multiplicative reasoning ability,” *Contemp. Educ. Psychol.*, vol. 40, pp. 55–71, Jan. 2015.

[2] A. Calvo-Morata, D. C. Rotaru, C. Alonso-Fernandez, M. Freire, I. Martinez-Ortiz, and B. Fernandez-Manjon, “Validation of a Cyberbullying Serious Game Using Game Analytics,” *IEEE Trans. Learn. Technol.*, pp. 1–1, 2018.

[3] B. Stokes, N. Walden, G. O’Shea, F. Nasso, G. Mariutto, and A. Burak, “Impact with games: A fragmented field,” pp. 1–28, 2015.

[4] A. Calderón and M. Ruiz, “A systematic literature review on serious games evaluation: An application to software project management,” *Comput. Educ.*, vol. 87, pp. 396–422, Sep. 2015.

[5] P. Long, G. Siemens, C. Gráinne, and D. Gašević, “LAK ’11 : proceedings of the 1st International Conference on Learning Analytics and Knowledge, February 27 - March 1, 2011, Banff, Alberta, Canada,” in *1st International Conference on Learning Analytics and Knowledge*, 2011, p. 195.

[6] M. Liu, J. Kang, S. Liu, W. Zou, and J. Hodson, “Learning Analytics as an Assessment Tool in Serious Games: A Review of Literature,” in *Serious Games and Edutainment Applications*, Cham: Springer International Publishing, 2017, pp. 537–563.

[7] The LimeSurvey Project Team, “LimeSurvey,” 2013. [Online]. Available: https://www.limesurvey.org/.

[8] C. Alonso-Fernández, R. Caballero Roldán, M. Freire, I. Martinez-ortiz, and B. Fernández-Manjón, “Predicting students’ knowledge after playing a serious game based on learning analytics data (under review),” *J. Comput. Assist. Learn.*, 2019.

[9] M. C. D.-M. Dan-Cristian Rotaru, Sara García-Herranz, Manuel Freire, Iván Martínez-Ortiz, Baltasar Fernández-Manjón, “Using Game Technology to Automatize Neuropsychological Tests and Research in Active Aging,” *GOODTECHS 2018 - 4th EAI Int. Conf. Smart Objects Technol. Soc. Good*, 2018.