

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE INFORMÁTICA



TESIS DOCTORAL

Contribuciones a la adopción de los juegos serios como herramientas
educativas mediante técnicas de analíticas de aprendizaje

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Iván José Pérez Colado

DIRECTORES

Iván Martínez Ortiz
Baltasar Fernández Manjón

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE INFORMÁTICA



TESIS DOCTORAL

CONTRIBUCIONES A LA ADOPCIÓN DE LOS JUEGOS SERIOS COMO
HERRAMIENTAS EDUCATIVAS MEDIANTE TÉCNICAS DE ANALÍTICAS DE
APRENDIZAJE

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Iván José Pérez Colado

DIRECTORES

Iván Martínez Ortiz

Baltasar Fernández Manjón

Madrid, 2023

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE INFORMÁTICA



TESIS DOCTORAL

**CONTRIBUCIONES A LA ADOPCIÓN DE LOS JUEGOS SERIOS COMO
HERRAMIENTAS EDUCATIVAS MEDIANTE TÉCNICAS DE ANALÍTICAS DE
APRENDIZAJE**

**MEMORIA PARA OPTAR AL GRADO DE DOCTOR EN EL
PROGRAMA DE DOCTORADO EN INGENIERÍA INFORMÁTICA**

PRESENTADA POR

Iván José Pérez Colado

DIRECTORES

Iván Martínez Ortiz

Baltasar Fernández Manjón

Madrid, 2023

*“Verás, existe una conexión fundamental entre lo que uno parece y lo que uno es.
Todos los niños Fata lo saben, pero vosotros, los mortales, no lo veis.
Nos contamos una historia sobre nosotros mismos. Siempre. Continuamente.
Esa historia es lo que nos convierte en lo que somos. Nos construimos a nosotros
mismos a partir de esa historia.”*

Patrick Rothfuss – El Nombre del Viento

Agradecimientos

Voy a intentar ser breve, porque, aunque tengo bastantes personas en mi cabeza por mencionar, no puedo mencionar a todos y pido disculpas por adelantado si me olvido de alguien. Si no te he mencionado a ti en específico, pues aprovecho para darte las gracias a ti también, que seguro que hemos compartido algunos buenos momentos juntos que me han llevado hasta este momento y lugar.

Empiezo agradeciendo a mis compañeros del Aula 16 y del doctorado en general, en especial agradecer a: Moh, Joaquín, Marti, Pablo, Miguel, y Jesús, con los que hemos compartido infinitos cafés, juegos, cotilleos, cenas, cervezas, e incluso algún viaje loco. También trasteando con la impresora 3D. La verdad es que fue increíble haber podido estar allí en el despacho con vosotros, recordaré esos días por siempre.

Sigo con Cristian, Gorka, Irene, Isma, Frank y Max, que sois buenos amigos míos y os he conocido gracias a la Universidad, y que tendría que dedicaros un párrafo a cada uno sobre las horas compartidas y paseos juntos. Os diré algo íntimo que compartís acerca de mi: todos me habéis visto llorar, y todos habéis cuidado de mi persona. Os estoy infinitamente agradecido por ayudarme, y empujarme hasta aquí.

A Alma gracias por sus pruebas hechas con todas las herramientas, Ana Rus con quien compartimos varias sesiones divertidas e interesantes de experimentos, y a Piotr que ayudó increíblemente a que uAdventure se convirtiese en una realidad gracias a la primera versión de los editores que hizo. Gracias por todas vuestras contribuciones.

A mis compañeros del grupo e-UCM: Toni, Cristi y Manu. Es difícil decir cuántísima ayuda me habéis dado en todo este tiempo. No solo por todos los proyectos en los que hemos trabajado juntos, muchas veces esforzándonos de manera heroica, sino especialmente por ayudarme a presentar los logros que íbamos consiguiendo. Vuestro apoyo y guía me ha hecho sentir arropado y seguro de que conseguiríamos completar todo lo que nos propusiéramos. Estoy infinitamente agradecido con vosotros.

A mis directores, Iván “*el Señor de las Máquinas*”, a quien recuerdo desde el principio con su camiseta de la “guardia de la noche” mientras me enseñaba fundamentos de la programación y quien es, probablemente, la persona que más me ha enseñado en la informática, y a Balta “*The Collejator*”, quien me ha hecho aprender y crecer como persona, enseñándome lo que debía exigirme a mí mismo, poniéndome en primera línea

y a su lado, y confiando en mí. A los dos, gracias por haber apostado por mí desde el principio, por guiarme a lo largo de todo el camino, y por enseñarme lo lejos que podía llegar. Gracias de verdad, por todo, y por haber hecho esta tesis posible.

A Milén, Sora, Javi y Chechu, mis amigos de toda la vida, y a Lalo, Wolf, Frejya, Aitor, Nacho y Eva, los más nuevos. Solo desearía más tiempo con vosotros, porque a vuestro lado he vivido las aventuras más increíbles de mi vida. Juntos hasta el final.

A *Rejvel*, por tu cariño y apoyo constante. Es tanto lo compartido que es difícil pensar en por qué darte las gracias sin echar una lagrima de ternura. Gracias por tu cariño infinito, y por siempre creer en mí.

A *Pincha*, por hacerme tan inmensamente feliz cada día. Los que hemos pasado juntos mientras escribía la tesis son algunos de mis favoritos de toda mi vida, compartirlos contigo ha sido un sueño, y despertarme contigo abrazándome a mi lado lo hacía realidad.

Y por supuesto, a mi familia, que me habéis dado todo desde antes de que tuviera conciencia. No solo es el apoyo infinito, es el amor incondicional lo que ha hecho que, aunque nuestra vida no fuese la más opulenta, para mí siempre fuese la más abundante. Y a ti, Victorma, porque eres el mejor *Player 2* con el que jugar a la vida. Gracias de corazón.

Sobre este documento

Este trabajo de tesis doctoral es una recopilación de publicaciones de acuerdo con lo expuesto en la Normativa de Desarrollo del Real Decreto 99/2011, del 28 de Enero (BOE 10/02/2011), por el que se regulan las enseñanzas oficiales de doctorado en la Universidad Complutense de Madrid¹ ².

A continuación, se enumeran los artículos presentados:

- ☰ Iván José Pérez Colado, Cristina Alonso Fernández, Manuel Freire Morán, Iván Martínez Ortiz y Baltasar Fernández Manjón (2018). **Game learning analytics is not informagic!** In *2018 IEEE Global Engineering Education Conference (EDUCON)*. IEEE, 1729-1737 <https://doi.org/10.1109/educon.2018.8363443>
- ☰ Iván José Pérez Colado, Dan Cristian Rotaru, Manuel Freire Morán, Iván Martínez Ortiz y Baltasar Fernández Manjón (2018). **Multi-Level Game Learning Analytics for Serious Games**. In *2018 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. IEEE, 1-4 <https://doi.org/10.1109/vs-games.2018.8493435>
- ☰ Iván José Pérez Colado, Antonio Calvo Morata, Cristina Alonso Fernández, Manuel Freire Morán, Iván Martínez Ortiz y Baltasar Fernández Manjón (2019). **Simva: Simplifying the Scientific Validation of Serious Games**. In *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 113-115 <https://doi.org/10.1109/icalt.2019.00033>
- ☰ Antonio Calvo Morata, Cristina Alonso Fernández, Iván José Pérez Colado, Manuel Freire Morán, Iván Martínez Ortiz y Baltasar Fernández Manjón (2019). **Improving Teacher Game Learning Analytics Dashboards through ad-hoc Development**. *JUCS - Journal of Universal Computer Science* 25(12): 1507-1530. <https://doi.org/10.3217/jucs-025-12-1507>
- ☰ Víctor Manuel Pérez Colado, Iván José Pérez Colado, Manuel Freire Morán, Iván Martínez Ortiz, Baltasar Fernández Manjón (2019). **uAdventure: Simplifying**

¹ <https://www.boe.es/eli/es/rd/2011/01/28/99/con>

² <https://edotorado.ucm.es/normativa>

Narrative Serious Games Development. In *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*. 2019. IEEE, 119-123 <https://doi.org/10.1109/icalt.2019.00030>

- ☰ Iván José Pérez Colado, Víctor Manuel Pérez Colado, Iván Martínez Ortiz, Manuel Freire Morán y Baltasar Fernández Manjón (2020). **Simplifying Serious Games Authoring and Validation with uAdventure and SIMVA.** In *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 106-108 <https://doi.org/10.1109/icalt49669.2020.00039>
- ☰ Cristina Alonso Fernández, Iván José Pérez Colado, Antonio Calvo Morata, Manuel Freire Morán, Iván Martínez Ortiz y Baltasar Fernández Manjón (2020). **Applications of Simva to Simplify Serious Games Validation and Deployment.** In *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje (Vol. 15, Issue 3, pp. 161–170)*. Institute of Electrical and Electronics Engineers (IEEE). <https://doi.org/10.1109/rita.2020.3008117>
- ☰ Iván José Pérez Colado, Víctor Manuel Pérez Colado, Manuel Freire Morán, Iván Martínez Ortiz y Baltasar Fernández Manjón (2021). **e-learning Standards in Game-Based Learning?** In *2021 International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 81-82 <https://doi.org/10.1109/icalt52272.2021.00032>

De acuerdo con la normativa, este documento incluye un índice con los títulos de los capítulos, un estudio del estado del arte, una descripción de los objetivos propuestos para esta tesis y una discusión sobre el contenido incluido en los artículos presentados.

Adicionalmente se presenta un capítulo con conclusiones y posibles líneas futuras de investigación relacionada con el tema.

Finalmente, se incluye una bibliografía con las referencias de los artículos mencionados en esta tesis doctoral.

Resumen

Los juegos serios son aquellos videojuegos con objetivos que van más allá del ocio, pudiendo aplicarse con multitud de finalidades y en diferentes sectores. En particular, en la enseñanza los juegos serios ofrecen al estudiante una manera atractiva e inmersiva de aprender. Permiten además simular situaciones donde el jugador puede poner a prueba su conocimiento experimentando y aprendiendo de sus errores de forma segura. Pese a su continuo crecimiento en distintos sectores, todavía presentan una escasa adopción en la educación y los colegios. Esto es debido a múltiples factores, incluyendo la desconfianza de profesores y organizaciones en su eficacia, o las dificultades tecnológicas para su uso e implantación; además, al ser frecuentemente distribuidos como herramientas de caja negra, resulta difícil para los profesores entender qué están haciendo los alumnos mientras juegan, y cómo y dónde está teniendo lugar el aprendizaje.

Para evaluar la eficacia de un juego serio, la validación formal se basa en medir los cambios que provoca en sus jugadores, ya se trate de incremento de conocimiento, mayor concienciación, u otros. La forma más frecuente de medir estos cambios consiste en pasar encuestas a los jugadores de forma previa y posterior a la sesión de juego; la diferencia entre encuestas, atribuible al juego, proporciona una medida de su efecto. Aunque la validación proporciona confianza de cara a la utilización de un juego, permite mejorarlo y alarga su ciclo de vida, sólo es un primer paso. Un profesor que se anime a usar un juego serio querrá ir más allá, saliendo del modelo de caja negra para poder entender mejor cómo aprenden sus alumnos y dar soporte a los alumnos que lo necesiten.

El proceso de aprendizaje de un estudiante con un juego puede estudiarse en base a la actividad del estudiante cuando juega, en lo que se denomina analítica de aprendizaje aplicadas a juegos. Esto requiere que los desarrolladores implementen toda la funcionalidad necesaria en el juego para registrar y enviar al sistema de analítica las acciones de usuario, y un sistema de analíticas que almacene, procese y presente los resultados. La literatura muestra una clara prevalencia de las soluciones ad-hoc para las analíticas de aprendizaje y con una escasa aplicación de estándares técnicos que faciliten la interoperabilidad de dichas soluciones. La complejidad y los altos costes de implementación de estos sistemas pueden reducirse a través del uso y adopción de sistemas genéricos, el uso de software libre y estándares abiertos. Pero la escasez de este tipo de soluciones frena su implantación. Además, la recolección y análisis de datos

presenta problemas de privacidad, lo cual tiene consecuencias tanto éticas como legales, a las que cualquier sistema debe dar una respuesta adecuada.

Esta tesis analiza cómo pueden aprovecharse las analíticas de aprendizaje de juegos para mejorar diversos aspectos del ciclo de vida de los juegos serios y finalmente generalizar su adopción en la educación. Estas mejoras están enfocadas en facilitar su implantación, generando modelos de integración que permiten sistematizar el proceso de desarrollo de analíticas, y que permiten mejorar etapas de evaluación de calidad (como se ha hecho en la participación en los proyectos europeos H2020 RAGE y Beaconing). Las analíticas de aprendizaje proporcionan nuevas capacidades de supervisión, evaluación y control de los docentes. Dada la importancia de las herramientas de autoría dentro del ciclo de vida de un juego, se estudia la integración de analíticas educativas en el propio desarrollo del juego, incorporando un modelo de analíticas educativas genérico en la herramienta de autoría uAdventure, lo que permite incluir GLA en los juegos automáticamente. Los estudios realizados han culminado en la creación de una plataforma de GLA orientada a validación de juegos serios llamada Simva, que facilita el diseño y desarrollo de experimentos. Esta plataforma permite crear grupos de usuarios pseudoanonimizados basados en *tokens*, códigos cortos fácilmente distribuibles físicamente, que identifican a los participantes de cada experimento para darles acceso tanto a encuestas online como al juego. Estos *tokens* permiten vincular los registros de sus sesiones de juego a los sistemas de GLA y además simplifican el cumplimiento de la normativa europea sobre privacidad (GDPR). Se ha incluido también el soporte de estándares educativos, como xAPI, LTI o CMI-5, que permiten y simplifican la integración de la herramienta con otras como campus virtuales, y la homogeneización y reutilización de los datos recogidos.

El uso de la plataforma Simva nos ha permitido también validar la propuesta de la tesis, simplificando la validación formal y el acceso a datos de analítica. El estudio de la automatización del proceso de GLA y validación utilizando uAdventure y Simva ha permitido a los usuarios medir y validar juegos de manera automatizada, integrando ambos procesos en el proceso de desarrollo del juego serio. En definitiva, mediante el enfoque planteado se da soporte a reducir el impacto de las barreras que frenan la adopción de los juegos serios con la que generar una oferta educativa alternativa.

Palabras clave: juegos serios, game learning analytics, visualización de datos, herramientas de investigación, herramientas de autoría.

Abstract

Serious games are video games with a purpose beyond entertainment and can be applied to different goals and sectors. Specifically in education, serious games offer students an attractive and immersive way of learning. They also allow simulating situations where players can test their knowledge by experimenting and learning from mistakes safely. Despite their continuous growth in different sectors, their adoption in education and schools is still low. This is due to multiple factors, including distrust of teachers and organizations in their effectiveness, or technological difficulties in their use and implementation; additionally, since games are often distributed as black box tools, it is difficult for teachers to understand what students are doing while playing, and how and where learning is taking place.

The effectiveness of a serious game can be evaluated using formal validation, which relies on measuring the changes it produces in its players, be it increased knowledge, increased awareness, or other. The most common way to measure these changes is to survey players before and after the game session; the difference between surveys, attributable to the game, provides a measure of its effect. While validation provides confidence when using a game, allows it to be improved, and extends its life cycle, it is only a first step. A teacher who is willing to use a serious game will want to go beyond the black box model to better understand how their students learn, and to provide support to students who need it.

The learning process of students within a game can be studied based on their activity while playing, in what is called game learning analytics. This requires developers to implement all the necessary functionality in the game to record and send user actions to an analytics system; and an analytics system to store, process and present the results. The literature shows a clear prevalence of ad-hoc implementations of learning analytics, with very limited use of standards. The complexity and high implementation costs of these systems can be reduced through the use and adoption of generic systems, free software, and open standards. But the scarcity of these types of solutions slows down their implementation. In addition, data collection and analysis present privacy issues, which have both ethical and legal consequences, to which any such system must provide an adequate response.

This thesis analyzes how game learning analytics (GLA) can be leveraged to improve various aspects of the life cycle of serious games and ultimately generalize their adoption

in education. These improvements are focused on facilitating their implementation, generating integration models that allow systematizing the analytics development process, and that allow improving quality evaluation stages (as exemplified in participation in the European H2020 projects RAGE and Beaconing). Learning analytics provide teachers with new capabilities for monitoring, evaluation and control. Given the importance of authoring tools within the life cycle of a game, this work studies the integration of learning analytics in the game development itself, incorporating a generic learning analytics model in the uAdventure authoring tool, which allows GLA to be included automatically into games. The studies conducted have culminated in the creation of a GLA platform oriented towards validation of serious games called Simva, which facilitates the design and development of experiments. This platform allows the creation of pseudo-anonymized user groups based on tokens, short codes which can be easily distributed physically to participants in each experiment, which identify them and provide access to both online surveys and the game. Such tokens allow linking the records of their game sessions to the GLA systems and also simplify compliance with European privacy regulations (GDPR). We have also included support for educational standards, including xAPI, LTI or cmi5, which allow and simplify the integration of the tool with other tools such as virtual campuses, and the homogenization and reuse of collected data.

The use of the Simva platform has also allowed us to validate the thesis proposal, simplifying formal validation and access to analytics data. The study of the automation of GLA and the validation process using uAdventure and Simva has allowed users to measure and validate games in an automated way, integrating both processes into the broader serious game development process. Ultimately, the proposed approach reduces the impact of barriers that hinder the adoption of serious games as a key ingredient of an improved educational experience.

Keywords: serious games, game learning analytics, data visualization, research tools, authoring tools.

Contenido

Agradecimientos	XVII
Sobre este documento	XIX
Resumen	XXI
Abstract	XXIII
Contenido	XXV
Índice de figuras.....	XXIX
Estructura de la tesis	XXXIII
Capítulo 1. Introducción y motivación	35
1.1. Breve reflexión personal sobre los juegos serios.....	35
1.2. Motivación.....	36
1.3. Línea de investigación	41
Capítulo 2. Estado del arte.....	43
2.1. Los juegos serios	43
2.2. Creación / Herramientas de autoría / modelos de desarrollo.....	47
2.3. Validación.....	52
2.4. Analíticas de juegos (GLA)	55
2.5. Estándares	60
2.6. xAPI y los perfiles de aplicación xAPI	68
2.7. A modo de conclusión	72
Capítulo 3. Objetivos.....	75
3.1. Mejoras en el proceso de diseño de analíticas.....	78
3.2. Sistematización del proceso de integración de GLA con herramientas de autoría y juegos serios	79
3.3. Desarrollo de una plataforma de GLA para la validación sistemática de juegos serios.....	81
3.4. Aplicación de estándares educativos	83
3.5. Validación del sistema y sus componentes.....	85
Capítulo 4. Discusión integradora y contribuciones.....	87

4.1. Mejoras en el proceso de diseño de analíticas	87
4.2. Sistematización del proceso de integración de GLA con herramientas de autoría y juegos serios	92
4.3. Desarrollar una plataforma de GLA orientada a la validación de Juegos Serios	101
4.3.1. Prueba de concepto y prototipo: SurveyManager.....	101
4.3.2. Plataforma de GLA y validación: Simva.....	105
4.3.3. Integración de Simva con herramientas de autoría de juegos serios: uAdventure.....	114
4.4. Estándares de e-learning, LTI y CMI-5.....	115
4.5. Pruebas con usuarios y casos de uso	118
4.5.1. Caso de estudio del Juego de Primeros Auxilios	120
4.5.2. Caso de estudio del proyecto Beaconing: analíticas multi-nivel.....	122
4.5.3. Pruebas de validación de Juegos Serios	123
Capítulo 5. Conclusiones y Trabajo futuro	129
5.1. Conclusiones y principales aportaciones	129
5.2. Aplicaciones Software.....	134
5.3. Trabajo futuro.....	134
Capítulo 6. Artículos presentados	138
6.1. Game Learning Analytics is not informagic!	138
6.1.1. Cita completa.....	138
6.1.2. Resumen original de la publicación.....	138
6.2. Multi-level Game Learning Analytics for Serious Games	148
6.2.1. Cita completa.....	148
6.2.2. Resumen original de la publicación.....	148
6.3. Simva: Simplifying the scientific validation of serious games	153
6.3.1. Cita completa.....	153
6.3.2. Resumen original de la publicación.....	153
6.4. Improving Teacher Game Learning Analytics Dashboards through ad-hoc Development.....	157
6.4.1. Cita completa.....	157
6.4.2. Resumen original de la publicación.....	157
6.5. uAdventure: Simplifying Narrative Serious Games Development	182

6.5.1. <i>Cita completa</i>	182
6.5.2. <i>Resumen original de la publicación</i>	182
6.6. Simplifying Serious Games Authoring and Validation with uAdventure and SIMVA	188
6.6.1. <i>Cita completa</i>	188
6.6.2. <i>Resumen original de la publicación</i>	188
6.7. Applications of Simva to Simplify Serious Games Validation and Deployment	192
6.7.1. <i>Cita completa</i>	192
6.7.2. <i>Resumen original de la publicación</i>	192
6.8. e-learning Standards in Game-Based Learning?	203
6.8.1. <i>Cita completa</i>	203
6.8.2. <i>Resumen original de la publicación</i>	203
Referencias	209

Índice de figuras

Figura 1. Distribución por sectores de los juegos serios entre los periodos de 1980-2001 con 953 juegos y 2002-2001 con 1265 juegos respectivamente	45
Figura 2. Tabla que representa los estándares educativos y sus características y propósitos, mostrando en la sección superior aquellos estándares con objetivos más específicos, y en la inferior aquellos que tratan de cubrir múltiples necesidades.....	66
Figura 3. Representación de una traza xAPI dividida en secciones, mostrando el actor en la parte superior, seguido del verbo, el objeto, así como el momento en el que ocurrió y su resultado.....	68
Figura 4. Planificación de los objetivos de la tesis.....	77
Figura 5. Relación entre LAM (rectángulo sombreado inferior) y LAS (rectángulo en parte superior) - el LAS proporciona el mecanismo, y el LAM la información de cómo proceder.	80
Figura 6. Representación del proceso de ejecución de GLA. El estudiante comienza jugando al juego que incluye un componente de seguimiento, llamado Tracker, que envía los datos de interacción al sistema de analíticas, donde se almacena, analiza, y prepara para su posterior visualización por profesores u otros usuarios para extraer conclusiones.	93
Figura 7. Proceso de aplicación del modelo de analíticas utilizando el editor de uAdventure donde se especifica que una escena se compone de múltiples alternativas a seleccionar, y cómo generar analíticas con ella.....	95
Figura 8. Recopilatorio de todos los eventos que genera la integración de GLA implementada en la herramienta de autoría de juegos serios uAdventure.	95
Figura 9. Editor de uAdventure para especificar qué preguntas son correctas e incorrectas, pudiendo facilitar dichos datos al sistema de analíticas para su análisis y visualización de los resultados.	96
Figura 10. Conjunto de editores de uAdventure para especificar <i>completables</i> en forma de tareas e hitos para dividir el juego permitiendo crear un modelo de progreso, además de permitir asociar puntuaciones a cada tarea.	96

Figura 11. Conjunto de visualizaciones genéricas que forman un panel para los juegos creados con la herramienta de autoría de juegos serios uAdventure. Este panel se incluye automáticamente sin necesidad de programar.	98
Figura 12. Paneles de visualizaciones ad-hoc desarrollados para (1) el juego Conectado, (2) el juego FormalZ, y (3) el juego WaterCooler que aplica el modelo Thomas-Kilmann.	100
Figura 13. La herramienta SurveyManager, mostrando: (A) un PDF imprimible para el usuario donde anotar los nombres de los participantes, asociados a cada token único recortable para su fácil distribución durante el experimento, y (B) captura de pantalla de la interfaz de una actividad ya creada, mostrando las diferentes encuestas, su estado, y los ficheros de trazas del juego recogidos, pudiendo descargar todos los resultados.	103
Figura 14. Proceso de GLA a través de SurveyManager, incluyendo al juego, al sistema de analíticas, y cómo profesor y estudiantes interactúan con el sistema durante el desarrollo del experimento.	104
Figura 15. Sistema de gestión de usuarios y grupos en Simva (B), mostrando las capacidades integradas para la creación de lotes de usuarios pseudoanonimizados con un simple formulario (A), junto a un PDF imprimible que representa el grupo creado (C).	107
Figura 16. Captura de Simva mostrando la interfaz de los estudios con dos ramas de estudio diferentes, una con solo <i>gameplay</i> y otra con encuestas <i>pre-post</i>	108
Figura 17. Actividades implementadas en Simva, incluyendo: Actividad de LimeSurvey para encuestas, Actividad de Gameplay para jugar al juego, Actividad manual donde el investigador marca manualmente, además de la actividad de LTI que conecta con una herramienta externa.	109
Figura 18 Captura de Simva mostrando la interfaz de los estudios con dos ramas de estudio diferentes, una con solo <i>gameplay</i> y otra con encuestas <i>pre-post</i>	111
Figura 19. Planificador de Simva ejecutando un estudio para el participante “o57j” donde, primero realiza la encuesta previa, y después realiza la sesión de juego. La transición entre las actividades la realiza dicho planificador, de manera automática, cuando identifica que las actividades se han completado.	112

Figura 20. Proceso de <i>data-science</i> utilizando Simva y TxMon, un proyecto que permite ejecutar el LAM por defecto desarrollado en Jupyter Notebooks.	113
Figura 21. Integración de Simva en uAdventure con un editor que automatiza el despliegue del experimento con una sencilla configuración incluyendo las encuestas, analíticas y los participantes.	114
Figura 22. Posibles conexiones que se pueden realizar entre Simva, un LMS y una herramienta que implemente el estándar de interoperabilidad LTI. La representación muestra a Simva con rol de herramienta utilizada por la plataforma LMS Moodle, que incluye a Simva en un curso. A su vez, Simva integra la herramienta LTI Pumva, repositorio de juegos serios de uAdventure donde delegar la ejecución del juego, manteniendo las ventajas del uso de Simva, facilitando y simplificando su uso.	116
Figura 23. Editores de Simva que utilizan LTI para añadir tanto plataformas habilitadas para utilizar Simva como herramienta, así como añadir herramientas para ser utilizadas en actividades mediante la plataforma Simva.	117
Figura 24. Fotografías de una de las primeras sesiones de prueba realizadas con el juego de Primeros Auxilios tomadas en 2017, donde se observa a los participantes jugar al juego y completar las encuestas. La fotografía de la esquina inferior izquierda muestra el proceso de evaluación de los paneles de analíticas desarrollados. En ellas aparecen los investigadores Antonio Calvo-Morata, Cristina Alonso-Fernández e Iván J. Pérez-Colado.....	121
Figura 25. Captura de pantalla de la integración ad-hoc de Simva en el juego Conectado. La imagen de la izquierda muestra el formulario donde introducir el código de acceso, dando acceso a la encuesta previa.	124
Figura 26. Fotografías de algunas de las sesiones realizadas durante los experimentos de validación del juego Conectado. La imagen de la izquierda muestra al investigador I. Perez-Colado realizando una encuesta al profesor acerca de los paneles de analíticas durante la sesión. A la derecha, junto a Calvo-Morata, los alumnos participan en el experimento jugando al juego.....	125

Estructura de la tesis

Esta tesis doctoral utiliza el formato de recopilación de publicaciones. Esta recopilación se ha integrado en el capítulo 6, al final de esta tesis. En los capítulos previos se proporciona la línea de investigación que se ha desarrollado, ofreciendo un contexto y un hilo conductor entre las diferentes contribuciones que se presentan.

El documento se estructura de la siguiente forma:

- Capítulo 1. Introducción y motivación
- Capítulo 2. Estado del arte
- Capítulo 3. Objetivos y planteamiento de la tesis
- Capítulo 4. Discusión integradora y contribuciones
- Capítulo 5. Conclusiones y trabajo futuro
- Capítulo 6. Artículos presentados

Al final del documento se incluyen las referencias bibliográficas completas que se han utilizado en este documento.

Capítulo 1. Introducción y motivación

Esta tesis busca simplificar la adopción de los juegos serios y contribuir a mejorar su ciclo de vida mediante la sistematización y automatización de las analíticas de aprendizaje.

Para contextualizar el trabajo se ha realizado un estudio del estado del arte de los juegos serios y las analíticas de aprendizaje, incluyendo los estándares educativos aplicables. A partir de dicho contexto, este trabajo propone una metodología genérica para la sistematización del proceso de desarrollo de las analíticas de aprendizaje en juegos o Game Learning Analytics (GLA). El enfoque se ha probado sobre varios juegos que usaban GLA ad-hoc, para posteriormente trabajar en su automatización, e incluso en la integración de un modelo de análisis genérico en la herramienta de autoría de juegos serios uAdventure. De este modo se cierra el ciclo vida de simplificación, ya que todos los juegos desarrollados con uAdventure incorporarán analíticas de aprendizaje automáticamente sin requerir de sus autores conocimientos de las mismas - aunque, de así deseárselo, pueden fácilmente ser adaptadas a cualquier necesidad. Esta automatización ha sido posible ya que previamente se desarrolló una plataforma de GLA basada en estándares (y en particular xAPI) llamada Simva, que no sólo permite simplificar el GLA para cualquier juego serio, sino que además permite gestionar el diseño experimental de validación y las actividades de despliegue, simplificando el trabajo de los investigadores. Finalmente, se plantea el uso de Simva en la automatización del proceso de validación formal de juegos serios, ya que permite relacionar la información de analíticas con los resultados de los formularios pre-post (todo ello compatible con la anonimización de datos y la protección de la privacidad del usuario para cumplir con la normativa europea GDPR). Este capítulo presenta una breve reflexión sobre los juegos serios (sección 1.1), la motivación de la tesis (sección 1.2) así como sus orígenes, presentando el trabajo que lo precede, y finalizando con la línea de investigación (sección 1.3) que sigue el trabajo realizado.

1.1. Breve reflexión personal sobre los juegos serios

Aunque he sido un estudiante con buenas notas la verdad es que siempre he tenido que hacer un gran esfuerzo para mantener la atención, sobre todo en las clases teóricas. A pesar de que pueda ser complejo e incluso no adecuado para todos los alumnos, yo personalmente hubiera preferido que al menos en algunos casos el proceso hubiera sido diferente.

Cuando trato de aprender algo nuevo prefiero descubrir cómo funcionan las cosas como si fuese un puzle por resolver y del que aprender casi por ingeniería inversa. Al final generalmente logro el objetivo que pretendía, pero aprendiendo además otras cosas que no esperaba en el proceso (muchas veces por los fallos que he cometido), y que han llevado a mi cabeza a pensar un paso más allá. Si bien puede resultar un proceso más lento y laborioso a efectos prácticos, cuando lo comparo con utilizar directamente un manual, elijo esta metodología de aprendizaje por el mismo motivo por el que me disgusta cuando un amigo me desvela el final de una película: disfruto con el desafío y descubriendo el conocimiento por mí mismo.

Creer con los videojuegos como una de mis aficiones principales me ha enseñado que son una de las maneras más inmersivas de descubrir y experimentar, despertando mi continuo interés, y logrando que me implicase por seguir avanzando hasta completarlos. Yo he aprendido sobre muchos temas y he “vivido” distintas experiencias con los videojuegos. Como yo, muchas otras personas sienten lo mismo cuando juegan a videojuegos. Por ello creo, que si pudiéramos abordar problemas complejos de la educación en el futuro a través de juegos estoy seguro de que despertaremos mucho más interés en los estudiantes. Pero para lograrlo, aún tenemos un arduo camino por delante, que implica tanto mejorar la imagen que se tiene de ellos para lograr que se vean como herramientas educativas confiables y eficaces, como simplificar su adopción y las formas para garantizar su calidad. Esto permitirá proporcionar más opciones a los educadores para elaborar una oferta educativa más motivadora que utilice juegos - la que me hubiera gustado tener a mi como alumno.

1.2. Motivación

Los videojuegos pueden ser utilizados con objetivos que van más allá del ocio. Tal y como planteó Abt en los años 70 del siglo pasado, un juego serio es aquel “cuyo propósito va más allá del mero entretenimiento” (Abt, 1970), pudiendo ser aplicados con múltiples finalidades y a diferentes sectores. Aunque entonces la tecnología era mucho más limitada, ya planteaba, por ejemplo, que los juegos eran una muy buena forma no solamente de aprender sino también de evaluar conocimientos complejos. Actualmente, algunos de los sectores donde más se utilizan estos juegos son la educación (De Freitas, 2018), la ecología (Liarakou, Sakka, Gavrilakis, & Tsolakidis, 2012), la salud (Kato, 2012) y la empresa (Larson, 2020). Aunque prácticamente podemos encontrar el uso de

juegos serios en cualquier tipo de sector como, por ejemplo, el turismo (Xu, Buhalis, & Weber, 2017). Su penetración en todos estos sectores se debe a que proporcionan una serie de beneficios cuando los comparamos con su alternativa tradicional. Algunos de estos beneficios son que consiguen captar la atención de sus jugadores (Ravenscroft, Lindstaedt, Delgado Kloos, & Hernández-Leo, 2012) y les permiten participar de forma activa en el proceso de aprendizaje, descubriendo y experimentando (DeKanter, 2005) en un entorno simulado controlado y seguro (Corti, 2006; Squire & Jenkins, 2003). Por ejemplo, utilizando un juego de aventura, los estudiantes pueden disfrutar explorando el entorno mientras aprenden porque necesitan aplicar conocimientos para resolver sus enigmas. Debido a sus beneficios, los juegos serios se han estudiado para multitud de propósitos tales como la adquisición de conocimientos en materias como inglés (Suh, Kim, & Kim, 2010) o ciencias sociales (Nishikawa & Jaeger, 2011), el desarrollo de habilidades tales como las técnicas de reanimación cardio-pulmonar (Creutzfeldt, Hedman, & Felländer-Tsai, 2012), la obtención de habilidades cognitivas tales como el desarrollo del pensamiento computacional (Kazimoglu, Kiernan, Bacon, & Mackinnon, 2012), o el desarrollo de habilidades afectivas y la concienciación como la prevención del ciberbullying (Calvo-Morata, Alonso-Fernández, Freire, Martínez-Ortiz, & Fernández-Manjón, 2021). Incluso han sido usados para fomentar cambios de comportamiento, por ejemplo para evitar el abuso de sustancias (Verduin, LaRowe, Myrick, Cannon-Bowers, & Bowers, 2013).

Pese a los múltiples puntos a favor de su uso y la alta aplicación que vemos en múltiples sectores (sobre todo en aquellos donde hay más financiación, como defensa, negocios o salud: desarrollar buenos juegos es caro), en la educación general vemos que en la práctica la adopción es todavía escasa (Barko & Sadler, 2013; Jean Justice & Ritzhaupt, 2015). Para que los educadores elijan un juego serio como herramienta educativa clave en el aprendizaje deben superar múltiples barreras, tales como la desconfianza en su eficacia (Muñoz-Cristóbal et al., 2018; Pishtari et al., 2020) o escollos tecnológicos (Jean Justice & Ritzhaupt, 2015) que dificultan su uso e implantación. Además, los juegos serios generalmente son difíciles de usar e integrar en plataformas educativas y suelen distribuirse como herramientas de caja negra (Horn et al., 2016; Moreno-Ger, Burgos, Martínez-Ortiz, Sierra, & Fernández-Manjón, 2008). Este modelo de caja negra se debe a que el juego no genera información acerca de lo que sucede durante su uso y, por tanto, es posible que el profesor no termine de entender por qué un estudiante obtiene sus

resultados, dificultando su adopción (Muñoz-Cristóbal et al., 2018). Esto implica que sería deseable que los desarrolladores e investigadores realizarán las pruebas necesarias para garantizar que los juegos serios cumplen con sus objetivos educativos. Pero esto habitualmente no sucede ya que este proceso de validación es complejo y puede ser incluso más costoso que la creación del propio juego.

Dada la variedad y diversidad de los juegos serios, el estudio de su eficacia es una tarea muy compleja. Existen múltiples formas de estudiar la eficacia de un juego serio (Chaudy, 2016), tales como la realización de estudios cualitativos, la realización de entrevistas personales, o la correlación de diversas fuentes de datos. Estos estudios permiten a los investigadores probar diversas hipótesis sobre los juegos, tales como su capacidad para la adquisición de los conocimientos o habilidades, con lo que es posible reducir la barrera de la desconfianza de los docentes hacia los juegos. Desde el punto de vista práctico, uno de los objetivos más interesantes que puede buscar un estudio sobre un juego es medir hasta qué punto el juego provoca el efecto deseado en sus jugadores (sea dicho efecto un incremento de conocimiento, de concienciación u otro). Este proceso se conoce como validación formal (Van der Kooij, Hoogendoorn, Spijkerman, & Visch, 2015; Yusoff, 2010), en el que se evalúan las capacidades formativas de un juego serio. El proceso de validación consiste en la observación de los cambios en las características de los jugadores, incluyendo conocimientos y aptitudes, para atribuir los resultados a la experiencia de uso del juego (Loh, Sheng, & Ifenthaler, 2015). Someter a los juegos serios a un proceso de validación formal aporta datos empíricos de su efectividad que fundamentan la confianza del profesor en las capacidades del juego para cumplir con sus objetivos educativos. De entre todas las metodologías, para la validación de juegos serios destaca el uso de cuestionarios pre-post, donde el participante realiza un cuestionario previo antes de comenzar el experimento y otro posterior al completarlo, lo cual permite comparar los cambios obtenidos en el periodo de juego atribuibles a su utilización (Bellotti, Kapralos, Lee, Moreno-Ger, & Berta, 2013). Dado el alto volumen de datos a manejar (que en la mayoría de los casos se realiza de forma manual) y el conocimiento experimental que requiere, el proceso de validación supone costes elevados difíciles de incluir en el presupuesto para el desarrollo de la mayoría de los juegos serios. Esto provoca que la mayor parte de los juegos serios no sean validados (o no de forma completa), afectando a la confianza en los mismos de los profesores y a su capacidad para justificar su uso en un programa educativo curricular.

Por otro lado, para aquellos juegos validados que sí ofrecen garantías para que el profesor pueda delegar en ellos ciertos objetivos educativos, el profesor necesita poder realizar una supervisión con la que evaluar los conocimientos adquiridos, así como conocer las dificultades que encuentran sus alumnos mientras aprenden, para darles soporte durante el proceso de aprendizaje. Para reducir el esfuerzo de educadores y romper las barreras que limitan su adopción, se está popularizando la investigación en el uso de técnicas de Analíticas de Aprendizaje, o Learning Analytics (LA) en el ámbito de los juegos serios (Alonso-Fernández, Calvo-Morata, Freire, Martínez-Ortiz, & Fernández-Manjón, 2019; Matcha, Uzir, Gasevic, & Pardo, 2020). Las Analíticas de Aprendizaje aplicadas a juegos aportan información acerca de lo que los jugadores hacen mientras juegan, “abriendo” el tradicional modelo de caja negra (A. Serrano-Laguna & Fernandez-Manjon, 2014). Con estos nuevos datos de juego, los investigadores pueden analizar el rendimiento de los jugadores, ya sea mediante paneles de control o mediante técnicas de inteligencia artificial o de minería de datos (Fayyad, Piatetsky-Shapiro, & Smyth, 1996; Romero & Ventura, 2010). Los datos de analítica permiten por tanto aumentar la confianza de los docentes, e incluso pueden ser usados para evaluar los conocimientos de los jugadores (Alonso-Fernández, Martínez-Ortiz, Caballero, Freire, & Fernández-Manjón, 2020). La analítica de aprendizaje propone “medir, recoger, analizar y presentar los datos sobre los alumnos y su contexto, con el fin de comprender y optimizar el aprendizaje y los entornos en los que este aprendizaje se produce” (Long & Siemens, 2011), y a través de su aplicación en juegos serios surge la disciplina conocida como Game Learning Analytics (Freire, Serrano-Laguna, Manero-Iglesias, & Martínez-Ortiz, 2016).

El proceso de extracción de la información de la sesión de juego necesita que el juego implemente la funcionalidad necesaria para registrar y comunicar las acciones del usuario (Alonso-Fernandez, Calvo, Freire, Martinez-Ortiz, & Fernandez-Manjon, 2017), además de requerir de un sistema de analíticas capaz de procesar los datos y presentar los resultados (visualizaciones o cuadros de mando - *dashboards*). Para realizar este proceso, la revisión de la literatura (Alonso-Fernández et al., 2019; Muslim, Chatti, & Guesmi, 2020; Samuelsen, Chen, & Wasson, 2019) presenta una clara prevalencia de las implementaciones ad-hoc, con una tendencia en los últimos años a popularizar el uso de estándares educativos para analíticas de aprendizaje que democratizen su uso. Los elevados costes de implementación de un sistema de analíticas pueden reducirse mediante la reutilización de sistemas de analíticas genéricos, aunque la escasez de este tipo de

sistemas destaca la necesidad de software libre, estándares abiertos y APIs abiertas, que simplifiquen los problemas de coste e interoperabilidad (Chatti, Muslim, & Schroeder, 2017) que dificultan su adopción.

En los últimos años, una de las líneas del grupo de investigación e-UCM está centrada en crear una implementación genérica para analíticas de aprendizaje sobre juegos, utilizando el estándar xAPI (Bakhouyi, Dehbi, Talea, & Hajoui, 2017). Esto incluye un proceso con el que generar analíticas de aprendizaje, así como una serie de componentes genéricas con los que elaborar una arquitectura que proporcione soporte al análisis de estas analíticas y que reduzca la tendencia a crear implementaciones ad-hoc. Una de las contribuciones del grupo a los juegos serios fue el desarrollo de un perfil de xAPI aplicado a juegos serios, xAPI-SG (Á. Serrano-Laguna et al., 2017), que permite a los desarrolladores generar una serie de trazas de juego siguiendo un formato estandarizado, proporcionando un modelo compatible con sistemas de analíticas genéricos (Hauge et al., 2014), y dando soporte a la creación de análisis y visualizaciones por defecto en dichos sistemas. El trabajo del grupo de investigación en analíticas de aprendizaje se centró en la creación de este sistema de analíticas genérico en el contexto del proyecto europeo H2020 RAGE.

Tras haber estado trabajando en la primera iteración del proyecto e-Adventure (Torrente, Del Blanco, Marchiori, Moreno-Ger, & Fernández-Manjón, 2010) sobre Unity (una herramienta de autoría para crear juegos serios de aventura) a modo de trabajo de fin de máster (I. J. Pérez Colado, 2016), el autor de esta tesis comenzó a explorar las analíticas de aprendizaje realizando una prueba de concepto en la que se integraba el modelo de xAPI para juegos serios dentro de dicha herramienta de autoría. Las herramientas de autoría son una interesante forma de acercar a los docentes al proceso de desarrollo de un juego serio, permitiéndoles adaptar y manipular los juegos serios a sus necesidades, reduciendo así la desconfianza que tienen sobre los contenidos (Marchiori, Torrente, et al., 2012). El objetivo de este prototipo de integración era comprobar que era posible generar trazas utilizando un modelo de analíticas genérico, permitiendo que los docentes conozcan lo que los estudiantes realizan mientras juegan al juego y, al mismo tiempo, evitando la implementación ad-hoc de analíticas para cada juego. En este punto, surgió la posibilidad de realizar una tesis doctoral continuando con la exploración del mundo de las analíticas de aprendizaje y la creación de sistemas genéricos con los que conseguir

democratizar y generalizar el uso de analíticas de aprendizaje, que comenzaba con la extensión del proyecto H2020 RAGE para dar soporte a estas nuevas analíticas.

1.3. Línea de investigación

Como expone la motivación, la adopción de los juegos serios actualmente se encuentra fuertemente lastrada por diversos motivos entre los que destacan la falta de información sobre lo que hacen los alumnos y la escasa prevalencia de la validación formal. En este sentido, el principal objetivo de esta tesis es mejorar el ciclo de vida y la adopción de los juegos serios mediante la integración de Game Learning Analytics (GLA) y la creación de herramientas de soporte a las GLA, que faciliten la validación de los juegos serios y su aplicación como herramientas de evaluación. Como punto de partida, se tomó la nueva herramienta uAdventure y la plataforma de analítica desarrollada para el proyecto europeo H2020 RAGE, como base para el desarrollo de analíticas de aprendizaje y la modelización de análisis y visualizaciones. Sobre estas plataformas se ha trabajado de forma sistemática para establecer una metodología con la que desarrollar e integrar analíticas de aprendizaje basadas en estándares, en concreto en xAPI, facilitando la labor de investigadores y desarrolladores. Además, con este nuevo enfoque basado en analíticas se han investigado diferentes oportunidades para simplificar el proceso de validación formal utilizando sistemas genéricos que fomentan un proceso de validación formal democratizado con soporte a estándares. Este proceso de validación no sólo permite que el juego pueda garantizar su efectividad, sino también que aquellas analíticas que genere también sean validadas garantizando que el docente pueda confiar en ellas como método complementario de supervisión y evaluación de los alumnos.

Para llevar a cabo nuestro objetivo de mejorar el ciclo de vida y adopción de los juegos serios mediante la integración de Game Learning Analytics (GLA) y de herramientas de soporte a las GLA, llevaremos a cabo los siguientes objetivos, desarrollados en mayor detalle en el Capítulo 3:

1. Mejorar el proceso de implementación de analíticas para Juegos Serios mediante la utilización y sistematización de Modelos de Analítica de Aprendizaje (Learning Analytics Models, LAMs).
2. Sistematizar el proceso de integración de Game Learning Analytics con herramientas de autoría y juegos serios.

3. Desarrollar una plataforma de Game Learning Analytics orientada a la validación de Juegos Serios mediante actividades.
4. Mejorar el proceso de integración de herramientas de Game Learning Analytics y *data-science* mediante la aplicación de estándares educativos.
5. Validar y realizar pruebas de efectividad de las herramientas y soluciones desarrolladas.

Capítulo 2. Estado del arte

Los juegos serios son herramientas que han sido aplicadas con éxito en contextos educativos tanto orientados a la formación como al entrenamiento de habilidades. También han sido aplicados en sectores como el de la salud (Kato, 2012), o el del turismo (Xu et al., 2017). Sin embargo, pese a los beneficios y los casos de éxito existentes, su aplicación todavía no se ha generalizado debido a múltiples retos como: la complejidad para crearlos o encontrar un juego serio adecuado al escenario educativo específico; la dificultad para integrar el juego serio como una actividad más dentro del programa educativo; la dificultad para analizar el impacto que ha tenido el juego en los estudiantes. Este último problema es de especial relevancia ya que los juegos serios son considerados herramientas de caja negra (Horn et al., 2016; Moreno-Ger, Sierra, & Fernández-Manjón, 2008) ya que suelen ofrecer una información centrada en proporcionar un resumen de los resultados obtenidos en las tareas y actividades llevadas a cabo como parte del juego. Pero que ofrecen escasa información acerca del proceso, es decir, cómo se han llegado a esos resultados. Esta situación hace que el juego sea poco transparente y que el docente tenga que confiar plenamente en la eficacia del juego. Esta situación se agrava ya que la validación formal de la eficacia de los juegos serios tampoco está generalizada, generando aún más desconfianza entre los docentes.

En este capítulo se presentará un análisis del estado del arte relacionado con los juegos serios. Centrándose en su creación, herramientas y metodologías existentes, y en el proceso de validación de los juegos serios, haciendo especial hincapié en las limitaciones de las aproximaciones existentes. Asimismo, también se analizará la aplicación de las analíticas de aprendizaje aplicadas a los juegos serios y los estándares y especificaciones técnicas aplicadas a las herramientas educativas. El objetivo es establecer las bases sobre las que se articula este trabajo de tesis para mejorar la adopción de los juegos serios en el ámbito educativo, facilitando y dando soporte técnico durante todo el ciclo de vida de desarrollo de los juegos serios, desde su desarrollo hasta su puesta en práctica en el aula.

2.1. Los juegos serios

Los videojuegos, así como los juegos serios, se originaron en la década de 1970 con la proliferación de ordenadores domésticos. A lo largo de las décadas posteriores, esta nueva forma de entretenimiento ha seguido una tendencia creciente. Tanto es así, que el análisis y predicción acerca de la evolución del sector del entretenimiento presentado en el

informe Global Entertainment & Media Outlook 2018-2022 (van Eeden & Chow, 2018), expone que en los próximos cinco años el mercado de los videojuegos será el segundo que más crecimiento mostrará, solo por detrás del mercado de las comunicaciones.

Los juegos serios han sido definidos como aquellos juegos “cuyo propósito va más allá del mero entretenimiento” (Abt, 1970), así como “cualquier uso significativo de recursos informáticos, o de recursos de la industria del juego, cuya misión principal no sea el entretenimiento” (Sawyer, 2007). En estas dos definiciones destacan dos elementos clave: entretenimiento y propósito adicional. El objetivo principal es que los juegos puedan utilizarse como un tipo de herramienta educativa más, aunque claramente diferenciada ya que introducen una serie de beneficios específicos como pueden ser, entre otras, la disponibilidad realimentación inmediata al jugador, proporcionar al jugador un entorno controlado donde el escenario simulado en el juego permite al jugador equivocarse y reflexionar sobre las acciones realizadas en un entorno sin peligro para el jugador. Estas son, entre otras, algunas de las razones por las que ha crecido el interés en los juegos serios.

Uno de los primeros juegos serios es The Oregon Trail³ (1971), producido por el Consorcio de Educación Computacional de Minnesota (MECC), creado para enseñar a niños acerca de la colonización de Norteamérica por los pioneros de la Ruta de Oregón. También producido por MECC, Lemonade Strand⁴ (1973) es un videojuego destinado a enseñar a niños los principios básicos de la gestión en los negocios usando un puesto de limonada como contexto y teniendo que abordar diferentes aspectos relacionados con la gestión del negocio tales como: la gestión de inventario de productos y materias primas, los precios o la inversión en anuncios. No obstante, pese a la existencia de múltiples ejemplos de juegos serios desde el nacimiento de los videojuegos, su popularización y crecimiento en su interés no llega realmente hasta la década de los 2000. El caso del juego serio America's Army⁵, desarrollado en 2005 para fomentar el reclutamiento de jóvenes estadounidenses, representa uno de los primeros hitos llegando a ser “el primer juego serio exitoso y bien ejecutado que ganó la atención del público” (De Freitas, 2018; Djaouti, Alvarez, Jessel, & Rampnoux, 2011).

³ Juego The Oregon Trail: <https://www.visitoregon.com/the-oregon-trail-game-online>

⁴ Juego The Lemonade Strand: https://archive.org/details/Lemonade_Stand_1979_Apple

⁵ Juego America's Army: <https://www.americasarmy.com/>

Según el estudio de Djaouti, Alvarez y Jessel (Djaouti et al., 2011) el número de juegos serios creados ha crecido regularmente a lo largo de los años. En el estudio se presenta que entre los años 1980 y 2001 fueron lanzados 926 juegos serios al mercado, lo que se triplicaría en los años posteriores al publicarse 1.265 juegos serios durante el periodo de 2002 a 2010; sugiriendo además una tendencia creciente en la velocidad de publicación de juegos serios de manera regular. Todo este crecimiento sugiere además que la educación basada en juegos, o Game Based Learning en inglés, podría estar siendo considerada y adoptada como alternativa o complemento a las metodologías educativas tradicionales.

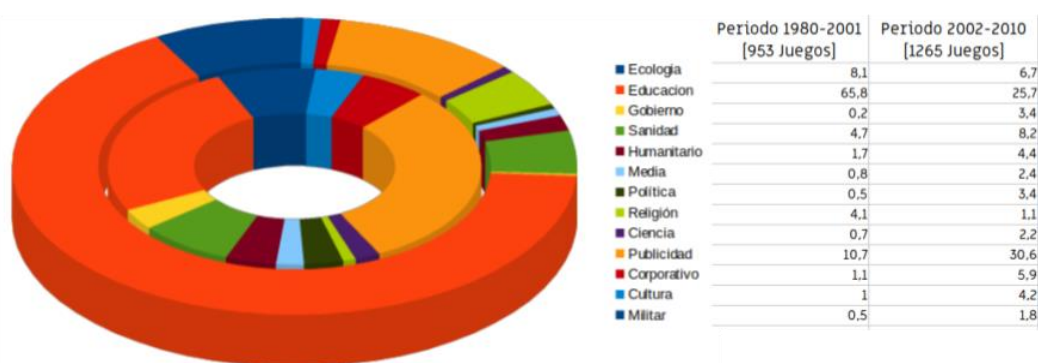


Figura 1. Distribución por sectores de los juegos serios entre los periodos de 1980-2001 con 953 juegos y 2002-2010 con 1265 juegos respectivamente

Es importante destacar que los juegos serios aparecen también en la literatura como videojuegos educativos o *educational games* los cuales “son juegos específicamente diseñados para entrenamiento o educación con objetivos educativos concretos” (Kato, 2010; Rus Cano, 2019). Debido a su similitud con respecto a la definición de juego serio acaban siendo considerados como sinónimos (Michael & Chen, 2005). Nótese que, aunque el sector educativo ha jugado un papel importante en la aplicación de los juegos serios como muestra **Figura 1** llegando a ser predominante en el primer periodo (1980-2001), la distribución mostrada en el segundo periodo (2002-2010) muestra crecimiento en todos los sectores de la sociedad.

La efectividad de los juegos serios está estrechamente relacionada con su naturaleza de videojuego, heredando algunos de los beneficios de este formato (Calvo-Morata, 2020). En primer lugar, promueven el aprendizaje mediante el descubrimiento. El ser humano desde temprana edad comienza a aprender jugando e interactuando con objetos de su entorno, desarrollando múltiples habilidades complejas como la resolución de conflictos,

así como la propia imaginación (Randel, Morris, Wetzel, & Whitehill, 1992; Szczurek, 1982). En segundo lugar, permiten representar cualquier situación, incluyendo situaciones difíciles o imposibles de simular en un entorno educativo real debido a peligrosidad, costes o no disponibilidad de recursos como, por ejemplo, personajes históricos u obras artísticas (Corti, 2006; Squire & Jenkins, 2003). En tercer lugar, pone al jugador en un rol activo, donde debe interactuar, recolectar información y usarla en el juego para tomar las decisiones adecuadas para poder progresar en el juego y alcanzar los objetivos (Oblinger, 2004). En cuarto lugar, captan la atención del jugador, pudiendo romper la “barrera de los 10 minutos de atención” al aplicar ciclos cortos en la resolución de objetivos o retos planteados en el juego. Además, estos objetivos o retos muchas veces son adaptables en dificultad a las capacidades del jugador, proponiendo retos asequibles y satisfactorios que permiten retener su atención durante periodos largos de tiempo (DeKanter, 2005; Ravenscroft et al., 2012). Por último, proporcionan los resultados y consecuencias de las acciones de manera rápida a los estudiantes, permitiendo al jugador aprender y retroalimentar su proceso de juego (DeKanter, 2005).

Asimismo, los beneficios obtenidos no se circunscriben únicamente a propósitos formativos, capacitantes o meramente informativos, pues los juegos serios también han destacado en desarrollo de habilidades sociales (Calderón & Ruiz, 2015), así como agente de concienciación social. Un ejemplo es el caso de Conectado (Calvo-Morata, Freire, Martínez-Ortiz, & Fernández-Manjón, 2018), juego serio dirigido a alumnos con una edad objetivo de entre 12 y 17 años con efectividad probada para la concienciación e identificación de casos de *bullying* y *ciberbullying*. En este sentido, Egenfeldt-Nielsen (Egenfeldt-Nielsen, 2006) analiza cómo los juegos serios pueden aplicarse a diversos métodos de enseñanza (conductismo, cognitivismo, construccionismo y sociocultural). Asimismo, cuando se usan juegos diseñados con objetivos educativos específicos para materias o asignaturas concretas, éstos tienen un mayor potencial de resultar efectivos en el proceso de aprendizaje (Anastasiadis, Lampropoulos, & Siakas, 2018), mientras que cuando se usan juegos no diseñados como juegos serios en origen, su efectividad es limitada (Blumberg, Altschuler, Almonte, & Mileaf, 2013).

Centrándonos en el sector educativo, el aprendizaje basado en juegos (Game-Based Learning) es una metodología educativa usada en los diferentes niveles educativos (colegios, universidades u otros centros educativos) que incluye la utilización de juegos serios en el proceso de formación de los alumnos (Prensky, 2001). Este modelo de

aprendizaje, comparado con métodos tradicionales de enseñanza como las clases magistrales, ha mostrado ser más efectivo y motivador (Kafai, 2001; Prensky, 2001), incentivado por sus beneficios y características, teniendo la peculiaridad de poner al alumno como centro del proceso de aprendizaje con un rol activo, mientras disponen un entorno estimulante, divertido e interesante.

Al igual que con las producciones audiovisuales existen diferentes géneros y formatos para transmitir el contenido de la manera más eficaz. Con la evolución de los videojuegos surgen diferentes estilos, cuyas mecánicas de interacción utilizadas permiten clasificarlos en una taxonomía comúnmente conocida como género de juego (Aslan, 2016). Algunos géneros destacables son: juegos de aventura, acción, rol, simulación, deporte, peleas y estrategia (Setya Murti, Dicky Hastjarjo, & Ferdiana, 2019). De todos ellos, el género de aventura resulta especialmente interesante en el ámbito educativo pues promueven la reflexión, ofreciendo un gran equilibrio entre los conocimientos presentados y su utilización para resolución de problemas (Dickey, 2006). En los juegos de Aventura, la historia del juego es su elemento principal y donde el jugador adoptará un rol concreto, construyendo el juego en torno a la narrativa. Pese a ello, no se ha encontrado un impacto directo entre la historia y la efectividad del juego (Hsu & Wang, 2010; Wouters, van Nimwegen, van Oostendorp, & van der Spek, 2013).

Estos y otros aspectos de diseño del juego serio son una parte fundamental para lograr maximizar su efectividad y garantizar tanto la viabilidad del desarrollo como la futura utilidad real que tendrá el juego serio para los docentes. En la siguiente sección analizaremos el proceso de creación de los juegos serios, profundizando en los diferentes enfoques y metodologías y qué oportunidades existen para mejorar la utilidad y adopción de los juegos serios.

2.2. Creación / Herramientas de autoría / modelos de desarrollo

Como ya se ha mencionado, los juegos serios pueden ser herramientas educativas útiles para la enseñanza, especialmente cuando han sido diseñados, desarrollados y validados para una serie de objetivos educativos pedagógicos concretos (De Freitas, 2018; Kato, 2012; Larson, 2020; Liarakou et al., 2012; Xu et al., 2017). Por tanto, para lograr un resultado efectivo, el proceso de diseño, desarrollo y validación de un juego serio requiere de una simbiosis entre el diseño pedagógico e instruccional, a través de un proceso de

desarrollo multidisciplinar (Blow, 2004). Este proceso requiere de expertos en la materia que aborda el juego, docentes y desarrolladores de juegos serios con el objetivo de lograr su máxima efectividad. Para lograr maximizar esta efectividad, el diseño del juego no sólo debe cubrir los objetivos pedagógicos, sino que las mecánicas de juego seleccionadas deben ser las más apropiadas desde el punto de vista pedagógico. Además, deben resultar atractivas e interesantes para el jugador, logrando un buen equilibrio entre su parte lúdica para captar la atención del jugador y los aspectos educativos (Moreno-Ger, Burgos, et al., 2008; Quinn, 1994). Durante este proceso de creación también se debe clarificar otro aspecto clave que determinará la futura utilidad del mismo, esto es, definir los diferentes resultados (como datos o métricas) que facilitará el juego al docente para ser útil en su toma de decisiones y como prueba de su efectividad.

Si bien este paradigma de desarrollo multidisciplinario, uniendo desarrollo de videojuegos y pedagogía es fundamental para su efectividad, puede llevar a que el proceso se complique en exceso debido a la gran cantidad de personas y sus diferentes perfiles, pudiendo afectar negativamente al resultado del desarrollo (Sloney & Murphy, 2008). Por ejemplo, pueden darse problemas a la hora de interpretar y trasladar los objetivos pedagógicos en mecánicas del juego, limitando la efectividad del mismo o bien aumentando el coste de desarrollo del juego. Para resolver este problema, y pese a que no existe aún una solución metodológica consensuada para este proceso (Rus Cano, 2019), se han abierto nuevas líneas de investigación en esta área con el objetivo de desarrollar una serie de metodologías y modelos que ayuden a fomentar un mejor entendimiento de la efectividad final del juego serio y su utilidad (Bellotti, Berta, De Gloria, & Primavera, 2009). Además de estas metodologías que buscan maximizar el impacto en los objetivos pedagógicos planteados durante el diseño, también existen diferentes enfoques técnicos aplicables al proceso de desarrollo que pueden afectar tanto a su efectividad como al coste final de su desarrollo. En la presente sección analizamos algunas de estas metodologías y enfoques que resultan interesantes o innovadores para contextualizar este trabajo de tesis.

El diseño instruccional es una de las disciplinas más populares a la hora de aplicar una metodología a nuestro desarrollo, proporcionando una serie de principios de diseño sobre los que fundamentar el proceso de aprendizaje del estudiante. El diseño instruccional se fundamenta en cuatro componentes (Faber, Dankbaar, & van Merriënboer, 2018; van Merriënboer & Kirschner, 2017): (1) las tareas de aprendizaje, o *learning tasks*, que buscan que el estudiante utilice habilidades que construyan y practiquen su conocimiento

de forma secuencial y progresiva; (2) la información procedural, o *procedural information*, que es aquella información intrínseca a la forma de jugar que permite representar la rutina que el jugador debe aprender y automatizar; (3) la información de apoyo, o *supportive information*, que ofrece soporte al jugador para desarrollar sus mecanismos de resolución de tareas; y (4) la práctica de tareas, o *part-task practice*, que puede utilizarse para aspectos recurrentes. A través de la aplicación de los principios de diseño instruccional se pueden estructurar las diferentes tareas y actividades del juego de forma que los actores involucrados en el desarrollo puedan comprender mejor los objetivos pedagógicos y las tareas a realizar para lograrlos. Otro de los aspectos aplicados durante para el desarrollo de juegos serios es el uso de los Principios de Diseño Universal, de sus siglas en inglés UDP (CAST, 2018; Hitchcock & Stahl, 2003). Diseñar un juego serio usando los UDP permite asegurar que el estudiante aprende de forma significativa, eficaz y estimulante, buscando que los estudiantes se sientan motivados. Finalmente, estos diseños pueden ser complementados con otros principios de diseño que mejoren la efectividad a través de mejorar la satisfacción del jugador a la hora de jugarlos, tales como el diseño centrado en el usuario (User-Centered Design) (Pagulayan, Keeker, Wixon, Romero, & Fuller, 2003) y los principios de diseño orientados a mejorar la accesibilidad de los juegos serios (Game Accessibility Guidelines) (Ablegamers Foundation, 2012).

Al utilizar una metodología de diseño en el desarrollo se logra que la definición de las diferentes partes del juego se describan de forma sistemática en torno a los objetivos educativos y las características del usuario final. Este diseño inicial puede además complementarse con el uso de otra serie de metodologías que nos permitan profundizar de forma sistemática desde el diseño de alto nivel, hasta el bajo nivel en el que se definirán las mecánicas de juego concretas en base a los objetivos pedagógicos. Por ejemplo, el modelo Learning Mechanics-Game Mechanics, o LM-GM, simplifica la elección de las diferentes mecánicas en función del tipo de enfoque pedagógico elegido (Lim et al., 2015). En este punto, resulta especialmente interesante la posibilidad de incluir detalles acerca de cómo podremos verificar que el jugador ha logrado los objetivos educativos al poder definir, por ejemplo, una serie de indicadores o trazas asociadas a las mecánicas de juego que representan cada objetivo educativo. De hecho, existe una variante del modelo LM-GM llamada Activity Theory-based Model of Serious Games (ATMSG) diseñada para poder modelizar la extracción de evidencias y analíticas (Callaghan, McShane, Eguíluz, & Savin-Baden, 2018). Al considerar la extracción de evidencias como parte del

diseño, no sólo se fomenta que se simplifique el proceso de validación formal, sino que además se pueden utilizar estas métricas durante todo el ciclo de vida del juego para mejorar su calidad final.

Una vez establecido un diseño de juego y sus mecánicas, la siguiente etapa en el proceso de creación es su implementación. Existen múltiples enfoques para realizar la implementación que se distinguen principalmente en la elección de la plataforma o herramienta sobre la que se construirá el juego. En función de la plataforma escogida, el desarrollador contará con mayor o menor libertad para implementarlo a cambio de obtener una serie de funcionalidades con las que simplificar el desarrollo del juego. Los principales enfoques aplicados a la hora de desarrollar un juego serio son: la utilización de motores de videojuegos de propósito genérico, la utilización de motores de propósito específico y el uso de herramientas de autoría.

El uso de motores de videojuegos de propósito general como por ejemplo Unity⁶, Unreal Engine⁷ o Godot⁸, permite que el desarrollador cuente con la máxima libertad para implementar el juego. Estos motores se caracterizan por ser capaces de representar todo tipo de mecánicas y estilos de juego, contando además con sistemas que simplifican la creación del juego tales como un sistema para la simulación de físicas, la creación de interfaces o la gestión de la interacción. Además, estos motores habitualmente cuentan con una potente comunidad de desarrollo y el soporte de una o varias empresas. Tanto el soporte como la comunidad de desarrollo facilita la obtención de recursos, librerías y contenidos que simplifican el desarrollo del juego. Además, el soporte empresarial del motor asegura que la inversión realizada en el desarrollo pueda amortizarse en un plazo mayor y aprovecharse del soporte multiplataforma que, normalmente, es uno de los requisitos en el desarrollo profesional de videojuegos.

Por otro lado, el uso de motores de propósito específico permite tanto simplificar cómo reducir los costes de desarrollo de videojuegos orientados a un género de juego o estilo visual específico. Algunos ejemplos más significativos de este tipo de motores son Adventure Game Studio⁹ que está especialmente indicado para la creación de aventuras

⁶ El motor de videojuegos Unity se encuentra disponible en: <https://unity.com>

⁷ El motor Unreal Engine es propiedad de la empresa Epic Games disponible en: <https://epicgames.com>

⁸ Motor de videojuegos libre disponible en: <https://godot.org>

⁹ Motor de aventuras gráficas: <https://www.adventuregamestudio.co.uk/>

gráficas, RPGMaker¹⁰ para la creación de juegos RPG y GameMaker¹¹ que permite la creación de juegos de forma simplificada al centrarse en el estilo visual 2D.

Finalmente, las herramientas de autoría se centran en simplificar la experiencia de usuario y reducir la barrera de conocimientos técnicos a la hora de crear los juegos serios. Estas características hacen que sean especialmente atractivas para que puedan ser utilizadas directamente por los propios docentes. Es común que, para lograr esta experiencia de usuario simplificada, las herramientas de autoría se centren en géneros de juego específicos. Por ejemplo, herramientas como Adobe Captivate¹² o uAdventure se centran en el género narrativo (Perez-Colado, Perez-Colado, Martinez-Ortiz, Freire-Moran, & Fernandez-Manjon, 2017). Además, estas herramientas también suelen incluir metodologías de diseño y funcionalidades orientadas a la educación con las que mejorar la efectividad y utilidad de los juegos creados. De hecho, algunas herramientas como uAdventure son capaces de generar datos relativos a la interacción del jugador con el juego y que pueden utilizarse para generar evidencias relacionadas con el proceso de aprendizaje.

Aunque existen múltiples enfoques para la creación de juegos serios, uno de los más habituales (y demandados por los docentes) es partir de un juego serio existente y adaptarlo a sus necesidades educativas concretas (Torrente, Del Blanco, Moreno-Ger, & Fernández-Manjón, 2012). El objetivo de los docentes es replicar su experiencia en la reutilización de contenidos educativos disponibles en un repositorio común (como el repositorio común del INTEF¹³) o como parte de una comunidad de docentes, donde utilizan estos contenidos como punto de partida y los adaptan a su escenario educativo concreto. En este escenario de reutilización, las herramientas de autoría destacan frente a los motores genéricos o específicos, ya que habitualmente es uno de los casos de uso que ya cubren y, por tanto, se pueden mantener controlados el coste y el esfuerzo de realizar la modificación del juego existente.

Por tanto, las herramientas de autoría son de especial relevancia y se encuentran mejor alineadas con el objetivo general de este trabajo de tesis ya que permite integrar de manera

¹⁰ Motor de juegos RPG disponible en: <https://www.rpgmakerweb.com/>

¹¹ Motor de videojuegos disponible en: <https://gamemaker.io/es/gamemaker>

¹² Herramienta de Adobe disponible en: <https://captivate.com>

¹³ Disponible en: <https://intef.es/recursos-educativos/>

más efectiva a los expertos del dominio (docentes), maximizar la utilidad del juego, facilitar la extracción de evidencias (en algunos casos como uAdventure) con un coste muy reducido. Asimismo, esta simplificación del proceso de desarrollo y contención de costes permite dedicar suficientes esfuerzos al proceso de validación de la eficacia del juego.

En la siguiente sección profundizaremos en el proceso de validación de los juegos serios con el que un juego puede probar su efectividad a la hora de lograr sus objetivos educativos.

2.3. Validación

Como se menciona en la introducción de este trabajo de tesis, una de las principales problemáticas a las que se enfrenta un docente a la hora de utilizar un juego es la falta de evidencias en los juegos serios con las que justificar su uso dentro de un programa educativo concreto. Para paliar esta limitación, los juegos serios deben pasar por un proceso de evaluación y validación que asegure que cumplen con sus propósitos y objetivos educativos para los cuales han sido diseñados. No obstante, por su naturaleza, estos procesos de validación son técnicamente complejos y costosos, dado que requieren de grandes volúmenes de datos, usuarios y pruebas con las que garantizar la efectividad del juego. Esto provoca que en la mayoría de los casos los juegos acaben sin ningún tipo de validación formal de su efectividad.

El proceso de validación de un juego consiste en verificar que la experiencia de juego genera cambios observables en las características de los jugadores, como su conocimiento o actitud (Loh et al., 2015). Para poder realizar la validación, es necesario poner en práctica un diseño experimental que evalúe las capacidades formativas del juego serio. Existen diferentes alternativas a la hora de realizar éste proceso, pero podemos organizarlas en las siguientes categorías (Chaudy, 2016):

1. *Investigación experimental*, en un entorno controlado con participantes elegidos de manera aleatoria, y con una manipulación activa de las variables. Al controlar el entorno, los factores de interés pueden ser bien identificados, y otros factores de influencia descartados. No obstante, este entorno más artificial puede influir en el comportamiento del participante.

2. *Investigación cuasi-experimental*, en un entorno natural, con participantes seleccionados por características concretas, pero sin manipular las variables. Como ventaja, al ser en un entorno natural, factores externos pueden influir, pero sus relaciones pueden ser identificadas. Sin embargo, este tipo de entornos reducen el control, perdiendo precisión, limitando la capacidad para identificar el comportamiento del participante.
3. *Encuestas, exámenes y cuestionarios*, en donde los propios participantes (en algún momento antes y/o después de realizar en el experimento) son los que comparten, mediante estos medios, toda la información acerca de lo que saben, sus propias características o cómo se comportan. Esta aproximación tiene la ventaja de que los datos se recogen de manera más sencilla y son más sencillos de comparar. Además, este tipo de experimentación simplifica la introducción de un grupo de control, que provea los datos, pero sin la influencia del objeto del experimento. Como desventaja, la precisión de los datos es cuestionable debido al propio mecanismo de obtención de datos; así como la posibilidad de que factores externos influyan en el comportamiento del participante.
4. *Entrevistas estructuradas*, donde se recoge información en profundidad acerca de los puntos de vista de los participantes, en un orden planificado donde se especifican las preguntas. Como consecuencia a esta estructuración, la información recogida es uniforme y comparable. Las principales desventajas son, la gran cantidad de tiempo que lleva hacerlas, o que la formalidad de una entrevista (así como el orden de las preguntas) puede no permitir al participante expresarse con libertad.
5. *Entrevistas no estructuradas*, donde también recogemos información en profundidad, pero hay libertad en el contenido y estructura de las preguntas. Esto permite obtener información mucho más detallada acerca de la experiencia del participante, y además el investigador puede adaptar las preguntas a cada entrevista. Este tipo de entrevistas también consumen mucho tiempo, y su principal desventaja es que la información recolectada no es comparable.
6. *Entrevistas grupales*, donde se recoge información de la opinión del grupo en una discusión grupal. Aprovechando estas dinámicas de grupo, algunas de las respuestas ayudarán a mejorar otras. Además, al no ser individual, consume

menos tiempo. La principal desventaja se ve influenciada por estas mismas dinámicas de grupo, donde algunos participantes pierden la oportunidad de expresarse, y se pierden los puntos de vista individuales.

7. *Investigación mediante la observación*, mediante el estudio del comportamiento en un entorno natural, con o sin intervención. Da la oportunidad de estudiar el comportamiento detalladamente y es menos disruptiva para el jugador, aunque también la presencia del investigador puede influir en el comportamiento del participante.
8. *Casos de estudio*, recolectando información en profundidad de una persona en concreto, o un grupo pequeño. Es útil cuando se explora un área nueva, pero debido a la limitada fuente de datos, la generalización de los resultados también resulta limitada.
9. *Investigación de archivo*, utilizando información ya existente y almacenada, y recogida con otros propósitos. Si esta información además es recogida utilizando estándares de seguimiento, permitirían mezclar información de múltiples fuentes e incluso ubicarla temporalmente. No obstante, los datos utilizados pueden ser poco precisos o incompletos.

De entre todas estas metodologías, la validación realizada mediante una investigación basada en encuestas, en concreto usando pre-test y post-test, también llamados cuestionarios pre-post, destaca por ser con diferencia la metodología de investigación más usada en juegos serios (Bellotti et al., 2013). Estos cuestionarios pre-post establecen una base para demostrar su efectividad al correlacionarse con el uso del juego mediante su análisis comparativo. En este proceso se comprueba, comparando las características del participante antes y después de la prueba, y si ha habido un cambio significativo entre las evaluaciones. Por ejemplo, si queremos validar si un juego serio es efectivo para enseñar conocimientos concretos de biología deberemos de realizar una comparación de los conocimientos de los usuarios antes y después de que tengan su experiencia de juego.

Loh et al. (Loh et al., 2015) destacan que los jugadores pueden tardar en completar los videojuegos, en algunas ocasiones, periodos largos (días o semanas), lo que pueden suponer un sesgo durante el proceso de validación ya que los usuarios podrían recibir influencia externas que afecten a los resultados del experimento. No obstante, los mismos autores remarcan que pese a que existen juegos serios de todo tipo de duración, los juegos

serios que se utilizan habitualmente en el entorno educativo ajustan su duración a la duración típica de una clase magistral, por lo que realizar el proceso de validación en éstos dará resultados más significativos debido a la poca influencia externa. En la literatura podemos encontrar múltiples juegos diseñados para ser utilizados en una única sesión de juego de entre 5 y 30 minutos de duración, adaptándose también al tiempo necesario para poder llevar a cabo las tareas relacionadas con el experimento (Loh et al., 2015).

La validación de los juegos serios no es sólo un factor determinante a la hora de justificar su adopción por parte de los docentes, sino que el propio proceso de desarrollo y la viabilidad del proyecto puede verse afectado por la falta de validación del juego o por una validación poco adecuada del mismo. Por tanto, es necesario mejorar el proceso de evaluación y validación de los juegos serios de modo que puedan resolverse las dudas de la eficacia del juego para todos los actores involucrados (Loh et al., 2015; Nickols, 2005).

Completado el proceso de validación, el juego serio resulta más atractivo para los docentes, pues pueden depositar su confianza y justificar su uso en el currículum docente. Sin embargo, en ocasiones, debido a la poca información proporcionada a los docentes, la confianza en este proceso de validación puede ser limitada (Alonso-Fernandez et al., 2017). Además, pese a que pueda estar validado, existen otros problemas como la pérdida de control del proceso de aprendizaje, o la falta de evidencias para poder evaluar a los estudiantes mediante la interacción del mismo con el juego serio (Jean Justice & Ritzhaupt, 2015).

2.4. Analíticas de juegos (GLA)

Los juegos serios son distribuidos con frecuencia como cajas negras (Horn et al., 2016; Moreno-Ger, Burgos, et al., 2008), es decir, son sistemas cerrados donde normalmente es complejo obtener información relativa a la experiencia educativa y la experiencia de juego (Torrente et al., 2010). Estas cajas negras, en el mejor de los casos, no ofrecen más que información acerca de sus resultados sin justificar los pasos que ha seguido el jugador hasta obtenerlos. Sin embargo, en otros sectores como el de los videojuegos para entretenimiento es común que se apliquen prácticas como la recopilación de telemetría que permiten comprender cómo juegan sus jugadores a través de la observación de la interacción del usuario. La analítica de juegos utiliza esta información de telemetría, para mejorar la efectividad comercial, extrayendo una serie de métricas que nos indiquen si un

usuario está avanzando correctamente a través de nuestro embudo o *funnel* comercial. Extrapolando esta idea a los juegos serios, es posible utilizar la interacción del usuario para conocer si el jugador está avanzando correctamente a través del proceso educativo.

Este proceso de recolección y análisis de datos se conoce como analíticas de juegos educativos, *Game Learning Analytics* o GLA (Freire et al., 2016), y tiene como objetivos "medir, recoger, analizar y presentar los datos sobre los alumnos y su contexto, con el fin de comprender y optimizar el aprendizaje y los entornos en los que este aprendizaje se produce" (Long & Siemens, 2011). El paradigma GLA se puede ubicar a caballo entre las analíticas de aprendizaje, *Learning Analytics* o LA, ya que son utilizadas para seguir el progreso del proceso de aprendizaje en plataformas educativas digitales como Campus Virtuales (Fundación Telefónica, 2020); y las analíticas de juegos, *Game Analytics* o GA, donde se utilizan para monitorizar el proceso de juego con el objetivo de mejorar la calidad final del juego, detectando posibles oportunidades de mejora o fallos que afecten al jugador y a su experiencia de juego.

El proceso de creación de un modelo de analíticas aplicable a un juego serio comienza con la creación del *Learning Analytics Model* (LAM). En el diseño de un LAM se analizan los objetivos educativos de un juego serio con el propósito de identificar métricas que permitan monitorizar el progreso en la consecución de dichos objetivos. El diseño del LAM es un proceso sistemático donde los objetivos educativos (que provienen de su diseño instruccional) se traducen desde el lenguaje natural, hasta el nivel de interacción más bajo basado en la secuencia de acciones que realiza el usuario dentro del juego. Para su identificación existen metodologías como la ATMSG (Callaghan et al., 2018) que plantea un acercamiento basado en cuatro fases de actuación hasta identificar las acciones necesarias para realizar el seguimiento de los objetivos educativos. Por ejemplo, el objetivo educativo de "activar una alarma de incendios en caso de fuego", será identificado al encontrar el punto del juego donde el jugador encuentra la alarma y la activa utilizando sus conocimientos disponibles. Los datos extraídos en base al LAM, gracias a su vinculación con los objetivos educativos, permiten no sólo evaluar el aprendizaje y monitorizar el progreso del estudiante durante la sesión de juego, sino que, una vez validado el juego, estos datos pueden ser suficientes predecir el efecto positivo en el estudiante tras la sesión de juego sin necesidad de tener que utilizar un cuestionario (Alonso-Fernández et al., 2020).

El instituto Global McKinsey¹⁴ define *Big Data* como “conjuntos de datos cuyo tamaño sobrepasa las capacidades para capturar, guardar, gestionar y analizar de las herramientas de bases de datos tradicionales” (Long & Siemens, 2011). Teniendo en cuenta el nivel de detalle y granularidad a la que podemos llegar en la recopilación de acciones del usuario dentro del juego, podemos clasificar las GLA dentro de la categoría de Big Data. Asimismo, este gran volumen de datos nos permite aplicar técnicas de minería de datos (DM, del inglés *Data Mining*), con el objetivo de poder extraer información interesante, interpretable, útil a partir de los datos (Fayyad et al., 1996). Al aplicar procesos de DM sobre los datos recogidos con GLA, buscando resolver problemas en el proceso de aprendizaje y la educación, esta minería de datos se categoriza como minería de datos educacional (EDM, del inglés *Educational Data Mining*). No obstante, pese a compartir similitudes entre LA y EDM y compartir el propósito común de mejorar la calidad de la educación mediante el análisis de grandes cantidades de datos, en la actualidad siguen siendo consideradas como disciplinas diferentes (Liñán, Alejandro, & Pérez, 2015).

En concreto, la EDM facilita que (Romero & Ventura, 2010):

1. Los estudiantes pueden recibir una experiencia educativa personalizada, recibiendo recomendaciones de actividades, recursos y tareas que podrían mejorar el proceso de aprendizaje, así como recordando y simplificando actividades menos relevantes.
2. Los profesores obtienen información que permite analizar el comportamiento de los estudiantes, pudiendo predecir su rendimiento, categorizarlos en grupos, encontrar patrones irregulares de comportamiento, así como encontrar los errores más comunes que permitan adaptar los recursos para una mayor efectividad.
3. Los diseñadores de una experiencia educativa (como un curso) obtienen información para, entre otras cosas, evaluar su estructura y efectividad con respecto a sus objetivos educativos, así como encontrar modelos y generar perfiles de estudiantes y profesores que permitan entender mejor el funcionamiento del mismo.

¹⁴ El instituto Global McKinsey es un instituto de investigación empresarial y económica que publica información basada en hechos, que ayudan a tomar decisiones de gestión y políticas. <https://www.mckinsey.com/>

4. Las organizaciones educativas obtienen información que ayuda a simplificar el proceso de toma de decisiones que permitan, alcanzar objetivos de la organización u optimizar el coste y retorno de inversión de los cursos.
5. Los responsables de las organizaciones educativas pueden obtener métricas para poder organizar de manera más efectiva los recursos de la institución (humanos y materiales) y su oferta educativa, permitiendo también evaluar a los profesores, sus planes curriculares y su efectividad.

Finalmente, pese a que la investigación en EDM comenzó varios años antes, se espera que ambos campos continúen expandiéndose de manera regular (Johnson, Becker, Cummins, Estrada, & Meira, 2012).

Debido a que tanto LA como EDM comparten similitudes, es importante establecer aspectos claros diferenciadores entre ambas. Los autores (Liñán et al., 2015; Siemens & Baker, 2012) organizan las diferencias entre ambas en cinco categorías:

1. Descubrimiento de resultados: En LA se utiliza del juicio de las personas que analizan los datos para obtener conclusiones, usando resultados automáticos como herramienta; mientras que en EDM se busca obtener el resultado automático, utilizando el juicio de las personas para este objetivo. En LA es frecuente la utilización de visualizaciones gráficas, por su facilidad de creación, para representar los datos con el objetivo de permitir extraer conclusiones a los usuarios finales de dichas visualizaciones (Chittaro, 2006).
2. Reducción y holística: las LA buscan entender los sistemas en su completitud, pudiendo utilizar un panel, o *dashboard*, lleno de visualizaciones gráficas para dar una retrospectiva que permita entenderlo; mientras que EDM se encarga de reducir estos sistemas a componentes, analizando tanto éstos como sus relaciones.
3. Adaptación y personalización: LA muestra un mayor énfasis en mostrar la información a los instructores y estudiantes para permitirles tomar decisiones y adaptarse, mientras que EDM tiene un mayor énfasis en adaptar automáticamente el contenido al estudiante eliminando poco a poco a las personas del proceso.
4. Orígenes: LA surge de la web semántica, los “*intelligent curriculum*” e intervenciones sistémicas, mientras que EDM surge del software educativo y sus perfiles de usuario, buscando predicciones automatizadas. Esto se observa tanto en los sistemas de gestión de aprendizaje (LMS, del inglés *Learning Management*

System); como en los cursos online masivos y abiertos o, por sus siglas en inglés MOOCs, siendo estos últimos los mayores precursores por la gran cantidad de alumnos que participan, haciendo inviable realizar un seguimiento individual mediante métodos tradicionales (Fundación Telefónica, 2020).

5. Técnicas y metodologías específicas: LA incluye múltiples análisis (social, de sentimientos, de influencia, de discurso, o de conceptos), predicciones del éxito del estudiante, así como modelos de sensibilización; Por contraparte EDM incluye: clasificación y *clustering*, modelos bayesianos, minería de relaciones y visualizaciones.

Pese a sus diferencias, LA y EDM trabajan de manera conjunta a la hora de mejorar el ciclo de vida de los juegos serios, y aunque hemos visto sus diferencias, la mención del uno implica, en la mayoría de los casos, la mención del otro. Por tanto, para simplificar la lectura del presente trabajo, utilizaremos la terminología GLA de ahora en adelante como paraguas en la aplicación de técnicas de LA y EDM aplicadas a los juegos serios.

Analizando las técnicas y metodologías utilizadas en GLA, encontramos que, han ido evolucionando a lo largo del tiempo, pudiendo identificarse dos periodos (Baker & Yacef, 2009; Liñán et al., 2015). Mientras que, en el primer periodo, 1995-2005, las técnicas más comunes incluyen: minería de relaciones (43%), modelos predictivos (28%), análisis exploratorio de los datos (17%) y *clustering* (15%); en el segundo periodo, 2008-2009, las más comunes incluyen: Modelos predictivos (42%), descubrimiento con modelos (19%), *clustering* (15%), análisis exploratorio (12%) y, pasando de la primera reduciendo mucho su aparición, minería de relaciones (9%).

Como se explica a lo largo de esta sección, GLA provee herramientas de especial relevancia para todos los roles que participan en el proceso educativo, esto es, estudiantes, profesores y desarrolladores de contenido educativo o, en este caso, de juegos serios. Para poder integrar analíticas, es necesario que no sólo el juego serio esté dotado de capacidades para generar registros de lo que está ocurriendo en el proceso de juego, también es necesaria una infraestructura capaz de almacenar y explotar los datos. Este es uno de los principales impedimentos de su adopción, pues son necesarias grandes inversiones adicionales puesto que, por un lado, es necesario de un conocimiento técnico especializado que todavía no se ha asentado en el sector comercial y una infraestructura compleja para el manejo de los datos (especialmente cuando se trata de cumplir con las

obligaciones legales de éstas como la GDPR). Por ello, se acaba fomentando que su desarrollo se realice de forma improvisada o ad-hoc, o en algunos casos no llegue a realizarse.

Esta tendencia en la aplicación de soluciones ad-hoc ha sido identificada en las principales revisiones exhaustivas de trabajos en esta área de investigación (Alonso-Fernández et al., 2019; Muslim et al., 2020; Samuelsen et al., 2019). Como resultado de esta tendencia, no se puede aprovechar al máximo el potencial de las analíticas ya que es difícil o muy costoso integrar estos desarrollos con las herramientas educativas existentes como un LMS o que sea difícilmente replicable la experiencia más allá del juego serio o plataforma específica utilizada en el caso de estudio. De hecho, autores (Chatti et al., 2017) destacan “la necesidad de software open-source, estándares abiertos y APIs abiertas para resolver el problema de interoperabilidad en este campo”.

2.5. Estándares

Cuando los docentes quieren utilizar una herramienta educativa (como un juego serio) en un aula, dos de los aspectos más importantes para ello son la efectividad de la herramienta y la facilidad para utilizarla. En el caso de los juegos serios, podemos probar la efectividad de éstos mediante la validación de los mismos aplicando alguna de las metodologías analizadas previamente. Sin embargo, para facilitar su uso en un aula, la transición entre las diferentes actividades educativas, incluidas el juego serio, debe suceder de manera fluida, con el objetivo de ajustar las diferentes actividades al tiempo de la clase. Asimismo, cuando los estudiantes juegan a un videojuego es probable que el profesor pierda el control de la actividad si no se le proporcionan las herramientas y mecanismos de supervisión y seguimiento que le permitan entender qué está sucediendo mientras cada estudiante juega (Jean Justice & Ritzhaupt, 2015). Por tanto, es necesario proporcionar medios para que los docentes puedan seguir utilizando las plataformas y herramientas que venían utilizando regularmente, como un LMS, e integrar tanto los juegos serios como las herramientas de soporte dentro de dichas plataformas.

Las Tecnologías de la Información y la Comunicación (TIC) han venido aplicándose en diferentes contextos educativos a lo largo del tiempo. De esta aplicación, han surgido numerosas especificaciones y estándares técnicos que fomentan y facilitan la interoperabilidad entre herramientas educativas. De este modo, es posible integrar desde el punto de vista más técnico, herramientas heterogéneas que potencialmente han sido

desarrolladas por diferentes proveedores en un único ecosistema. Sin embargo, en el caso de los juegos serios, es habitual que sean herramientas fuera del contexto habitual y que éstos proporcionen una información limitada dentro de su propia plataforma generando así silos de datos (Barra, Gordillo, Gallego, & Quemada, 2013). Esta falta de información no permite saber qué está haciendo el usuario, convirtiéndolos, como ya se mencionó en la sección anterior, en *cajas negras* que impiden que el profesor pueda responder a preguntas cómo: ¿qué decisiones o elecciones ha tomado el estudiante?, ¿en qué orden? o ¿cuáles son los resultados y consecuencias de dichas acciones? (Moreno-Ger, Sierra, et al., 2008). Esta sección se centra en aquellos estándares y especificaciones técnicas educativas que faciliten la interoperabilidad entre las diferentes herramientas educativas con el objetivo de mejorar la usabilidad de los juegos en aula, y que también permitan que los juegos sean más transparentes para los docentes con el objetivo de aumentar su confianza en ellos.

Como hemos mencionado previamente, la aplicación de estándares y especificaciones entre herramientas educativas no es completamente nueva. En el pasado, estas iniciativas se han dirigido a facilitar la interoperabilidad de contenidos educativos reutilizables, denominados *Learning Objects*, o LOs (Polsani, 2005). Es importante su mención ya que, como parte del ecosistema de LOs, han surgido diferentes estándares y especificaciones para promover la interoperabilidad y su reutilización.

Asimismo, autores como del Blanco et al. y Torrente et al. (del Blanco, Serrano-Laguna, Martínez-Ortiz, Fernández-Manjón, & Stanescu, 2013; Torrente, Moreno-Ger, Martínez-Ortiz, & Fernandez-Manjon, 2009) han identificado problemáticas y retos similares a los que se enfrentan los LOs para el caso particular de los juegos serios, identificando los siguientes retos:

1. Cómo obtener el juego serio (descubrimiento). Tras haber identificado la necesidad de un recurso educativo que enseñe un contenido, el educador puede: disponer de algún recurso ya creado o desarrollar un nuevo recurso que cumpla con los requisitos. Si se utiliza un recurso ya existente se reducen drásticamente los costes de obtención, por ello, es importante que los educadores dispongan de

métodos para encontrar y compartir sus recursos educativos, pudiendo crear repositorios¹⁵ donde compartir todos estos recursos.

2. Cómo dar acceso al juego serio a los estudiantes (empaquetado y distribución). Una vez el educador ha obtenido el recurso, los estudiantes autenticados deben de ser capaces de acceder a él de la manera más sencilla. Adicionalmente, y debido a que los juegos serios son recursos que pueden ser ejecutados por los estudiantes en múltiples plataformas y tecnologías, sumado a la posibilidad de que tengan que utilizar sus propios dispositivos *Bring Your Own Device* o BYOD (Afreen, 2014), presenta un posible entorno heterogéneo donde se remarca la importancia de facilitar la distribución y ejecución del mismo de la manera más transparente posible.
3. Cómo simplificar la integración y recopilación de la información del proceso de juego con el resto de componentes educativos (comunicación y modelado de datos). Finalmente y antes mencionado, son necesarios mecanismos de control, supervisión y seguimiento para que, no sólo el educador sea consciente del proceso de aprendizaje de manera detallada, sino que además todos aquellos componentes educativos interesados que forman parte del proceso también deben de recibir dicha información para poder realizar procesos descritos en la sección de analíticas, cómo la adaptación del contenido educativo a las capacidades del estudiante, o algo tan simple como realizar la evaluación.

Debido a la constante evolución y la aparición de nuevos estándares y especificaciones, es necesario volver a revisar y analizar de manera crítica la aplicabilidad de los mismos, tanto de manera aislada como de manera conjunta con la finalidad de facilitar el proceso de interoperabilidad de los juegos serios con el resto de herramientas educativas existentes, en especial los LMS.

Para abordar el primer reto, se puede aplicar *IEEE Learning Object Metadata*, o LOM (Learning Technology Standards Committee of the IEEE, 2002) como describir las características principales de un LO como metadatos, que pueden ser posteriormente utilizados para la mera documentación del LO como para facilitar y guiar el proceso de

¹⁵ Algunos repositorios son: European Schoolnet (<http://re.eun.org/>), la iniciativa del gobierno español Agrega2 (<http://agrega.educacion.es/>), o el proyecto español EDIA (<https://cedec.intef.es/proyecto-edia/>) creado por la comunidad.

búsqueda de los LOs dentro de un repositorio. Existen numerosos repositorios de LOs tanto a nivel internacional como a nivel nacional que aprovechan estos metadatos (Palavitsinis, Manouselis, & Sanchez-Alonso, 2014). Dada su amplia adopción y su flexibilidad, LOM es utilizado como parte de otros estándares como es el estándar *Shareable Content Object Reference Model*, o SCORM, y también el estándar *IMS Content Packaging*, o IMS CP (IMS Global Learning Consortium, 2007). Además de LOM, y para resolver algunas de sus limitaciones, surge como una nueva familia de estándares *ISO/IEC 19788 Metadata for Learning Resources* (Pons Betrián, Hilera González, & Pagés Arévalo, 2011), el cual, aunque sigue evolucionando y siendo mantenido, no ha conseguido tanta adopción como LOM.

Con respecto al segundo reto, centrado en el empaquetamiento y distribución, IMS CP es el estándar que habitualmente es utilizado para para la empaquetación de contenido (IMS Global Learning Consortium, 2007), permitiendo distribuir un LO como un único archivo .zip que empaqueta un contenido educativo que puede ser básico o una estructura más compleja (e.g. una única actividad o una lección con múltiples actividades) además de un fichero índice que permite describir la estructura y finalidad del LO. IMS CP ha sido ampliamente adoptado por la gran mayoría de LMSs tanto de código abierto como comerciales y ha sido utilizado de manera efectiva para el empaquetado de recursos educativos distribuibles a través de la web. En este sentido, podemos asumir que los juegos serios que utilicen tecnologías web como HTML5 *Canvas* o WebGL (y en el pasado Java o Flash), puedan ser empaquetados y distribuidos como paquetes IMS CP. No obstante, esta especificación tiene sus limitaciones ya que no todos los juegos pueden distribuirse como contenido web, y además complica las labores de mantenimiento de los juegos (distribución de actualizaciones) y que los paquetes IMS CP deben de ser paquetes autocontenidos.

También relacionado con el segundo reto, pero en relación con la distribución de los recursos educativos, la especificación IMS Learning Tools Interoperability, o LTI, es el mecanismo de-facto para facilitar la integración de herramienta externas en una plataforma LMS (Bakhouyi et al., 2017). IMS LTI, proporciona un mecanismo de comunicación bidireccional entre un LMS y una herramienta externa. Por ejemplo, cuando el usuario accede a la herramienta externa desde el LMS, éste envía la información relativa al perfil del usuario (e.g. identificador único y roles que tiene en el LMS) para evitar tener que gestionar nuevas cuentas y permisos para cada LMS y curso donde se

quiera utilizar. Por otro lado, también es posible enviar métricas desde la herramienta externa al LMS, por ejemplo, enviar resultados de evaluación y si el estudiante ha completado la actividad. Para terminar, LTI dispone de una extensión denominada *Submission Review* el docente pueda acceder no solo a la nota que ha obtenido el estudiante en la actividad realizada en la herramienta externa, sino que también puede acceder a una vista detallada donde sería posible mostrar el proceso o registro de las acciones realizadas dentro de la herramienta externa que han llevado a conseguir dicha nota.

El tercer reto se puede abordar mediante el uso de un formato flexible que permita realizar el seguimiento de todas aquellas interacciones significativas, o eventos, que suceden mientras que el estudiante utiliza la herramienta. El estándar xAPI que provee tanto un modelo de datos para realizar el seguimiento de flujos de actividad, como una *Application Program Interface*, o API para el uso de *Learning Analytics* (Bakhouyi et al., 2017). Las trazas generadas con xAPI (también llamadas *statements*) incluyen información relativa al actor que ha llevado a cabo la interacción, al verbo que representa el tipo de acción realizada, y también al objeto que recibe la acción, con opción de incluir además los resultados y consecuencias de la acción, pudiendo ser notas u otros datos personalizados. Vale la pena destacar que xAPI puede ser extendido utilizando perfiles de aplicación que permiten definir un vocabulario común específico para una comunidad de práctica. En la siguiente sección se analizan algunos perfiles de aplicación de xAPI que son de especial relevancia con el presente trabajo de tesis.

La aplicación de xAPI como formato de datos extensible y una API simple para la recopilación de los datos ha promovido algunos casos de estudio de interés como el llevado a cabo por Samuelsen, Chen y Wasson (Samuelsen et al., 2019) que plantean una arquitectura para un ecosistema de aprendizaje que utiliza analíticas multiplataforma (Mangaroska, Vesin, & Giannakos, 2019), o el caso de *Learning Pulse* (Di Mitri et al., 2017) que utiliza *machine learning* sobre datos de múltiples fuentes que utilizan el estándar xAPI para predecir rendimiento en sistemas autorregulados de enseñanza que utilizan datos multimodales

También en relación con este tercer reto, existe la especificación IMS Caliper, que comparte unos objetivos similares a xAPI. No obstante, el modelo de datos de Caliper no

es tan flexible por lo que no goza del mismo grado de adopción que xAPI (Alonso-Fernández et al., 2019; Muslim et al., 2020; Samuelsen et al., 2019).

Una reciente revisión sobre LA (Muslim et al., 2020), muestra cómo cada vez hay más adopción de estos estándares educativos, siendo utilizados entre otras cosas para el modelado de la información recolectada. En este análisis se consideran un total de 137 herramientas, de las que 110 implementan su propio modelo de datos para manejar la información y el resto utilizan estándares o especificaciones relacionada con la recopilación de analíticas. Más concretamente, 22 de ellas utilizan xAPI, 2 utilizan Caliper, 2 utilizan *Activity Streams* y las 2 restantes utilizan *Learning Context Data Model*. Encontramos adicionalmente otra revisión de la literatura (Alonso-Fernández et al., 2019) que muestra que, de un total de 87 estudios analizados, aquellos que especifican un formato para la recolección de datos (16 de 87) muestran que: hay un total de 3 estudios utilizan xAPI, mientras que 3 utilizan XML, 3 utilizan *string ad-hoc*, 1 utiliza CSV, y el resto (6) utilizan tablas. Asimismo, en dicha revisión no se muestra ningún caso de uso de Caliper (Samuelsen et al., 2019).

Si bien no está claro por qué xAPI y Caliper todavía no tienen un alto grado de adopción, se apunta a una posible carencia de conocimientos sobre el uso de estas especificaciones (Samuelsen et al., 2019). Esta carencia se ve reflejada en usos incoherentes o inconsistentes de la especificación. Aun así, también remarcan que son buenos puntos de partida para integración de datos en LA. Otro de los motivos apunta a los problemas que los investigadores tienen al pasar de utilizar sistemas alternativos o ad-hoc, a adoptar y utilizar estos formatos estandarizados. Este proceso obliga a realizar un proceso de transformación de los datos del formato original al estandarizado, pudiendo existir *plugins* para realizar esta labor en algunas herramientas digitales que ayudarían en este proceso (Samuelsen et al., 2019). De esta manera se remarca la importancia de mejorar y desarrollar conjuntos de datos específicos para EDM y GLA, que sean portables (Romero & Ventura, 2020) y que consideren principios éticos, de privacidad, de protección de datos y de consentimiento (Ferguson, Hoel, Scheffel, & Drachsler, 2016).

Los estándares analizados abordan por separado las tres problemáticas antes mencionadas. No obstante, existen otras especificaciones y estándares que agrupan algunas de las especificaciones y estándares previamente analizados, con el objetivo de

abordar más de uno de los retos identificados previamente, con el objetivo de proporcionar una visión más integradora y global a la hora de abordar los retos.

La **Figura 2** muestra un resumen de las especificaciones y estándares individuales (filas superiores), y de los compendios de estándares que tratan de abordar más de un reto (filas inferiores).

	Característica					
	Descubrimiento	Acceso		Comunicación		
	Etiquetado	Empaquetado	Distribución	Reportes	Modelo de datos	Protocolos
LOM	X	-	-	-	-	-
CP	LOM	X	-	-	-	-
LTI	-	-	X	X	-	-
xAPI	-	-	-	-	X	-
SCORM	LOM	CP	-	X	-	X
CC	LOM	X	LTI	LTI	-	-
CMI-5	Basic	X	X	X	xAPI	X

Figura 2. Tabla que representa los estándares educativos y sus características y propósitos, mostrando en la sección superior aquellos estándares con objetivos más específicos, y en la inferior aquellos que tratan de cubrir múltiples necesidades

El primero de ellos es el *Shareable Content Object Reference Model*, o SCORM. SCORM nace como una colección de estándares y especificaciones para *Learning Objects* desarrollados por la iniciativa Alternative Distributed Learning, construida partiendo de los estándares IMS LOM para etiquetado estándar e IMS CP para empaquetado, añadiendo además una capa sencilla con mecanismos para la comunicación estableciendo protocolos para facilitar reportes de la actividad que incluyen información como la nota o su duración. Asimismo, también incluye el estándar IMS Simple Sequencing, o IMS SS, que secuenciación (ordenación) de múltiples actividades que pueda formar parte de una lección o curso. Pese a su gran adopción para su uso en LMS, donde las versiones 1.2 y 2004 de SCORM han sido las más populares y adoptadas, presenta algunas limitaciones. La principal es que, pese a que SCORM permite que LO pueda reportar información al LMS, el modelo de datos es muy limitado (si se ha superado la actividad y una puntuación). Por tanto, pese a que SCORM ofrece un punto de partida y ejemplo acerca de cómo abordar los retos de manera holística, no permite que pueda ser utilizado de

manera general para aprovechar todo el potencial de los juegos serios a la hora de recopilar las interacciones posibles. Además, SCORM está diseñado para funcionar utilizando principalmente en un entorno web, limitando considerablemente las tecnologías y plataformas con las que podríamos desarrollar los juegos serios. Finalmente, es importante mencionar que SCORM adolece de algunas limitaciones que podrían ser explotadas para hacer trampas (Papazoglou Papazoglakis, 2013).

El segundo de ellos es IMS Common Cartridge, una especificación similar a SCORM pero desarrollada por IMS y que también utiliza IMS CP para el empaquetamiento del contenido, pero de manera más flexible (Bakhouyi et al., 2017). Además, provee de mejores herramientas para construir paquetes a partir de módulos, facilitando la reutilización del contenido en un paquete. Finalmente, también se pueden incluir enlaces de LTI a modo de accesos directos a herramientas externas que pueden ser configuradas al importar el paquete IMS CC en el LMS.

Finalmente, ADL junto al Aviation Industry Computer-Based Training Committee, o AICC, presentan la especificación CMI-5, el que es considerado en algunos aspectos el “sucesor espiritual” de SCORM, resolviendo algunas de las limitaciones identificadas en SCORM, además integrar nativamente xAPI para la recopilación de analíticas de aprendizaje (Bakhouyi et al., 2017). Los paquetes de CMI-5 incluyen un fichero XML que describe los elementos internos utilizables en un curso, incluyendo *Assignable Units*, o AUs, las cuales representan actividades xAPI, con bloques encajables que pueden crear estructuras jerárquicas. Asimismo, uno de los factores más importantes para la aplicación de CMI-5 a juegos serios es el mecanismo que se utiliza a la hora de iniciar la AU y que puede ser aplicado no sólo en contextos web, sino que también permite saltar fuera del navegador e integrarse con una aplicación nativa (tanto para smartphones como en una aplicación de escritorio). De este modo, aunque el punto de entrada para iniciar la AU es un LMS compatible con CMI-5 podemos romper la barrera del navegador y trasladar la experiencia educativa a un juego externo, pudiendo posteriormente volver al LMS de manera fluida. Es importante destacar que, pese a que CMI-5 fue lanzado en 2016, sigue sin tener soporte por parte de la mayoría de LMS.

Por otro lado, CMI-5 pone en práctica la flexibilidad de xAPI ya que CMI-5 proporciona un perfil de aplicación de xAPI para poder recopilar los eventos de más alto nivel en la actividad educativa (inicio, fin, progreso) y promueve el uso de un perfil de aplicación

específico para recopilar los eventos de grano más fino dentro de la actividad educativa lanzada, por ejemplo, un juego serio.

En la siguiente sección se analiza en mayor detalle el concepto de perfil de aplicación xAPI y algunos perfiles de aplicación concretos que son de especial relevancia para este trabajo de tesis, como es el perfil de aplicación de xAPI para juegos serios

2.6. xAPI y los perfiles de aplicación xAPI

Como se ha mencionado en el apartado anterior, el estándar xAPI ofrece un formato muy flexible, que incluso podría considerarse como demasiado laxo, ya que impone muy pocas restricciones. Esta flexibilidad nos facilita que xAPI pueda ser aplicados en diversos contextos y herramientas facilitando exclusivamente la recolección y almacenamiento de datos. Como puede observarse en la **Figura 3** una traza xAPI se componen de cuatro componentes principales:

1. El actor que realiza la acción, normalmente el estudiante, aunque podría ser el docente para poder representar una intervención durante el proceso educativo.
2. El verbo que especifica qué acción se ha realizado dentro de la actividad educativa.

```
{
  "actor": { "name": "Dan" },
  "verb": {
    "id": "http://adlnet.gov/expapi/verbs/completed"
  },
  "object": {
    "id": "1",
    "definition": {
      "type":
      "https://w3id.org/xapi/seriousgames/activity-types/level"
    }
  },
  "timestamp": "2017-09-18T14:55:56.688Z",
  "result": {
    "success": true,
    "extensions": {
      "time": 120
    }
  }
}
```

Figura 3. Representación de una traza xAPI dividida en secciones, mostrando el actor en la parte superior, seguido del verbo, el objeto, así como el momento en el que ocurrió y su resultado.

3. El objeto concreto que recibe la acción, pudiendo incluir un identificador, así como una definición donde se puede concretar el tipo de ese objeto, por ejemplo, el tipo de actividad educativa (e.g. cuestionario, juego, vídeo, etc.).
4. Por último, el resultado de la acción. El resultado de la acción permite recopilar datos de relevancia como la duración de la actividad (si la traza representa que el estudiante ha terminado la actividad), puntuaciones, aciertos o fallos, así como variables personalizables por cada herramienta.

No obstante, esta flexibilidad también tiene la contrapartida de que cada herramienta podría construir su formato específico dentro del marco que proporciona xAPI. En este sentido, lograríamos compatibilidad con el estándar xAPI, pero la utilización de los datos fuera de la propia herramienta es limitada ya que la semántica sería específica para la herramienta. Además, tampoco podríamos aplicar otras herramientas, por ejemplo, para analizar los datos recopilados debido al mismo problema.

Los perfiles de aplicación para xAPI nacen con el propósito de paliar esta situación, permitiendo definir un vocabulario común y una semántica asociada dentro de una comunidad de práctica. El objetivo de esto es permitir, facilitar y fomentar que, en una comunidad, como la de los desarrolladores de juegos serios, puedan compartir un vocabulario común, de modo que sea posible crear un ecosistema de herramientas que faciliten la recopilación y análisis de los datos que generan los juegos serios. A continuación, analizaremos brevemente los perfiles de aplicación para CMI-5 y juegos serios por ser los más relevantes para este trabajo de tesis.

El perfil de aplicación xAPI para CMI-5, se encuentra definido dentro de la propia especificación CMI-5. Este perfil de aplicación, que es de obligado uso en la interacción entre un LMS y una AU CMI-5, tiene dos objetivos principales. En primer lugar, tiene como objetivo poder trasladar información relativa al contexto educativo y al usuario del LMS a la AU y, en segundo lugar, poder recopilar unas analíticas básicas para cualquier AU CMI-5. Más concretamente, se recopila el comienzo y la finalización de la actividad, así como el progreso del usuario en la AU. No obstante, dentro de la especificación CMI-5 se indica claramente cómo se puede combinar el perfil de aplicación xAPI para CMI-5 con otro perfil de aplicación diferente para proporcionar unas analíticas más detalladas para una AU, por ejemplo, el perfil de aplicación xAPI para juegos serios si la AU representa un juego serio.

El perfil de aplicación xAPI para juegos serios nace como una iniciativa para establecer un vocabulario común y una semántica asociada, en el dominio de los juegos serios (Á. Serrano-Laguna et al., 2017). En este perfil de aplicación se introducen una serie de constructos para los que puede ser relevante recopilar analíticas, definiendo con claridad, los verbos, objetos y tipos de actividad que podemos utilizar para cada uno de ellos. Los principales elementos dentro de este perfil de aplicación son:

1. Los *Completables* son elementos dentro del juego que un jugador puede empezar, sobre lo que puede progresar hasta completarlo, quizás múltiples veces. Los verbos asociados incluyen *initialized*, *progressed*, así como *completed*. Por otra parte, los tipos de objetos que se definen son: *game*, *session*, *level*, *quest*, *stage*, *combat*, *storynode*, *race* y *completable*.
2. Los *Accesibles*, como un espacio virtual dentro del mundo del juego donde un usuario puede acceder o saltar una o múltiples veces. Los verbos que incluye son *accessed*, y *skipped*. Los tipos de objetos que define son: *screen*, *area*, *zone*, *cutscene* y *accesible*.
3. Las *Alternatives*, que son decisiones a las que el jugador debe de enfrentarse a lo largo del juego, donde podrá elegir entre diferentes opciones, las cuales pueden ser desbloqueadas. Los verbos específicos para este constructo son *selected* y *unlocked*. Los tipos de objetos que define son: *question*, *menu*, *dialog*, *path*, *arena* y *alternative*.
4. Los *GameObjects*, agrupan de manera general cualquier elemento del juego con el que el jugador puede interactuar. Como verbos específicos para este constructo tenemos *interacted* o *used*, y definiciones de objetos como *enemy*, *NPC*, *item*, o *gameobject*.

Es importante destacar que el perfil de aplicación xAPI para juegos serios (xAPI-SG) permite definir un *Learning Analytics Model* reutilizable, de modo que se posibilita que se puedan crear herramientas específicas para el soporte de GLA para diversos juegos serios siempre que se utilice este perfil de aplicación como marco de referencia. Estas herramientas pueden ser utilizadas tanto para facilitar la recopilación de las analíticas, exploración y visualización de los datos recopilados, así como para poder definir análisis genéricos asociados a los conceptos definidos dentro del xAPI-SG.

El último perfil que destacar está relacionado con el auge de los juegos serios geoposicionados. Con la llegada de Pokémon GO, un juego gratuito de para smartphones, que combina actividades de exploración en la calle y videojuegos combinándolos creando una experiencia de Realidad Aumentada (de Souza e Silva, 2017), surgen también juegos serios geoposicionados y herramientas para la creación de los mismos (V. M. Pérez Colado, 2022). Con el mismo objetivo con el que se creó el perfil xAPI-SG surge un nuevo perfil de aplicación para poder representar analíticas de relevancia en actividades educativas geoposicionadas (Pérez-Colado, Rotaru, Freire-Morán, Martínez-Ortiz, & Fernández-Manjón, 2018). Este nuevo perfil de aplicación se utiliza normalmente en combinación con el perfil xAPI-SG, para complementar y precisar aquellos aspectos relacionados el posicionamiento del usuario o de los elementos de juego dentro del juego serio. Los principales elementos que se definen en este perfil de aplicación son:

1. *Location*, que describe la posición a nivel mundial del jugador expresada en formato *GeoPoint*¹⁶. Gracias a esta extensión, cuando realizamos el análisis podemos observar dónde (físicamente) se produce la acción, pudiendo extraer conclusiones cómo que, si el jugador abre muchas veces el menú, quizás esté físicamente perdido fuera del área de juego y necesite ayuda.
2. *Orientation*, que permite conocer hacia dónde está mirando el jugador, y que resulta muy útil para identificar casos en los que el jugador ha llegado a una conclusión equivocada que le lleve en dirección opuesta al objetivo.
3. *Guide*, que permite incluir en caso de haberla, la guía de puntos que está siguiendo el jugador en ese momento, y que ayuda a contextualizar mucho más detalladamente las trazas para un mejor análisis.

Los juegos serios geoposicionados son un caso donde las analíticas y la existencia de un sistema de soporte a las GLA juegan un papel fundamental. Esto es debido a que los juegos serios geoposicionados acentúan mucho más el problema de monitorización y seguimiento de la actividad educativa ya que es mucho más frecuente que el docente y los estudiantes se encuentren separados físicamente y que, por tanto, se requiera un mayor

¹⁶ Tipo de datos para almacenar coordenadas geoposicionadas definido por ElasticSearch B.V y disponible online en: <https://www.elastic.co/guide/en/elasticsearch/reference/current/geo-point.html>

esfuerzo de coordinación y monitorización del progreso de la actividad (V. M. Pérez Colado, 2022).

Los perfiles de aplicación de xAPI remarcan la flexibilidad del estándar para ayudar a los investigadores, ya que permiten establecer las bases para una sólida representación de los datos, establecer una semántica compartida en un dominio concreto (comunidad de práctica) y posibilitan la reutilización con herramientas de soporte al GLA. La aplicación de estos perfiles de aplicación facilita el desarrollo multidisciplinario donde desarrolladores de juegos serios, científicos de datos, e investigadores pueden beneficiarse de un sustrato común con el que aplicar GLA con éxito.

2.7. A modo de conclusión

Los juegos serios son herramientas educativas muy interesantes al aunar aprendizaje y entretenimiento. Los juegos serios promueven el aprendizaje activo, donde el alumno puede aplicar el conocimiento aprendiendo de los errores y permiten representar cualquier situación de forma controlada y motivadora. Sin embargo, pese a su alto potencial para mejorar la enseñanza su uso en la educación general es todavía limitado. Los educadores deben enfrentarse a diferentes retos para usar eficazmente estos juegos, entre los cuales se encuentran:

- Conseguir o construir los juegos serios. Ya sea desarrollando un juego serio desde cero, o partiendo y personalizando uno existente mediante una herramienta de autoría.
- Asegurarse de que realmente son juegos útiles, es decir, que su eficacia esté probada mediante la validación de los mismos. Además, los juegos, deben proporcionar información y métricas suficientes como para que puedan ser utilizados de manera adecuada en el proceso educativo.
- Distribuir el juego entre los estudiantes y monitorizar el progreso de las actividades de juego.

Para abordar estos problemas, es necesario aplicar una aproximación holística que contemple el ciclo de vida completo del juego serio, empezando desde su creación, pasando por su uso y la verificación de los resultados de su aplicación. Como hemos visto, aquellos juegos diseñados en torno a objetivos educativos específicos tienen ventajas a la hora de probar su eficacia y de facilitar el proceso de validación necesarios para mejorar

la confianza de los profesores en ellos. Una de las formas de medir esta eficacia es a través del uso de analíticas de aprendizaje, que pueden aplicarse durante todo el ciclo de vida del juego. Por un lado, facilitando a los desarrolladores la creación y abstracción de métricas para verificar que el juego cumple las expectativas de los usuarios y por otro permitiendo el diseño de experimentos para validar la eficacia de dicho juego serio. Pero las analíticas también son útiles para que los profesores entiendan realmente lo que pasa en la clase cuando se despliegan los juegos y los estudiantes están usándolos.

En esta tesis nos centraremos especialmente en las oportunidades que nos ofrecen las analíticas de aprendizaje, haciendo un especial énfasis en su aplicación en el proceso de validación de los juegos serios, y en construir un ecosistema de herramientas que faciliten su uso. Es importante remarcar que el proceso de validación puede realizarse de múltiples maneras (e.g. observación, entrevistas personales, encuestas pre-post) por lo que una herramienta de validación debe ser flexible para soportar múltiples casos de estudio, así como la creación de grupos de control. También debe permitir la gestión de usuarios, posiblemente anonimizados, y sus datos. Se deben considerar factores como las normativas de privacidad y protección de datos y que esto sea compatible con su acceso a los investigadores para realizar labores de Game Learning Analytics y Educational Data Mining. Esto permite funcionalidades como extraer evidencias para el proceso de validación e incluso se puede usar como mecanismo de evaluación de los jugadores.

El uso de estándares educativos puede facilitar diversos aspectos relacionados con la adopción de los juegos serios. Por ejemplo, su despliegue en el aula y la generalización de analíticas, fomentando la construcción de herramientas interoperables basadas en estándares que puedan ser reutilizadas. Entre los estándares analizados destaca xAPI y su combinación con el perfil de aplicación para juegos serios como mecanismo para facilitar la creación de un ecosistema de herramientas para el soporte de GLA en la comunidad de juegos serios. Asimismo, IMS LTI y CMI-5 permiten abordar otros aspectos más técnicos pero necesarios para facilitar el despliegue de los juegos en el aula y su integración eficaz con otras herramientas educativas como son los LMSs.

Capítulo 3. Objetivos

A lo largo del capítulo anterior se han expuesto algunos de los beneficios de los juegos serios como son su impacto como herramienta motivacional, o que proporcionan una experiencia inmersiva y, a la vez, un escenario controlado y seguro donde poder simular situaciones educativas que en un entorno real podrían ser poco seguras o muy costosas. Pero, pese a estos beneficios, su adopción todavía sigue siendo limitada en la educación general. Asimismo, en el capítulo anterior se han analizado las diferentes limitaciones y problemas relacionados con el ciclo de vida de los juegos serios que afectan gravemente a su proceso de desarrollo y a su puesta en práctica de manera extensiva.

Una gran parte de estos problemas están relacionados con el hecho de que, tradicionalmente, los juegos han sido percibidos como una *caja negra* (Horn et al., 2016; Moreno-Ger, Sierra, et al., 2008). El impacto del juego serio en los estudiantes (y su potencial beneficio) sólo se podía observar de manera limitada, normalmente a través de una puntuación, un porcentaje de progreso o a través de hitos conseguidos durante el juego. Si se trata al juego como caja negra, no es posible observar el proceso que siguió el estudiante para lograr dicha puntuación o progreso, por lo que es necesario confiar plenamente en mediciones externas sobre la fiabilidad y efectividad del juego. Pese a que es posible realizar una validación formal de su efectividad para justificar su uso en un currículum educativo, esta falta de detalle lastra la confianza de los docentes en la propia herramienta. Además, limita las oportunidades y escenarios de aplicación del juego serio al contexto de uso específicamente considerado por parte de los creadores del mismo.

Para paliar esta limitación, creemos que es necesario mejorar la transparencia del proceso para saber realmente qué sucede dentro del juego durante las sesiones de juego. Conocer cómo juegan los estudiantes permite que los docentes no sean meros actores pasivos, incluso permite que puedan participar activamente en las sesiones de juego como facilitadores. Esto requiere que seamos capaces de proporcionarles las herramientas adecuadas para que comprendan qué está pasando mientras sus estudiantes juegan. En este sentido, creemos que las analíticas de aprendizaje aplicadas a los juegos serios pueden jugar un papel clave en este proceso de monitorización del progreso de los estudiantes durante las sesiones de juego. También pueden contribuir a proporcionar las evidencias necesarias para que los docentes puedan evaluar a los estudiantes de manera

más eficaz y transparente. De esta forma se puede evaluar no sólo el resultado final sino también el proceso que han seguido los estudiantes para alcanzar sus resultados finales.

Asimismo, la implementación de las analíticas de aprendizaje usando estándares técnicos educativos proporciona más generalidad a las soluciones software y una mayor interoperabilidad. Esto también contribuye a simplificar el proceso de validación de la efectividad de los juegos serios ya que reduce el coste de las analíticas al simplificar la reutilización de herramientas y componentes genéricos compatibles con estas especificaciones técnicas (e.g. *tracker* basado en xAPI, uso de LRS estándar).

Con todo ello, se plantea como objetivo global de este trabajo de tesis mejorar el ciclo de vida y adopción de los juegos serios mediante la integración de Game Learning Analytics (GLA). Para ello se desarrollarán herramientas de soporte a las GLA, que faciliten su uso con juegos serios, por ejemplo, para mejorar su validación. Para abordar este objetivo general, se plantean los siguientes objetivos específicos:

- O1. Mejorar el proceso de implementación de analíticas para juegos serios mediante la utilización y sistematización de Modelos de Analítica de Aprendizaje (LAMs).
- O2. Sistematización del proceso de integración de analíticas de aprendizaje en juegos (Game Learning Analytics) con herramientas de autoría y juegos serios.
- O3. Desarrollar una plataforma de analíticas de aprendizaje en juegos orientada a la validación de juegos serios.
- O4. Mejorar el proceso de integración de las herramientas de ciencia de datos y de analíticas de aprendizaje en juegos mediante la aplicación de estándares educativos.
- O5. Validar y probar la efectividad de las herramientas y soluciones desarrolladas.

Por tanto, y como elemento vertebrador de estos objetivos en este trabajo de tesis, se diseñará y desarrollará una plataforma de GLA de propósito general con capacidades para realizar actividades de analíticas y que permitan validar juegos serios. Estos objetivos se organizan para su desarrollo en la planificación mostrada en la **Figura 4**. En ella, además de mostrarse el orden en el que se abordan los objetivos, también se indican las diferentes fases de prototipado, mejora y mantenimiento de esta plataforma de GLA. Desde el punto de vista temporal, se comenzará con el primer objetivo (O1) y se realizarán las fases de prototipado de O2 (diseño de plataforma), empezando a desarrollar e integrar componentes con el sistema de analíticas. A mitad del desarrollo de O2, se abordará O3

(plataforma), comenzando con el prototipado de un sistema capaz gestionar encuestas para validación formal de juegos serios, para después ampliarlo hasta la construcción de una plataforma de GLA para validar juegos. Esta plataforma de validación será capaz de integrarse con las herramientas desarrolladas en el segundo objetivo. Como parte del proceso de desarrollo de este O3, se incluirá el desarrollo de O4 (integración), dando soporte a varios estándares educativos relacionados tanto con las analíticas de aprendizaje, como a la integración de herramientas educativas en plataformas educativas existentes (e.g. Moodle, SAKAI, Canvas, etc.). El quinto objetivo (O5, validación y pruebas) se desarrollará durante todo este trabajo de tesis para asegurar el correcto funcionamiento de todos los componentes software desarrollados, validar su efectividad y resolver las limitaciones identificadas durante la validación.

La metodología utilizada es una metodología ágil basada en el desarrollo de prototipos incrementales. Estos prototipos se aplican y evalúan de modo que permite construir y ampliar los resultados obtenidos en fases previas, para lograr herramientas finales más probadas y funcionales.

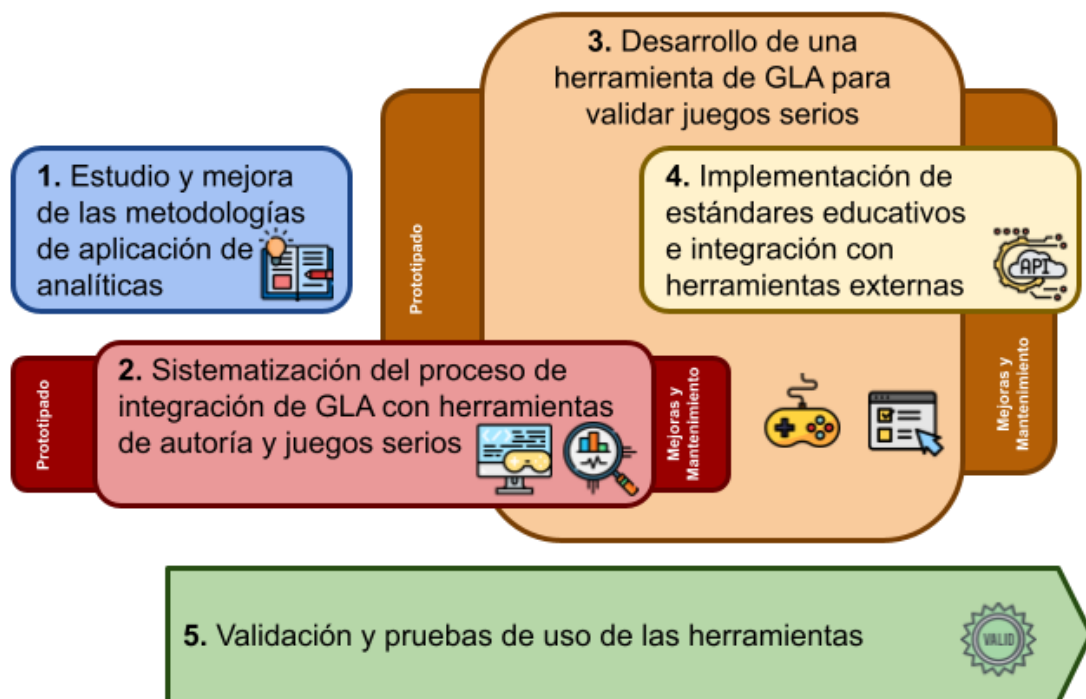


Figura 4. Planificación de los objetivos de la tesis.

A continuación, se detallan los objetivos, explicando brevemente las limitaciones identificadas en el estado del arte que han servido como elementos motivadores, así como

las características, funcionalidades y componentes que se desarrollarán para abordarlos de manera efectiva.

3.1. Mejoras en el proceso de diseño de analíticas

Habitualmente, los docentes combinan diferentes actividades y contenidos educativos dentro de una plataforma educativa para abordar las diferentes competencias y habilidades que forman parte del programa educativo. Estas actividades pueden ser muy variadas como, por ejemplo, presentaciones, videos, cuestionarios, actividades externas o incluso juegos serios. Mientras que las actividades más simples pueden ofrecer resultados y métricas claras con las que los docentes pueden evaluar el progreso de los estudiantes, en el caso de los juegos serios nos encontramos limitados por las métricas y resultados que los diseñadores del juego han identificado como potencialmente relevantes y que, además, hayan proporcionado al docente un medio para poder consultar dichas métricas. Pese a que los juegos serios pueden ser adecuados para la experiencia educativa, las métricas obtenidas pueden no ser adecuadas para los docentes, dificultando su trabajo o incluso la justificación de este tipo de herramienta como un recurso educativo más.

Las analíticas de aprendizaje aplicadas a los juegos serios permiten mejorar la comprensión del proceso de aprendizaje que siguen los estudiantes y, además, devolver el control a los docentes durante las sesiones de juego. A través de las analíticas, los docentes pueden monitorizar y supervisar la sesión de juego para corroborar que los estudiantes realmente están siguiendo y aprovechando la actividad, e incluso intervenir durante las sesiones si algún estudiante se queda estancado en alguna parte del juego. No obstante, para que un profesor pueda llegar a beneficiarse de estas ventajas, el juego tiene que ser capaz de identificar y tomar nota de aquellas acciones que está realizando el jugador que sean relevantes para el proceso educativo. Para ello, los desarrolladores e investigadores deben definir un modelo de analíticas de aprendizaje, o *Learning Analytics Model* (LAM) (Hauge et al., 2014), donde se describan todas aquellas acciones significativas que permitan identificar y evaluar la satisfacción de los objetivos educativos abordados en el juego serio, así como aquellas acciones que enriquezcan el conjunto de datos y permitan una mejor visión del proceso de juego y aprendizaje. De esta manera, la inclusión de un LAM en un juego serio proporciona una base sólida sobre la que extraer conclusiones significativas acerca de la efectividad del juego. La identificación y elección de aquellas acciones y métricas significativas con las que

construir el LAM será determinante para su utilidad durante el resto del ciclo de vida del juego serio.

Puesto que, por lo general, no existe un uso generalizado y sistematizado de las GLA, en este objetivo se pretende abordar el proceso de construcción del LAM para mejorar la calidad de las analíticas finales y sistematizar su proceso de selección, secuenciación y detalle. Como resultado de este proceso, se generará un modelo de analíticas con el que implementar un sistema de analíticas que permita medir la efectividad del juego y la integración de estos resultados en el proceso de evaluación del estudiante de cara al programa docente.

Por otro lado, aunque la inclusión de un LAM en un juego permite la extracción de conclusiones para el proceso de evaluación de los estudiantes como una actividad aislada, los juegos serios (al igual que otros contenidos educativos) suelen formar parte de un conjunto de actividades y contenidos educativos organizados a modo de curso en una plataforma de aprendizaje, para los que también es posible o deseable recopilar analíticas de aprendizaje. En este sentido, las diferentes actividades y recursos educativos, incluido el juego serio, contribuyen al progreso de los diferentes objetivos educativos planteados para el curso. Para facilitar este tipo de analíticas de mayor granularidad planteamos un modelo de analíticas agregadas, que llamaremos meta-LAM, y que permite asociar los resultados obtenidos en el juego serio a los diferentes objetivos educativos de un programa docente específico. Este meta-LAM también es de utilidad en situaciones donde un juego serio integre otros minijuegos que han sido desarrollados de manera independiente, pero que en su conjunto forman una actividad educativa unificada que aborda una serie de objetivos educativos concretos

3.2. Sistematización del proceso de integración de GLA con herramientas de autoría y juegos serios

Una vez completado el objetivo anterior, habremos establecido un proceso que nos permite construir LAMs en juegos serios con los que medir y evaluar su efectividad e incluso contribuir al progreso y analíticas agregadas para medir el progreso del proceso educativo de manera global. No obstante, para poder llegar a extraer dichas conclusiones de los datos de juego, los datos generados basados en el LAM deben ser explotados a través de un *Learning Analytics System* (LAS) (Hauge et al., 2014; Serrano Laguna, 2012), que se encarga de digerir los datos crudos y ofrecer al usuario final los análisis,

visualizaciones y cuadros de mando (*dashboards*) con los que facilitar el proceso de toma de decisiones o incluso la propagación automática de las conclusiones a otros sistemas. Por ejemplo, sería posible que el LAS propague los diferentes resultados obtenidos en el juego a la plataforma de aprendizaje donde el docente tiene organizadas las diferentes actividades del curso. Los componentes del LAM y LAS y sus relaciones se presentan en la Figura 5.

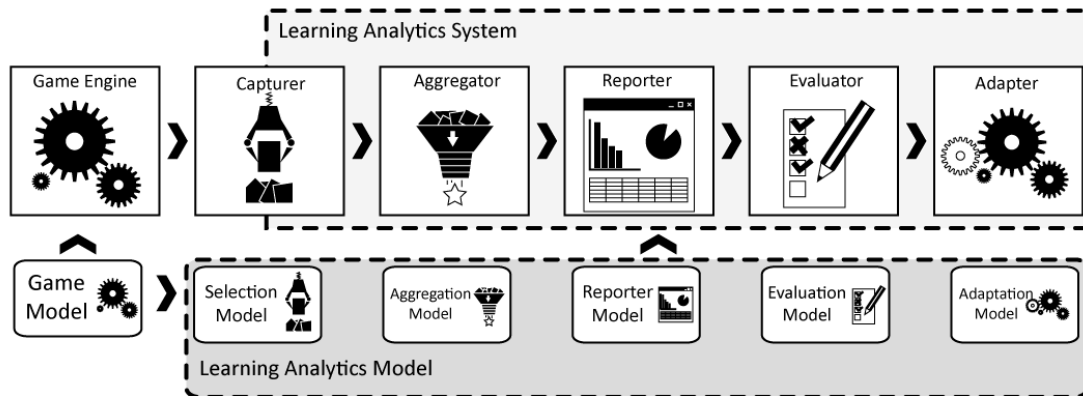


Figura 5. Relación entre LAM (rectángulo sombreado inferior) y LAS (rectángulo en parte superior) - el LAS proporciona el mecanismo, y el LAM la información de cómo proceder.

Si bien la construcción de un LAS a medida del LAM sería la forma directa de lograr este proceso, la construcción de un LAS no es un proceso sencillo al alcance de la mayoría de los desarrolladores de juegos serios. Este hecho ha promovido que, en muchos casos, se descarte el uso de analíticas o se improvisen soluciones mucho menos flexibles que, a la larga, impiden una explotación efectiva de los datos recopilados.

Para simplificar la explotación de un LAM, se proporcionarán soluciones que permitan sistematizar el proceso de integración de GLA durante todo el ciclo de vida del juego serio. Estas soluciones tomarán como base la experiencia adquirida en los proyectos europeos del H2020 RAGE y BEACONING, y proporcionará un LAS con el objetivo de lograr un proceso de integración sistemático y genérico capaz de adaptarse a múltiples LAMs. Este LAS contará con un recolector de datos capaz de almacenar los eventos de relevancia recopilados en los juegos serios, un agregador encargado de realizar los procesos computacionales de análisis sobre los datos recibidos, así como un cuadro de mandos capaz de mostrar las visualizaciones de los resultados de los análisis.

En este objetivo se hará especial énfasis en dos de las partes del proceso de integración que más afectan a los desarrolladores de juegos serios: el proceso de generación y recolección de las trazas y el proceso de análisis y visualización de las mismas. Asimismo, este objetivo se abordará aplicando xAPI como estándar de-facto para analíticas de aprendizaje, con el objetivo de que sirva como formato para facilitar otras de las tareas relacionadas en el procesamiento de analíticas como la agregación, análisis y almacenamiento. Como resultado, los componentes desarrollados podrán ser interoperables con otras herramientas desarrolladas fuera del marco de este trabajo de tesis.

Como parte de este objetivo, se validará la aproximación con dos casos de estudio concretos. Por un lado, se explorará el proceso de integración utilizando varios juegos serios previamente existentes del grupo e-UCM, así como el juego serio FormalZ que forma parte del proyecto ERASMUS+ IMPRESS. Por otro lado, se explorará la integración de un LAM genérico en la herramienta de autoría uAdventure. Esta integración permitirá que los desarrolladores puedan disponer de un modelo de analíticas y visualizaciones por defecto. Esta integración se realizará basándose en el perfil de xAPI para juegos serios (xAPI-SG), de modo que esta integración no sólo sea de utilidad para los otros objetivos de este trabajo de tesis, sino también para que pueda ser compatible con otras herramientas como un *Learning Record Store* (LRS) genérico compatible con el estándar xAPI

3.3. Desarrollo de una plataforma de GLA para la validación sistemática de juegos serios

Uno de los retos principales en la validación de los juegos serios está relacionado con la complejidad a la hora de poner en práctica el diseño experimental. La necesidad de gestionar un gran volumen de usuarios en diferentes grupos de estudio, facilitar el despliegue de los experimentos en el aula, etc., hace que el proceso sea tan costoso que no esté al alcance de muchos de los equipos de desarrollo de juegos serios.

En este objetivo se busca sistematizar el proceso de validación para reducir su complejidad, fomentando por tanto la validación de los juegos serios. Si bien existen múltiples metodologías para la experimentación y validación, la metodología basada en cuestionarios previo y posterior a la sesión de juego (cuestionarios *pre-post*) es la más utilizada para la validación de juegos serios (Bellotti et al., 2013). Si se identifica un

cambio positivo significativo comparando los resultados en ambos cuestionarios, se puede inferir que dicho cambio es debido a la eficacia del juego serio.

Una de las claves de la simplificación consiste en automatizar gran parte de este proceso, no solo para reducir la carga de trabajo de las personas involucradas en el experimento, sino también para minimizar los riesgos de posibles errores humanos en este tipo de experimentos donde, cuando se producen errores, no siempre es posible volver a tener acceso a los estudiantes que participan en el experimento. Por ejemplo, tradicionalmente, el proceso de validación implica preparar e imprimir múltiples cuestionarios en papel, y distribuirlos entre los participantes, lo que consume grandes cantidades de tiempo del profesor, especialmente cuando hay muchos participantes. Por otro lado, aunque se podría optar por utilizar una herramienta para cuestionarios en línea para reducir la carga de gestión, normalmente este tipo de herramientas no están pensadas para poder organizar múltiples cuestionarios cumplimentados por un mismo usuario (*pre* y luego *post*), ni asociar la disponibilidad de cuestionarios a condiciones externas (*pre* sólo antes de jugar, *post* sólo después). Por tanto, como parte de este objetivo, buscamos simplificar el proceso de realización de cuestionarios y facilitar su análisis comparativo posterior.

Pese a que el foco de este objetivo se centrará en la validación en formato pre-post, la plataforma ha de ser flexible y amoldarse al concepto de estudio. Un estudio podrá estar compuesto por múltiples actividades en una secuencia concreta pero configurable para cada grupo de participantes. Las posibles actividades pueden ser una entrevista, una sesión de juego, o la visualización de un vídeo, abriendo las puertas a otros estilos de validación. Otro de los aspectos a explorar será la posibilidad de realizar experimentos de validación con grupo de control. Los participantes del grupo de control realizarán una actividad alternativa (por ejemplo, jugar a un juego de duración similar pero no relacionado con el tema del estudio), pudiendo así descartar que la mejora de conocimiento no sea derivada del uso del juego, puesto que los cambios que se produzcan en las encuestas previa y posterior del grupo de control pueden contrastarse con los resultados del resto de participantes. Finalmente, se podrían llevar a cabo experimentos con intervenciones espaciadas temporalmente, midiendo las variaciones de las puntuaciones de los cuestionarios con días o semanas de diferencia en las actividades.

Adicionalmente, los cuestionarios pre-post pueden combinarse con GLA para que los investigadores también puedan extraer conclusiones (Freire et al., 2016) que permitan

validar el juego a través del análisis de las *trazas* y los resultados. Después de haber trabajado en los diferentes elementos que componen un LAS mediante el desarrollo e integración a medida realizado en el segundo objetivo, podemos generar una arquitectura que reutilice el mayor número de partes posibles en una nueva plataforma de validación unificada con GLA. Esta solución deberá cumplir los siguientes requisitos. En primer lugar, deberá proporcionar un sistema de control y autenticación de usuarios seguro, que además sea capaz de mantener su anonimato (utilizando identificadores anónimos) para así cumplir con el reglamento general de protección de datos. En segundo lugar, debe integrar un sistema de almacenamiento y gestión de los datos que mantenga su consistencia, los etiquete correctamente, y evite sobrecargas. En tercer lugar, debe ser un sistema confiable, permitiendo que los datos puedan ser replicados en copias de seguridad, y resiliente durante los experimentos puesto que en la mayoría de los casos se cuenta con oportunidades muy limitadas para realizar experimentos, así como con franjas de tiempo muy estrictas. En cuarto lugar, debe ser escalable de manera sencilla para poder añadir más capacidad de procesamiento y así poder amoldarse a grandes volúmenes de usuarios y experimentos de forma simultánea. Por último, su despliegue debe ser sencillo, resultando así accesible para todo aquel interesado en validar un juego serio.

Además de esta plataforma, el componente recolector, o *tracker*, incluido en el juego que se encargará de conectar el juego con la plataforma también debe de ser ampliado en consecuencia durante la realización de este objetivo para poder complementar estas características. Por ejemplo, para hacer frente a problemas de conectividad, debe de ser capaz de almacenar una copia de seguridad de los datos - de forma que puedan ser enviados automáticamente cuando se restablezca la conectividad o, en el caso peor, recolectados manualmente por los investigadores más adelante.

3.4. Aplicación de estándares educativos

Uno de los problemas identificados en el estado del arte es la dificultad que experimentan los usuarios a la hora de utilizar juegos serios en clase, puesto que es común que estos no se integren bien con las herramientas educativas que se utilizan para organizar el proceso de enseñanza (por ejemplo, un LMS como Moodle). Habitualmente, los juegos serios no implementan ningún tipo de mecanismo de comunicación para recolectar los resultados que ha logrado el estudiante en el juego y almacenarlo en el LMS para que esté accesible con el resto de resultados de otras actividades. Para lidiar con esta problemática de

integración, los estándares técnicos educativos ofrecen diferentes protocolos y mecanismos para fomentar la interoperabilidad de los recursos educativos. Estos estándares técnicos educativos abordan diferentes aspectos de la interoperabilidad, entre otros: el empaquetamiento y distribución, la identificación del usuario dentro del recurso educativo, o la comunicación bidireccional entre el recurso y las plataformas educativas. Uno de los estándares más ilustrativos y comúnmente utilizado es SCORM que, entre otras cosas, permite empaquetar algunos juegos serios como contenidos educativos para que sean distribuidos desde el LMS (Á. Serrano-Laguna et al., 2017), y reportar un resultado al LMS.

Partiendo de la plataforma capaz de realizar GLA y validar juegos serios del objetivo anterior, este objetivo busca mejorar la situación de la integración de los juegos serios con los procesos educativos a través de la utilización de estándares educativos que fomentan la interoperabilidad entre las herramientas educativas. Para ello, el presente objetivo se divide en tres fases. En primer lugar, requiere un análisis de los estándares educativos existentes y las posibilidades y limitaciones que estos ofrecen. Este análisis se presenta como parte del estado del arte, en la Sección 2.5. En segundo lugar, requiere soporte en la plataforma de GLA, para permitir una comunicación bidireccional tanto con un LMS (u otras plataformas educativas) como con los propios juegos que permita la identificación del usuario durante todo el flujo de la actividad y el retorno de resultados entre las diferentes plataformas. A través de este mecanismo de intercambio, los docentes podrán conectar la plataforma GLA con su LMS y realizar actividades con juegos serios, incluyendo además la posibilidad de realizar actividades de validación. Esta comunicación también permite ampliar las capacidades de evaluación de la plataforma para que no sólo sea capaz de evaluar juegos serios sino todo tipo de plataformas educativas compatibles con los estándares soportados. Como caso de ejemplo de este tipo de validación, se integra la plataforma con el juego/plataforma FormalZ del proyecto ERASMUS+ IMPRESS.

La última parte de este objetivo ofrece una solución a la actual problemática de soporte a actividades con analíticas de aprendizaje en las principales plataformas educativas. Actualmente, el estándar CMI5 permite el lanzamiento de actividades que utilicen analíticas de aprendizaje. Sin embargo, su adopción en las principales plataformas educativas es aún muy limitada, lastrando la normalización del uso de actividades con GLA. Para solucionarlo, la plataforma GLA deberá ofrecer soporte a lanzamiento CMI5

y con ello, combinados a sus otros mecanismos de integración, permitirá que se pueda acceder a GLA desde LMSs tradicionales sin ningún esfuerzo adicional.

3.5. Validación del sistema y sus componentes

Para verificar que el sistema de GLA y sus componentes funcionan correctamente y cumplen con los requisitos de diseño, es necesario validarlos, idealmente poniéndolos a prueba en distintos casos de uso. La herramienta de autoría de juegos serios uAdventure desarrollada en el grupo e-UCM brinda la oportunidad de reutilizar, reeditar y/o personalizar juegos creados con esta herramienta o su predecesora (eAdventure). Para este objetivo, se explota esta capacidad integrando un LAM genérico que dota de analíticas por defecto a los juegos serios creados con uAdventure. Esto se realiza tras el desarrollo e integración del componente recolector de datos que registra las acciones que hace el usuario a lo largo del juego, así como su integración con la plataforma de analíticas RAGE Analytics, que procesa los datos y los pone a disposición de los usuarios mediante paneles de visualizaciones. A través de esta experiencia, se generará el modelo de analíticas para un juego existente para uAdventure y se implementará usando el prototipo de sistema de analíticas con el objetivo de probar y mejorar los componentes desarrollados. Se diseñarán además cuestionarios para los alumnos que realizarán antes y después de la sesión de juego, ayudando no sólo a validar el propio juego serio sino también el proceso de validación.

Los resultados obtenidos de los pilotos realizados con un primer prototipo permitieron sentar las bases para el desarrollo de la plataforma de GLA orientada a la validación del objetivo 2. Una vez completado su desarrollo e integración según el objetivo 3, nuevas pruebas tanto con uAdventure como con juegos independientes permitieron validar todas las capacidades integradas en el sistema: gestión de usuarios, recopilación de analíticas y soporte a la validación de juegos serios. Finalmente, para validar el objetivo 4, tras implementar los estándares educativos de interoperabilidad, se realizó la integración de la plataforma en un campus virtual.

Capítulo 4. Discusión integradora y contribuciones

A lo largo del presente capítulo se describe el trabajo realizado para alcanzar los objetivos de la tesis. Como la tesis está compuesta por un compendio de publicaciones, las descripciones sólo hacen referencia a los resultados principales de las publicaciones correspondientes, si bien los detalles están disponibles en las publicaciones completas incluidas en la propia tesis. Además, se proporciona una discusión integradora sobre cómo dichos trabajos han contribuido a lograr los resultados finales de la tesis.

A continuación, se presenta un subapartado por cada uno de los cinco objetivos concretos y se relaciona con las correspondientes publicaciones para describir cómo se han abordado para alcanzar una solución a las problemáticas planteadas.

4.1. Mejoras en el proceso de diseño de analíticas

Un elemento crucial para lograr analíticas efectivas sobre juegos es disponer de un modelo que relacione el diseño educativo con las analíticas concretas del juego. En el capítulo anterior se enfatiza la función de un LAM, o *Learning Analytics Model* (Hauge et al., 2014), donde se definen los requisitos necesarios para poder realizar analíticas que den soporte al ciclo de vida de un juego, pero esta definición es genérica y poco específica. Algunos autores (Mislevy, 2005) plantean crear un modelo de evaluación basado en la evidencia, o *evidence-centered assessment design* (EDC), donde se define la información que debe de ser registrada, cómo y cuándo se deben generar registros, y cómo interpretar los datos resultantes. De este modelo, resulta especialmente interesante la capa de evaluación conceptual, *Conceptual Assessment Network* (CAF), que relaciona las variables del modelo (qué se está midiendo), el entorno donde los estudiantes completan las tareas (dónde se miden), la metodología observar los datos y analizarlos (cómo se miden), y finalmente el modelo de ensamblado que unifica todas estas tareas en una única característica (cuánto se debemos de medirlos). Todas estas cuestiones pueden ser abordadas por un LAM, y, tras el estudio y trabajo realizado acerca de cómo secuenciar y sistematizar éste proceso en el proyecto BEACONING (Freire, Martínez-Ortiz, Alonso-Fernández, Pérez-Colado, & Fernández-Manjon, 2015), se plantea la siguiente lista de elementos que debe incluir dicho documento:

1. Los objetivos educativos específicos, definidos en un listado antes de realizar ninguna tarea relacionada con analíticas. Con este listado obtenemos una idea

específica del diseño educativo del que partir como base para generar las analíticas, siendo necesaria la vinculación de este diseño con las analíticas educativas (Bakharia et al., 2016).

2. Los objetivos del juego, incluyendo sus misiones y niveles, y su correspondencia con los objetivos educativos. Esta interrelación no tiene por qué ser biyectiva, de tal manera que un objetivo educativo puede satisfacerse con múltiples objetivos del juego, y de manera inversa, un sólo objetivo del juego puede completar varios objetivos educativos. Con los objetivos del juego obtenemos un diseño del juego específico adaptado a los objetivos educativos.
3. Las trazas a enviar, a definir por los desarrolladores del juego serio basándose en los diseños educativos y del juego definidos anteriormente. Se define la información necesaria que debe ser registrada y enviada por el juego para que se pueda identificar que los objetivos del juego se cumplen. Se propone el uso del estándar de xAPI para juegos serios (Á. Serrano-Laguna et al., 2017) pues permite escoger las interacciones a partir del vocabulario base planteado, ya que su diseño orientado a los juegos serios unifica la terminología para el desarrollo multidisciplinario de las analíticas. Es importante destacar que se pueden añadir otras trazas que se consideren importantes para los usuarios estén o no relacionadas directamente con los objetivos educativos. Esta información extra es útil para otros procesos de depuración y puede ser posteriormente explotada con técnicas de minería de datos.
4. El modelo de análisis, que defina cómo las trazas deben de ser analizadas e interpretadas. En general, el proceso de análisis consiste en mantener una estimación actualizada del estado de los objetivos educativos, especificando cómo cada una de estas trazas que se genera va aportando y modifica dicho estado. La introspección educativa más valiosa se obtiene teniendo en cuenta los objetivos educativos y cómo se reflejan en los objetivos del juego.
5. Las visualizaciones y gráficas que representen los resultados del análisis de manera adecuada para los interesados (e.g. *dashboards*, paneles o cuadros de mando). Cabe destacar que, si se usa un modelo de analíticas genérico, es probable que se puedan reutilizar muchas de las visualizaciones genéricas, pudiendo requerir únicamente modificaciones menores. Si las genéricas no son suficientes, se deben especificar las nuevas visualizaciones a desarrollar y añadir en el panel

Pese a que es habitual que los análisis se realicen después de jugar al juego, no tiene por qué ser el caso, ya que muchos sistemas pueden actualizar sus visualizaciones casi en tiempo real. Esto nos permite crear alertas o avisos para los profesores permitiendo solucionar posibles problemas o interactuar con alumnos cuando el juego lo indique. Por ejemplo, cuando un estudiante se haya quedado atascado en un nivel, o haya avanzado mucho más rápido de lo esperado y necesite atención, o incluso cuando parezca que pueda estar haciendo trampas. Es posible incluir estos casos dentro del modelo de analíticas, pudiendo especificar un tipo de visualización en concreto que se muestre únicamente en los paneles de gráficas cuando sea necesario. Estas situaciones pueden ser especificadas junto a las visualizaciones estáticas tradicionales en la misma fase de diseño. De esta manera, las alertas y avisos le dan al profesor la oportunidad de recibir información inmediata durante la sesión de juego.

Una vez descrito aquello que debe incluir un LAM, se puede ejemplificar utilizando como objeto de estudio un juego serio, una aventura gráfica *point-and-click*, de nuestro grupo de investigación que fue validado con éxito para enseñar técnicas de primeros auxilios (Marchiori, Ferrer, et al., 2012). Este juego fue desarrollado con e-Adventure (Torrente et al., 2010), lo que implica que necesitaba Java¹⁷ para poder ser ejecutado. El ciclo de vida de este juego puede prolongarse sustancialmente si, utilizando su herramienta sucesora uAdventure, desarrollada en Unity¹⁸, se importa el juego en ella y vuelve a ser compilado utilizando las capacidades multiplataforma de la herramienta. Además de importarlo se le incorporaron las analíticas de aprendizaje. Para la creación de este modelo de analíticas se utiliza el vocabulario estandarizado para juegos serios, xAPI-SG, presentado en (Á. Serrano-Laguna et al., 2017), maximizando la reutilización y compatibilidad de las trazas generadas. El modelo de analíticas desarrollado se encuentra descrito en la publicación Perez-Colado, I., Alonso-Fernandez, C., Freire, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2018). **Game learning analytics is not informagic!** *2018 IEEE Global Engineering Education Conference (EDUCON)*, 1729–1737, donde además se expanden muchos de los conceptos y buenas prácticas acerca del modelo de

¹⁷ Java es un lenguaje de programación y una plataforma informática comercializada por Sun Microsystems en 1995 y que, aunque sigue siendo muy utilizado en programación en general en el navegador ha sido sustituido por otros lenguajes (al no soportar los navegadores nativamente los applets java). Sitio web online: <https://www.java.com/>

¹⁸ Unity es un motor de videojuegos con capacidad para generar aplicaciones multiplataforma de manera sencilla. Sitio web online: <https://unity.com/>.

analíticas aquí estudiados. Con este modelo podemos estudiar la efectividad de uso de un modelo de analíticas a la hora de sistematizar el proceso de GLA, y validar el LAM como documentación necesaria que ayude a los desarrolladores a implementar todos los componentes.

Por otra parte, en el anterior capítulo también se menciona la posibilidad de generar un meta-LAM o modelo de analíticas superior que sea capaz de aplicar el proceso antes mencionado, pero sobre un curso *gamificado*, o sobre un juego serio compuesto por múltiples minijuegos. Esto es debido a que, incluso cuando los juegos utilizados como bloques de construcción disponen de un modelo de analíticas, si se usan varios juegos estos modelos fallan a la hora de expresar cómo evaluar el conjunto en su totalidad. Por ejemplo, una característica frecuente que miden las analíticas en un juego serio es el progreso del jugador en el mismo o la completitud del juego. Si se intenta aplicar esta idea al conjunto de varios juegos se pueden plantear escenarios donde las analíticas no representen correctamente el estado del juego porque no exista un modelo capaz de entender el juego serio como un todo y analice los objetivos educativos correctamente. Por ello se propone el uso de un modelo de analíticas para juegos multinivel llamado meta-LAM, que realice la agregación de los modelos de analíticas individuales en un modelo unificado. Por ejemplo, si conocemos el momento del juego donde se lanza un minijuego, podemos representar dicho momento como objetivo del juego asociado al objetivo educativo que enseñe el minijuego. De esta manera podemos mantener ambos, los LAMs independientes de cada juego, y el meta-LAM que describe cómo se conectan entre ellos. Con todo ello, se propone que los meta-LAM, de manera similar a los LAM para los juegos, deben definir el conjunto de objetivos educativos del conjunto completo y cómo éstos se alcanzan mediante los juegos o componentes. Como punto diferenciador con respecto a un LAM estándar, el meta-LAM debe enfatizar la jerarquía y flujo de la información, lo que incluye las agregaciones y análisis que involucren a más de un nivel. Este tipo de estructuras pueden ser representadas como árboles, donde las hojas representan los minijuegos o actividades a realizar con otros recursos (Westera, Nadolski, Hummel, & Wopereis, 2008), pudiendo hacer que la completitud de las hojas contribuya al progreso.

Tras realizar una revisión de los estándares y las especificaciones utilizadas en *e-learning* se decidió utilizar como referencia de soporte la especificación IMS Simple Sequencing,

descrita en la especificación SCORM 2004 4th Edition Sequencing and Navigation¹⁹ (SN). Esta especificación define un árbol de actividades (AT) como la estructura jerárquica principal de las actividades educativas. Este estándar incluye algunas características específicas como el mecanismo de agregación de información llamado *Rollup*, que permite establecer reglas para medir la completitud de las actividades, o añadir pesos a las actividades que componen un nodo del árbol y así establecer cuánto avance supone cuantitativamente completar una actividad hoja al objetivo educativo del que depende. Más detalles acerca del proceso de diseño del meta-LAM, y qué debe componerlo se detalla también en la publicación anteriormente mencionada (**Game learning analytics is not informagic!**), donde se explica el meta-LAM teórico propuesto para el proyecto europeo H2020 BEACONING.

Si bien el meta-LAM apunta a definir los mismos elementos que un modelo de analíticas estándar, puede tener complejidades añadidas si no se tiene acceso al código fuente del juego completo para poder añadir las trazas. Para rebajar estos requisitos técnicos y que siga siendo efectivo, es posible plantear un conjunto de características unificadas que todas las actividades deban compartir con sus nodos superiores en la arquitectura. Por ello, el meta-LAM enfatiza en:

1. Definir los objetivos educativos del juego multi-nivel, relacionándolos con las actividades educativas que lo componen, pudiendo esta relación no ser biyectiva al igual que para un modelo de analíticas estándar.
2. Definir la estructura y jerarquía de las actividades, especificando qué actividades están compuestas por otras. En general una estructura sencilla ayuda al posterior análisis.
3. Definir el modelo de agregación para que el análisis global se pueda generar a partir de los progresos y puntuaciones de los objetivos educativos mediante la agregación de los nodos que los componen. Se pueden definir modificadores como límites mínimos de puntuación que deben alcanzar las actividades para que contribuyan al progreso, o incluso definir un modelo de agregación distribuido

¹⁹ Enlace a la especificación disponible en: ADL. (2009). SCORM 2004 (4th Edition). <http://www.adlnet.org/adl-research/scorm/scorm-2004-4th-edition/>

donde las actividades pueden contribuir a múltiples objetivos educativos en mayor o menor medida.

Este trabajo acerca de analíticas aplicadas a juegos multi-nivel se expone en el artículo Perez-Colado, I. J., Rotaru, D. C., Freire-Moran, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2018). **Multi-level game learning analytics for serious games**. *2018 10th International Conference on Virtual Worlds and Games for Serious Applications, VS-Games 2018 - Proceedings*. Además, se describe el panel de visualizaciones y las modificaciones realizadas sobre el sistema de analíticas para que éste fuese capaz de realizar analíticas multi-nivel.

4.2. Sistematización del proceso de integración de GLA con herramientas de autoría y juegos serios

El LAM y el meta-LAM previamente presentados describen cómo realizar el diseño de un modelo de analíticas, es decir, el modelo para sistematizar el GLA. Pero finalmente es necesario implementar dicho modelo y ponerlo en práctica sobre un sistema de analíticas para poder explotarlo de modo que todos los participantes del proceso educativo puedan beneficiarse de las ventajas de las analíticas. Esta implementación es un proceso complejo, frágil y propenso a fallos, que requiere la participación de expertos multidisciplinares además de implicar altos costes económicos. En el epígrafe anterior se presentó el desarrollo para generar analíticas sobre un juego ya existente (juego de primeros auxilios), no sólo con el objetivo de alargar su ciclo de vida, sino con el propósito de estudiar cómo ayuda un LAM a la sistematización del proceso de analíticas. Para abordar los aspectos de la implementación, primero se debe entender cómo funciona el proceso de analíticas de manera simplificada.

En la **Figura 6** se muestra el juego, con un *tracker* que envía los datos, y el sistema de analíticas que almacena, analiza, y procesa las trazas generadas hasta mostrarlas en un panel de visualizaciones accesible para el profesor. En esta primera fase, únicamente se realizan las modificaciones necesarias para que el juego sea capaz de comunicar los eventos que suceden al sistema de analíticas, sin tener en cuenta al resto de componentes del sistema de analíticas (*Learning Analytics System, LAS*).

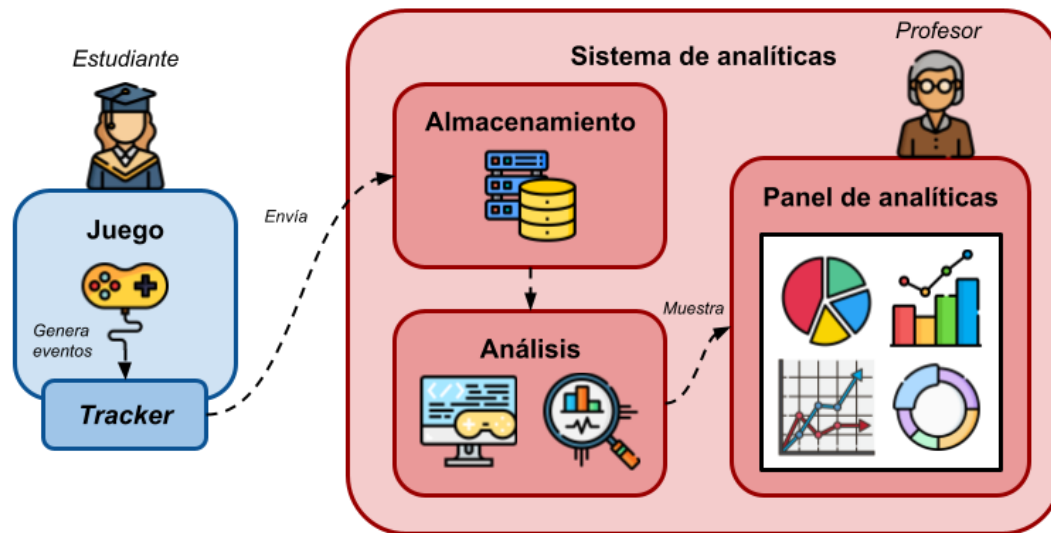


Figura 6. Representación del proceso de ejecución de GLA. El estudiante comienza jugando al juego que incluye un componente de seguimiento, llamado Tracker, que envía los datos de interacción al sistema de analíticas, donde se almacena, analiza, y prepara para su posterior visualización por profesores u otros usuarios para extraer conclusiones.

Debido a que el juego de primeros auxilios se encuentra altamente ligado a uAdventure, para su implementación se plantean dos alternativas: (a) implementar únicamente el código necesario de manera *ad-hoc*, o (b) generalizar el modelo y modificar el editor de uAdventure para poder incluir los requisitos del modelo de analíticas. Aunque la primera opción es más sencilla, mediante la generalización del modelo y añadiendo los editores se puede conseguir sistematizar completamente el proceso de analíticas para todos los juegos serios generados con uAdventure. Añadir esta característica mejora su ciclo de vida, y haciendo que, de manera sencilla y gratuita, aquellos juegos generados con uAdventure sean capaces de generar analíticas de manera automática. Esto además ayuda a incluir el LAM en el proceso de desarrollo gracias a esta integración en la herramienta de autoría. Para la generalización del modelo creado para el juego de primeros auxilios, se realizó un estudio sobre cómo extraer todos aquellos eventos comunes aplicables a un amplio rango de juegos del género de aventura gráfica *point-and-click*, y como se reflejaban en el modelo de datos de uAdventure. Gracias a este modelo genérico, en fases posteriores del desarrollo se pueden crear visualizaciones genéricas reutilizables. Para poder representar los datos de forma genérica, éstos deben de cumplir un lenguaje común y estar estandarizados. Por ejemplo, si una visualización tiene que medir la puntuación de una prueba, los datos deben estar representados siempre en el mismo formato, ya que la puntuación podría medirse de distintas formas, por ejemplo, numéricamente de 1 al 10 o

alfabéticamente desde la F hasta la A. En (Á. Serrano-Laguna et al., 2017) se presenta un modelo y un vocabulario específico que utiliza xAPI, llamado xAPI-SG, para la comunicación y estandarización del proceso de analíticas específico para juegos serios. Este modelo permite, entre otras cosas, la reutilización y comunicación de elementos que constituyen el sistema de analíticas mediante un lenguaje común. Para el desarrollo e implementación del modelo genérico de analíticas en uAdventure se ha utilizado xAPI-SG. El trabajo realizado sobre uAdventure para alcanzar la meta de poder generar juegos que incorporen GLA automáticamente se presenta en el artículo Pérez-Colado, V. M., Pérez-Colado, I. J., Freire-Morán, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2019). **UAdventure: Simplifying narrative serious games development**. *Proceedings - IEEE 19th International Conference on Advanced Learning Technologies, ICALT 2019*, 119–123. A continuación, se detallan todos los elementos importantes del modelo genérico, incluyendo una tabla con todos los eventos de interacción generados, así como capturas de los editores añadidos a uAdventure y su propósito.

Para poder generalizar el modelo, éste tiene que ser capaz de aplicarse a un amplio rango de aventuras *point-and-click* que típicamente se componen de múltiples escenas interconectadas por una narrativa donde el jugador interactúa con los personajes o elementos de la escena mediante *clicks* de ratón para avanzar en la narrativa. El objetivo es que todos los eventos significativos para GLA que ocurren dentro del juego puedan ser descritos con xAPI y su vocabulario de juegos serios. En resumen, dicho vocabulario define: cómo localizar la posición del jugador en el juego mediante *Accesibles*, cómo categorizar los elementos de una escena mediante *GameObjects*, qué elecciones toma el jugador mediante *Alternatives*, y cómo medir el progreso del juego mediante *Completables*. Todos los eventos que genera uAdventure se describen en la **Figura 8**, donde se incluyen además los once tipos de interacción que se registran de manera automática en todos los juegos generados con uAdventure.

No obstante, uAdventure no disponía de editores para poder incluir todas las características que requiere el modelo de analíticas. Por ejemplo, cómo se indica si una respuesta es correcta o incorrecta, o qué progreso ha alcanzado el alumno en cada punto del juego. Por lo que ha sido necesario implementar una serie de editores que permiten añadir este tipo de datos al modelo de uAdventure.

Event	Cause	xAPI Type xAPI Verb	Target	Result R: response, S: success, Ext: extensions
NPC Interaction	Player opens NPC actions menu	Character Interacted	NPC name	Ext: Action name
Item Interaction	Player opens item actions menu	Item Interacted	Item name	Ext: Action name
Scene access	Player enters a scene	Accessible Accessed	Scene id	
Cutscene start	Player starts a cutscene	Cutscene Accessed	Cutscene id	
Cutscene skip	Player presses skip	Cutscene Skipped	Cutscene id	Ext: Percent watched
Exit selection in alternative type scenes	Player selects an exit in current scene, for menus or visual choices	(Alternative, Question Menu or Path) Selected	Exit Id	R: Arriving scene S: Based on exit conditions
Dialog choice	Player selects one dialog option	Alternative Selected	Question Id	R: Response S: Correctness
Task start	Player reaches a milestone	Completable Started	Completable Id	
Task progress	Player reaches one of the milestones	Completable Progressed	Completable Id	Ext: Milestone progress value
Task finish	Player reaches a milestone or completes all the steps	Completable Completed	Completable Id	R: Score from variable S: Based on conditions Ext: Time
Game start	Player visits title	Game Started	Game name	
Game progress	Accomplishment of any of the levels	Game Progressed	Game name	Ext: Progress as percent of levels (tasks) completed
Game end	Milestone or all levels completed	Game Completed	Game name	R: Avg. score of all levels S: Based on conditions Ext: Time

Figura 8. Recopilatorio de todos los eventos que genera la integración de GLA implementada en la herramienta de autoría de juegos serios uAdventure.

Para dar soporte a las analíticas de aprendizaje mediante las trazas generadas en el juego, se han desarrollado una serie de editores que permiten insertar metadatos de analíticas como, por ejemplo, identificadores asociados a personajes, objetos, escenas o preguntas. También se ha incluido un editor que permite definir un modelo basado en hitos para poder medir el progreso del jugador. El primer editor presentado en la **Figura 7** permite modificar la definición de cualquier elemento del juego de manera simbólica. En ella se representa una escena con cuatro alternativas gráficas a elegir, siendo formalmente equivalente una pregunta con diferentes respuestas.

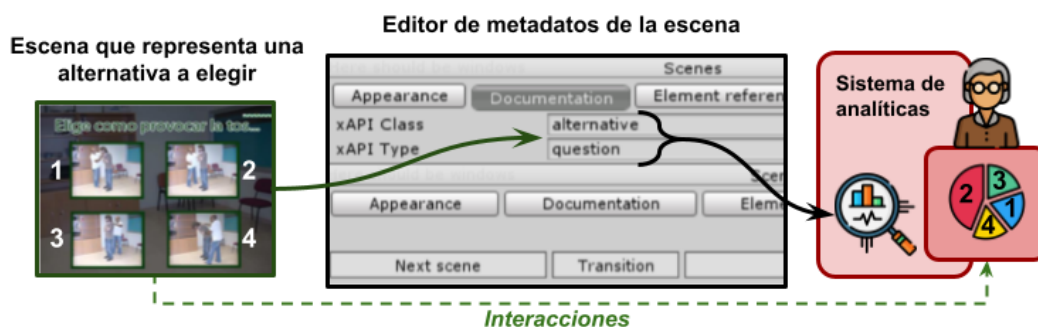


Figura 7. Proceso de aplicación del modelo de analíticas utilizando el editor de uAdventure donde se especifica que una escena se compone de múltiples alternativas a seleccionar, y cómo generar analíticas con ella.

Otro editor desarrollado permite añadir información sobre las preguntas y respuestas que se realizan en el juego. Como se muestra en la **Figura 9**, dicho editor permite establecer un identificador único de pregunta para que el sistema de analíticas pueda categorizarla, además de incluir una casilla marcada para indicar si la opción es correcta o incorrecta. Con esta información, el análisis puede identificar fácilmente el grado de acierto de los alumnos en las preguntas, y evaluar así el conocimiento de los alumnos.

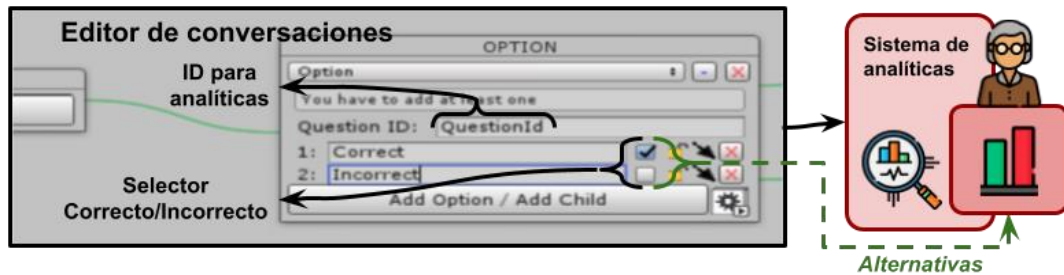


Figura 9. Editor de uAdventure para especificar qué preguntas son correctas e incorrectas, pudiendo facilitar dichos datos al sistema de analíticas para su análisis y visualización de los resultados.

El último editor a mostrar permite definir un modelo de progreso basado en hitos que se cumplen cuando en el juego se dan diversas condiciones o situaciones como, por ejemplo, que ha visitado un escenario, o que un objeto significativo ha sido recogido. Mediante este editor genérico, se puede dividir el juego en pequeñas tareas o bloques, que además pueden tener una puntuación y progreso asociados para su evaluación. Esto incrementa no sólo el control del estado de completitud del juego en cada momento, sino que además da información extra sobre cómo de eficaces están siendo los alumnos a la hora de resolver los problemas y aplicar el conocimiento aprendido. En el editor de la **Figura 10**

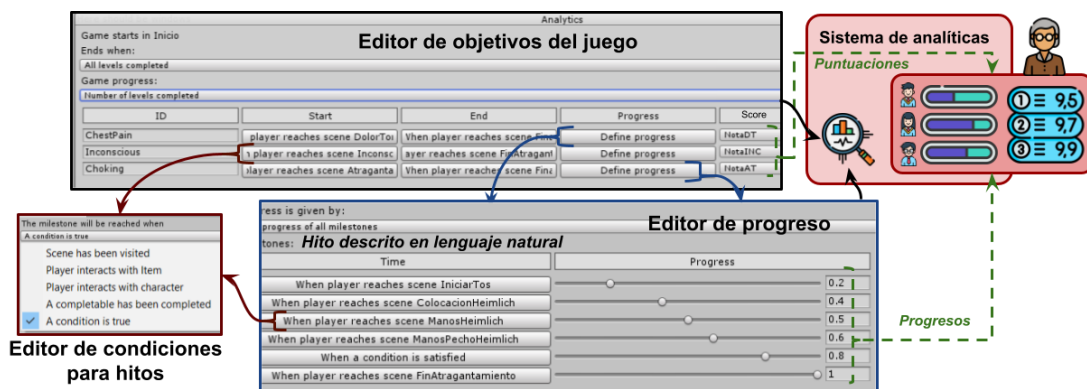


Figura 10. Conjunto de editores de uAdventure para especificar *completables* en forma de tareas e hitos para dividir el juego permitiendo crear un modelo de progreso, además de permitir asociar puntuaciones a cada tarea.

se muestra el proceso de creación de *completables* para el juego, pudiendo crear el modelo de progreso con el que generar las trazas para que el sistema de analíticas conozca tanto el progreso de los alumnos en el juego como sus puntuaciones.

Con todos estos editores, además de algunos otros incluidos y detallados en otras publicaciones de uAdventure (Perez-Colado, Perez-Colado, Martinez-Ortiz, et al., 2017) se realizaron las modificaciones sobre el juego de primeros auxilios. Esto permitió volver a compilarlo, habiendo incluido en el juego todos los requisitos especificados en el LAM, e incluyendo un *tracker* capaz de enviar los datos generados al sistema de analíticas. Esto lleva al proyecto al siguiente paso, donde se tienen que realizar las modificaciones necesarias en el LAS para que éste sea capaz de entender, analizar y mostrar los datos. Estas modificaciones se incluyen como parte del trabajo presentado en Perez-Colado, I., Alonso-Fernandez, C., Freire, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2018). **Game learning analytics is not informagic!** *2018 IEEE Global Engineering Education Conference (EDUCON)*, 1729–1737, incluyendo capturas de algunas de las visualizaciones generadas.

El sistema de analíticas utilizado para procesar las trazas generadas por los juegos es el desarrollado para el proyecto H2020 RAGE Analytics, al que posteriormente se han realizado aportaciones y mejoras. Este LAS se caracteriza por estar enfocado a los juegos serios y ser capaz de trabajar con trazas en formato xAPI, pero quizás su parte más distintiva es la de incluir características específicas para analíticas en tiempo real. Sin embargo, el sistema no incluía soporte para el vocabulario xAPI-SG de modo que no disponía de análisis por defecto, por ejemplo, para poder saber si los objetivos educativos se han satisfecho o no. A través del análisis del LAM del juego de primeros auxilios, y su generalización e integración en uAdventure, se definen una serie de análisis y visualizaciones a desarrollar para poder realizar GLA. Éstas incluyen las presentadas en la **Figura 11** donde se muestran nueve visualizaciones genéricas que construyen un panel con el que identificar los objetivos educativos y evaluar a los participantes. Por ejemplo, la visualización A nos permite saber el grado de acierto de un participante al aplicar el conocimiento y la visualización G muestra las puntuaciones conseguidas en cada tarea. Además de los objetivos educativos, estas visualizaciones nos permiten medir tiempos medios por tarea con la visualización H, identificar el grado de progreso con las visualizaciones D, E, e I.

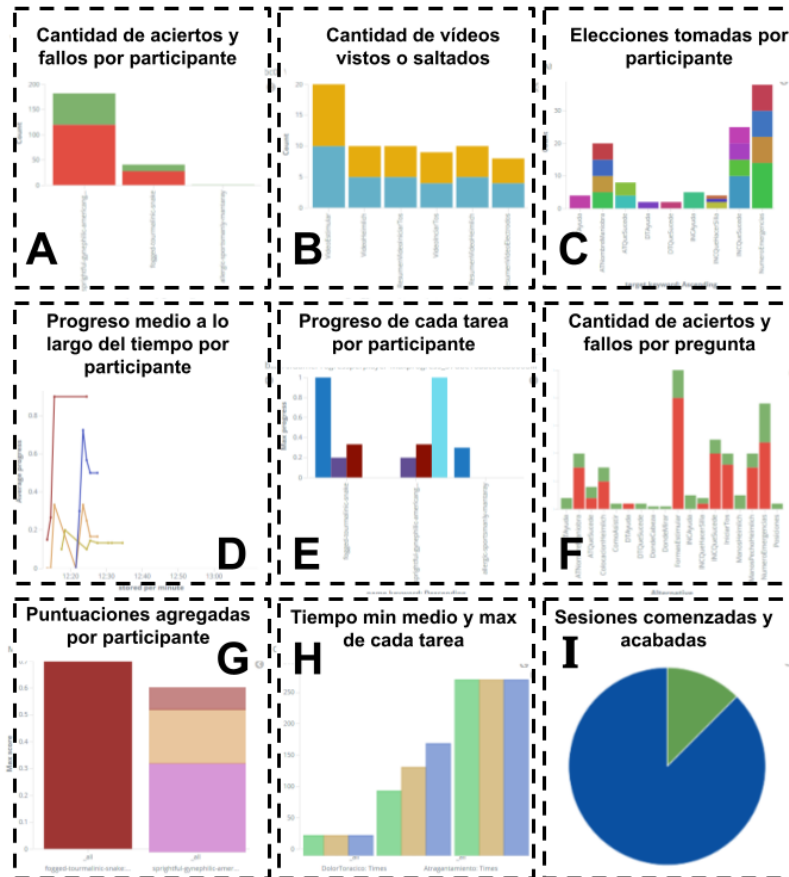


Figura 11. Conjunto de visualizaciones genéricas que forman un panel para los juegos creados con la herramienta de autoría de juegos serios uAdventure. Este panel se incluye automáticamente sin necesidad de programar.

Como resultado, se ha conseguido automatizar el proceso de analíticas para aquellos juegos que se generen con uAdventure. Para lograrlo se ha ampliado uAdventure aplicando el modelo genérico de analíticas para que genere eventos de interacción, se han incluido los editores para poder añadir metadatos relacionados con analíticas al modelo, y se ha conectado mediante un *tracker* con un sistema de analíticas y, finalmente, se ha desarrollado un panel genérico reutilizable con un conjunto de visualizaciones genéricas. Mediante el uso de un proceso sistemático, centrado en el uso de un modelo de analíticas, la herramienta de autoría junto con el sistema de analíticas son capaces de ofrecer resultados en tiempo real para evaluar un juego serio. Aunque la opción de tiempo real es más compleja de desplegar y es más frecuente que el análisis se haga con posterioridad al juego.

Las visualizaciones genéricas pueden ser una herramienta muy útil, ya que son capaces de mostrar información, como tiempos y puntuaciones a la hora de completar tareas que

pueden ser comparadas con valores de referencia para medir el rendimiento de los estudiantes. Además, al ser genéricas, suponen un coste muy bajo de implementación al no tener que desarrollarse (a lo sumo configurarlas en los editores). Sin embargo, las visualizaciones genéricas no son muy habituales, ya que la mayoría de sistemas mostrados y analizados en revisiones de la literatura (Alonso-Fernández et al., 2019; Muslim et al., 2020; Samuelsen et al., 2019) utilizan sistemas desarrollados *ad-hoc*, debido a que los sistemas a medida de cada juego permiten un mayor nivel de detalle en la observación y entendimiento del proceso de juego que ayuda a la extracción de conclusiones. El uso de un LAM ayuda a sistematizar una implementación genérica de analíticas para una herramienta de autoría, consiguiendo dar soporte automático a todos los juegos creados con dicha herramienta. No obstante, si sólo pudiera usarse conjuntamente con uAdventure su aplicación sería muy limitada. A continuación, se evalúa la sistematización de las GLA para implementaciones *ad-hoc* utilizando el mismo proceso basado en el LAM.

Se presentan tres implementaciones *ad-hoc*, en colaboración con sus autores originales quienes marcan las necesidades educativas guiando el diseño del LAM y su implementación, realizadas utilizando el procedimiento sistematizado para la integración e implementación de GLA. Son proyectos desarrollados en el grupo e-UCM completamente o mediante colaboración con otros grupos en proyectos europeos: (a) el juego *Conectado* (Calvo-Morata, Freire, et al., 2018) que busca concienciar a los alumnos acerca del *cyberbullying*, (b) el juego *Water Cooler* (Hollins, Humphreys, Yuan, Sleightholme, & Kickmeier-Rust, 2017) que utiliza un modelo de analíticas basado en el modelo de resolución de conflictos Thomas-Kilmann (Thomas & Kilmann, 1976) y, (c) el juego *FormalZ* (Prasetya et al., 2019) que enseña especificaciones formales con un juego tipo *tower-defense*. En todos los casos, el desarrollo e implementación del juego serio ha sido realizado por equipos externos, o al menos de manera separada al desarrollo del sistema de analíticas. Para cada caso se ha diseñado un LAM en colaboración con los autores originales, distribuido a los diferentes miembros del equipo con el objetivo de sistematizar y coordinar las partes del desarrollo de GLA. En todos ellos se especifica el uso de xAPI como estándar para el formato de las trazas, así como xAPI-SG como vocabulario específico. Para la implementación del LAS se reutiliza RAGE Analytics, plataforma sobre la que se continúan realizando mejoras. Los modelos de analíticas generados para los tres juegos serios, así como imágenes y descripciones de los paneles

de analíticas desarrollados en base a los requisitos necesarios para validar los objetivos educativos de cada juego se presentan como parte del trabajo publicado en Calvo-Morata, A., Alonso-Fernández, C., Pérez-Colado, I. J., Freire-Morán, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2019). **Improving Teacher Game Learning Analytics Dashboards through ad-hoc Development.** *Journal of Universal Computer Science*, 25(12), 1507–1530. En la Figura 12 se muestran los tres juegos antes mencionados, junto a los paneles que se detallan en la publicación.

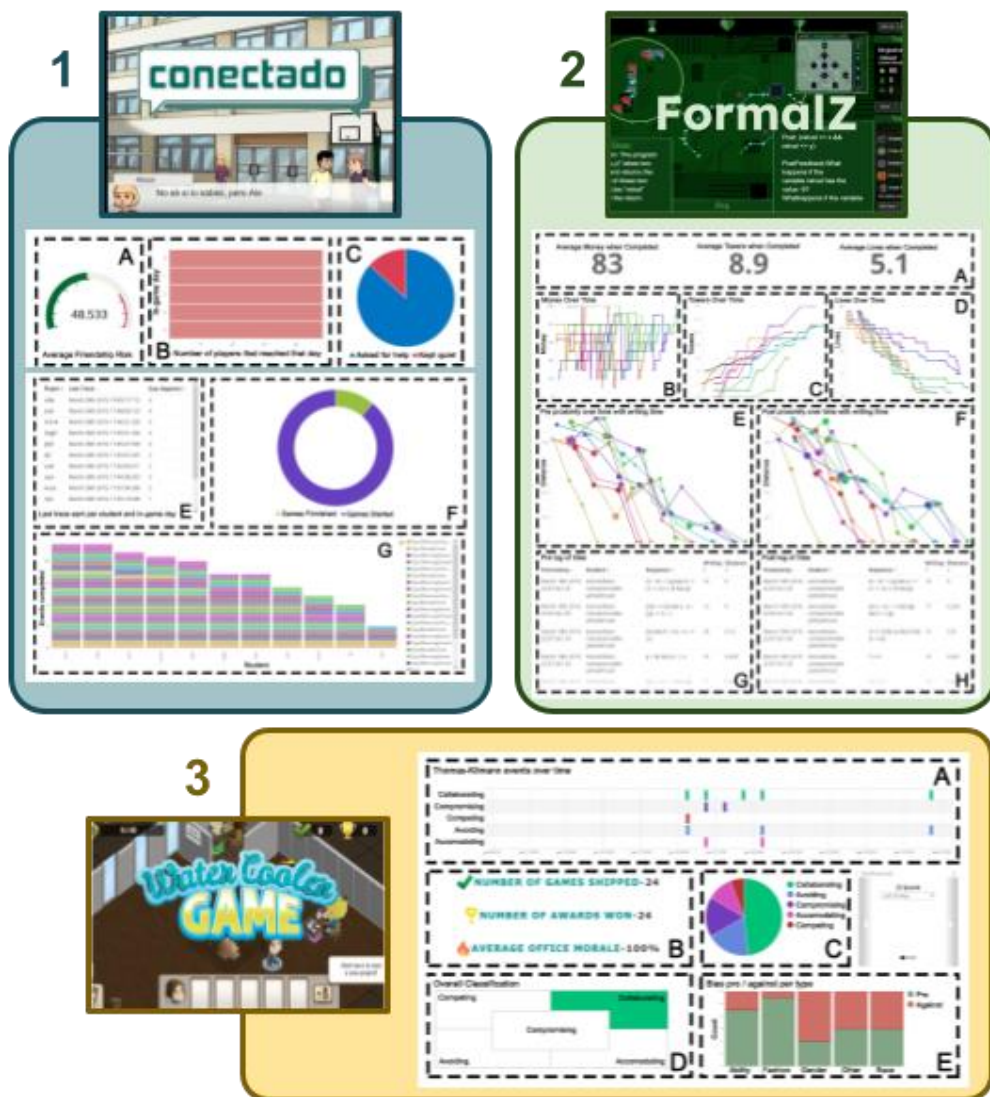


Figura 12. Paneles de visualizaciones ad-hoc desarrollados para (1) el juego Conectado, (2) el juego FormalZ, y (3) el juego WaterCooler que aplica el modelo Thomas-Kilmann.

4.3. Desarrollar una plataforma de GLA orientada a la validación de Juegos Serios

La información sobre el proceso de aprendizaje que se consigue al aplicar GLA ha demostrado ayudar en distintas fases del ciclo de vida del juego. Por un lado, a educadores para aplicar juegos serios en actividades educativas pudiendo controlar y evaluar la sesión, como a investigadores para mejorar el juego analizando los datos recogidos y utilizarlos para dar soporte a su proceso de validación. El GLA contribuye a la adopción de juegos serios al transformar un modelo de caja negra en una fuente de información.

Asimismo, una validación formal que demuestre que el juego serio cumple sus propósitos educativos contribuye a que los educadores puedan adoptar y utilizar los juegos confiando en su eficacia. En esta sección se presentan las contribuciones realizadas para simplificar el proceso de validación de los juegos serios. Para alcanzar la solución final, se han realizado varias iteraciones sobre el proyecto que ha permitido refinar el producto final, mejorando la flexibilidad de la herramienta a la hora de adaptarse a múltiples diseños experimentales

4.3.1. Prueba de concepto y prototipo: SurveyManager

Actualmente, la forma más aceptada para validar juegos serios consiste en el uso de cuestionarios externos para los jugadores. Comúnmente involucra dos cuestionarios: un cuestionario previo antes de realizar el juego, para medir las características iniciales del participante, y un cuestionario posterior a la sesión de juego, usado para medir los cambios que se producen durante el juego (aunque ambos pueden solicitar otra información de utilidad como, por ejemplo, demográfica). Si se identifica un cambio estadísticamente relevante, la validación formal se considerará exitosa. Este diseño experimental debe tener en cuenta también otros aspectos como, por ejemplo, la protección de datos o la privacidad de los participantes (sobre todo en el caso de ser menores de edad o jugadores con algún tipo de limitación cognitiva).

El objetivo es automatizar este proceso para simplificarlo, creando una herramienta capaz de gestionar actividades para la realización de experimentos basados en encuestas pre-post. Una opción es crear un portal donde los investigadores puedan acceder, gestionar y crear estas actividades mediante un simple formulario con tres elementos: 1) las encuestas anterior y posterior codificadas en algún lenguaje de marcado compatible, 2) el

número de participantes, para los que se generan identificadores de acceso anonimizados, o tokens, y 3) el juego objeto del experimento, para poder agrupar los datos de todas las actividades asociándose al mismo caso de estudio. Finalmente, el sistema debe estar integrado con alguna plataforma de GLA para facilitar la recolección de datos, e incluir de manera adicional, un almacenamiento *backup* donde los juegos puedan guardar datos del proceso de juego de manera independiente al sistema de analíticas (para evitar pérdidas de datos ya que los procesos experimentales son muy caros y complejos, de modo que si se produce un error en muchos casos no sería posible repetir el experimento).

Con los requisitos citados se ha desarrollado una herramienta prototipo, con nombre SurveyManager (SM), disponible con licencia de código libre en el repositorio: <https://github.com/Synpheros/SurveyManager>. Esta herramienta, integra y utiliza la plataforma de gestión y uso de encuestas online LimeSurvey²⁰ (LS), la cual dispone de un API llamada *RemoteControl* con la que automatizar su manejo, pudiendo crear o borrar encuestas, añadir participantes o descargar los resultados. Por otra parte, SM facilita interfaces web para gestionar las actividades del diseño experimental, con las fases de encuesta y juego, incluyendo la gestión de las encuestas de LS de manera simplificada y el almacenamiento para las analíticas de los juegos dentro de ellas. También controla el estado y secuenciación del experimento, para que los juegos puedan limitar el acceso al juego si no se ha realizado la encuesta, y así asegurar el desarrollo correcto del experimento.

La **Figura 13**, donde se muestran varias de las funcionalidades descritas anteriormente. Además, describe la funcionalidad de crear un PDF imprimible con códigos (A) con el que ayudar al desarrollo presencial del experimento, que permita repartir los identificadores anónimos a cada uno de los participantes, asociándose con su nombre. Mediante el uso del formato físico se aíslan los datos que permiten identificar a la persona, manteniendo la anonimización de los datos digitales para los investigadores, pero permitiendo a profesores realizar la evaluación y soporte en la sesión (pseudo-anonimización en origen). También se muestra la interfaz web desarrollada para representar una actividad (B) donde se asocian estos *tokens* al estado de las encuestas, junto a enlaces para descargar las respuestas. Para la creación de las actividades, se

²⁰ Plataforma para la gestión y uso de encuestas online disponible en el sitio web: <https://www.limesurvey.org/>

configura un número de participantes y una pareja de cuestionarios en formato *lss* compatible con LimeSurvey. Finalmente, en la figura se muestra en la columna de la derecha los registros de actividad generados por el juego, representados como *traces*, los cuales son almacenados como copia de seguridad en caso de que la conexión con el sistema de analíticas no se realice correctamente. Estos datos son descargables por los investigadores.

Figure 13 illustrates the SurveyManager interface. Part (A) shows a PDF table for 'Clase EICaton 1A' with columns for 'No.', 'Nombre', and 'Código'. The 'Código' column contains tokens like GYRJ, WEFF, YEYT, ZMBL, WSFJ, MLBT, KGMV, IBAH, MPGH, XZGZ, DQRA, VVEU, BLCN, TTTQ, and KWL. Below the table are icons for participants GYRJ and ZMBL, and a 'LimeSurvey' logo. Part (B) shows a screenshot of the SurveyManager interface for 'Clase EICaton 1A'. It displays a list of surveys with columns for 'Code', 'Conectado (PrePost)', and 'Traces'. The 'Traces' column contains links for downloading results. Below the list are icons for 'Pre', 'Post', and 'Datos' (Data).

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
6		MLBT MLBT MLBT MLBT
7		KGMV KGMV KGMV KGMV
8		IBAH IBAH IBAH IBAH
9		MPGH MPGH MPGH MPGH
10		XZGZ XZGZ XZGZ XZGZ
11		DQRA DQRA DQRA DQRA
12		VVEU VVEU VVEU VVEU
13		BLCN BLCN BLCN BLCN
14		TTTQ TTTQ TTTQ TTTQ
15		KWL KWL KWL KWL

Figura 13. La herramienta SurveyManager, mostrando: (A) un PDF imprimible para el usuario donde anotar los nombres de los participantes, asociados a cada token único recordable para su fácil distribución durante el experimento, y (B) captura de pantalla de la interfaz de una actividad ya creada, mostrando las diferentes encuestas, su estado, y los ficheros de trazas del juego recogidos, pudiendo descargar todos los resultados.

Asimismo, SurveyManager no solo incluye soporte para las encuestas, sino que también se ha integrado el LAS anteriormente utilizado en RAGE Analytics. Por tanto, dispone de soporte del LAM genérico previamente desarrollado de modo que se automatiza el proceso de analíticas de la actividad experimental a realizar. Esta implementación actúa como puente que prepara el LAS, crea y gestiona los usuarios para que puedan mandar sus datos de analíticas al LAS, y crea un código de seguimiento para distribuirlo con los investigadores que lo integran con el juego (el juego se particulariza para su uso en este diseño experimental).

Con este prototipo, se han realizado una serie de pruebas de usabilidad en el grupo de investigación (Alonso-Fernandez, 2017). En concreto, se ha integrado el juego de primeros auxilios y se ha revalidado en colegios. Mediante estas pruebas se identificaron una serie de carencias para generar la nueva versión mejorada de la arquitectura y la plataforma. Estas carencias son las siguientes:

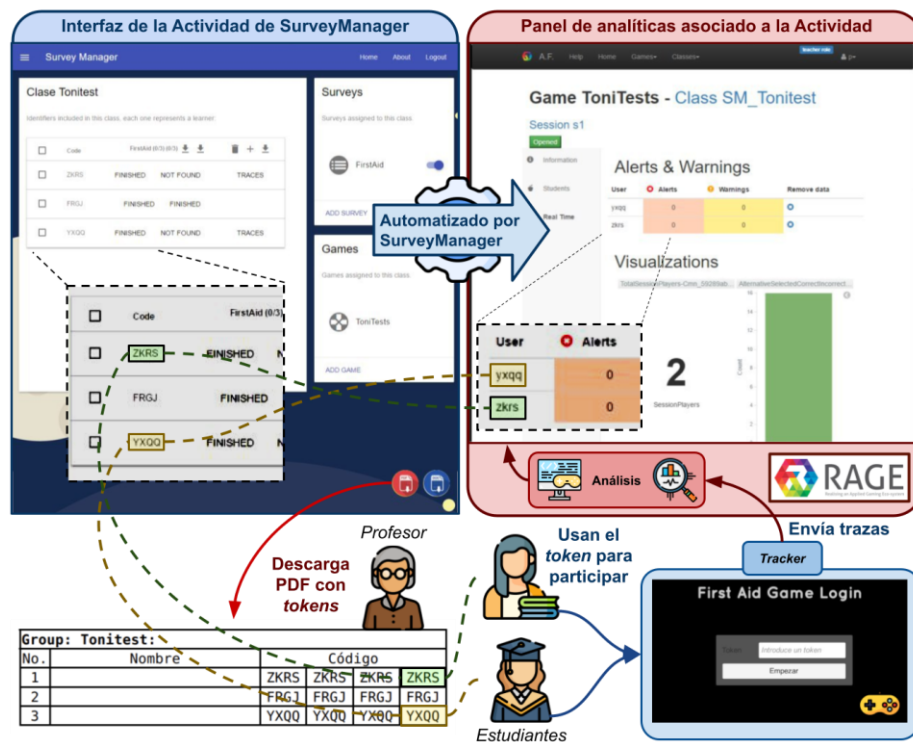


Figura 14. Proceso de GLA a través de SurveyManager, incluyendo al juego, al sistema de analíticas, y cómo profesor y estudiantes interactúan con el sistema durante el desarrollo del experimento.

- El sistema de participantes y *tokens* creado, pese a permitir identificar al usuario, es poco seguro pues no presenta ningún nivel de autenticación, dificulta la reutilización de usuarios y grupos entre experimentos al encontrarse dentro de la propia actividad. Esta funcionalidad se reemplaza por un sistema de gestión de usuarios completo en la siguiente versión.
- El sistema de actividades es limitado pues solo permite un único caso de diseño experimental (pre-test, juego, post-test), limitando la opción de crear un grupo de control, realizar pruebas con ramas (*A/B testing*), o integrar otros tipos de actividades de distinta complejidad. Por ejemplo, una actividad que consista en leer un recurso educativo para validarlo, simplificando el concepto de actividad con objetivo de utilizarlas como bloques de construcción de una estructura superior, denominada estudio.
- El sistema de ejecución del experimento necesita una implementación *ad-hoc* para su correcto funcionamiento. Si bien se remarca que, por ejemplo, no permite empezar a jugar a un juego sin haber realizado la encuesta previa, es labor del juego consultar a SurveyManager. Se reemplaza esta funcionalidad, añadiendo

además un portal web que se encargue de comprobar el estado y redirigir automáticamente al usuario a través del desarrollo del experimento.

- El sistema debe ser fácilmente instalable. Debido al carácter de prototipo del proyecto, el proyecto depende de otros componentes como el sistema de analíticas y una base de datos. El mantenimiento y ejecución de estos sistemas no deben suponer un quebradero de cabeza para el investigador, automatizando el despliegue de la infraestructura al completo, integrando todos los componentes necesarios.

Con el cierre del proyecto SurveyManager se completa el prototipo, y la primera iteración del proyecto para el desarrollo de una plataforma de GLA orientada a la validación. En siguientes secciones se aborda el desarrollo de la versión final, con detalles de implementación, así como los cambios y mejoras que ha sufrido respecto a SurveyManager.

4.3.2. Plataforma de GLA y validación: Simva

La herramienta prototipo SurveyManager, que surge para simplificar la validación mediante encuestas previa y posterior, y su análisis comparativo, aunque ha resultado un éxito que cumple con las expectativas iniciales, también presentaba una serie de carencias. No tenía un sistema de gestión de usuarios y grupos, y no disponía de la flexibilidad para aplicar diferentes diseños experimentales que no utilicen el ya mencionado *pre-post*. Se planteo generalizar la solución haciéndola mucho más potente y flexible y dio lugar a Simva, proveniente del acrónimo de *Simple Validator*. En esta iteración del proyecto se han identificado los siguientes requisitos finales para la plataforma Simva.

- Sistema para gestionar estudios. Los estudios engloban la gestión e implementación del diseño experimental, pudiendo incluir diferentes ramas de prueba (e.g. grupo de control, *A/B testing*), y se construyen utilizando como bloques básicos las diferentes actividades que lo componen. Este sistema también permite administrar la organización de los participantes en el estudio, incluyendo un sistema que permita asignar a los participantes a las diferentes ramas de estudio (aleatoriamente o de algún otro modo que considere el investigador).
- Sistema de actividades flexible y desacoplado. Las actividades representan las tareas a realizar en las etapas de la evaluación, pudiendo incluir juegos,

cuestionarios, o actividades online. Es necesario que las actividades automaticen y simplifiquen la configuración de aquellas herramientas externas que sean necesarias, con una arquitectura que separe la funcionalidad de cada actividad del núcleo de la herramienta. Esto permite adaptar mejor el comportamiento de los estudios a las actividades que lo componen y sus requisitos individuales, lo que mejora las capacidades y facilita el uso de la herramienta para investigadores.

- El sistema de actividades debe incluir las siguientes actividades obligatorias: (1) Actividad de *Gameplay*, que supone el centro del experimento, direcciona al usuario al juego, habilita para el profesor paneles automáticos para la explotación de *Game Learning Analytics* y dispone un almacén datos (en vocabulario xAPI, un *Learning Record Store* -LRS) para el juego serio donde almacenar las trazas. (2) Actividad de Encuesta, que gestione las encuestas y maneje LimeSurvey, realizando el despliegue, la gestión de participantes, el control del estado y la obtención de las respuestas de manera automática. (3) Actividad *online*, que controla el investigador de manera manual, permitiendo validar otro tipo de recursos educativos como presentaciones online o vídeos con lecciones que se deseen validar.
- Sistema de gestión de usuarios y grupos. Se añade un sistema para el manejo flexible de usuarios, incluyendo participantes y administradores. Se añade un punto de acceso unificado que requiera autenticación y autorización, centrado en la anonimización de los participantes, así como la creación y reutilización de grupos. Se incluyen dos funcionalidades específicas: creación de usuarios con identificadores anónimos que permitan cumplir con la normativa europea de privacidad y protección de datos GDPR, así como descargar un PDF imprimible con los códigos de anonimización que facilite el desarrollo de la sesión experimental presencial.
- Sistema de ejecución del experimento. Con todo definido, es necesario facilitar herramientas para guiar a los participantes a través del flujo del estudio, que permita controlar su estado, limitando a los participantes para evitar que realicen actividades antes de tiempo sin completar las previas. Este sistema incluye un API donde los distintos servicios consultan y utilizan el plan generado por Simva, así como una interfaz Web que, con un enlace de participación y un token permite ejecutar todo el experimento a través de un navegador.

La plataforma Simva de Game Learning Analytics orientada a la validación desarrollada se presenta en el artículo Perez-Colado, I. J., Calvo-Morata, A., Alonso-Fernández, C., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2019). **Simva: Simplifying the scientific validation of serious games**. *Proceedings - IEEE 19th International Conference on Advanced Learning Technologies, ICAALT 2019*, 113–115. A continuación, se amplía el trabajo presentado en la publicación para mostrar con mayor detalle el trabajo realizado, las contribuciones y las capacidades de la herramienta desarrollada.

4.3.2.1. *Sistema de gestión de usuarios y grupos.*

Simva incluye un sistema de gestión de usuarios y grupos orientado a la realización de experimentos. Implementa las funcionalidades básicas de un sistema de gestión de usuarios y grupos, pero añadiendo otras características que facilitan la preparación y desarrollo de experimentos. Simva mejora dos funcionalidades ya existentes en el prototipo como son la anonimización de usuarios y la generación automática de PDFs imprimibles. Y además añade un sistema de gestión de propietarios para los grupos, pudiendo compartir grupos con otros investigadores y así tener acceso a los estudios relacionados. El investigador también puede añadir manualmente nuevos usuarios con un nombre de usuario, rol y correo electrónico, permitiendo además añadir usuarios existentes, como *tokens* de otros grupos ya existentes, o investigadores que deseen participar con su propio usuario. Finalmente, este sistema permite añadir un lote de nuevos usuarios aleatorios. Esto se representa en la **Figura 15** donde se muestra un grupo en Simva (B), incluyendo formulario para añadir el lote de participantes (A), el grupo creado con dicho formulario, y el PDF imprimible (C). Este sistema de gestión de usuarios

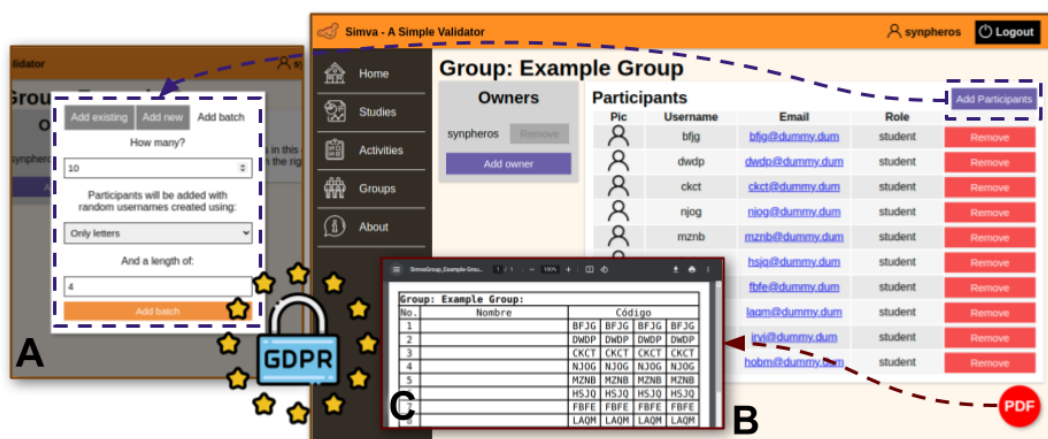


Figura 15. Sistema de gestión de usuarios y grupos en Simva (B), mostrando las capacidades integradas para la creación de lotes de usuarios pseudoanonimizados con un simple formulario (A), junto a un PDF imprimible que representa el grupo creado (C).

presenta grupos desacoplados de los estudios, reutilizables y permitiendo compartirlos entre investigadores, así como facilita la autenticación y autorización de los participantes mediante un sistema de autenticación unificado (*single sign-on* o SSO).

4.3.2.2. Sistema de gestión de estudios.

El elemento principal del diseño experimental en Simva son los estudios que representan un experimento para realizar una validación. Están compuestos por diferentes ramas de prueba, o *test*, entre las que se distribuyen a los participantes. Una rama de prueba se compone por una secuencia de actividades a realizar. Estas actividades son unidades independientes y atómicas, y se utilizarán como bloques de construcción en las ramas de prueba pudiendo estructurar una rama acorde a las necesidades del estudio concreto.

Las ramas de estudio permiten la creación de casos de estudio más completos que permitan contrastar los resultados de validación. Por ejemplo, que utilicen un grupo de control, un esquema alternativo de validación (*A/B testing*), un análisis cruzado, o estudios con más ramas en las que distribuir a los participantes. Cada rama representa la ruta que realizará cada participante. Éste suponía uno de los puntos más limitantes del prototipo SurveyManager, donde el diseño basado en encuestas *pre-post* era la única configuración disponible para los estudios. Además, ahora se permiten añadir nuevas ramas de estudio para casos en los que existan nuevas versiones del juego o se usen encuestas diferentes. Esto permite mantener toda la información unificada en un mismo estudio, pero diferenciando los resultados de cada versión y rama, a la vez que posibilita una flexibilidad total para controlar las diferentes actividades de cada rama de prueba del estudio. En la **Figura 16** se presenta un ejemplo más complejo con dos ramas distintas.

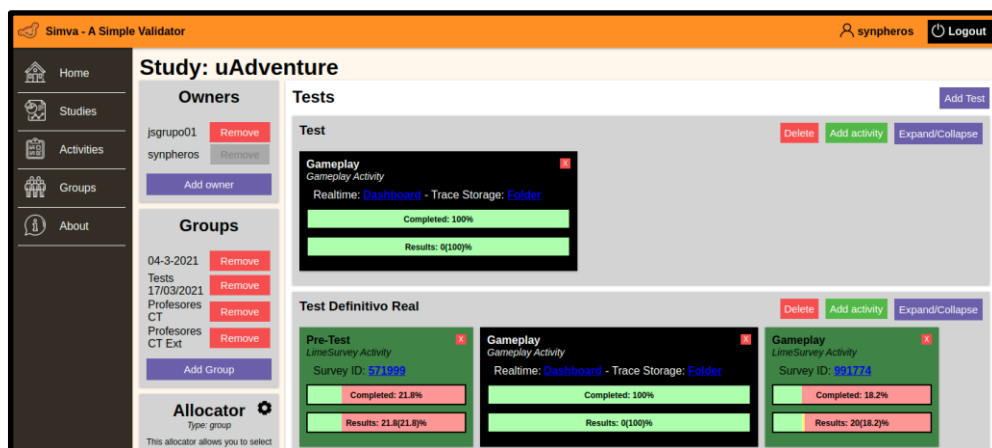


Figura 16. Captura de Simva mostrando la interfaz de los estudios con dos ramas de estudio diferentes, una con solo *gameplay* y otra con encuestas *pre-post*.

Una vez configurado el experimento, Simva permite asignar a los estudios los grupos deseados lo que permite reutilizar el mismo estudio para múltiples grupos, manteniendo unificados todos los resultados de todos los grupos que hayan participado en el experimento. Finalmente, Simva incluye un sistema de asignación o *allocation* que permite gestionar de forma flexible a los usuarios, grupos y estudios. Se han incorporado varios tipos de asignadores, incluyendo: el asignador por grupo que permite asignar a cada grupo participante del estudio a una rama de prueba y el asignador básico para asociar manualmente a cada participante. Se incluye también la funcionalidad de un asignador que permita la distribución ponderada o una distribución aleatoria.

4.3.2.3. Sistema de gestión de actividades y actividades implementadas.

Simva también automatiza el uso y configuración de otros sistemas para realizar diferentes actividades durante los experimentos. Las actividades son los elementos atómicos que forman parte de la secuencia de un estudio. Este nuevo sistema de actividades surge para aislar las funcionalidades independientes de cada tarea a realizar, sea jugar a un juego, hacer una encuesta, o ver un vídeo educativo. Asimismo, el carácter atómico de las actividades requiere que éstas configuren de forma automática aquellos componentes y sistemas conectados que sean necesarios para su ejecución, notificando también a los sistemas externos la información de los participantes para que puedan controlar el acceso. De esta manera, su ejecución es transparente para el usuario final

La **Figura 17** muestra esquemáticamente las principales actividades implementadas en Simva que se presentan con más detalle a continuación:

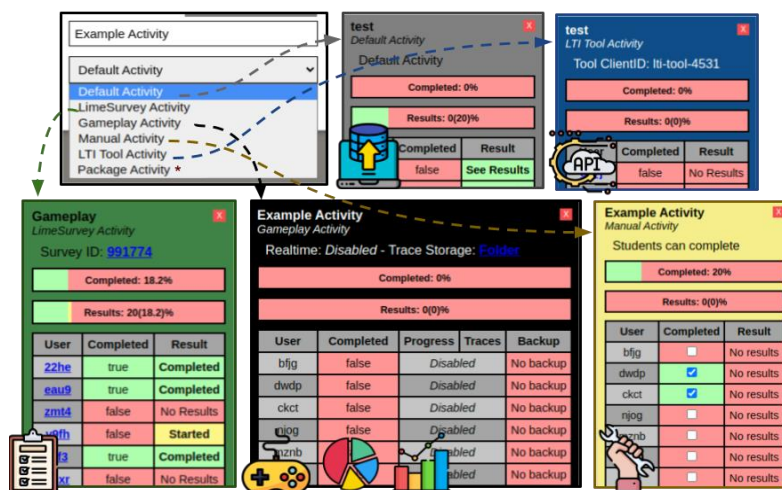


Figura 17. Actividades implementadas en Simva, incluyendo: Actividad de LimeSurvey para encuestas, Actividad de Gameplay para jugar al juego, Actividad manual donde el investigador marca manualmente, además de la actividad de LTI que conecta con una herramienta externa.

1. Actividad de Encuesta: integra y maneja las encuestas en LimeSurvey. Permite a los investigadores subir encuestas o modificar las existentes y crear la actividad Simva correspondiente a estas encuestas mediante una interfaz amigable. Además, muestra el progreso de completitud de las encuestas, y facilita un enlace para dar acceso a los estudiantes a la encuesta correspondiente.
2. Actividades de sesiones de juego con GLA: representan una sesión de juego, y permiten configurar las analíticas. Las analíticas pueden ser de dos tipos: *Realtime* que usa RAGE Analytics o *Data-Science* que almacena la analítica de juego en la nube interna de Simva (aunque también puede hacerlo en otro tipo de nubes externas como Amazon S3²¹ o Minio²²). Esta actividad facilita una API recolector que automatiza el despliegue de los sistemas configurados, sin necesidad de realizar cambios o modificaciones en el juego, que únicamente debe mandar los datos en formato xAPI.
3. Actividades de configuración manual: diseñadas para poder validar otros elementos que no sean tan interactivos o capaces de comunicar su estado tales como presentaciones, videos, u otros recursos, donde ya sea el participante o el investigador tienen que indicar manualmente la completitud de la actividad para poder continuar. La flexibilidad de la interfaz de Simva permite adaptar las columnas de participantes para incluir elementos gráficos interactivos con los que comunicar los resultados.
4. Actividades paquete: permite la gestión de juegos empaquetados mediante estándares su almacenamiento en Simva. Este tipo de actividad ha requerido un desarrollo adicional dada la cantidad de estándares de empaquetado existentes mencionados en el estado del arte y su distinto soporte e implementación por terceros. Actualmente se ha implementado soporte de IMS CP o CMI-5. Esta actividad surge de la necesidad de encontrar un lugar donde almacenar el juego serio para poder distribuirlo como paquete estándar. Este tipo de estándares permiten a Simva distribuir los juegos, autenticar su acceso, así como facilitar datos del estudiante al juego, además de permitir GLA facilitando medios para añadir la configuración de un almacenamiento de analíticas (LRS).

²¹ Sistema de almacenamiento escalable en la nube de Amazon: <https://aws.amazon.com/es/s3/>

²² Alternativa libre que implementa el API de Amazon S3 para la creación de una nube de almacenamiento escalable autoadministrada. Disponible online en <https://min.io/>

5. Actividades con herramientas externas: permite incluir otro tipo de herramientas o plataformas externas como LMS, o campus virtuales externos, donde se puede desarrollar un curso a validar. A diferencia de las actividades paquete, esta actividad entre herramientas autónomas necesita que ambas herramientas soporten interoperabilidad (e.g. LTI) e intercambien previamente claves de seguridad necesarias. De esta manera los usuarios pueden realizar la transición sin perder su autenticación, así como enviar estado y resultados de las actividades de manera segura. Esta actividad engloba toda esta funcionalidad, permitiendo gestionar las herramientas externas, añadir nuevas, y vincular la actividad con una actividad perteneciente a la herramienta externa.
6. Actividad por defecto ad-hoc: Esta actividad está diseñada para facilitar a investigadores que no dispongan de ningún recurso en sus juegos o herramientas interactivas para generar analíticas, no tengan un *tracker* para hacer el seguimiento, o no implementen el estándar xAPI, un lugar en Simva donde comunicarse con peticiones web para añadir resultados en texto plano, y marcar que la actividad ha sido completada.

Con las actividades desarrolladas, es sencillo replicar el diseño experimental utilizando en SurveyManager, centrado en utilizar encuesta previa, sesión de juego, y encuesta posterior. En la **Figura 18** se muestra una captura de un experimento en ejecución, incluyendo enlaces para acceder a los resultados de todas las actividades. En la captura se observan diferentes colores, verde, rojizo y amarillo, que ayudan a la fácil identificación del estado del experimento. Así de forma gráfica es sencillo detectar dónde hay problemas y ayudar a los participantes que lo requieran.



Figura 18 Captura de Simva mostrando la interfaz de los estudios con dos ramas de estudio diferentes, una con solo *gameplay* y otra con encuestas *pre-post*.

4.3.2.4. Sistema de ejecución del experimento o planificador.

Una vez creado y configurado un estudio Simba dispone de un sistema para gestionar la ejecución del experimento, denominado planificador o *scheduler*. Es el sistema encargado de gestionar el estado de completitud de las diferentes actividades del experimento para cada uno de los participantes. Pone en uso a los asignadores o *allocators* cuando un usuario comienza a participar en el experimento, dirigiendo al participante a la actividad que tiene que realizar, sin tener que pasar por ninguna pantalla de acceso o identificación.

Simva incluye un sistema para controlar la ejecución del experimento, que facilita tanto una interfaz web que direcciona a los usuarios a través de las actividades, como una interfaz API donde los juegos pueden consultar qué actividad es la que tienen que realizar. La vista web se representa en la **Figura 19** muestra el planificador ejecutando un estudio para el participante “o57j”. El planificador se muestra en la barra superior, y controla el estado de la actividad de manera automática. En este caso, cuando el participante completa la encuesta, automáticamente el planificador carga el juego, realizando la transición entre sistemas mientras mantiene su acceso y autorización.

Por otra parte, la interfaz API permite integrar los juegos directamente, abriendo la posibilidad de integrar Simva dentro del propio juego. Esto permite integrar juegos que no implementen tecnologías web, pero que soliciten al usuario realizar encuestas de la misma manera.

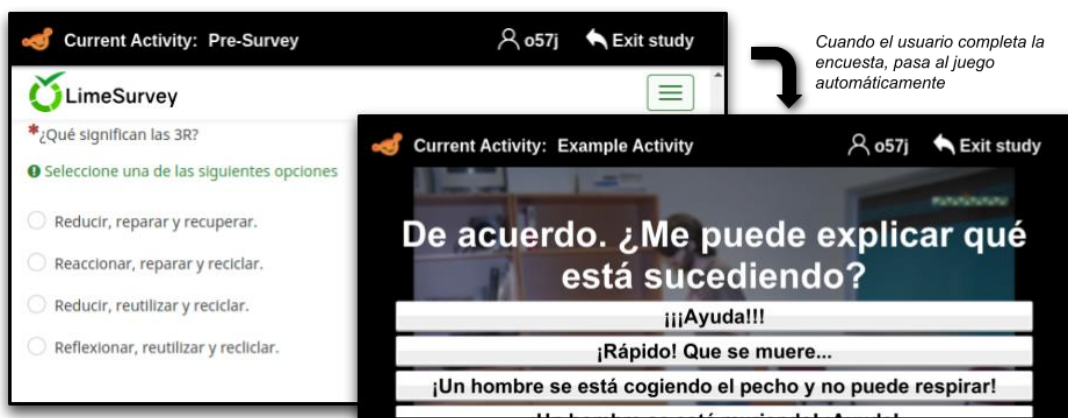


Figura 19. Planificador de Simva ejecutando un estudio para el participante “o57j” donde, primero realiza la encuesta previa, y después realiza la sesión de juego. La transición entre las actividades la realiza dicho planificador, de manera automática, cuando identifica que las actividades se han completado.

4.3.2.5. Una arquitectura escalable y abierta: conexión con sistemas de ciencia de datos.

Simva tiene una arquitectura software abierta basada en sistemas interconectados para delegar la funcionalidad más crítica en plataformas confiables y escalables y para facilitar la integración con otros proyectos desarrollados por terceros. Entre los sistemas integrados en la arquitectura se incluyen al sistema de inicio de sesión unificado *single sign-on* SSO (KeyCloak²³), el sistema de colas de procesamiento (Kafka²⁴), o el sistema de almacenamiento en la nube (Minio). Esta arquitectura se presenta en el artículo Pérez-Colado, I. J., Pérez-Colado, V. M., Martínez-Ortiz, I., Freire, M., & Fernández-Manjón, B. (2020). **A Scalable Architecture for One-Stop Evaluation of Serious Games.** *Lecture Notes in Computer Science, 12517 LNCS*, 69–78. En esta publicación se presentan detalles de todos los componentes de la arquitectura, con sus funciones y relaciones.

Otra de las grandes ventajas de esta arquitectura es que se puede conectar fácilmente con otros sistemas externos. Por ejemplo, es posible conectarlo y exportar los datos para ser analizados con herramientas genéricas de ciencia de datos como Jupyter Notebooks²⁵. Esto ha permitido conectarlo con otro sistema desarrollado en el grupo e-UCM

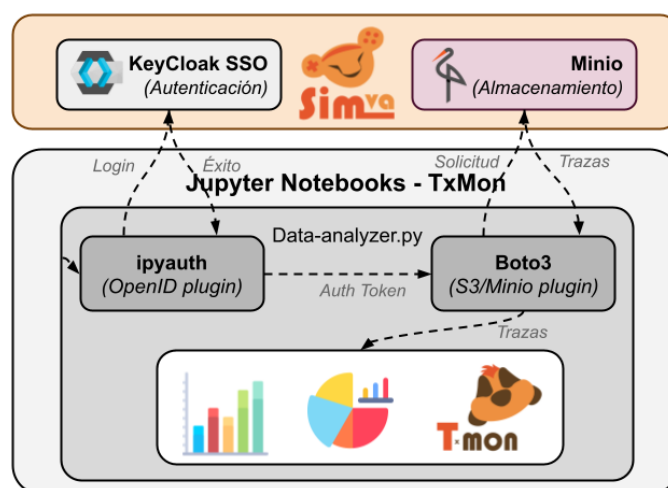


Figura 20. Proceso de *data-science* utilizando Simva y TxMon, un proyecto que permite ejecutar el LAM por defecto desarrollado en Jupyter Notebooks.

²³ Gestor de identidad y acceso de código abierto disponible en: <https://www.keycloak.org/>

²⁴ Proyecto de intermediación de mensajes, utilizado para gestionar colas de procesamiento: <https://kafka.apache.org/>

²⁵ Proyecto libre para ejecutar *scripts* y hacer *data-science*: <https://jupyter.org/>

denominado TxMon (sistema disponible en el repositorio de GitHub: <https://github.com/e-ucm/t-mon>). El proyecto TxMon surge para poder facilitar un análisis sencillo de la información recopilada de analíticas a los investigadores que no quieran utilizar *Realtime* y RAGE Analytics. Ofrece un conjunto de análisis similares a los presentados y desarrollados utilizando el LAM por defecto. Pero en este caso el análisis de las trazas se realiza de manera externa en Jupyter Notebooks, bajo demanda y con control completo de los analistas de datos que pueden incluir nuevos análisis o modificar los existentes sin depender de Simva. En la **Figura 20** se presenta el proceso de explotación de Simva utilizando TxMon, incluyendo la autorización y acceso a los registros de juego, así como las interacciones de ambos proyectos

4.3.3. Integración de Simva con herramientas de autoría de juegos serios: uAdventure

Simva simplifica el procedimiento habitual de validación de juegos serios automatizando el proceso y permitiendo el diseño y despliegue de un experimento mediante la creación de casos de estudio basados en actividades con el juego. En el caso más habitual incluye una encuesta previa, la actividad de juego y otra encuesta posterior. Pero todavía es posible llevar el proceso un paso más allá mediante la integración con una herramienta de autoría de juegos serios. Se ha integrado Simva con el sistema de autoría de juegos serios uAdventure.

Continuando el trabajo realizado en uAdventure, se han realizado una serie de modificaciones y mejoras sobre la herramienta para añadir nuevos editores que permitan

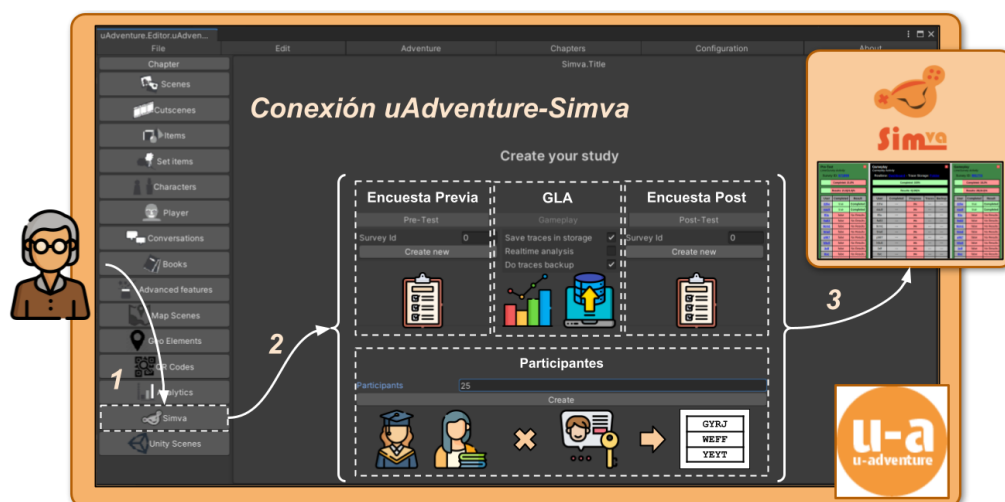


Figura 21. Integración de Simva en uAdventure con un editor que automatiza el despliegue del experimento con una sencilla configuración incluyendo las encuestas, analíticas y los participantes.

configurar y desplegar de manera sencilla un diseño experimental. Esta nueva funcionalidad se representa en la **Figura 21** donde, a través de un formulario guiado paso por paso, el usuario diseña el experimento basándose en la plantilla de experimento que utiliza encuestas *pre-post*. Se deben configurar también la cantidad de participantes y los servicios que desea incluir y utilizar de manera automática como, por ejemplo, si se quieren usar las analíticas en tiempo real o un sistema de almacenamiento para posteriormente analizar los datos en un entorno externo de análisis de datos.

La integración realizada mediante el desarrollo del editor comunica a uAdventure y Simva de manera que el usuario de la herramienta de autoría no necesita abandonar ésta para configurar todo lo necesario en Simva. Esto incluye obtener los recursos necesarios para el desarrollo del experimento (e.g. hojas con tokens de identificación), incluyendo en la compilación del juego, todo lo necesario para que el juego se comunique con el planificador de Simva y sea capaz de gestionar tanto las analíticas. Además, incluye controles para evitar que el juego pueda usarse fuera del flujo del diseño experimental requiriendo la identificación de los usuarios con un token de identificación.

La integración de Simva con uAdventure, así como ejemplos de los estudios generados, se encuentra publicada en el artículo Pérez-Colado, I. J., Pérez-Colado, V. M., Martínez-Ortiz, I., Freire-Moran, M., & Fernández-Manjon, B. (2020). **Simplifying Serious Games Authoring and Validation with uAdventure and SIMVA**. *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*, 106–108.

4.4. Estándares de e-learning, LTI y CMI-5

En la tesis se han desarrollado una serie de herramientas que simplifican el ciclo de vida de los juegos serios permitiendo mediante las analíticas que los participantes (e.g. desarrolladores, investigadores y profesores) entiendan qué es lo que pasa realmente en el proceso educativo con juegos serios. Con el objetivo de mejorar las herramientas desarrolladas en la implementación siempre que ha sido posible se han tratado de usar estándares. Esto permite extender las capacidades de interoperabilidad con otros sistemas y, a medio plazo, facilitará su acceso a nuevos investigadores que deseen explotar las ventajas de Simva para realizar actividades *gamificadas* con GLA o validar sus propios juegos serios.

En el estado del arte se han analizado los estándares educativos aplicables al ámbito de los juegos serios que presentan la oportunidad de mejorar y alargar el ciclo de vida, así como facilitar el uso de las herramientas desarrolladas. Aunque Simva ya implementaba desde su primera versión algunos estándares de e-learning como, por ejemplo, un recolector de GLA preparado para procesar trazas que utilicen el formato xAPI y estándares de empaquetado como IMS CP que permite incluir nuevos recursos educativos en modo de paquete IMS (incluidos juegos). No obstante, se realizó un análisis más amplio, seleccionando otros estándares a implementar que se consideraban los más prometedores para extender las funcionalidades de Simva. Los principales resultados de este trabajo se presentan en el artículo Perez-Colado, I. J., Perez-Colado, V. M., Freire, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2021). **e-learning Standards in Game-Based Learning? 2021 International Conference on Advanced Learning Technologies (ICALT)**, 81–82.

De entre todos los estándares, se han elegido aquellos que por su gran potencial permiten mejorar el ciclo de vida de Simva y su interoperabilidad. En primer lugar, se elige IMS LTI centrado en la comunicación de plataformas, permitiendo la transición de usuarios autenticados entre servicios sin solicitar credenciales, así como el reporte de resultados de uso para la evaluación de las sesiones. Además, se propone una jerarquía de plataforma-herramienta (Platform-Tool) para relacionar los sistemas, donde la plataforma

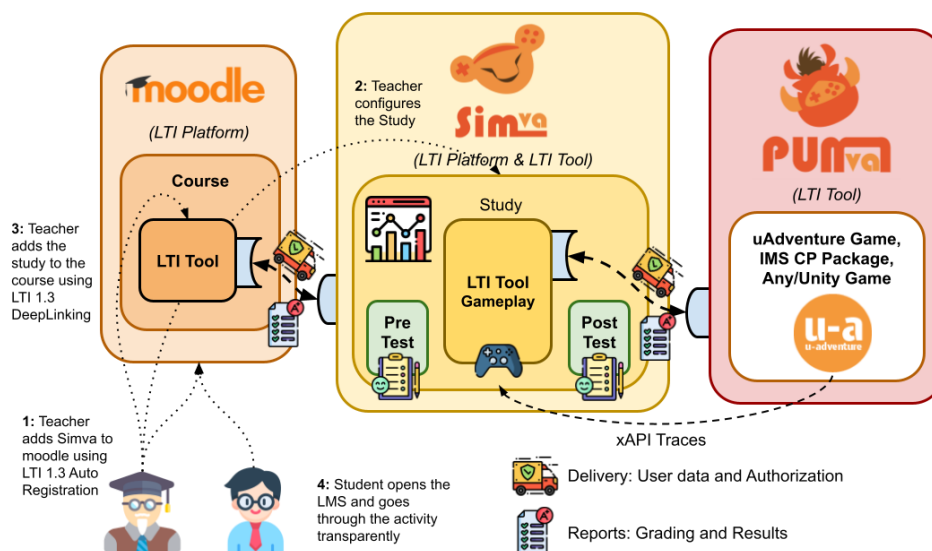


Figura 22. Posibles conexiones que se pueden realizar entre Simva, un LMS y una herramienta que implemente el estándar de interoperabilidad LTI. La representación muestra a Simva con rol de herramienta utilizada por la plataforma LMS Moodle, que incluye a Simva en un curso. A su vez, Simva integra la herramienta LTI Pumva, repositorio de juegos serios de uAdventure donde delegar la ejecución del juego, manteniendo las ventajas del uso de Simva, facilitando y simplificando su uso.

utiliza a la herramienta para delegar una función educativa concreta. Esta relación se representa en la **Figura 22**, donde se muestran las posibles relaciones que podría establecer Simva con otros sistemas mediante la implementación de LTI.

Para realizar la integración de LTI en Simva, se ha incluido un manager para añadir y gestionar las plataformas que tendrán acceso para utilizar Simva como herramienta. De este modo se puede lanzar un estudio desde un LMS habilitado en Simva. También se ha incluido un nuevo tipo de actividad, conocida como Actividad con Herramienta Externa, capaz de gestionar en su creación el manejo de herramientas externas mediante LTI, permitiendo añadir al diseño experimental sistemas externos. Esta integración permite especificar no sólo la herramienta a utilizar, sino qué parte de la herramienta se desea utilizar. Esta integración con LTI se representa en la **Figura 23**, junto a los editores que permiten añadir tanto plataformas como herramientas de LTI.

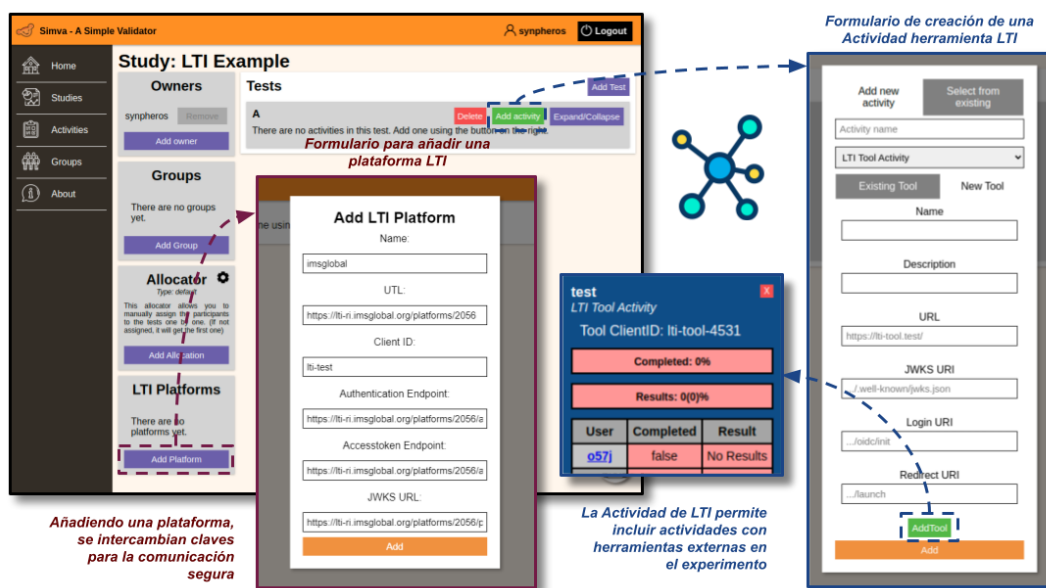


Figura 23. Editores de Simva que utilizan LTI para añadir tanto plataformas habilitadas para utilizar Simva como herramienta, así como añadir herramientas para ser utilizadas en actividades mediante la plataforma Simva.

Por tanto, con la integración de LTI en Simva se logra mejorar las posibilidades de interconectividad de la plataforma con otros sistemas educativos, como LMS o plataformas de aprendizaje. Sin embargo, LTI presenta una serie de limitaciones intrínsecas a su funcionalidad, dado que está diseñado para interconectar sistemas y plataformas autosuficientes, las cuales disponen de sus propios recursos y entorno de ejecución para funcionar.

Si bien los juegos serios pueden presentarse también como parte de una plataforma, es más frecuente que se distribuyan como programas independientes que se ejecutan en el ordenador del usuario. Normalmente se hace mediante un paquete que incluye los recursos del juego junto a un archivo ejecutable de manera nativa o mediante un navegador web que utilice tecnologías como HTML5 *Canvas*. Para facilitar la adopción y distribución de este tipo de contenido educativo, incluyendo vídeos formativos o presentaciones además de los juegos serios, se implementan estándares de empaquetado y distribución, y además de IMS CP se ha considerado también CMI-5. Aunque las funcionalidades de ambos son similares hay ciertas diferencias. IMS CP permite una integración rápida y sencilla pero limitada por la propia especificación. CMI-5 permite una integración más completa de los contenidos, pues además del empaquetado y distribución, también permite gestionar otros aspectos como la configuración de un LRS compatible con xAPI donde el juego puede enviar los datos de analíticas, o posibilitar una interfaz para que los juegos puedan obtener datos acerca del participante para poder personalizar la sesión de juego.

La integración de los estándares de empaquetado en Simva se realiza mediante la Actividad Paquete que permite en su creación seleccionar un paquete que contenga contenido educativo, y que incluya en su interior la documentación necesaria para poder realizar la importación. Simva da soporte a paquetes que utilicen IMS CP o CMI-5. En el proceso de creación de la actividad, una vez se ha realizado la subida del paquete, Simva realiza el análisis de su contenido y configura el entorno para poder recibir trazas para GLA.

4.5. Pruebas con usuarios y casos de uso

La metodología de desarrollo elegida durante el proceso de la tesis se basa en la realización de iteraciones incrementales y pruebas tempranas utilizando prototipos. Por lo tanto, para la validación de las herramientas desarrolladas durante las iteraciones se han realizado múltiples pruebas a lo largo del desarrollo utilizando los resultados obtenidos como retroalimentación para las siguientes iteraciones. Estas pruebas se han dividido en múltiples casos experimentales enfocados a poner en práctica el trabajo desarrollado, incluyendo

- Caso de estudio del Juego de Primeros Auxilios: se usó como objeto de experimentación el juego serio previamente desarrollado con eAdventure que

permitió realizar una prueba de concepto del proceso implementado de GLA. Para la realización de estas pruebas fue necesario incluir un conjunto de herramientas prototipo que permitían la ejecución del proceso completo de GLA y su validación. Además, requirió crear toda la documentación necesaria referente al LAM (que no existía previamente) y decidir cómo se aplicaba, tanto al juego, como a la herramienta uAdventure.

- Caso de estudio del proyecto Beaconing para el análisis multi-nivel: Mediante la integración de GLA en el proyecto Beaconing se realizó una implementación de analíticas multi-nivel. Esto permitió probar el proceso de generación y aplicación de un modelo de analíticas multi-nivel, que utilice un meta-LAM, en un sistema con estructuras de actividades *gamificadas* con una organización jerárquica.
- Casos de validación de juegos serios: con objetivo de validar y retroalimentar los prototipos de SurveyManager y Simva se ha colaborado con otros investigadores del grupo de investigación para realizar la validación y proceso de GLA en dos juegos serios: el juego Conectado, diseñado para la concienciación del *cyberbullying*, y el juego serio geoposicionado El Señor de los Arbustos, que enseña el proceso de reciclaje a jóvenes.
- Pruebas para la automatización del proceso de validación y GLA: también se han realizado pruebas mediante la integración realizada en uAdventure que, como se ha detallado previamente, comprende tanto el desarrollo de un modelo de análisis educativo genérico capaz de automatizar el proceso de GLA como la integración de Simva en la herramienta para el despliegue automatizado del diseño experimental con el que realizar la validación. Las pruebas realizadas comprenden un caso de estudio en el que un grupo de alumnos de la asignatura de Grado de Desarrollo de Videojuegos, en la asignatura Juegos Serios diseñan y desarrollan un juego en uAdventure, donde en última instancia, los alumnos realizan la validación de sus juegos mediante las herramientas integradas con el editor. También se comprende otro caso de estudio donde un grupo de alumnos de la asignatura e-Learning del Máster en Ingeniería informática, usan uAdventure y aplican el proceso de GLA utilizando un modelo de análisis educativo sobre un juego serio que ellos mismos crean.

A continuación, se describen con mayor detalle estas pruebas y casos de uso.

4.5.1. Caso de estudio del Juego de Primeros Auxilios

Para la validación del primer prototipo desarrollado, y con el objetivo de poner en prueba y evaluar el proceso de implementación de GLA en un juego serio, además de poner en prueba el proceso de validación integrado y digitalizado mediante el uso de encuestas online, se presenta este caso de estudio, el cual plantea en la reutilización un juego serio desarrollado con eAdventure por el grupo de investigación como objeto de trabajo. Sobre este juego se realiza el proceso de desarrollo de analíticas utilizando y diseñando un modelo de análisis educativo LAM. Además, se realiza la validación computerizada del producto terminado mediante el uso de cuestionarios previo y posterior usando plataformas de encuestas online.

El juego elegido es el juego para enseñar técnicas de Primeros Auxilios (Marchiori, Ferrer, et al., 2012), y utilizando uAdventure, se realiza la importación del juego sobre el que se aplican e implementan las mejoras necesarias para aplicar el modelo de análisis hasta poder presentar los datos en visualizaciones. El primer prototipo de la herramienta uAdventure forma parte del trabajo previo a la tesis, desarrollado como parte del Trabajo de Fin de Máster de Pérez-Colado, I. J. (I. J. Pérez Colado, 2016), y que incluye ya algunas mejoras como un *tracker* capaz de generar trazas para GLA. Este modelo de análisis forma parte del trabajo presentado en Perez-Colado, I., Alonso-Fernandez, C., Freire, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2018). **Game learning analytics is not informagic!** *2018 IEEE Global Engineering Education Conference (EDUCON)*, 1729–1737, y que complementa el trabajo realizado en el Trabajo de fin de Máster de Alonso-Fernandez (Alonso-Fernandez, 2017) en el que se aplican técnicas de minería de datos sobre los datos extraídos con GLA.

Para la implementación del LAM, y con objetivo de mejorar la herramienta uAdventure para poder sistematizar y automatizar el proceso de GLA, se realizan una serie de modificaciones en la herramienta para que ésta implemente un modelo de analíticas genérico basado en el perfil xAPI-SG para definir las interacciones que se deben registrar para poder realizar la evaluación. Este modelo se complementa de una serie de editores con los que se realiza la personalización del modelo de analíticas a las necesidades específicas de cada juego, permitiendo añadir modelos de medición del progreso y puntuaciones para facilitar la evaluación. Detalles acerca del modelo y los editores implementados para su modificación se incluyen en la publicación (Perez-Colado, Perez-Colado, Freire-Moran, Martinez-Ortiz, & Fernandez-Manjon, 2017) donde se evalúa el

proceso de integración de GLA en una herramienta de autoría. Finalmente, la automatización del proceso de GLA como parte del ciclo de vida de un Juego Serio forma parte del trabajo presentado en Perez-Colado, V. M., Pérez-Colado, I. J., Freire-Morán, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2019). **UAdventure: Simplifying narrative serious games development**. *Proceedings - IEEE 19th International Conference on Advanced Learning Technologies, ICALT 2019*, 119–123, donde se explora cómo mediante el uso de uAdventure además se simplificar el desarrollo de un juego serio, integrando funcionalidades para realizar GLA de manera semiautomatizada.

Para la realización de los experimentos se implementa y utiliza la primera versión prototipo de la plataforma de validación desarrollada en la presente tesis, que integra y utiliza plataformas online para la realización de encuestas con la que recoger los datos de los participantes, y añadiendo tanto un sencillo control de acceso y autorización basado en identificadores anónimos o tokens, como herramientas que automatizan la integración y despliegue de todo lo necesario para realizar GLA. En la **Figura 24** se incluyen fotografías de una sesión de experimento realizada con el prototipo completo donde se observa a los participantes jugando al juego y completando encuestas, y mostrando en la



Figura 24. Fotografías de una de las primeras sesiones de prueba realizadas con el juego de Primeros Auxilios tomadas en 2017, donde se observa a los participantes jugar al juego y completar las encuestas. La fotografía de la esquina inferior izquierda muestra el proceso de evaluación de los paneles de analíticas desarrollados. En ellas aparecen los investigadores Antonio Calvo-Morata, Cristina Alonso-Fernández e Iván J. Pérez-Colado.

fotografía inferior izquierda una entrevista realizada a un profesor quién evalúa el panel de analíticas desarrollado, así como los beneficios de aplicar GLA.

Se completan los experimentos, concluyendo el éxito de la prueba de concepto para los procesos y herramientas diseñados para implementar y aplicar GLA y realizar la validación mediante cuestionarios online. Las herramientas diseñadas han permitido alcanzar conclusiones similares a las alcanzadas durante el proceso de validación del juego original, demostrando su eficacia a la hora de enseñar técnicas de primeros auxilios, pero reduciendo altamente el esfuerzo de procesamiento de los datos del experimento, e incluyendo analíticas en tiempo real con alertas específicas para dar soporte a evaluadores e investigadores. Por otra parte, la iteración concluye también con la necesidad de mejorar la herramienta acorde a nuevos requisitos que mejoren la flexibilidad, facilidad de uso, y robustez del sistema, además de añadir más funcionalidades.

Éste caso de estudio ha sido continuado por miembros del grupo de investigación (Alonso-Fernández et al., 2020) para evaluar la posibilidad de aplicar un modelo predictivo capaz de identificar patrones en el proceso de juego mediante el análisis de las trazas con el que predecir el resultado de la sesión.

4.5.2. Caso de estudio del proyecto Beaconing: analíticas multi-nivel

A través del grupo de investigación, se presenta la oportunidad de realizar la integración de GLA en el proyecto europeo H2020 Beaconing. Este proyecto incluye una herramienta de autoría de juegos que utiliza juegos existentes a modo de plantillas, y presentados a modo de *Gamified Lesson Plans*. Mediante su personalización se permite adaptar el contenido de los juegos que componen las lecciones *gamificadas*, su desarrollo y diálogos, para satisfacer necesidades educativas de cada plan. Además, esta herramienta permite que cada juego pueda lanzar y personalizar a su vez el contenido de minijuegos genéricos con los que delegar objetivos educativos que se superan al completar los minijuegos con éxito. Mediante esta integración, los juegos generados son capaces de generar analíticas de manera automática para el usuario.

La organización jerárquica de los juegos compuestos por minijuegos en el proyecto H2020 Beaconing se organiza de manera similar a un LMS que compone sus cursos con múltiples recursos educativos. Para aplicar analíticas en este proyecto, se desarrolla un modelo de analíticas multi-nivel, con un panel de analíticas capaz de filtrar los resultados de cada uno de los minijuegos del juego, y con el que identificar de manera sencilla los

objetivos educativos y su grado de satisfacción. El modelo de analíticas multi-nivel desarrollado supone un caso de aplicación exitosa de un meta-LAM y forma parte del trabajo presentado en Perez-Colado, I. J., Rotaru, D. C., Freire-Moran, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2018). **Multi-level game learning analytics for serious games**. *2018 10th International Conference on Virtual Worlds and Games for Serious Applications, VS-Games 2018*, donde se incluye el panel de analíticas desarrollado junto a la interfaz que permite la navegación entre los niveles de la jerarquía del juego.

Para su implementación se realiza una integración de la plataforma de GLA del proyecto RAGE Analytics con los servicios de Beaconing, incluyendo el desarrollo de un análisis capaz de procesar trazas de múltiples actividades de manera unificada, y que aplica un proceso conocido como “*rollup*” en las trazas en donde cada nodo de la estructura, después de procesar los datos, los envía al nodo superior hasta que alcanza la raíz, donde finaliza su procesamiento.

4.5.3. Pruebas de validación de Juegos Serios

A continuación, se presentan dos casos de estudio realizados, que prueban las capacidades de la plataforma para la validación de juegos, tanto en la creación de un diseño experimental, como en la ejecución de los experimentos. Algunos de estos experimentos se describen en el artículo Alonso-Fernandez, C., Perez-Colado, I. J., Calvo-Morata, A., Freire, M., Ortiz, I. M., & Fernandez-Manjon, B. F. (2020). **Applications of Simva to Simplify Serious Games Validation and Deployment**. *IEEE Revista Iberoamericana de Tecnologías Del Aprendizaje*, 15(3), 161–170

4.5.3.1. La validación del juego Conectado.

El primer caso de estudio de validación de un juego serio realizado con Simva es el juego serio Conectado (Calvo-Morata, Rotaru, et al., 2018). Este trabajo se realizó en colaboración con el autor principal del juego Antonio Calvo-Morata quien realizó el diseño experimental usando Simva. El juego tiene el propósito de concienciar a niños de entre 12 y 17 años sobre el problema de la violencia escolar (*bullying* y *ciberbullying*). El juego está diseñado para ser completado en una sesión de clase tutelada por un profesor, con un grupo de alumnos de tamaño medio. El diseño experimental se plantea para ajustarse a estas sesiones, utilizando encuestas previa y posterior para la validación, diseñadas por el autor.

Para la ejecución del experimento, se realiza una integración *ad-hoc* del juego con la plataforma Simva, integrando los servicios de acceso de usuario, y controlando la ejecución del experimento mediante la API de planificación. La **Figura 25** muestra dos capturas del propio juego donde se pide el código de acceso de Simva al participante, el



Figura 25. Captura de pantalla de la integración *ad-hoc* de Simva en el juego Conectado. La imagen de la izquierda muestra el formulario donde introducir el código de acceso, dando acceso a la encuesta previa.

cual tras ser introducido correctamente abre automáticamente el navegador para acceder a la encuesta inicial. No se permite de nuevo el acceso y avanza en el juego hasta que se completa esta encuesta inicial.

Para desarrollar el experimento y aplicarlo en la plataforma de validación Simva, se genera un grupo de casos de estudio que divide a los participantes por escuelas y grupos. Algunos de los requisitos específicos a remarcar del experimento incluyen el caso de uso de un grupo de control, así como un caso especial donde los profesores que participan en el experimento también rellenan una encuesta adicional con preguntas específicas con las que se evalúan otros elementos como los paneles de analíticas o la facilidad de uso del juego. A lo largo del experimento participaron un total de 1004 estudiantes en dos grupos de experimentos, el primero incluyendo 257 usuarios. Del total de participantes, 902 completaron los dos cuestionarios, incluyendo además los registros de GLA recolectados de cada sesión de juego.

Los resultados de la validación realizada fueron publicados por su autor (Calvo-Morata et al., 2021) y presentan resultados significativos siendo una validación formal exitosa. Estos resultados respaldan la validez de Simva para adaptarse a las necesidades del investigador a la hora de diseñar un experimento y llevarlo a cabo, permitiendo su ejecución de manera sencilla y facilitando la obtención y utilización de los datos. Se puede considerar por tanto que la plataforma Simva cumple con las expectativas ya que ha ayudado a simplificar significativamente el proceso de validación del juego en un experimento con un gran número de usuarios.



Figura 26. Fotografías de algunas de las sesiones realizadas durante los experimentos de validación del juego Conectado. La imagen de la izquierda muestra al investigador I. Perez-Colado realizando una encuesta al profesor acerca de los paneles de analíticas durante la sesión.

Para el desarrollo del experimento, se realizan labores de soporte y mantenimiento de la infraestructura durante la ejecución, participando también en la supervisión de algunas de las sesiones, y donde se aprovecha para evaluar parcialmente los paneles de analíticas desarrollados. En la **Figura 26** se incluyen dos fotografías de los experimentos realizados, donde se observan tanto a los alumnos participando y jugando al juego, como el panel de analíticas que muestra datos del sistema procesados en tiempo real.

4.5.3.2. La validación del juego El Señor de los Arbustos.

El segundo caso de estudio de validación de un juego serio se realiza con el proyecto El Señor de los Arbustos en colaboración con su desarrollador Víctor M. Pérez-Colado (V. M. Pérez Colado, 2022). En este caso se usa Simva para validar la efectividad de un juego serio geoposicionado a la hora de reemplazar una actividad gymkhana física de contenido similar. El juego tiene como objetivo la concienciación sobre el reciclaje, así como el concepto de la recogida de basura como “la cuarta R” del reciclaje y plantea una aventura de fantasía en la que los participantes utilizan sus conocimientos acerca de los procesos

de reciclaje para completar una serie de misiones con las que acabar con el villano. El juego incluye elementos de realidad aumentada al ser geoposicionada e incorpora el uso de la cámara y códigos QR.

El juego desarrollado con la herramienta de autoría uAdventure explota las capacidades de la herramienta para la explotación de GLA, generando registros de actividad de las sesiones de juego. Para la integración de las herramientas de validación y la conexión con Simva se integran en uAdventure un conjunto de componentes que comunican a la herramienta y juegos con la plataforma. Asimismo, se desarrollan una serie de editores en la herramienta que simplifica el despliegue del diseño experimental que permite la validación del juego. Aunque esta integración supone un mayor esfuerzo de implementación, añade de forma permanente en la herramienta de autoría un conjunto de herramientas con las que automatizar el proceso de validación. Esta integración se incluye como parte del trabajo presentado en Pérez-Colado, I. J., Pérez-Colado, V. M., Martínez-Ortiz, I., Freire-Moran, M., & Fernandez-Manjon, B. (2020). **Simplifying Serious Games Authoring and Validation with uAdventure and SIMVA**. *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*, 106–108.

Tras crear los casos de estudio en Simva se realizan tres pruebas piloto con usuarios seleccionados lo que permite lograr mejoras de estabilidad en el juego y en el sistema de analíticas geoposicionadas. Los resultados de los pilotos presentan una recepción del juego positiva, con una gran acogida por los participantes, incluso en los casos en los que el participante no logró completar el juego. Una vez completados los pilotos con el juego se preparan los casos de estudio necesarios para la realización de los experimentos con usuarios reales.

Los experimentos se realizan con dificultades debido a condiciones meteorológicas y luego a la pandemia del Covid-19, con la cuarentena domiciliaria, así como las duras medidas para reducir los contagios y evitar riesgos en la salud de los participantes, que frenan la realización de los experimentos hasta que la situación lo permita. Finalmente, se consiguió programar un nuevo experimento en el que se recogieron datos de las encuestas realizadas para la validación de 22 participantes. Los resultados resultan insuficientes para realizar la validación dado el limitado conjunto de datos recogidos durante los experimentos realizados. Sin embargo, la plataforma Simva se adaptó

completamente a las necesidades del investigador, comportándose adecuadamente durante la participación de los usuarios en el experimento.

4.5.3.3. Pruebas para la automatización del proceso de validación y GLA.

Con objetivo de evaluar las posibilidades de las herramientas de autoría de juegos serios para automatizar el proceso de validación y generalizar las GLA se utiliza uAdventure como herramienta para los alumnos de las asignaturas Juegos Serios y e-Learning previamente mencionadas. Los alumnos tienen que implementar un juego serio en uAdventure y aplicar el proceso de GLA y validación usando sus funcionalidades. En las pruebas de los juegos desarrollados participan el resto de compañeros de la asignatura, aprovechando el potencial de uAdventure y Simva para agilizar y simplificar estos procesos.

En ambas asignaturas, como parte de las prácticas, los alumnos realizaron el desarrollo y validación de los juegos, en la mayoría de los casos de forma exitosa, produciendo como resultado varios juegos serios con analíticas incorporadas y sobre los que se pueden hacer pruebas. Se realiza un especial énfasis en cómo establecer los objetivos educativos y cómo identificar que el estudiante ya los ha satisfecho mediante el seguimiento que se realiza de la sesión de juego, especificando los eventos importantes que se deben registrar. De los productos resultantes, se han elegido dos casos especialmente relevantes por sus características, que incluyen: el juego llamado La Mansión Paranormal, experiencia narrativa que utiliza minijuegos basados en la resolución de puzzles de lógica para avanzar en la aventura, y el juego titulado Guerra Civil UCM. El caso de estudio de La Mansión Paranormal plantea un diseño experimental un poco diferente donde en lugar de evaluar a los alumnos, se evalúa la perspectiva de uso de los profesores acerca del juego y su utilización en clase, recolectando los cuestionarios previo y posterior, y permitiendo incluir en ellos su opinión acerca de su utilidad y futura aplicabilidad. Además, el diseño de GLA del juego incluye analíticas personalizadas para identificar las interacciones realizadas en los minijuegos. Por otra parte, el caso de uso del juego Guerra Civil UCM es un juego serio geoposicionado que trata el tema de la guerra civil en el campus de la UCM, y que presenta mecánicas de juegos de búsqueda de tesoros para progresar a lo largo de la trama, la cual se desarrolla hablando con personajes de ambos bandos. El juego desarrollado presentó carencias durante la realización de los pilotos, las cuales se identifican gracias al proceso de validación realizado, trabajándose en una nueva versión que incluye mejoras como una mejor narrativa de la historia o la posibilidad de realizar

la aventura de manera no lineal. La validación concluye con una prueba donde los participantes consiguen completar el juego serio geoposicionado de forma completamente autónoma.

En todos los casos, los esfuerzos necesarios para realizar la validación se han visto reducidos drásticamente, permitiendo a los alumnos diseñar y poner en ejecución el diseño experimental sin invertir esfuerzo extra en la recolección y digitalización de los datos. Además, se realiza de manera segura y anonimizada para los participantes lo que ha contribuido a que los alumnos conozcan las posibilidades de las GLA. Los resultados obtenidos muestran como mediante la integración de herramientas para simplificar el proceso de GLA y validación en un juego en la propia herramienta utilizada para crearlo se simplifican de forma significativa el esfuerzo que implican estas técnicas.

Capítulo 5. Conclusiones y Trabajo futuro

En este capítulo se resumen las principales conclusiones y contribuciones de este trabajo de tesis doctoral, así como algunas líneas de trabajo futuro que se han abierto como resultado del mismo.

Cabe destacar que el presente trabajo no solo ha abordado los objetivos desde el punto de vista puramente de investigación, sino que las conclusiones y aportaciones también se encuentran soportadas por el desarrollo de herramientas software. Estas herramientas nos han permitido poner en práctica y validar los objetivos propuestos para este trabajo de tesis doctoral. Como principal resultado software, se ha creado una herramienta gratuita y de código abierto, denominada Simva (acrónimo de *Simple-Validator*) que está disponible para su uso por cualquier interesado.

5.1. Conclusiones y principales aportaciones

Como se ha descrito previamente, el objetivo global de este trabajo de tesis es mejorar el ciclo de vida y la adopción de los juegos serios mediante la integración de Game Learning Analytics (GLA) y de herramientas de soporte a las GLA que faciliten la validación de los juegos serios y su aplicación como herramientas de evaluación. Desde el punto de vista tecnológico, la principal aportación ha sido el diseño de una arquitectura genérica y el desarrollo de un ecosistema de herramientas que permite disminuir el coste de aplicación de GLA durante el ciclo de vida de los juegos serios. Este ciclo completo incluye desde la etapa de diseño y validación sistemática de los juegos, hasta la etapa de despliegue con alumnos en un escenario educativo real. A continuación, se presentan de manera más detallada las aportaciones de esta tesis doctoral:

- **Proceso sistemático de aplicación de GLA.** Para alcanzar tanto el objetivo general como los objetivos concretos (descritos en el Capítulo 1), se ha desarrollado un proceso sistemático para aplicar GLA basado en el uso de un Learning Analytics Model (LAM). Este modelo facilita la extracción y el análisis de datos significativos de los juegos serios que permiten validar sus objetivos educativos.
- **Simva, herramienta de soporte a las GLA.** El LAM propuesto está soportado tecnológicamente mediante una herramienta de soporte a las GLA denominada Simva (ver 4.3.2). Esta herramienta facilita la validación de los juegos serios (e

incluso la de otras herramientas o recursos educativos altamente interactivos) mediante el concepto de estudio. Este estudio simplifica el diseño experimental ya que permite aglutinar y automatizar tanto la realización de la sesión de juego como la realización de cuestionarios en línea *pre-post* para permitir realizar un análisis comparativo que sustente una validación formal del juego serio. Nótese que Simva permite realizar estas tareas mediante la generación de identificadores pseudoaleatorios (*tokens*) con el objetivo de facilitar el cumplimiento con la normativa europea de privacidad y protección de datos (GDPR).

- **LAM para proporcionar información de las analíticas.** Mediante la aplicación de un Learning Analytics Model se puede proporcionar información a los distintos actores a partir del seguimiento de las acciones significativas del usuario en el juego. Por ejemplo, para los educadores estas analíticas pueden generar visualizaciones con información relevante sobre la experiencia de aprendizaje de los estudiantes. Para los creadores del juego se pueden proporcionar métricas relacionadas con su eficacia y para facilitar su validación. Tras dicha validación formal, se podrían usar para evaluar el impacto del juego en los alumnos sin necesidad de utilizar cuestionarios pre-post.

La definición de un LAM es clave durante el ciclo de vida del juego serio ya que encapsula qué datos se deben recolectar, en qué formato, cómo se deben analizar, procesar y mostrar, con el propósito de medir los objetivos educativos del juego. El LAM debe fundamentarse en el diseño educativo e, idealmente, definirse antes que el propio juego, para posteriormente dirigir y adaptarse en las siguientes etapas de desarrollo del juego. En la primera parte de la sección 4.1 se han incluido el proceso y todos los detalles para generar y explotar dicho modelo.

- **LAM y meta-LAM para juegos complejos.** El modelo definido en un LAM es genérico y contempla juegos serios complejos. Es posible generar un meta-LAM para aquellos casos en los que los juegos forman parte de estructuras más complejas (por ejemplo, juegos con minijuegos, cursos que incluyen juegos como parte del material). En este modelo genérico se contemplan las necesidades y requisitos que deben cumplir, su estructura y sus componentes, y se establece la relación entre los objetivos educativos y los elementos que lo componen. En la parte final de la sección 4.1 se propone un meta-modelo que organiza los juegos de manera jerárquica en estructura de árbol, basado en SCORM Simple

Sequencing. Cada juego contribuye a alcanzar uno o varios objetivos de aprendizaje, definiendo modelos de progreso y evaluación basada en la agregación ponderada de los resultados de cada juego. A través de un caso de estudio (descrito en la sección 4.5.2) se implementa un caso de analíticas multi-nivel, con un panel de visualizaciones que permite explorar los resultados del árbol de actividades.

- **Desarrollo de analíticas ad-hoc.** Mediante la aplicación del proceso de GLA basado en el diseño de un LAM, se desarrollan con éxito analíticas ad-hoc con múltiples casos de uso (ver 4.2). Dicho proceso ha demostrado ser eficaz a la hora de ayudar a equipos multidisciplinares a conectar los componentes de un sistema de analíticas con el juego de manera exitosa. Cabe destacar que la disponibilidad de unas visualizaciones por defecto ha facilitado la comprensión por parte de los usuarios finales del potencial de las analíticas y fomentado la discusión y colaboración dentro de los equipos multidisciplinares para detallar el LAM específico y sus visualizaciones ad-hoc. El uso de xAPI ha facilitado la definición e implementación de los LAM. También ha permitido la reutilización de componentes genéricos de un sistema de analíticas y agilizado el desarrollo, por ejemplo, reutilizando un Learning Record Store (LRS) existente como almacén para las analíticas del juego. En todos los casos, este proceso ha conseguido implementar GLA de manera exitosa, simplificando su aplicabilidad y, por tanto, mejorando la adopción de analíticas de aprendizaje en los juegos serios.
- **Integración de Simva con uAdventure.** Como demostración de su aplicación y para simplificar la creación de juegos serios con analíticas por defecto, Simva se ha integrado con la herramienta de autoría de juegos serios uAdventure aplicando un LAM genérico (ver 4.2 y 4.3.3). Esta integración se basa en una serie de asistentes de alto nivel que permiten personalizar el LAM para cada juego concreto. Esta integración con uAdventure permite crear, con poco esfuerzo para el autor, juegos que tienen integradas un conjunto mínimo de analíticas y su recopilación automatizada y centralizada a través de Simva. Además, Simva proporciona análisis y visualizaciones sobre este LAM genérico, permitiendo hacer una explotación básica de los datos. Esta integración se estudia a través del juego de Primeros Auxilios (ver 4.5.1).

El modelo de integración entre Simva y uAdventure está basado en la aplicación del perfil de aplicación de xAPI para juegos serios (xAPI-SG) desarrollado en el grupo e-UCM. La aplicación xAPI y xAPI-SG permite desacoplar el juego creado con uAdventure de Simva, de modo que es posible utilizar cualquier otra herramienta compatible con el estándar xAPI para almacenar las analíticas o para procesar las analíticas recopiladas. Por ejemplo, los datos recopilados con Simva pueden ser procesadas con la herramienta TxMon (<https://github.com/e-ucm/t-mon>), que es una herramienta genérica (basada en Jupyter Notebook) de procesamiento de trazas xAPI compatible con el perfil xAPI-SG.

- **Simplificación de los procesos de validación formal de juegos serios.** El proceso de validación formal de un juego serio tradicionalmente utiliza cuestionarios pre-post, difíciles de relacionar con la información de analíticas. Este proceso se puede simplificar y automatizar mediante el uso de nuestro modelo y de las herramientas desarrolladas. Simva gestiona el acceso (anónimo) al juego, orquesta las actividades de los participantes en un experimento (pre-test, sesión de juego, post-test), recopila los datos de ambas encuestas y de la actividad del juego bajo un identificador único anónimo por participante, reduciendo drásticamente el esfuerzo y los costes asociados a la validación. Mediante el desarrollo de la plataforma Simva (ver 4.3.1 y 4.3.2) se facilita la creación de un diseño experimental (estudio en Simva) basado en actividades, que permite la creación de actividades de cuestionarios y actividades de juego con soporte a analíticas. Además, Simva permite la monitorización del progreso de los participantes en cada una de las actividades (completado, en progreso) y proporciona un almacén de información unificado vinculando todos los datos de un usuario proveniente de las diferentes fuentes (cuestionarios, datos de interacción de un juego). Simva también controla el progreso del participante en cada una de las actividades durante el propio experimento con el objetivo de poder monitorizar y ayudar a los usuarios si fuera necesario. El uso de Simva en varios casos de estudio (ver 4.5.3) ha simplificado la validación de juegos serios para los investigadores de manera exitosa.
- **Automatización de los experimentos de validación de juegos serios.** Es posible automatizar los experimentos de validación de juegos serios mediante la integración de Simva en una herramienta de autoría. Ampliando el trabajo

realizado para integrar analíticas (y un LAM genérico) en uAdventure, se ha extendido la integración con Simva (ver 4.3.3), con el objetivo de gestionar el diseño experimental desde el propio uAdventure. Esta integración se ha basado en la creación de un nuevo asistente orientado a configurar los datos mínimos para probar y validar los juegos, esto es, el número de participantes y los cuestionarios pre-post a utilizar. El proceso automatizado de validación de un juego serio mediante uAdventure y Simva se ha puesto en práctica en varios casos de estudio (ver 4.5.3) donde alumnos desarrollan un juego serio utilizando uAdventure y realizan su validación mediante Simva de manera automatizada

- **Implementación basada en estándares e-learning.** Además del uso de xAPI, el soporte de otros estándares educativos orientados a la interoperabilidad como LTI o CMI-5 simplifican las tareas de distribución de un juego serio y de integración de una plataforma de GLA en otras herramientas educativas ya existentes como un Learning Management System (LMS) como Moodle, Sakai, etc. Mediante la aplicación de estos estándares, se minimiza el coste de utilizar juegos serios y GLA, agilizando la experiencia de usuario y facilitando su adopción. Mediante la implementación de estos estándares en Simva (ver 4.4) es posible
 - Integrar Simva en un LMS como una herramienta externa utilizando LTI, minimizando la carga administrativa de tener que crear las cuentas y los permisos de usuario, como paso previo a llevar a cabo un experimento con usuarios, permitiendo lanzar una sesión de juego que, de manera transparente para el usuario, incluya la recolección y almacenaje de analíticas de aprendizaje
 - Almacenar y desplegar a través de Simva juegos serios empaquetados como unidades de asignación en terminología CMI-5 (que pueden ser generados con uAdventure) en un LMS. Pese a que CMI-5 nació con el objetivo de facilitar la integración de experiencias educativas que generen analíticas de aprendizaje compatibles con xAPI, su grado de adopción es todavía bajo. En este sentido, Simva facilita esta compatibilidad, proporcionando una interfaz LTI al LMS y la compatibilidad CMI-5 con el juego educativo. Simva no solo almacena las analíticas xAPI, sino que

permite configurar otro LRS externo, de modo que permite hacer de puente entre diferentes herramientas educativas.

- Integrar en Simva otras herramientas externas compatibles con LTI, como otros LMS, permitiendo realizar la validación de estas herramientas o algún contenido específico de ellas (por ejemplo, un curso) a través de Simva

5.2. Aplicaciones Software

La plataforma implementada Simva se encuentra publicada con licencia de código libre en el repositorio de GitHub: <https://github.com/e-ucm/simva>, aunque para facilitar su ejecución, se recomienda utilizar el proyecto anexo que incluye la infraestructura desarrollada para trabajar con Simva.

La infraestructura se encuentra publicada en el repositorio de GitHub: <https://github.com/e-ucm/simva-infra>. Esta infraestructura integra el resto de servicios necesarios para el correcto despliegue y funcionamiento con la ayuda de un pequeño *script* ejecutable que automatiza su instalación. Esto simplifica la instalación del software como un proyecto llave en mano.

Por otra parte, el trabajo realizado en uAdventure, con su soporte de GLA y su integración con Simva, forma parte del código de dicha herramienta la cual se encuentra publicada en el repositorio: <https://github.com/e-ucm/uAdventure>.

5.3. Trabajo futuro

Una vez completado el trabajo de la tesis, habiendo alcanzado los objetivos planteados para la misma, se han identificado algunas futuras líneas de investigación como continuación del presente trabajo. En este sentido, se presentan tres líneas de trabajo, relacionadas con la mejora de (1) la adopción de analíticas de aprendizaje, (2) la validación de juegos serios, (3) la propia herramienta Simva, dentro de cada una las cuales se contemplan diferentes actividades.

En primer lugar, para mejorar la adopción de las analíticas de aprendizaje en juegos, se han identificado las siguientes actividades:

- Si bien en las analíticas de juegos los desarrollos *ad-hoc* son los más habituales, se puede contribuir a la adopción de analíticas genéricas mejorando el soporte de

estándares (por ejemplo, para más vocabularios de xAPI). El presente trabajo ha utilizado como hilo conductor en el desarrollo dar soporte al perfil de aplicación de juegos serios xAPI-SG. Sería deseable extender el soporte para otros perfiles de aplicación de xAPI, incluyendo más tipos de análisis y visualizaciones para ellos, lo que permitiría extender los resultados obtenidos en el presente trabajo de tesis a otras áreas de aplicación dentro del contexto educativo. Además, este proceso podría beneficiarse no solo de la experiencia metodológica para dar soporte a xAPI-SG, sino también reutilizar y extender el software desarrollado como parte de este trabajo de tesis.

- Aunque el análisis realizado con la plataforma de GLA del proyecto H2020 RAGE Analytics nos ha permitido disponer de un entorno capaz de procesar y mostrar trazas procedentes de juegos en tiempo real, este tipo de infraestructuras son costosas de mantener y requieren de un servidor o clúster capaz de procesar grandes cantidades de datos. Parte del trabajo realizado en Simva ha permitido limitar el coste de desarrollo y mantenimiento de una plataforma de soporte a GLA, con el objetivo de facilitar a los investigadores extraer conclusiones de los datos aplicando técnicas de *data-science* mediante herramientas como Jupyter Notebooks. En este sentido Simva se convierte en una fuente de datos compatible con la API S3 de AWS que puede ser consumida desde la mayoría de herramientas de *data-science*. Basándose en esta capacidad, la herramienta TxMon, permite proporcionar un conjunto de análisis y visualizaciones por defecto para trazas de xAPI compatibles con el perfil de aplicación de xAPI-SG como un conjunto de Jupyter notebooks. El desacoplamiento entre la recopilación y el procesamiento de las GLA permite la experimentación y exploración en ambas áreas de manera más ágil con el objetivo de probar nuevas aproximaciones y aplicaciones de las GLA. Estos desarrollos iniciales se deberían completar, extender y madurar para obtener herramientas fiables y que fueran más útiles para todos los implicados en el proceso (desarrolladores, investigadores, profesores, e incluso los propios alumnos).

En segundo lugar, si bien las analíticas y los cuestionarios pre-post han demostrado ser herramientas que juegan un papel fundamental en la validación de los juegos serios, se plantean las siguientes actividades relacionadas con la mejora de la validación de los juegos serios:

- Además del método de validación basado en cuestionarios pre-post, a lo largo del estado del arte se presentan otros tipos de metodologías para la validación como la investigación de archivo (reutilizando datos existentes pero facilitando su incorporación en el sistema), o la realización de entrevistas individuales o grupales (facilitando mecanismos para sistematizar el proceso de entrevista), que aunque sean menos frecuentes en el ámbito de los juegos serios, pueden resultar más apropiadas en algunos casos, proponiendo como trabajo futuro dar soporte a más metodologías de validación en Simva.
- Automatizar aún más el análisis comparativo en Simva para aquellos experimentos basados en cuestionarios pre-post. Si bien se han realizado validaciones utilizando Simva con éxito, el proceso final de contrastar las encuestas y medir el incremento sigue siendo labor del investigador por lo que se podría avanzar en una mayor automatización y soporte a los investigadores.
- Uso de técnicas de inteligencia artificial. Por ejemplo, entrenar un modelo predictivo mediante machine learning utilizando el resultado del pre-post de validación junto a los registros del juego para posteriormente ser capaces de predecir el efecto del juego serio utilizando exclusivamente las analíticas de aprendizaje recopiladas durante de la sesión de juego (es decir, sin necesidad de realizar los cuestionarios). Estos modelos predictivos ayudarían además a que los evaluadores puedan identificar una sesión de juego anómala que permita realizar una intervención temprana. El auge de la inteligencia artificial y el machine learning supone una oportunidad para explotar el potencial de esta tecnología con el objetivo de simplificar las tareas de análisis a los investigadores, e incluso a los profesores (por ejemplo, para la evaluación de sus estudiantes)

Para terminar, todas las líneas de trabajo anteriormente descritas proponen mejoras que se construyen sobre la herramienta actual Simva. Las actividades que se proponen son las siguientes:

- Aumentar los casos de prueba implementados con Simva, ampliando los juegos educativos probados y validados con la herramienta, mientras se refina y simplifica aún más la experiencia de usuario en la integración con los LMS.
- Utilizar Simva en la asignatura de Juegos Serios del Grado en Desarrollo de Videojuegos como complemento a la aplicación de uAdventure. El objetivo es

proporcionar las herramientas necesarias a los estudiantes para que puedan diseñar y plantear un LAM para un juego serio, desarrollar el juego con uAdventure y posteriormente ponerlo en práctica mediante Simva.

Capítulo 6. Artículos presentados

En este capítulo se incluyen los artículos aportados como parte fundamental de esta tesis doctoral.

6.1. Game Learning Analytics is not informagic!

6.1.1. Cita completa

Iván José Pérez Colado, Cristina Alonso Fernández, Manuel Freire Morán, Iván Martínez Ortiz y Baltasar Fernández Manjón (2018). **Game learning analytics is not informagic!** In *2018 IEEE Global Engineering Education Conference (EDUCON)*. IEEE, 1729-1737 <https://doi.org/10.1109/educon.2018.8363443>

6.1.2. Resumen original de la publicación

Game learning analytics has a great potential to provide insight and improve the use of games in different educational situations. However, it is necessary to clearly establish what the learner's requirements are and to set realistic expectations about the learning process and outcomes. Application of game learning analytics requires pedagogically informed policies that settle the learning goals and relate them to analysis and visualization; and a supporting infrastructure that provides the mechanism on top of which it is executed. Both concerns can be addressed separated: on the one hand, there is a Learning Analytics Model (LAM) which describes how the analysis is carried out, interpreted as learning, and presented to stakeholders; and on the other hand, an underlying analytics system that concentrates on performance, security, flexibility and generality. An important advantage of this separation is that it allows LAM authors to concentrate on their area of expertise, limiting their exposition to the actual mechanism used underneath. However, LAMs built for a single game fail to account for the frequent case where games and their analytics are aggregated into larger, overarching plots, games or courses. This work describes an extension to an existing game learning analytics system, used in RAGE and BEACONING H2020 projects, which manages multilevel analytics through improvements to both policy and mechanism; and introduces meta-Learning Analytic Models, which characterize learning in hierarchical structures.

Game Learning Analytics is not *informagic*!

Ivan Perez-Colado, Cristina Alonso-Fernandez, Manuel Freire, Ivan Martinez-Ortiz, Baltasar Fernandez-Manjon
Dept. Software Engineering and Artificial Intelligence
Universidad Complutense de Madrid, Facultad de Informática
C/ Profesor Jose Garcia Santesmases, 9 28040 Madrid, Spain
{ivanjper, crisal03}@ucm.es {manuel.freire, imartinez, balta}@fdi.ucm.es

Abstract—Game learning analytics has a great potential to provide insight and improve the use of games in different educational situations. However, it is necessary to clearly establish what the learner’s requirements are and to set realistic expectations about the learning process and outcomes. Application of game learning analytics requires pedagogically informed policies that settle the learning goals and relate them to analysis and visualization; and a supporting infrastructure that provides the mechanism on top of which it is executed. Both concerns can be addressed separated: on the one hand, there is a Learning Analytics Model (LAM) which describes how the analysis is carried out, interpreted as learning, and presented to stakeholders; and on the other hand, an underlying analytics system that concentrates on performance, security, flexibility and generality. An important advantage of this separation is that it allows LAM authors to concentrate on their area of expertise, limiting their exposition to the actual mechanism used underneath. However, LAMs built for a single game fail to account for the frequent case where games and their analytics are aggregated into larger, overarching plots, games or courses. This work describes an extension to an existing game learning analytics system, used in RAGE and BEACONING H2020 projects, which manages multilevel analytics through improvements to both policy and mechanism; and introduces meta-Learning Analytic Models, which characterize learning in hierarchical structures.

Keywords—learning analytics; serious games; learning analytics models; multi-scale games; xAPI;

I. INTRODUCTION

One of the main goals of Learning Analytics (LA) is to allow educators to better understand how learners apply domain knowledge when using an e-learning system, helping educators to improve educational outcomes [1]. However, learning analytics needs several inputs at different levels before it can provide such understanding. The definition of these inputs can be described in a Learning Analytics Model (LAM). In the particular case of e-learning using serious games (SGs), which allows more immersive and engaging experiences for students than traditional contents [2], LAMs have the potential for providing very good insights into learning – at a corresponding cost in terms of LAM complexity. However, it is necessary to clearly establish what the LAM requirements are and to set realistic expectations. Too often, users of analytics expect them to be presented with such insights based exclusively on “shallow” interaction data, assuming that the system can infer the educational game design and what exactly counts as “learning” on its own. This is not game learning analytics: it is an unrealistic expectation that we call *informagic*.

Although many stakeholders may benefit from the information obtained from game learning analytics (such as game developers, game designers, and researchers), this paper mostly focuses on teachers, due to their central role in guiding the learning process of students. We consider two main approaches towards providing analytics to teachers: first, when delivered at near-real time, analytics can allow teachers to

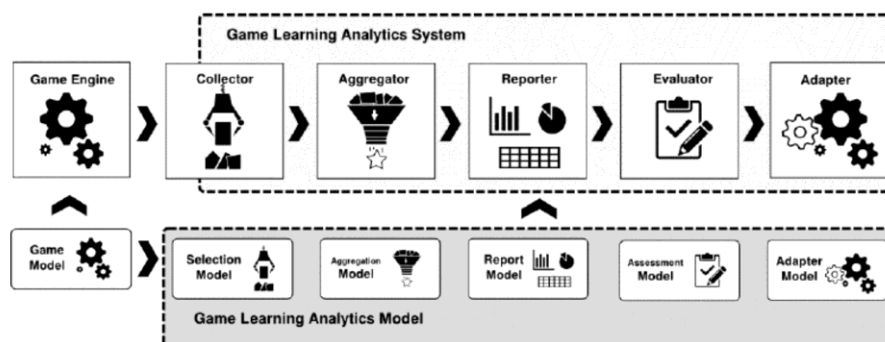


Figure 1. General architecture of a Game Learning Analytics Platform including the Learning Analytics Model and the Learning Analytics System [5].

maintain awareness of student actions in-game, and then help students that get stuck while playing [3]; and second, asynchronously, analytics can provide additional information to evaluate students and reflect or elicit comments on what was learnt during the game-play session [4]. Different stakeholders may be interested in other types of analytics; for example, managers could be interested in comparing the performance of groups against each other, rather than educational interventions. LAMs describe, for each stakeholder, what interaction information is gathered, how that interaction information is related to the educational design and how that correlation is interpreted and presented in the corresponding dashboards.

Educational games can be part of a more complex game learning approach. For example, a game may contain minigames; or a geolocalized game may launch different games depending on the location. Games that are not fully stand-alone, and are instead a part of a greater whole, require additional LAM-related information that describes how the users' learning and progress within each of these parts is to be considered into the context that contains them. The next section describes in greater detail the concept and advantages of LAMs, together with other formulations of the same basic idea. Section III then deals with meta-LAMs, that is, LAMs for multiple levels (multilevel analytics); and describes changes to analytics systems that allow meta-LAMs to be supported. Section IV details changes to data collection, and Section V focuses on analysis and visualization for LAMs and meta-LAMs. Finally, Section VI provides conclusions and future work.

II. LEARNING ANALYTICS MODEL

As described in [5], a LAM provides the models on how information should be tracked, aggregated and reported to a Learning Analytics System (LAS). This relation inside a Game Learning Analytics platform is depicted in Figure 1. LAMs isolate learning analytics users from the implementation details of the underlying LAS. This allows both systems to evolve independently as long as the interface between the model and the system is well-defined and represents policy and mechanism respectively.

The LAS underlying a serious game needs zero knowledge of game concepts or learning; and instead concentrates on blindly applying the LAMs for each of the games that report to it. This allows building generic analytics systems, where the allowable type of analytics is only constrained by the expressiveness of LAMs. Greater expressive power has the downside of increasing the difficulty of authoring in LAMs and supporting them at execution time. On the other hand, simplistic LAMs may prove too constraining to provide useful feedback. A reasonable compromise between both extremes is to provide a simple default LAM that delivers basic information on game completion and progress; and allow analytics' users to replace or extend it on a case-by-case basis. This way, (crude) analytics are available at zero cost, but more advanced insights can be obtained by investing additional effort where needed.

In [6], Chatti et al. propose a LAM-compatible reference model for learning analytics based on four dimensions: the data the system gathers, manages and uses for analysis (*what?*), the



Figure 2. The LAM-compatible Learning Analytics Reference Model proposed in [6] describes four dimensions: data, objectives, methods and stakeholders.

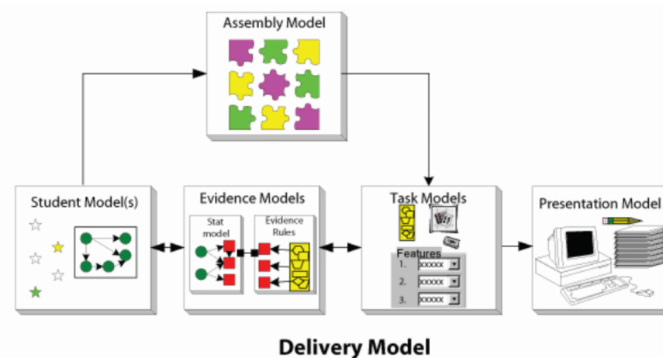


Figure 3. The Conceptual Assessment Framework layer of the evidence-centered assessment design described in [7] relies on student, evidence, task, assembly and presentation models.

stakeholders targeted by the analysis (*who?*), the reasons for the system to analyze the collected data (*why?*), and the way in which the analysis is performed on the collected data (*how?*). This reference model is illustrated in Figure 2. LAMs as defined in this paper attempt to answer 3 of these 4 questions, although motivating the specific answers for a LAM is welcome but non-mandatory: LAM authors must provide an executable definition of *what*, *how*, and *who* for the LAS to work on; the *why*, however, is only useful as documentation, since the LAS cannot interpret it. Of course, well-described LAMs will be much easier to maintain and adapt to evolving needs.

A related framework that focuses on describing educational assessment can be found in [7]. This evidence-centered assessment design (ECD) defines the information that should be tracked, how this tracking should occur and how the resulting data should be interpreted. Of particular interest is the Conceptual Assessment Framework (CAF) layer, depicted in Figure 3, which links the variables in the student model (*what are we measuring*), the environment in which students complete the task in the task models (*where do we measure it*), the observations to be made depending on the purposes and context in the evidence model (*how do we measure it*), and the assembly model that merges multiple tasks into a single feature (*how much do we need to measure it*) [7]. While LAMs may opt for simple approaches to many of CAF's models, any CAF model can be represented by a suitable LAM.

The interplay between the LAM, the underlying analytics system, and some of the main stakeholders is depicted in Figure 4, where blue boxes represent the mechanism (LAS), and red boxes represent the policy as described in the LAM. When students play games or mini-games, the tracker component embedded in those games sends traces (in our proposal, as xAPI statements) that describe actions and outcomes, which are then collected by the analytics system for storage, analysis and visualization.

Completely defining a LAM for a particular game (or mini-game) requires several decisions, depicted in Figure 5 together

with the main stakeholder in charge of their definition and the expected results of each activity. The key decisions that a LAM must address are [8]:

1. **Learning goals** to be achieved in the game (e.g. specific knowledge, procedures, and tasks) must be defined before any other LA procedure. These learning goals will result in a specific learning design that, as pointed out by Bakharia et al. in [9], need to be linked to LA to provide it with a semantic frame of reference.
2. **Game goals** (e.g. tasks, levels) that correspond to the learning goals. The correspondence may not be bijective: a single learning goal may be furthered only once several game goals are completed (e.g. several levels to achieve some knowledge); while a single game goal may contribute to multiple learning goals (e.g. a level that teaches various skills). The set of game goals will result in a specific game design.
3. **Traces to be sent**, as defined by game developers based on the established learning and game designs. The information to be traced and sent by the game should therefore be necessary and sufficient to, once analyzed, inform of the degree to which the game goals are being met. In our proposal, traces must follow the xAPI-SG Model [10], as described later in detail (see the Data Collection Section). Notice that it is also possible to define traces with any information that may be of interest to stakeholders, even if this information is not specifically required to gain insight regarding learning goals.
4. **Analysis model**, defining how traces from step 3 should be analyzed and interpreted. In general, this will require keeping an updated estimation of the extent to which each learning goal is being met every time a new trace is received. The most valuable educational insight is obtained from analysis that take into account both the learning goals and how those goals were reflected into the game goals.

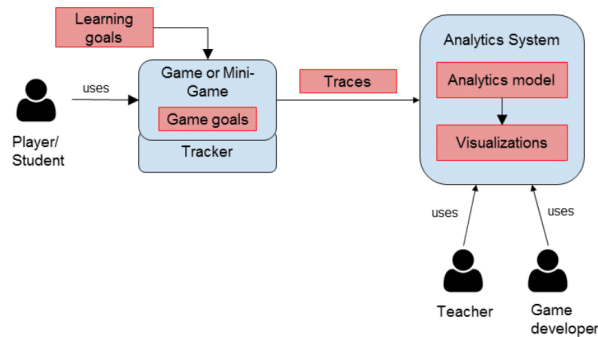


Figure 4. The Learning Analytics Model describes the information sent by the tracker, the analyses that will be performed and the visualizations later used to display results.

5. **Visualizations** that adequately represent the results of analysis, in the form of per-stakeholder dashboards. If a default LAM is being extended, then this LAM’s default visualizations may be sufficient, or require only minor modification. Otherwise, entirely new visualizations may be required.

While analytics is usually performed after the game has been played, this does not need to be the case, since many systems can update their visualizations in near-real time. This opens the door to real-time alerts and warnings, which once configured can notify teachers of possible problems or interaction opportunities that arise during game-play, such as a learner becoming stuck in a particular level or advancing much more quickly than

expected. From a LAM perspective, real-time alerts and warnings can be seen as a specific type of visualizations which only present themselves in dashboards when triggered, and can be specified together with traditional always-on visualizations when designing those dashboards. In this sense, alerts and warnings focus on providing immediate information for teachers to act on during game-play time.

A. LAM example

To further clarify the concept of a LAM, this subsection describes some parts of the LAM used for the First Aid Game [11], designed to teach players between 12 and 14 years old first-aid techniques in three particular situations: chest pain, unconsciousness and choking. The following paragraphs

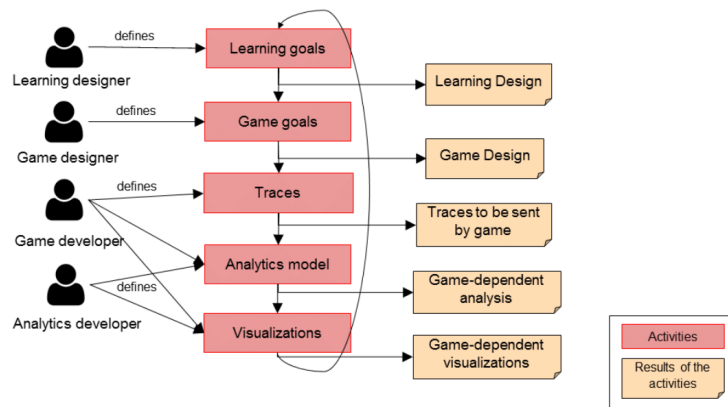


Figure 5. Learning Analytics Model steps, including the stakeholder in charge of each definition and the result of the activity carried out in each step.

describe a subset of the game's learning goals (LG), game goals (GG), traces to be sent, analyses and visualizations. This description covers only the chest pain situation, three of its main learning goals, and details of the consequent activities of the LAM that are of interest for those three learning goals in particular.

The main learning goal (LG1) of this game was to "*learn how to react in a first aid emergency situation*". There were also three learning goals related with each of the three game levels: for instance, for the first of them, the learning goal was to "*learn first aid techniques for chest pain situation*." (LG2). Another lower-level learning goal was to "*learn the emergency phone-number*" (LG3). These three learning goals were reflected in three game goals: "*successfully complete all three levels of the game*." (GG1, to satisfy LG1), "*successfully complete chest pain level*." (GG2, to satisfy LG2) and "*successfully answer the question about the emergency number*" (GG3, to satisfy LG3).

Among the traces defined to be sent by the game, the ones that corresponded to the previous learning and game goals are those related to progress in each level (given as a real number from 0 to 1) and to answers provided for specific questions. Three types of traces were sent for progress for each level: the *initialized* trace at the beginning, the *progressed* trace with each progress from 0 to 1, and the *completed* trace when level is finished; including the level's score as an extension [3]. For specific questions, a *selected* trace was sent, together with the actual option that was selected, and whether it was correct or not. A more in-depth description of what *initialized*, *progressed* and *completed* means can be found in the Section on Data Collection.

The analysis model of these traces defined GG2 (and therefore, LG2) as achieved if the *completed* trace of the level had an extension with a score greater than or equal to five points; so that the progress of the whole game increased by a third of the total with each successfully completed level. GG1 (and therefore, LG1) was considered achieved when all three levels had been successfully completed. Finally, GG3 (and therefore, LG3) was achieved once the *selected* trace of the specific question had a successful result.

To visually communicate the results for LG1 and LG2 to teachers, the default visualization that reports progress in each of the three levels (*completables*) and in the complete game (also a *completable*), together with the scores obtained, are considered to be enough. The visualization that reports progress consists of a bar chart with all students in the x-axis and, for each of them, and in the y-axis, four bars representing their progress in each of the three levels and the complete game; with each bar ranging from empty (not started) to full (level or game completed). To communicate LG3, another bar chart showing the answers (y-axis) grouped by students (x-axis) to each question (individual bars within each student) was used.

To help teachers keep control of the class and help students having difficulty completing each level, a personalized warning with the message "*the user has failed the Chest Pain game mode*" was triggered whenever a student completed the level but failed it, to indicate that the knowledge had not yet been acquired. Another warning was set with the message "*the user has failed the Emergency Number question*" when the specific

question was answered incorrectly. Apart from those specific warnings, general warnings indicated which students had been inactive for over a given period of time, or could be stuck in a particular section could be easily enabled. In this particular case, students were detected as stuck when the teacher-specified expected level completion time was exceeded. A more robust "stuck" indication could have used the degree of deviation from the norm, once a sufficient number of gameplay sessions was collected.

III. META-LAM

When a game is composed of multiple games, their individual LAMs that focus on each of these games in isolation fails to describe the larger game as a whole. For example, a geolocalized game may involve launching different context-appropriate games depending on the location; and it makes perfect sense to ask for analytics on the overall progress in the overarching game. However, once several games are part of a larger aggregation, the problem of progress and completion as indicators arises: if there are several paths along the game, and some of them may require complex conditions that may or may not occur, it becomes very difficult to measure how much of the game remains to be completed. Analytics for composite games is also expected to include dashboard visualizations that show global information about one student, across several classes and/or games. To address the need for aggregating results, progress/completion calculation, and new aggregate-aware meaning, we have proposed the use of Learning Analytics Models for multi-scale games: meta-LAMs that stitch together the individual game LAMs into a larger whole. For example, returning to the geolocated game example, we can envision a main game that proposes tasks to players that require them to reach certain areas; which, when reached, cause different mini-games or activities to be launched. Each game or activity would retain its own LAM, and the meta-LAM would describe how these are to be joined together.

Any proposed meta-LAM, similarly to particular LAMs for games or mini-games, should define a set of learning goals to be achieved by the whole game, how they are going to be achieved by its different games or components, how the information is going to be traced and what analyses and visualizations should be performed on the data to have a complete understanding of the learning process in the complete platform. Once defined, the meta-LAM should allow access to essential information such as overall student progress in the general higher-level game.

For this meta-LAM to be correctly defined, both the hierarchy and the information flows must be clearly specified. The information flow includes transfer, aggregation and analyses that involve more than one level. For instance, a typical approach will be to have a tree structure where the general game is decomposed in several mini-games (as leaves) [12]. These leaves will then contribute towards the progress of their parent node as they are completed.

A. Meta-LAM structure proposal

Although the meta-LAM will depend on the structure of the hierarchy of games, a standard meta-LAM can be considered for certain general structures, and in particular games that are structured as simple hierarchies (trees). In the case of graph

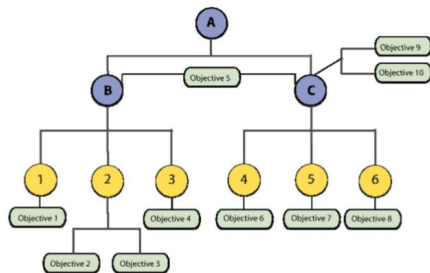


Figure 6. Activity Tree with Learning Objectives shared. Figure retrieved from ADL SCORM 2004 4th Edition Sequencing and Navigation (SN) specification [13].

structures, a tree overlay could be defined to allow the use of this standard meta-LAM. However, meta-LAM should not be limited to this particular scenario, and be flexible enough to allow any reasonable model to be implemented. Notice that when it comes to systematizing LAMs for multi-scale games, we have found no standard or widely accepted model on the literature that covers this issue.

After an exhaustive revision of specifications used on similar fields like e-learning, we decided to build upon a greatly simplified version of the IMS Simple Sequencing specification as described in SCORM 2004 4th Edition Sequencing and Navigation (SN) [13] specification. SCORM Simple Sequencing SN defines an Activity Tree (AT) as the main hierarchical structure of learning activities, which may or may not correspond to the actual internal organization of activities. Clusters are defined as learnings activities with immediate sub-activities, and they can contain both leaf activities (which are not clusters) and other clusters.

Learning activities in SN (activities for short) are nodes in the AT. Activities have completion (with start and finish) and mastery conditions, and they can contain sub-activities to any depth. Any effort to complete an activity is called an attempt. Attempts may be suspended and later restarted, while abortion removes all attempt information. Learning Objectives have no specific semantics: they may refer to competencies, masteries, shared values, etc. There is no direct correspondence between activities and objectives: activities may have more than one objective associated and multiple activities may reference the same objective. Figure 6 shows an example AT with the learning activity "A" as root. Some learning objectives are also defined: all objectives are local to their associated activity, except objective 5 that is shared between activities "B" and "C" [13]. Some limit conditions may be established for activities to determine when they are not allowed to be delivered. In SN, the only mandatory condition is the Attempt Limit, which will correspond to the maximum number of attempts to complete an activity.

In our proposal, the AT will correspond to our meta-LAM, which is the highest level structure (i.e. the composed multi-

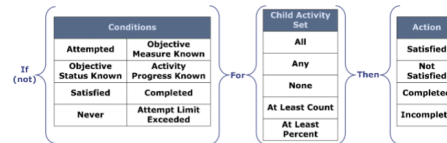


Figure 7. Rollup Rule Child Activity Set, Conditions and Actions. Figure retrieved from ADL SCORM 2004 4th Edition Sequencing and Navigation (SN) specification [13].

scale game or the geolocalized-game both with mini-games). Activities in the AT will correspond to games, so that every activity will have an associated LAM as long as it has an associated learning content to track and analyze. If so, the associated LAM for those activities will define the conditions for its successful (or unsuccessful) completion. Notice that a single LAM may be sufficient to cover several games: it is very frequent to find related game activities, or even the same activity with slightly different configuration data being used in multiple places.

Learning Objectives, as defined in Simple Sequencing, could easily correspond to learning goals to be achieved via playing. As with activities and objectives in Simple Sequencing, there is no direct correspondence in our proposal between games and learning goals. Simple Sequencing defines *Rollup* as the process of passing information from children nodes to parent nodes in the AT. Rollup rules define how progress for cluster activities is to be evaluated, and consist of a set of child activities to be considered, conditions to be evaluated against them and actions to be taken, as depicted in Figure 7. By default, all children activities are included in parent rollup, unless they are not tracked, or do not contribute to rollup (not mentioned in any activity set). Leaves activities are not affected by rollup rules. Weighted combinations of the information from child activities are used to determine progress, completion, and objectives met.

For the meta-LAM definition, rollup rules are essential as they define how information from single activities (games) is to be aggregated up for general multi-scale games. Following the Simple Sequencing proposal, we again consider that parent game nodes could define their progress and objectives as linear weighted combination of the progress and objectives achieved in the corresponding children game nodes.

This hierarchy also allows for multi-level status storage. This is required if players start a multi-scale game that launches other games in turn (as sub-activities in the AT), as the general status of the player, across all levels, should be stored. Similarly, multi-game variables may also be required. If they are global in the AT, they can be used and changed across games. Otherwise, communication of variables between games will be required for consistency. That means that there should be a clear and unique way to identify what a specific user is doing in the different games.

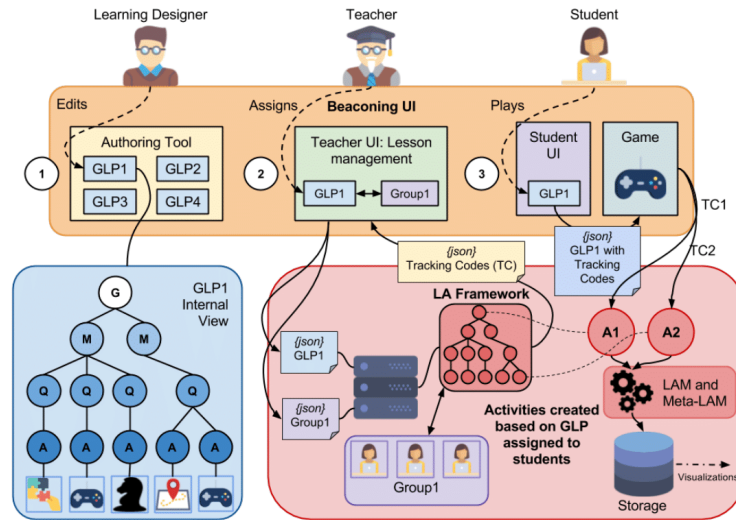


Figure 8. H2020 Beaconing Project structure proposal to apply LAM and meta-LAM using a model similar to the one proposed in SCORM Simple Sequencing SN. Game designers and data visualizations (dashboards) are omitted in this representation for simplicity.

The previously described structure for the meta-LAM is being considered for the case of the H2020 BEACONING Project, where a hierarchical structure of games and mini-games is required. In particular, learning designers define gamified lesson plans (GLP) that consist of missions, quests and activities. These activities may be games or mini-games. Figure 8 describes the first stages of the process: analytics definitions based on LAMs and meta-LAM, GLP creation and assignment to students, and students playing. For clarity, the role of game designers is being omitted in the figure.

IV. DATA COLLECTION

As defined in the previous chapters, the data collected from SGs is to be sent via traces for its later analysis and visualization. In our proposal, these traces must follow the xAPI standard. The Experience API (xAPI) is an e-learning standard used to track information from learning activities using *statements* composed of three main fields: an actor, a verb and an object. Additional fields may be included such as the context of the activity or its results.

Together with ADL [14], leaders of the community that created xAPI, the e-UCM Research Group developed the xAPI Serious Games Vocabulary (xAPI-SG) [15]. This profile, defined in detail in [10], defined concepts such as *completables* (e.g. tasks, quests or mini-games), *alternatives* and general *variables* to track interactions in the domain of SGs. It provides

a standard set of verbs (e.g. *initialized*, *progressed* and *completed* for *completables*; *selected* or *unlocked* for *alternatives*; etc.), activity types (e.g. *area*, *enemy*, *level*, *question*, etc.) and extensions (e.g. *progress*) general enough to cover the interactions of the player in most SGs but concrete enough to provide meaningful information of interest for the possible stakeholders. With its implementation in the xAPI standard, the interaction model provides a general, game-independent format for traces to model most of the players' interactions within a SG.

As required by the H2020 BEACONING project, an extension of the xAPI-SG model that vocabulary for geolocated games has been proposed, which includes verbs such as *entered*, objects such as *area* or *point-of-interest*, and an *orientation* extension that can be added to specific xAPI-SG statements.

While game developers using the xAPI-SG Profile can create their own trackers, e-UCM has also released a range of open-source tracker implementations for different platforms. They include: JavaScript¹, C#², Dot NET³, Unity (adapted from the C# tracker)⁴, Unity (specifically developed in Unity)⁵ and LibGDx (Java) (outdated)⁶.

V. ANALYSIS AND VISUALIZATION

This section briefly introduces the analytics system developed, as part of the H2020 RAGE Project, and tested,

¹ <https://github.com/e-ucm/js-tracker>

² <https://github.com/e-ucm/csharp-tracker>

³ <https://github.com/e-ucm/dotnet-tracker>

⁴ <https://github.com/e-ucm/unity-tracker>

⁵ <https://github.com/e-ucm/unity-tracker/tree/0.5.0>

⁶ <https://github.com/e-ucm/libgdxd-tracker>

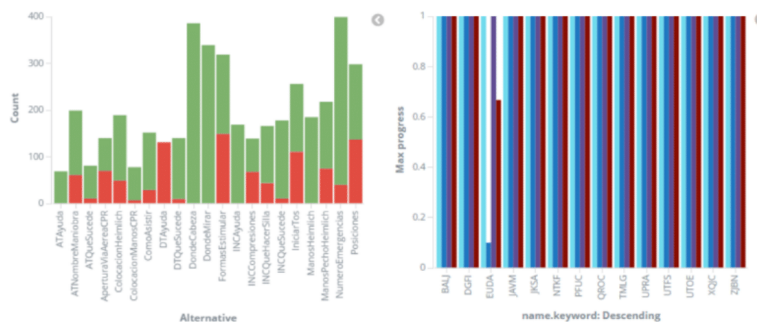


Figure 9. Bar chart of number of correct and incorrect answers in each alternative (left) and bar chart with progress (with 1 representing 100%) in each level and in the complete game by each player (right). Names of players have been replaced with random identifiers.

extended and improved as part of the H2020 BEACONING Project. Full documentation of the current state of the system can be found at the Official GitHub Wiki Page⁷. The meta-LAM described in this work is, as of this writing, currently being implemented into this system, with validation soon to follow.

When it comes to analyses and visualizations, this LAS provides a fixed set of analyses and visualizations corresponding to the “default LAM”, which is used for all new games until replaced or extended with a game-specific LAMs. While the default LAM is mainly focused on teachers, some of its default visualizations are targeted at game developers. The current set of analyses and visualizations for teachers provides information including scores, progress, alternatives selected, correct and incorrect answers, active sessions, and games started and completed; which are depicted in different types of visualizations, such as bar, line, or pie charts. Figure 9 shows two of these visualizations: number of correct and incorrect alternatives chosen by each player, and the maximum progress achieved by each player in each of 3 game levels and in the complete game; both for the First Aid Game example from Section II.

The previously defined set of analyses and visualizations is provided for both teachers and game developers. Meanwhile, students may require a different type of feedback that does not break the game flow and does not make them feel observed while they play (which may change their behavior). To avoid breaking the game flow of the students, the system provides an API that allows games to query analytics directly. Games wishing to use this API to provide in-game visualizations for students only have to make the corresponding calls to the analytics system; or, if they rely on the open-source tracker implementations described in section IV, simply request the information through the tracker itself. Finally, learning dashboards for students also present problems such as not providing recommended actions or focusing more on competition rather than on goal achievement [16]

The system allows game-specific LAMs for particular games. Their development must involve game developers, which can link in-game actions (mechanics) with game goals and to the information that is to be analyzed and displayed. As previously stated, the most valuable educational insight is obtained from analyses that take into account both the learning goals and how those learning goals are related to the game goals; and this is simply not possible for the default LAM, which must necessarily be generic. Note that game-specific LAMs do not need to address all of the steps depicted in Figure 5 to be useful; for example, choosing a visualization that makes player actions easier to relate to learning goals, without changing incoming data or how it is analyzed, can already provide significant improvements to insight as compared to a generic visualization.

Mirroring the distinction between the default LAM and game-specific LAMs, work is currently under way to provide a default meta-LAM for composite games, based on the structure proposed in Section III, and reflecting commonly used metrics such as progress or score; while particularly composite games will still be able to replace or extend it with their own specific meta-LAMs. These analyses must be aligned with the hierarchical structure of the game, and describe exactly how aggregations are to be performed between the different levels of the hierarchy. In the simple but common case of a tree structure, parent nodes may define their fields as the aggregation or average of their children’s fields. However, other options are possible; and whichever is chosen, must be unambiguously described in order to correctly define analyses and visualizations at the meta-LAM level.

VI. CONCLUSIONS AND FUTURE WORK

Learning Analytics Models encapsulate and describe how data from serious games is to be gathered, analyzed and displayed. When used correctly, they provide relevant information about the students’ learning experience to teachers and other stakeholders; feedback that is key to guide the correct application of serious games in education. Effective LAMs should be developed based on the learning design, and their

⁷ <https://github.com/e-ucm/rage-analytics/wiki>

development should ideally start even before that of the game itself, clearly establishing their requirements and their (realistic) expected outcomes. If the information from interactions obtained is limited and does not relate to the learning design, expecting analytics to provide insights into learning is to believe, as this paper's title states, in *informagic*.

Development of LAMs should follow a linear process where each step bases its development on previously established definitions and outputs, and preventing cases where late steps such as analysis or visualization are expected to provide learning insights from insufficient or context-free data.

While standard LAMs fit the needs of a single game, they fail to provide adequate insight in cases where games are part of more complex structures. For these cases, meta-Learning Analytics Models need to be defined to fulfill the needs of those structures. In our meta-LAM structure proposal, based on SCORM Simple Sequencing, games follow a tree structure where they can launch other mini-games, which can then launch further mini-games in a hierarchical fashion. Learning goals may be shared across different sub-games; and similarly, single sub-games may help achieve multiple learning goals. Aggregation from sub-games to their parent games consists of linear weighted combination of the fields of their children, which may include scores or progress.

Future work includes the improvement of the current default LAM, including both its default analyses and visualizations; and implementing and testing the proposed meta-LAM structure as part of current projects that require support for complex hierarchies of games and mini-games. Results of the assessment could also be used for adaptation in games, as authors have pointed out in [17].

Finally, it is also important to notice that the adoption of learning analytics can greatly benefit from its direct integration into game authoring tools that simplify costs and knowledge required for its application [18].

ACKNOWLEDGMENT

This work has been partially funded by Regional Government of Madrid (eMadrid S2013/ICE-2715), by the Ministry of Education (TIN2013-46149-C2-1-R, 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] D. Gašević, S. Dawson, and G. Siemens, "Let's not forget: Learning analytics are about learning," *TechTrends*, vol. 59, no. 1, pp. 64–71, 2015.
- [2] T. M. Connolly, E. A. Boyle, E. MacArthur, T. Hainey, and J. M. Boyle, "A systematic literature review of empirical evidence on computer games and serious games," *Comput. Educ.*, vol. 59, no. 2, pp. 661–686, Sep. 2012.
- [3] M. Liu, J. Lee, J. Kang, and S. Liu, "What We Can Learn from the Data: A Multiple-Case Study Examining Behavior Patterns by Students with Different Characteristics in Using a Serious Game," *Technol. Knowl. Learn.*, vol. 21, no. 1, pp. 33–57, Apr. 2016.
- [4] C. S. Loh, Y. Sheng, and D. Ifenthaler, *Serious Games Analytics*. Cham: Springer International Publishing, 2015.
- [5] M. Freire, Á. Serrano-Laguna, B. M. Iglesias, I. Martínez-Ortiz, P. Moreno-Ger, and B. Fernández-Manjón, "Game Learning Analytics: Learning Analytics for Serious Games," in *Learning, Design, and Technology*, Cham: Springer International Publishing, 2016, pp. 1–29.
- [6] M. A. Chatti *et al.*, "Learning Analytics: Challenges and Future Research Directions," *E-Learning Educ.*, no. 10, 2015.
- [7] R. Mislevy, "Evidence-Centered Assessment Design: Layers, Structures, and Terminology," *Princ. Assess. Des. Inq. Tech. Rep.* 9, vol. 9, no. July, p. 46, 2005.
- [8] M. Freire, I. Martínez-Ortiz, C. Alonso-Fernández, I. J. Pérez-Colado, and B. Fernández-Manjón, "D4.7 Game analytics and adaptation components reference implementation - H2020 BEACONING Project Deliverable." 2017.
- [9] A. Bakharia *et al.*, "A Conceptual Framework linking Learning Design with Learning Analytics," *Proc. LAK16 6th Int. Conf. Anal. Knowl.* 2016, no. April, pp. 329–338, 2016.
- [10] Á. Serrano-Laguna, I. Martínez-Ortiz, J. Haag, D. Regan, A. Johnson, and B. Fernández-Manjón, "Applying standards to systematize learning analytics in serious games," *Comput. Stand. Interfaces*, vol. 50, pp. 116–123, Feb. 2017.
- [11] e-ucm, "First Aid Game," 2012. [Online]. Available: <http://first-aid-game.e-ucm.es/>.
- [12] W. Westera, R. J. Nadolski, H. G. K. Hummel, and I. G. J. H. Wopereis, "Serious games for higher education: A framework for reducing design complexity," *J. Comput. Assist. Learn.*, vol. 24, no. 5, pp. 420–432, 2008.
- [13] ADL, "SCORM 2004 (4th Edition)," 2009. [Online]. Available: <http://www.adlnet.org/adl-research/scorm/scorm-2004-4th-edition/>.
- [14] ADL, "ADL." [Online]. Available: <https://www.adlnet.gov>. [Accessed: 22-Mar-2016].
- [15] eUCM Research Group, "xAPI Serious Games Profile," 2016. [Online]. Available: <http://w3id.org/xapi/seriousgames>. [Accessed: 24-Apr-2017].
- [16] I. Jivet, M. Scheffel, H. Drachler, and M. Specht, "Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017, vol. 10474 LNCS, pp. 82–96.
- [17] V. Shute, F. Ke, and L. Wang, "Assessment and Adaptation in Games," in *Instructional Techniques to Facilitate Learning and Motivation of Serious Games*, P. Wouters and H. van Oostendorp, Eds. Cham: Springer International Publishing, 2017, pp. 59–78.
- [18] I. Pérez Colado, V. Pérez Colado, I. Martínez-Ortiz, M. Freire, and B. Fernández-Manjón, "uAdventure: The eAdventure reboot - Combining the experience of commercial gaming tools and tailored educational tools," in *IEEE Global Engineering Education Conference (EDUCON)*, 2017, pp. 1754–1761.

6.2. Multi-level Game Learning Analytics for Serious Games

6.2.1. *Cita completa*

Iván José Pérez Colado, Dan Cristian Rotaru, Manuel Freire Morán, Iván Martínez Ortiz y Baltasar Fernández Manjón (2018). **Multi-Level Game Learning Analytics for Serious Games**. In *2018 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. IEEE, 1-4 <https://doi.org/10.1109/vs-games.2018.8493435>

6.2.2. *Resumen original de la publicación*

Serious games are usually used or deployed in an educational setting as an isolated or individual activity, disconnected from other curricular activities. However, to really increase the adoption of serious games in different educational scenarios, the combination and integration of games into the educational flow should be simplified. We envision Serious Games as new type of educational activity that can be combined as parts of other games (e.g. minigames integrated in larger games), integrated into other online activities, or even mixed with both game and non-game activities. In addition, if we want to make the most from serious games, a learning analytics system must be in place to harvest and analyze interactions, providing metrics and insights to instructors regarding the gameplay sessions. Moreover, if a course-level learning analytics strategy is designed, it must be aligned with the game learning analytics. This approach requires communication between games and educational activities used during the educational experience. From a game learning analytics standpoint, gaining insights from these integrated experiences introduces new requirements within potentially complex multi-level or hierarchical activities. Moreover, the analysis required to generate these metrics should be both efficient and provide insight in an understandable way and for different stakeholders. This paper describes an approach to multilevel game learning analytics from the perspectives of data model, implementation architecture, and result visualization in teacher-oriented dashboards.

Multi-level Game Learning Analytics for Serious Games

Ivan J. Perez-Colado
Dept. of Software
Engineering and AI
Facultad de Informática,
Universidad Complutense
de Madrid
Madrid, Spain
ivanjper@ucm.es

Dan Cristian Rotaru
Dept. of Software
Engineering and AI
Facultad de Informática,
Universidad Complutense
de Madrid
Madrid, Spain
drotaru@ucm.es

Manuel Freire-Moran
Dept. of Software
Engineering and AI
Facultad de Informática,
Universidad Complutense
de Madrid
Madrid, Spain
manuel.freire@fdi.ucm.es

Ivan Martinez-Ortiz
Dept. of Software
Engineering and AI
Facultad de Informática,
Universidad Complutense
de Madrid
Madrid, Spain
imartinez@fdi.ucm.es

Baltasar Fernandez-Manjon
Dept. of Software
Engineering and AI
Facultad de Informática,
Universidad Complutense
de Madrid
Madrid, Spain
balta@fdi.ucm.es

Abstract— Serious games are usually used or deployed in an educational setting as an isolated or individual activity, disconnected from other curricular activities. However, to really increase the adoption of serious games in different educational scenarios, the combination and integration of games into the educational flow should be simplified. We envision Serious Games as new type of educational activity that can be combined as parts of other games (e.g. minigames integrated in larger games), integrated into other online activities, or even mixed with both game and non-game activities. In addition, if we want to make the most from serious games, a learning analytics system must be in place to harvest and analyze interactions, providing metrics and insights to instructors regarding the gameplay sessions. Moreover, if a course-level learning analytics strategy is designed, it must be aligned with the game learning analytics. This approach requires communication between games and educational activities used during the educational experience. From a game learning analytics standpoint, gaining insights from these integrated experiences introduces new requirements within potentially complex multi-level or hierarchical activities. Moreover, the analysis required to generate these metrics should be both efficient and provide insight in an understandable way and for different stakeholders. This paper describes an approach to multilevel game learning analytics from the perspectives of data model, implementation architecture, and result visualization in teacher-oriented dashboards.

Keywords—game learning analytics; multi-level games; teacher-oriented dashboards; location-based games; learning analytics models; xAPI; serious games;

I. INTRODUCTION

Serious games defined as “video games whose primary goal is not that of entertainment” are becoming popular in different educational and training settings. As a medium, they build upon the popularity of video games, which are played by a large and growing percentage of the population, well beyond their initial teenage-male demographic. Games present advantages for education, since they have greater potential engagement and interactivity and immersion [1]. We consider that a key element in a more general serious game adoption is the inclusion of Game Learning Analytics (GLA), where in-game user interactions are collected and analyzed to gain insights and simplify game usage in different domains [2].

However, games are not frequently integrated in e-learning systems. Games are not easily composable with other educational activities. For instance, it is complex to include

games in hierarchies as the ones found in Learning Management Systems (LMS) that support highly structured contents such as SCORM [3] packages. Also, it is complex to communicate the games with the rest of the educational content and include them in the analytics. But from the learning analytics perspective the Experience API (xAPI) [4] specification simplify to track rich interaction data, generated by any educational activity including games. However, xAPI is mainly a format for structuring and collecting data, but it does not impose nor provide guidance or default strategy about what kind of analysis and visualizations can or should be performed with the data once collected.

To address these identified limitations, we have extended our previous serious games analytics system to support multi-level games. This evolved model supports new educational scenarios where, for example, a geolocated game can launch different minigames depending on the player’s position; or even launch another game composed of minigames. Moreover, this approach supports other low-interaction and even teacher-reported offline activities. We have built our approach based on the concept of Learning Analytics Models (LAMs) [5], that model how the inputs/signals/events generated from an activity are selected, aggregated, reported and evaluated. We also adopted the concept of Meta-LAMs [6] which are analytics models built using individual, per-activity LAMs, as building blocks. Fig. 1 briefly illustrates their relationship.

This paper describes a working implementation of multi-level analytics. Section II describes the conceptual model. Section III provides an overview of the system’s architecture, Section IV describes the application of the described conceptual model, and Section V presents the resulting visualizations available for teachers when using the system. Finally, we Section VI contains our conclusions and outlines future work.

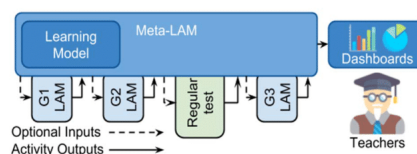


Fig. 1 Learning Model built using a Meta-LAM where multiple LAMs provide outputs as assessment resources and dashboards for teachers.

II. CONCEPTUAL MODEL

As defined in [2], a Learning Analytics Model (LAM) is a multidisciplinary effort, involving educators, designers, and developers to describe everything needed in order to assess a serious game. LAMs specify the player information that will be sent during gameplay as traces to the analytics server, and how these traces should be interpreted (analyzed) and displayed in different dashboards for different stakeholders. A Meta-LAM generalizes this approach and describes how a LAM should be applied to a multi-level game structure.

Meta-LAMs are built using outputs of LAMs as building blocks as illustrated in Fig. 1. This allows games to act as black boxes. To achieve this goal, outputs and aggregations of each level should be defined. Meta-LAMs should provide answers to the following questions:

1. What are the learning goals of the multi-level game? The Meta-LAM uses the outputs of the (mini)games as its inputs to satisfy learning goals.
2. How is the multi-level game hierarchy structured? Analysis needs a structure to aggregate learning goal progresses. The more complex the structure, the harder to use to combine individual game results, to design, and understand.
3. How should the data be aggregated through the structure, and what consequences can be extracted from this aggregation? Determining, when an activity is considered successful, its contributions and the common language. In a hierarchy, parent nodes are designed to receive updates from their children and its handling should be defined in the Meta-LAM.

We have built up our default Meta-LAM from a simplified version of SCORM 2004 4th Edition Sequencing and Navigation (SN) [3], which already includes similar concepts of aggregation. SN can be very briefly characterized by its use of a tree of activities, where each activity communicates with its parents, which aggregate activity results to update the LOs in a process termed rollup.

Answering question 1, designed Meta-LAM limits LOs to only learning goals and competencies to simplify the analysis. As Meta-LAM is generic, LOs and competencies are defined by the aggregation of the ones from every LAM.

Answering question 2, our adaptation maintains the tree structure, simplifying the resulting aggregation process as relationships are well-defined and easy to handle.

Answering question 3, our analysis uses SN rollup rules, where results for any activity node are propagated through its ancestors up to the root. This allows parent nodes to determine their progress and completion status. Also, analysis is parametrized with meta-data that includes optional *limits* and *contributions*. Where *limits* include thresholds that allow to accommodate the difficulty as minimum requirements of an activity and *contributions* include a LOs dictionary that describes how much each one should be incremented. Finally, regarding language, we use xAPI as a standard for all interaction traces, building up on pre-existing glossaries of terms related with Serious Games [7] and Location-based games [8].

III. ARCHITECTURE FOR PERFORMING MULTI-LEVEL ANALYTICS

This section describes the architecture that we use to implement the previously described Meta-LAM. Our LA server used a standalone, generic single-game analysis, called Default Analysis (DA). The DA processes xAPI traces from the actual gameplay according to the xAPI-SG [7] format, and generates a number of metrics and dashboards as a result.

Even through the conceptual model, described in section II, manages the multi-level requirements, it does not store or generate results for the gameplay of each game. We have combined the existing DA with a new Multi-Level Analysis (MLA) that implements the conceptual model of Section II.

To successfully build a Meta-LAM using independent analyses as building blocks, the Meta-LAM architecture designer must answer a few questions:

1. What does the Meta-LAM require from each activity? E.g. games sending xAPI-SG traces require DA, higher level analysis require an overall-storing-analysis.
2. What analysis-to-analysis communication is used? e.g. queue, messaging system, notifications, common storage, etc.

Answering question 1, regarding activity results, the combination of the DA and the MLA will generate, per each of the activities, a results file that contains all its gameplay information along its successors gameplays, meta information related with the achievement of game goals, and results generated based on MLA-generated traces.

However, some of the learning activities may have their own LAM and can use their own analysis to generate the results. To maintain coherency along the multi-level ecosystem, these LAM traces should be stored inside their own analysis results, and translated for their ancestor's analyses to understand them.

Answering question 2, regarding the communication method, all analyses use a common queue to obtain the traces. This way, instead of sending a trace, it is queued with the target node as destination, ensuring that the server is not overloaded. Additionally, every node stores their results into separated documents, causing the root node document to include the whole analysis results.

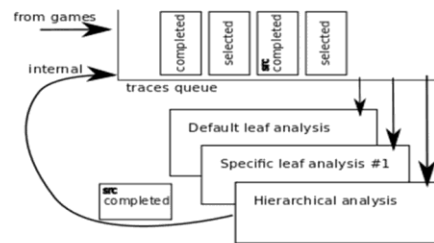


Fig. 2 Diagram of the multi-level game analysis consuming and bubbling traces. The hierarchical analysis propagates "completed" traces to the parent of each originating node, until the root of the hierarchy is reached. Completed trace label "scr" means it is an internal trace.

IV. MULTI-LEVEL ANALYSIS CONCEPTUAL MODEL APPLICATION

Our system follows a Kappa Architecture [9] by using a stream-based analysis, as we analyze traces sent from games in near real-time. When games send traces to the LA they're added to a queue. Analyses read them from this queue being processed once and only once by all analyses associated with the corresponding game. Analyses can maintain their state by reading and writing from and to datastores, and can add new created statements to the queue. Statements in the queue can be either game-generated (GGS) or analysis-generated (AGS).

By allowing analyses to feedback into themselves, we enable a message system where each statement is targeting a node. GGS are always targeted to leaves, while AGS are used to inform intermediate nodes of knowledge propagation within the hierarchy, bubbling up from leaves from each node to its parent until the root is reached. This process is illustrated in Fig. 2.

To feed our default multilevel dashboard, we perform several types of leaf aggregation: when leaf nodes are considered completed, LOs and competencies gets increased. For each leaf, computes the minimum, maximum and average scores of each player to easily contrast the results. Time-to-completion, and completion history: average time to complete missions and quests by players. Player choices, and choice accuracy: The amount of correct and incorrect alternatives is computed and stored, together with a list of choices over time. Player progress is stored, and an average calculated, so that we can later display average player progress. Performance of students is computed from the score and aggregated by years, months and weeks. When a child trace is completed the parent node progresses accordingly.

When a child node is completed, the parent node progresses accordingly. Furthermore, when all the children have been completed, the parent node completes itself, increasing its related contributions.

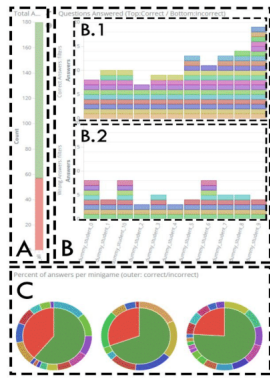


Fig. 4 Answers section showing aggregated answers. Includes total received answers, filtered by correct and incorrect view (A). Specific answers stacked per student (B), where top (B.1) are the correct and bottom (B.2) are the incorrect. And percent of answers received per sub-game (C)

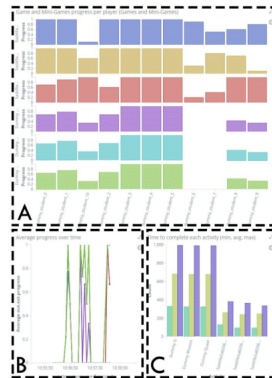


Fig. 6 Progress and duration section, including: a matrix of vertical bar progress visualization where X is student and Y is the activity (A); The average progress of each activity over time (B) and the min, avg and max duration per activity (C).

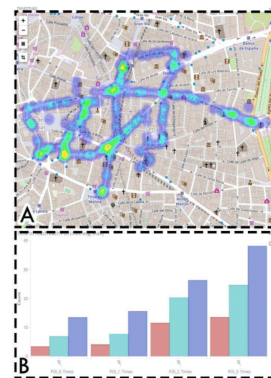


Fig. 5 Location-Based game section. Includes two visualizations: a heatmap of every position tracked by any activity (A); and the minimum, average and maximum duration of the travel between a POI and next POI (B)

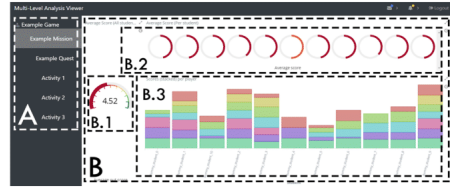


Fig. 3 Web tool used to navigate through game-tree dashboards. Menu list (A) allows to select dashboard while the main frame (B), displays the dashboard, in this case MLA dashboard. Main frame is displaying Scores section with aggregated scores, composed of Average Score of all students (B.1), Average Score per student (B.2) and Stacked scores per player (B.3).

V. USER INTERFACE

Regarding data visualization, one important difference between a regular LAM and the Meta-LAM is going from one single dashboard to multiple dashboards. This forces developers to add an upper layer in the UI allowing to navigate through the multi-level game tree. Our approach depicted in Fig. 3, consists of an admin-panel view where the left part (Fig. 3 A) depicts a menu that allows users to select the desired dashboard, while the dashboard are displayed in the main frame (Fig. 3 B).

The root (and in some cases non-leaf node) dashboard has been designed following the conceptual model described in Section II. Aims to be as generic as possible, mostly adhering to xAPI-SG, implying that every game that successfully implements xAPI-SG will fit the Meta-LAM and its data will be correctly analyzed for display in the Meta-LAM dashboard.

By design, every visualization is intended to be meaningful both unfiltered and filtered-by-student, with only few exceptions to this rule. Our approach aims to be generic enough; however, some cases are not covered when using specific LAM. The more heterogeneous the variety of traces, the more diverse the sections

included in the dashboard. Even though we could add further visualizations it would likely overwhelm teachers, worsening the user experience. Dashboards can easily be overloaded with too much information, so every visualization must have multiple purposes. Our solution is composed of five sections:

The *General information and configuration section* has been designed to allow teachers to quickly access information regarding whether the lesson is working properly; and allowing them to filter and personalize the dashboard. Includes four visualizations: a student list for dashboard filtering, a session counter for quick comparisons with the expected number of learners, a completion counter to check how many students have already finished, and a time-frame selector.

The *Scores section*, illustrated in Fig. 3, has been created to easily determine if a learner has succeeded or failed, and how well have each learner performed in each multi-level game node. As seen in the previous section, most GLA models can determine if a game has succeeded, and even determine a score that describes how well the learner has played. In the end, this allows educators to transform LA results into grade marks.

The *Answers section*, illustrated in Fig. 4, includes information about the choices that learners have made throughout all sub-games. It is meant to allow the teacher to specifically detect where learners have failed or succeeded, to personalize feedback and provide better assessment. Almost every game allows to make choices, and its answers are usually given with a correct or incorrect result. Those answers can be from a quiz, or can from a potion-mixing minigame, meaning that the potion have been successfully crafted or not.

The *Progress and duration section*, illustrated in Fig. 6, is designed to provide control of the current state of the students by providing and overview of the progress of each student in each activity, their progression, and its duration. It explores the possibilities of MLA by showing non-leaf nodes progress and duration, which are calculated by the analysis (as explained in section IV). It also allows to check when the progress has been made, showing the students' play periods, and providing feedback to understand the homework window available to the teacher, or adapt future games to more appropriate durations.

The *Location-Based games section*, illustrated in Fig. 5, satisfies a frequent use-case of multi-level games, which are Location-based gymkhana-like games, where multiple stops are presented, and each stop requires the player to pass a minigame to continue. This section provides, an overview of where the players have physically been, along the time it took to them to reach each stop. This is made using a heatmap, displaying the position of the players. It helps to understand the most-used routes to travel from stop to stop, together with the time spent.

VI. CONCLUSIONS AND FUTURE WORK

In the previous sections, we have identified and addressed the issues that arise from using multi-level learning analytics for multi-level / hierarchical tree-like activities, focusing on multi-level game scenarios.

In addition, we have also briefly described the main components of our solution, which is designed to be modular

and could potentially support several different analyses side-by-side using entirely different Meta-LAMs, by changing the chosen answers to the questions of Section II.

The implementation of the conceptual analysis can be done in several ways, however the driving questions stated in section III lead us to a modular architecture where we defined a default analysis that works at the individual game level; and a multi-level analysis (MLA) that performs all calculations for the intermediate levels. The combination of these two analyses satisfies the needs of Meta-LAM as described in IV.

We have developed a dashboard that displays aggregated results from the MLA, providing a useful tool to track and evaluate multi-level game scenarios. The dashboard is divided into five sections, one for management, three that gather information on aspects needed to assess learners, and a last one for location-based games as a usual multi-level game as in V.

Although an initial qualitative evaluation of the current dashboards has been positive, based on previous work with non-multilevel teacher dashboards, we feel that many teachers are already overwhelmed with relatively simple dashboards, and multi-level ones may produce additional cognitive overload. We will be collecting additional teacher feedback to lower the cognitive demands of multi-level dashboards.

ACKNOWLEDGMENTS

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

REFERENCES

- [1] T. M. Connolly, E. A. Boyle, E. MacArthur, T. Hainey, and J. M. Boyle, "A systematic literature review of empirical evidence on computer games and serious games," *Comput. Educ.*, vol. 59, no. 2, pp. 661–686, 2012.
- [2] M. Freire, Á. Serrano-Laguna, B. M. Iglesias, I. Martínez-Ortiz, P. Moreno-Ger, and B. Fernández-Manjón, "Game Learning Analytics: Learning Analytics for Serious Games," in *Learning, Design, and Technology*, 2016, pp. 1–29.
- [3] ADL, "SCORM 2004 (4th Edition)," 2009. [Online - Last accessed, April 2018]. Available: <http://www.adlnet.gov/research/scorm/scorm-2004-4th-edition/>.
- [4] B. Y. M. Torrance and C. Wiggins, "What Is xAPI?," *TD Talent Dev.*, vol. 70, no. 2, pp. 28–31, 2016.
- [5] J. M. Baalsrud Hauge et al., "Learning Analytics Architecture to Scaffold Learning Experience through Technology-based Methods," *Int. J. Serious Games*, vol. 2, no. 1, 2015.
- [6] I. J. Perez-Colado, C. Alonso-Fernández, M. Freire-Moran, I. Martínez-Ortiz, and B. Fernández-Manjón, "Game Learning Analytics is not informagic!," *IEEE Glob. Eng. Educ. Conf.*, 2018.
- [7] Á. Serrano-Laguna, I. Martínez-Ortiz, J. Haag, D. Regan, A. Johnson, and B. Fernández-Manjón, "Applying standards to systematize learning analytics in serious games," *Comput. Stand. Interfaces*, vol. 50, pp. 116–123, Feb. 2017.
- [8] V. M. Perez-Colado, D. C. Rotaru, F. Muel, M.-O. Ivan, and B. Fernandez-Manjon, "Learning analytics for location-based serious games," *IEEE Glob. Eng. Educ. Conf.*, 2018.
- [9] J. Kreps, "Questioning-the-Lambda-Architecture," 2014. [Online - Last accessed, April 2018]. Available: <https://www.oreilly.com/ideas/questioning-the-lambda-architecture>.

6.3. Simva: Simplifying the scientific validation of serious games

6.3.1. Cita completa

Iván José Pérez Colado, Antonio Calvo Morata, Cristina Alonso Fernández, Manuel Freire Morán, Iván Martínez Ortiz y Baltasar Fernández Manjón (2019). **Simva: Simplifying the Scientific Validation of Serious Games**. In *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 113-115 <https://doi.org/10.1109/icalt.2019.00033>

6.3.2. Resumen original de la publicación

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.

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 irreversibly. 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

Clase EICaton 1A					
Code	Conducta (PrePostQ1ver)				
<input type="checkbox"/>	Q1UJ	FINISHED	FINISHED	FINISHED	TRACES
<input type="checkbox"/>	WEFF	FINISHED	FINISHED	STARTED	TRACES
<input type="checkbox"/>	YEYT	FINISHED	NOT FOUND	NOT FOUND	TRACES
<input type="checkbox"/>	ZHSL	FINISHED	STARTED	NOT FOUND	TRACES

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 Automate Neuropsychological Tests and Research in Active Aging," *GOODTECHS 2018 - 4th EAI Int. Conf. Smart Objects Technol. Soc. Good*, 2018.

6.4. Improving Teacher Game Learning Analytics Dashboards through ad-hoc Development

6.4.1. Cita completa

Antonio Calvo Morata, Cristina Alonso Fernández, Iván José Pérez Colado, Manuel Freire Morán, Iván Matínez Ortiz y Baltasar Fernández Manjón (2019) **Improving Teacher Game Learning Analytics Dashboards through ad-hoc Development**. *JUCS - Journal of Universal Computer Science* 25(12): 1507-1530. <https://doi.org/10.3217/jucs-025-12-1507>

6.4.2. Resumen original de la publicación

Using games for education can increase the motivation and engagement of students and provide a more authentic learning environment where students can learn, test and apply new knowledge. However, the actual (serious) game application in schools is still limited, partly because teachers consider their use as a complex process. To increase game adoption, the integration of Game Learning Analytics (GLA) can provide teachers a thorough insight into the knowledge acquired by their students and usually presented through a visual dashboard. Although it is possible to provide a useful general metrics and a prefab dashboard, it may not fully cover teachers' expectations. In this paper, we study the ad-hoc adaptation of generic dashboards to increase their effectiveness through three case-studies. In these experiences, we adapt dashboards for teachers to include detailed information for more-focused analysis. With the positive results obtained from these scenarios, we have identified a methodological process to create ad-hoc GLA dashboards and extracted some lessons learned for dashboard development: simple but useful dashboards can provide a higher added value for stakeholders compared with more complex dashboards; teachers and game developers should be involved in dashboard design for better results; and, if possible, ad-hoc developed dashboards should be used as they have proved to be more effective than generic dashboards.

Improving Teacher Game Learning Analytics Dashboards through ad-hoc Development

Antonio Calvo-Morata

(Complutense University of Madrid, Madrid, Spain
acmorata@ucm.es)

Cristina Alonso-Fernández

(Complutense University of Madrid, Madrid, Spain
calonsofernandez@ucm.es)

Iván J. Pérez-Colado

(Complutense University of Madrid, Madrid, Spain
ivanjper@ucm.es)

Manuel Freire

(Complutense University of Madrid, Madrid, Spain
manuel.freire@fdi.ucm.es)

Iván Martínez-Ortiz

(Complutense University of Madrid, Madrid, Spain
imartinez@fdi.ucm.es)

Baltasar Fernández-Manjón

(Complutense University of Madrid, Madrid, Spain
balta@fdi.ucm.es)

Abstract: Using games for education can increase the motivation and engagement of students and provide a more authentic learning environment where students can learn, test and apply new knowledge. However, the actual (serious) game application in schools is still limited, partly because teachers consider their use as a complex process. To increase game adoption, the integration of Game Learning Analytics (GLA) can provide teachers a thorough insight into the knowledge acquired by their students and usually presented through a visual dashboard. Although it is possible to provide a useful general metrics and a prefab dashboard, it may not fully cover teachers' expectations. In this paper, we study the ad-hoc adaptation of generic dashboards to increase their effectiveness through three case-studies. In these experiences, we adapt dashboards for teachers to include detailed information for more-focused analysis. With the positive results obtained from these scenarios, we have identified a methodological process to create ad-hoc GLA dashboards and extracted some lessons learned for dashboard development: simple but useful dashboards can provide a higher added value for stakeholders compared with more complex dashboards; teachers and game developers should be involved in dashboard design for better results; and, if possible, ad-hoc developed dashboards should be used as they have proved to be more effective than generic dashboards.

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

Categories: K.3, L.5.1

1 Introduction

Games characteristics such as their engaging and motivating nature make games especially adequate for education [Connolly, Boyle, MacArthur, Hainey and Boyle 2012, Denis and Jouvelot 2005]. Despite these probed advantages, the adoption of serious games is still poor in formal education, partly due to a lack of standards for development, validation, and deployment in schools [Loh, Sheng and Ifenthaler 2015]. We consider that collecting and analyzing student gameplay is one of the keys to increase serious game adoption in schools. Collected analytics data can provide insight and improve all steps in the process, and the resulting insights can themselves constitute an important selling point supporting the use of games in the classrooms. In educational settings, Learning Analytics (LA) is used to provide insight into learners' actions to improve some aspects of the learning process and contexts. When applied to games, Game Learning Analytics (GLA) focuses on information gathered from players via in-game interactions.

Fig. 1 represents a generic Game Learning Analytics pipeline, focusing on 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 analytics stakeholders, such as students themselves, game developers, and academic officials, can be presented with their own dashboards.

The analysis of data has not only risen dramatically in the field of games [Bauckhage, Drachen and Thureau 2015]; it is a key trend across many fields and domains, characterized by collecting and analyzing large amounts of data and, therefore, sometimes referred to as *big data*. For instance, business analytics has been identified as one of the key technology trends in the 2010s [Chen, Chiang and Storey 2012]. Following the analysis of data, the results are typically presented visually, for different purposes and target audiences. The visual analytics field also has a wide range

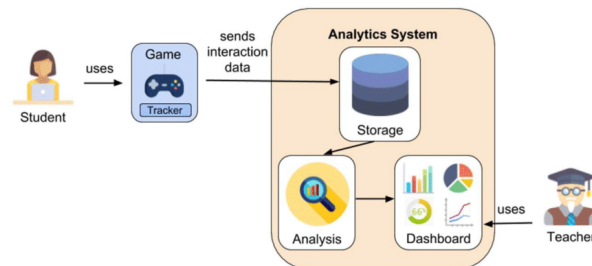


Figure 1: Game Learning Analytics pipeline: a student plays a game that includes a tracker component. The tracker sends interaction data to the Analytics System for storage, analysis and visualization. Teachers, among other stakeholders, interact with dashboards to gain insights into student gameplay and educational outcomes.

of possible application areas, including information synthesis, the discovery of unexpected results, and performing and communicating assessments in calls for action [Keim et al. 2008]. It is, therefore, no exception that, also for serious games, dashboards are the usual way to communicate information to stakeholders [Shoukry, Göbel and Steinmetz 2014]. 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. Besides, data visualization via dashboards is the most commonly used method in Game Analytics [Drachen, El-Nasr and Canossa 2013] to communicate insights and provide reusable reports for different stakeholders: visualizations can provide in a simple way a variety of information, such as heatmaps to display places where level crashes occurred.

It may appear that once a serious game implements tracking and event-reporting standards, it should be simple just to add Game Learning Analytics (GLA) to provide pedagogical insights that teachers can use in their classrooms. However, even though generic visualizations (and dashboards) provide a means to obtain information from commonly-present traces generated from GLA data, establishing the relationship between the visualization (or the event) and the consequence (e.g., the knowledge acquired) is not an easy task. When designing these generic visualizations and dashboards, pedagogy experts must help to determine how the Learning Goals (LGs) of a game can be identified in such simple visualizations built from standardized traces. They must participate in the creation of the Learning Analytics Model (LAM) [I. Perez-Colado, Alonso-Fernandez, Freire, Martinez-Ortiz and Fernandez-Manjon 2018] for the generic dashboard, as their knowledge helps to identify the suitable LGs that every game must-have for teachers to be able to evaluate and assess later.

These generic visualizations present relevant data such as times and scores, which can be compared against baselines to determine student performance. In a class, teachers can see whether whole tasks are working as intended or diagnose problems. How to display this information is far from standardized: a recent literature review on analytics in serious games [Alonso-Fernández, Calvo-Morata, Freire, Martínez-Ortiz and Fernández-Manjón 2019] found no consensus for the display and visualization of performance metrics for serious games. We have the same experience as we have found no standard patterns for this process, relying instead on different ad-hoc methods to visualize specific outcomes and general progress.

While generic dashboards might be enough to satisfy the basics, ad-hoc, game-specific dashboards allow a focused and higher level of introspection and understanding of the gameplay and its conclusions, as they are developed for a specific game and its stakeholders.

The stakeholders in GLA include students, teachers, game developers, administrators, and researchers, among others. 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 it is teachers that decide if games will be used in their classrooms. More narrowly, the last step of *Fig. 1* depicts the teacher using the dashboard, which requires teacher dashboards to be understandable, as described in [Schwendimann et al. 2017].

Generating 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. For example, school administrators will be interested in comparing summaries of the performance of each class; while teachers will want insights into individual students as compared to their classmates. Data collection and analysis should be as transparent as possible. The final step, the visualization of the information, is where game learning analytics can finally provide value. In this sense, when used by teachers, dashboards need to be evaluated as pedagogical tools, taking into account their goals, affect and motivation, and usability [Jivet, Scheffel, Drachsler and Specht 2018].

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 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 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 the number of attempts, that if possible, should be 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 (generally hierarchic) set of activities.

Nevertheless, a recent literature review [Jivet, Scheffel, Drachsler and Specht 2017] regarding the use of learning analytics dashboards in education points out several flaws and improvements for them, including that dashboards usually lack a secondary goal beyond awareness and reflection, and that focusing on each student's needs instead of using the common technique of comparing each student with their peers may be a better

pedagogical technique to promote motivation. Although default out-of-the-box visualizations are helpful, the real powerful insight is obtained when creating custom visualizations [Drachen et al. 2013].

Section 2 of this paper describes how our architecture, and specifically our teacher dashboards, tackle the above issues; and details how dashboards can be designed working with designers and teachers. Section 3 describes three specific use cases where we have developed teacher dashboards. Section 4 describes the lessons learned in teacher dashboard design. Finally, Section 5 summarizes our conclusions and Section 6 outlines future work.

2 Addressing the teacher dashboards issues

As mentioned before, the process of collecting, analyzing, and displaying data from game interactions on dashboards useful to the teacher is a complex task. Next, we will talk about the main decisions at the level of the analytical system, data collection and the process to create the dashboards that will show the analyzed data.

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 set of interactions that are usual in serious games, as detailed in [Serrano-Laguna et al. 2017]. 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. These 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 store no backup of the token-to-student correspondence table. We then rely exclusively on the 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, which they can use to trace the identity of students in the dashboard.

By using pseudonymous tokens, we guarantee that if data is stolen or lost, only the owners of the pseudonym-to-name conversion table will be able to identify the students who played. Not only that, but we open the possibility of sharing the pseudonymous data with game-developers or researchers, which can benefit from real gameplay data to improve the lifecycle of games by identifying issues and interesting patterns. Given the increased awareness of privacy risks (as reflected in the EU's GDPR), the use of pseudonyms is not the only precaution that has been taken. Our databases are secured to only allow internal connections, and APIs are secured through API gateways. Additionally, if students want their data to be deleted, we allow teachers to permanently delete data for specific students (again, identified via a pseudonymous token) from the system to guarantee their right to be forgotten.

When conducting experiments in schools, we require the school to sign a consent form that describes the type of data collection to be performed and acknowledging that the data is collected under pseudonyms that we will not be able to track back to specific students. Additionally, at the beginning of each gameplay session within an experiment, students are informed that their gameplay data is being collected for research purposes and can opt-out from this data collection.

From our work 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 [Alonso-Fernandez, Calvo, Freire, Martinez-Ortiz and Fernandez-Manjon 2017, I. Perez-Colado et al. 2018]. A simplification of this architecture is depicted in Fig. 2. To understand the workflow, first, the student starts to play the game. Inside the game, there is a component (a tracker) that handles student authentication with the analytics server, and which continuously sends information about what is happening inside the game. To ease integration of the tracker into games, it is available for multiple frameworks and programming languages, including Unity, JavaScript, and Java; and implements the xAPI-SG standard. With a single line of code, developers can link an event inside the game with a trace that will be sent to the analytics backend. Once the analytics backend receives the trace, it is enqueued in Apache Kafka to be consumed by an Apache Storm streaming analysis. This analysis performs the needed calculations and stores both the results and the raw traces into Elasticsearch, which provides storage and querying, for later use by Kibana, which can generate visualizations from Elasticsearch. Teachers can then visualize real-time results through the Analytics Frontend, which builds dashboards from Kibana visualizations, to assess their students.

The tracker communicates game events to the backend by sending traces in xAPI format, which are then categorized and further analyzed. The generic analysis performs simple aggregations, such as counting correct/incorrect answers or calculating the highest score obtained in a specific part of the game. However, some analyses can be more complex, for example by pre-calculating averages or time-clustering the results. An example of the possibilities of an analysis is the Multi-Level Analysis [I. J. Perez-Colado, Rotaru, Freire-Moran, Martinez-Ortiz and Fernandez-Manjon 2018], in which a hierarchy of analyses is provided, and traces are filtered, modified, and re-enqueued into the system to provide, for example, the progress and completion status of each

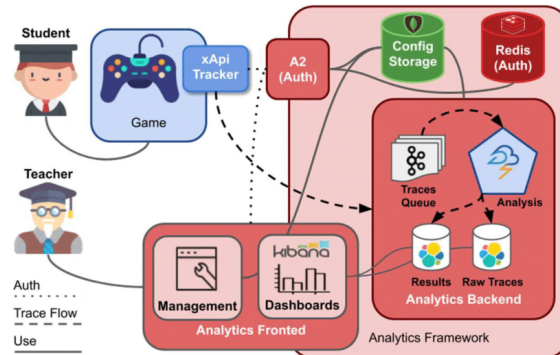


Figure 2: Game Learning Analytics pipeline for our analytics framework. A more specific explanation of what is described in Fig. 1. including technologies used.

node in an activity tree, where nodes can be games that include other games inside (see issue *e* in the previous section).

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, generic 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. On the other hand, game-specific dashboards can provide much better insights, since, for example, keeping track of a game's specific goals in a dashboard where each goal is represented by a specific visualization is much more intuitive than having to rely on one-size-fits-all generic dashboards.

Additionally, since our dashboards are updated in near real-time, with delays of few seconds between receiving interaction data (in xAPI-SG format) 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, besides 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 architecture allows for 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 3.1 and 3.2 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 specific dashboards. However, it still makes sense to provide a global dashboard to monitor progress across all the minigames, and to be able to switch granularity on demand. Indeed, this functionality was a key piece of the H2020 BEACONING EU Project. The resulting multi-level dashboards provide insights on player progress at any level of the activity hierarchy, which could encompass both serious games and more traditional classroom activities.

Building dashboard visualizations for teachers requires collaboration between both game designers/developers and teachers. Developers are closest to the game's requirements and characteristics and in charge of designing the interactions that will

yield observable data and report it to the analytics server. Subsection 2.1 describes this process in greater detail. On the other hand, their efforts will be in vain if the resulting dashboards are not useful to teachers. Subsection 2.2 describes teacher involvement in the teacher dashboard creation process. Note that communication between teachers and developers should occur as early as possible, since the design of the game and its choice of what to report when determines much of what can be achieved with the subsequent analysis and presentation in the form of dashboards. Several iterations may be needed before a good teacher dashboard is created.

In addition to these two roles, pedagogical experts are necessary to ensure that the game is pedagogically sound, and the game's analytics dashboard is an important tool to ensure this soundness. Pedagogical experts should be involved in the design of the game, and part of this design includes coming up with good ways to measure and assess learning – which should be embodied in the dashboard. After the game is finalized, pedagogical experts can be valuable to help teachers to build dashboards that “bring up the learning”, by emphasizing the links between in-game actions and actual pedagogical outcomes.

2.1 Working with game designers and developers to create dashboards

Game designers and developers choose the game's mechanics and contents and are in the best place to map this knowledge of its internal workings into meaningful interactions to be reported to the analytics server. Therefore, their involvement is essential when designing a dashboard's individual visualizations, to ensure that the specific data that they rely on is correctly collected, analyzed and presented. This process can be divided into the following steps:

1. Dashboard goal, and choice of general layout and contents: The first step is to define a draft of the desired dashboard including the visualizations that we would like to show. This includes the general layout and contents of visualizations, encompassing both the information to be shown and the desired visual mapping. For instance, possible visualization contents could be “total player activity over time, as a line chart”, “number of correct and incorrect answers per player, using stacked bars”, or “average score per player and players who have completed the game, as a sorted scatterplot”. Teacher involvement (see steps 2 and 3 in subsection 2.2) is important here.
2. Choose the data to fill the dashboards: Once the information to be shown in the dashboard is clearly defined, the next step is to define the specific variables that will be used to report this information. These report variables (encoded into different sections of xAPI statements) will be fed from data from the game. This step encompasses both the choice of xAPI encoding and that of choosing the game data from which it will be assembled. For instance, to show the number of correct and incorrect answers per player, the game would have to send, for each answer and player, whether the response is correct or not. To isolate game developers from as many xAPI details as possible, we provide easy-to-use libraries in several programming languages. Note that developers can choose to send additional analytics information, even if it does not make it into the final dashboards: this provides flexibility until dashboard design is finalized since it allows different dashboard ideas to be tested out with existing, previously-unused data. Indeed, they are encouraged to at least report

on all major xAPI-SG events; this is enough to feed the default analysis and dashboard (described above, and addressing issues *b* and *c*), and provides a good baseline on top of which game-specific improvements can be made.

3. Define how the data reported from the game will be analyzed. Building on the previous example, to show the number of incorrect and correct answers per player, given xAPI statements from different players stating whether each of their answers was correct or not, the analysis would have to aggregate results for each player and store it for later display.
4. Create the dashboard: following the previous steps, a new version of the dashboard can be created, which can be tested with actual game data, either by playing it anew or by replaying previously-collected data. Continuing with the previous example, we could choose to depict the number of correct and incorrect answers per player with a stacked bar chart where each bar corresponds to a player, and contains two smaller bars on top of the other, representing correct and incorrect answers, and encoded as green and red. The height of each bar would reflect answer counts. The resulting dashboard would then be evaluated by teachers, and possibly refined – see step 5 in the next section.

2.2 Working with teachers to create dashboards

Early teacher involvement is critical to the success of a teacher dashboard. While game designers or researchers may think that a given dashboard prototype is highly informative and useful, it is easy to forget that most of the teachers that will use the dashboards will not be experts in the game, and will not be familiarized with high-density visualizations favored by visualization experts; involving some non-designer teachers early during dashboard design helps it to be much more usable.

Beyond usability, teacher involvement allows dashboards to support teachers in the activities that surround their specific use of the game in the classroom. The process has the following steps:

1. Show the game to the teacher: Teachers should play and understand the game, its goals, how it is played, and the relationship between game mechanics and learning outcomes. Early teacher input will allow visualizations to be focused on their intended use of the game, making it more likely that the resulting dashboards will be both usable and informative. Without playing or understating the game, teachers cannot build an opinion on what they would like to see.
2. Define teacher goals: Once teachers have played, we can ask them to describe what they want to learn from how their students play. For example, teachers may be interested in knowing what parts of the game students are playing, whether they have problems in certain places or with certain questions, or the order in which they are choosing to solve the game's challenges.
3. Create a draft dashboard: given a set of goals, we can draft a possible dashboard that attempts to satisfy them. Each dashboard visualization will require different analyses, and data to feed them, which is the focus of the next step.
4. Ensure availability of underlying data: For each visualization, a suitable analysis must be available to produce the data which will be displayed, and

this analysis, in turn, requires the game to send the necessary interaction data to feed it.

5. Evaluate the dashboard with teachers: The draft dashboard can now be implemented and shown to teachers to evaluate its usability and usefulness. Based on teacher feedback, visualizations may be added or removed from the dashboard, prompting a new prototype/provision/develop/test cycle.

Note that the dashboard development processes for game designers and developers, on the one hand, and for teachers, on the other, are interrelated: steps 2 and 3 above feed into designer/developer step 1 (choice of dashboard layout and contents); and step 5 follows designer/developer step 4 (dashboard implementation). Furthermore, step 4 above can only be fulfilled with developer collaboration, since only developers will be aware of the exact data being sent, and how it is being analyzed. Finally, steps 2-5 are meant to be iterated. See *Fig. 3* for an illustration.

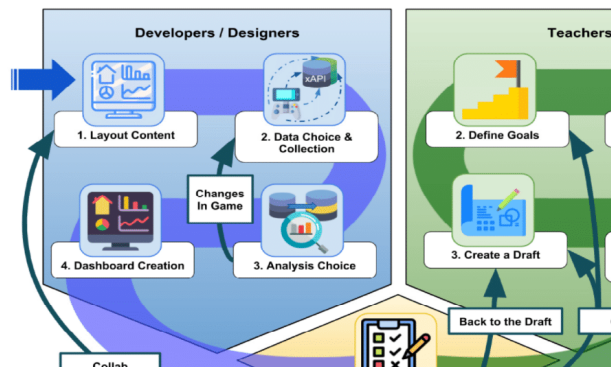


Figure 3: Dashboard creation process.

3 Use cases

We have used teacher dashboards and game learning analytics in multiple experiments, with goals that included the validation of the underlying analytics architecture, improving the serious games that we were analyzing, and evaluating game use and performance. In this section, we summarize three experiences in the creation and use of game learning analytics and teacher dashboards.

Note that all the dashboards are designed to be visualized in a web-browser in a traditional monitor, with a mouse used for interaction. Using the dashboards on a mobile device may result in poor performance since the frontend relies on CPU-intensive Kibana to display the real-time visualizations.

3.1 A custom dashboard for a serious game on cyberbullying awareness

Conectado [Calvo-Morata et al. 2018] is a serious game to raise awareness against cyberbullying by placing players in the role of a transfer student that suffers bullying and cyberbullying after arriving at a new school. Players experience the life of this student in first person, during each of 5 in-game days, while being exposed to feelings of helplessness and increased (in-game) social isolation. The game keeps track of the level of the 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. The risk is encoded as an integer between 0 and 100, where higher values correspond to higher risk, that is, worse social standing. There are indicators of risk for each character as well. The decisions that players make during the game, including whether players decide to ask for help to 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 *Figure 4*. From left to right, and from top to bottom:

- 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). In general, the lower the risk, the better, since this means that the students are keeping good relations with their in-game classmates by choosing good dialogue options.
- The 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-

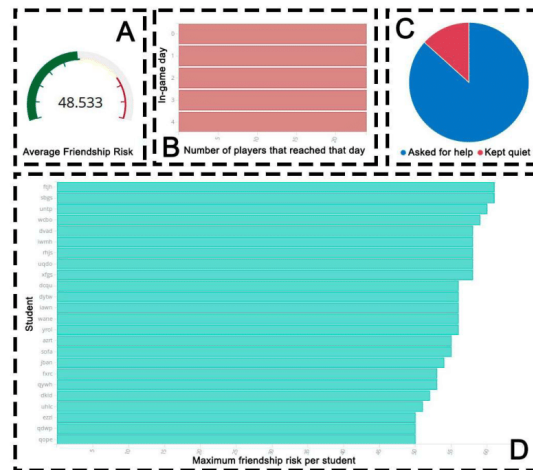


Figure 4: Some of the game-dependent visualizations of the game *Conectado*.

time how many students have played through each in-game day and help students who are too far behind their classmates.

- c. The number of players that have taken each possible action that determines the ending: this pie chart compares the number of players who have decided to complain about bullying (have asked for help) to the in-game parents and teachers vs. those that have decided to remain silent. Note that communicating that the harassment is happening is the first step to solve it.
- d. Maximum friendship risk per student: this visualization provides a more detailed view that complements the general metric provided in visualization (a) and allows teachers to quickly identify which students are doing best and worst in the game. Since the visualization is sorted, it also provides an overview of the distribution of risk scores throughout the class. This visualization is key for a quick response of the teacher in order to help the proper students, as students with higher risk need more help.

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, our dashboards display individual visualizations in two rows, with each visualization of the same shape and size. As seen in *Figure 4*, this can be modified to realize almost any design, by providing control over size, shape, and location of individual visualizations.

In the *Conectado* case, and as a part of our experiments in schools, we also asked teachers for feedback on the utility and usability of their dashboards while the videogame was being played in their classrooms. Teacher feedback resulted in a more complex dashboard, found in *Fig. 5*. This dashboard shows, in real-time, what players are doing, the percentage that has completed each phase, and their answers in key points in the story:

- e. The last trace sent: displays, in a list, for each player, the time at which that player last interacted with the game; together with the player's general progress in the game. This allows teachers to quickly see which users are inactive or are further behind, in order to take action and notify inactive users that they must play.
- f. Games started vs games completed: allows teachers to assess the general progress of all players. If all players have finished, then the pie chart will display two equally-large portions, one purple (games started) and another green (games completed). The ratio between both portions reflects the difference between these numbers. Ideally, the lesson should end with the same number of started and completed.
- g. Scenes started: displays, for all players, the game scenes that they have visited in their games, stacked from oldest (bottom) to newest (top). This allows teachers to spot trends and understand not only how far along they are, but also how they reached that point. This visualization is more useful after the lesson has ended, as it allows teachers to see the exact sequence of scenes played by each student, and might relate a bad performance with a set of scenes.

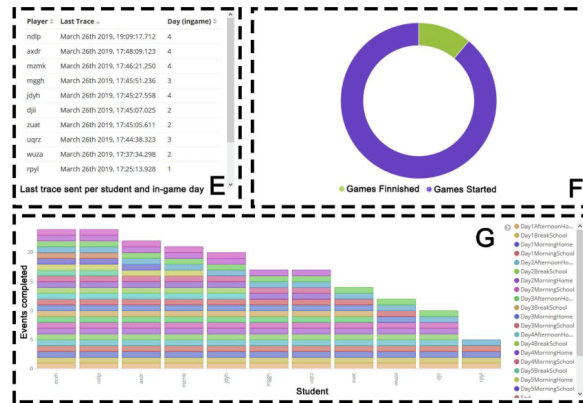


Figure 5: Game-specific visualizations added to the teacher dashboard from Fig. 3, based on teacher feedback.

We have developed the Conectado dashboard while also developing the game itself, including also game design and its pedagogical design. Concentrating all these roles greatly simplifies dashboard development since the points from sections 2.1.1 and 2.2.1 that refer to the game developer and teacher involvement can when handled by the same team that has also built the analytics system, require drastically less communication than what would be needed in less ideal circumstances. We later conducted several experiments to validate the dashboard's usability and usefulness.

We carried out experiments both to validate Conectado and its analytics dashboard, sending traces to the analytics framework described in Section 2. Details on the Conectado validation experiments are available in [Calvo-Morata et al. 2018], including participants, methodology, and results. Dashboard-wise, the main steps were as follows:

1. Participating teachers supervised the gameplay session in their classes using the dashboard, and used it to assess their students and understand what was happening in the class while using it. For example, they could identify inactive students using the "last traces sent" visualization (Fig. 5.E).
2. Once that all students finished the game, teachers were asked to fill a technology acceptance model (TAM) survey, which we used to evaluate the dashboard. The survey contained items on:
 - Previous technology-related experience
 - Previous analytics-related experience
 - Usability (how easy-to-use it is)
 - Usefulness (increased control individually and by group, supports assessment, etc.)

- Intention to use

Validation dashboard experiments took place in two different schools of Spain with a total of 5 different teachers. Each teacher lead multiple game validation sessions. The results of these experiments show that most teachers were satisfied with the dashboard and its usability and usefulness, therefore validating its design. Results also confirmed that teachers that had previous experience in dashboard or analytics usage quickly understood and took advantage of the dashboard, reporting a positive experience in their feedback; while the teachers that reported little or no experience with dashboards and analytics often experienced problems during their first steps and understood individual visualizations worse. Teachers also provided feedback on UI-related issues such as font size and the use of clearer titles to increase usability (this is already apparent in Fig. 5). Finally, most of the teachers pointed out that, if available, they would gladly use these tools and dashboards with more games to support student assessment.

3.2 A custom dashboard and analysis for a serious game on workplace interaction

This dashboard, where again both analysis and visualizations were developed ad-hoc, was built for a game centered on workplace interactions developed within the EU H2020 RAGE project. In the game, named *Watercooler* [Hollins, Humphreys, Yuan, Sleightholme and Kickmeier-Rust 2017], the player must work as an office assistant in a simulated game studio to improve the teamwork between employees. Its goal is that, through improving, enabling, prompting and challenging the attitudes, values and social skills, the player leads the company to success. Information regarding game validation experiments is available in [Hollins et al. 2017, TUGraz 2018, 2019a].

The game design included a requirement to use the Thomas–Kilmann Conflict Mode Instrument (TKI) [Thomas and Kilmann 2008] 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*; see element (e) in Fig. 6. Specific analysis and visualization were 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. 6 displays the seven visualizations developed for this dashboard:

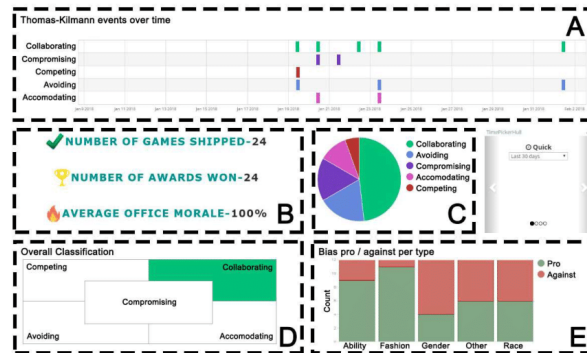


Figure 6: Some of the game-specific visualizations developed for the Thomas-Kilmann Conflict Mode Instrument.

- Thomas-Kilmann classification of a specific student's answers over time. The swim lanes receive and sort the type of the answers received, providing an overview of progress and showing, at a glance, whether the game is working properly.
- Three badges that provide a running score: Games shipped, a measure of team productivity; awards won, a measure of team quality; and Office morale, a percentage indicating the degree to which the player's in-game coworkers are happy with the player's choices. These scores help to maintain student engagement by allowing them to gauge how well they are doing compared to their classmates.
- 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, and should be interpreted as having 2 axes. The horizontal axis corresponds to *cooperativeness*, while the vertical axis represents *assertiveness*. It is split into 5 areas, one for each TKI category. When the visualization of "a" gets too complex, this visualization quickly identifies the category.
- Pie chart displaying the distribution of answers according to the 5 TKI categories. In addition to "c", this helps to identify if the distance between categories is close or far, making more significant the result displayed in c.
- 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.

Our role in this scenario was limited to dashboard developers, as we had minimal input regarding game knowledge or learning goals. Communication between dashboard development and game/learning design took place through email and videoconference.

At the beginning of the development, dashboard developers received a hand-drawn mockup of the dashboard. After several prototypes and developing the requested game-specific dashboard visualizations, we arrived at the dashboard displayed in Fig. 6.

The usability and usefulness of the improved dashboards was then evaluated by the stakeholders by a reviewing process where they evaluated the whole Game Analytics Suite [TUGraz 2019b]. Some of the evaluation aspects included: requirements fulfillment, usefulness, or impact on the effectiveness and added value. The review from the stakeholders showed overall good results, emphasizing in its usefulness speeding up the game development. Reviewers noted that using it is a non-frustrating experience, being easy to use. Finally, the results also pointed out that stakeholders were satisfied with the product and it provides added value for the game.

3.3 A custom dashboard with parametrized analysis for a serious game on formal specification

The last presented scenario is another ad-hoc dashboard created for the integration between the FormalZ [Prasetya et al. 2019] game and the analytics framework, as part of the IMPRESS Erasmus+ EU project. This game falls within the genre of “tower defense”, where players seek to protect a base by placing defensive towers that must destroy wave after wave of enemies launched against it, with the waves increasing in difficulty as the players upgrade their defenses. In FormalZ, each level corresponds to an electronic circuit in where the circuit chips must transform the values of incoming particles into certain output values. However, the chips are not working properly, so the student must specify preconditions and postconditions that must hold before the particle is allowed into the chip, and on the transformed particles once they exit. Allowing invalid particles into the chip, or generating invalid outputs, resulted in loss of in-game “lives”. These conditions are specified in natural language, which the student must transform into logic asserts to be interpreted by towers that destroy unwanted incoming particles.

To assess the students, game designers and analytics developers identified several indicators to track during gameplay, including money, the number of towers, and the number of remaining lives. These values were reported by the game to the analytics server in real-time, using xAPI-SG statements. Each wave initialization includes details about the conditions used, how close students are to the solution (proximity), and how much time the student took to write them. The last and most important event contains the player’s choices of preconditions and postconditions. Every time the student decides to apply a new condition, an interaction event is reported to the server, mentioning how long the student took to write it. Time-dependent visualizations display all lines as starting at the start of the game session of each player. While not all players started their sessions at the exact same time, using the start of each session as zero makes comparisons between time-lines much easier to perform.

- d. Number of lives over time, per student. Like b and c, but in this case, the higher the lives, the better the performance. If a student loses lives quickly, something must be wrong, and assistance is needed immediately.
- e. Distance between correct and player-selected precondition over time per student: The closer to zero, the closer to the correct answer. Additionally, dot size is used to represent the actual time the students spent writing the conditions, with bigger dots corresponding to longer time intervals. The filtered-by-student view is much more useful, as in b. Ideally, the closer and faster the student gets to zero distance, the better the performance.
- f. Distance between correct and player-selected postcondition over time, per student: as in (e), but for the postcondition.
- g. Detailed log of the preconditions as written by students, including the distance, writing time, timestamp, student, and the exact response given. This visualization is key to evaluate as it shows the exact answer given by the student. Also, it is useful to backtrack, as in this visualization, the evolution of the precondition can be seen.
- h. Exact log of the post conditions given by the students, like (g) but for postconditions.

The dashboard can be used not only to assess the students but also to assess the difficulty of individual tower defense levels. For example, if too many towers need to be used, or students keep running out of budget or lives, despite having already shown knowledge of the level, this would indicate that the level needs to be revised. Conversely, having all students breeze through a level would indicate that that level lacks significant challenge and that students are probably being bored by the exercise.

While developing this dashboard, we also performed full analytics integration: we explained the possibilities of the xAPI-SG standard to the game's developers, shared a list of example visualizations, and created and discussed mockups while meeting with the game team and evaluators. After several rounds, the dashboard reached its final version shown in *Fig. 6*. We also created a gameplay-emulator to populate functional dashboard mockups with example data – indeed, the results displayed in *Fig. 6* were populated using this emulator, as the main experiments have not taken place yet. As of this writing, the dashboard and game are still in validation, but the early results and feedback received by both game developers and teachers are very promising.

4 Lessons learnt in teacher dashboard design

Each of the case-studies described in the previous section has led to improvements in the analytics platform and teacher dashboards. Additionally, we have tested the default teacher dashboard (not displayed above), including the alerts and warnings, in experiments with over 600 students and 150 teachers. Based on our experience, we make the following recommendations to analytics dashboard authors:

- Consider the specific requirements of teachers when designing visualizations, and be aware that the teachers may not be aware of these requirements until they start to see the possibilities and limitations of the actual dashboards.
- Start with small, simple dashboards, and let the teachers propose additional visualizations before including them. An overly-complex dashboard will lead to information overload and decrease usability and teacher engagement.

- Ensure that all visualizations include helpful, descriptive labels; and include mechanisms that allow teachers to see more detailed explanations on demand.
- Allowing teachers to move and reorganize dashboard visualizations on demand also leads to greater use; additionally, we have observed that teachers appreciate aesthetic use of colors, and having beautiful visualizations drives engagement.

To make visualizations more understandable, we have completely reworked their names and labels, with a focus on clarity; and include customizable pop-up descriptions to explain the more complex visualizations. For instance, in *Fig. 6 (g)*, the visualization of biases exhibited/averted 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 in *Figs. 4-7*, we now allow dashboards to combine visualizations of different sizes so that more visually complex visualizations can be rendered in larger areas. Additionally, while in the initial teacher dashboards the positions of individual visualizations were not fixed, we have now made positioning entirely controllable. This has considerable advantages when comparing the dashboards for different experiments based on the same game, for example. On a related note, we also allow teachers to reorder and hide visualizations; this allows them to configure dashboards to display only what they find most useful, and facilitates comparisons in visualizations that share common axes.

The display of alerts and warnings has also been improved, by making 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. 8* depicts these improvements: 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, avoiding an extra interaction to display details.

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 [Verbert, Duval, Klerkx, Govaerts and Santos 2013]. 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.



Figure 8: Previous (top) and updated (bottom) alerts & warnings view. The previous display only displayed 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 this context switch.

The entire process of creating visualizations, analyzing, and collecting game data requires a complex and flexible infrastructure underneath. This combination of complexity and flexibility is potentially fragile, and can give rise to a variety of problems: it may not support the volume of interaction traces sent by games; or receive types of data that break specific analyses, causing run-time errors that leave visualizations without data to display; or authors may remove a game for which data is still being received; among a host of possible problems. It is therefore important to continue working on simplifying the infrastructure itself and increasing the system's resiliency while streamlining the process of creating and using dashboards. The goal here is to allow this type of infrastructure to be deployed as easy-to-manage stand-alone servers that allow non-expert users to take full advantage of reliable but powerful learning analytics.

5 Conclusions

Teachers are key to increasing the adoption of serious games by schools, and Game Learning Analytics should, therefore, prioritize 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 can still 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, support for custom-built dashboards (we present three case-studies), a flexible alerts and warnings system, and hierarchical dashboards. We have also provided a detailed step-by-step process for teacher dashboard authoring in collaboration with game designers and developers; if possible, we recommend involving both stakeholders in the process, as they provide complementary inputs from their different points of view: developers and designers from their perspective closer to the game; and teachers closer to the actual use of dashboards.

In our experiments using these dashboards, we have identified recommendations for dashboard authors, together with several improvements to our dashboards that make them more understandable and usable. Many of these improvements are already implemented, and others will be added on subsequent iterations, to be tested and validated using the *Conectado* serious game and other games from RAGE and BEACONING H2020 projects.

6 Future Work

While we have done significant work to develop the analytics framework that collects, analyzes, and displays data from different games, we still have much to do if we wish to allow non-specialists to create good dashboards by themselves. Improving the process of dashboard creation will require two key steps: first, we will have to add simple editors to configure both visualizations and, much more difficult, the associated analyses from where they will receive their data. Second, we must provide a quick method to test whether these analyses and visualizations are performing as expected since non-specialists are more likely to use a trial-and-error approach to development due to their lack of development expertise. One approach is to allow the upload of small amounts of xAPI-SG or CSV-formatted interaction traces to test with, even if these have not been collected by running a real game (as illustrated in Section 3.3). The platform would process, analyze and show visualizations of these, as usual, allowing dashboards to be tinkered with before applying them to an actual game. In the same sense, decoupling sets of traces from dashboards should allow trying out a new dashboard on data from an existing game, and keeping it if the teachers consider it better than the old version; or even allowing them to choose the dashboard that they feel most comfortable with.

In addition to this, we will continue to test and evaluate the platform and its dashboards in further experiments with new games.

Acknowledgements

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 (RAGE H2020-ICT-2014-1-644187, BEACONING H2020-ICT-2015-687676, Erasmus+ IMPRESS 2017-1-NL01-KA203-035259) and by the Telefónica-Complutense Chair on Digital Education and Serious Games.

References

- [Alonso-Fernández, Calvo-Morata, Freire, Martínez-Ortiz and Fernández-Manjón 2019] Alonso-Fernández, C., Calvo-Morata, A., Freire, M., Martínez-Ortiz, I., Fernández-Manjón, B.: 'Applications of data science to game learning analytics data: A systematic literature review'; *Computers and Education*, Vol. 141 (2019). <https://doi.org/10.1016/j.compedu.2019.103612>
- [Alonso-Fernandez, Calvo, Freire, Martinez-Ortiz and Fernandez-Manjon 2017] 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). IEEE (2017), pp. 1111–1118. <https://doi.org/10.1109/EDUCON.2017.7942988>
- [Baukhage, Drachen and Thureau 2015] Baukhage, C., Drachen, A., Thureau, C.: 'The Age of Analytics'; *IEEE Transactions on Computational Intelligence and AI in Games*, Vol. 7, No. 3 (2015), pp. 205–206. <https://doi.org/10.1109/tciaig.2015.2467166>
- [Calvo-Morata, Rotaru, Alonso-Fernandez, Freire, Martinez-Ortiz and Fernandez-Manjon 2018] Calvo-Morata, A., Rotaru, D. C., Alonso-Fernandez, C., Freire, M., Martinez-Ortiz, I., Fernandez-Manjon, B.: 'Validation of a Cyberbullying Serious Game Using Game Analytics'; *IEEE Transactions on Learning Technologies* (2018), pp. 1–12. <https://doi.org/10.1109/TLT.2018.2879354>
- [Chen, Chiang and Storey 2012] Chen, H., Chiang, R. H. L., Storey, V. C.: 'Business Intelligence and Analytics: From Big Data to Big Impact'; *MIS Quarterly*, Vol. 36, No. 4 (2012), pp. 1165–1188.
- [Connolly, Boyle, MacArthur, Hainey and Boyle 2012] 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'; *Computers & Education*, Vol. 59, No. 2 (2012), pp. 661–686. <https://doi.org/10.1016/j.compedu.2012.03.004>
- [Denis and Jouvelot 2005] 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), pp. 462–465.
- [Drachen, El-Nasr and Canossa 2013] Drachen, A., El-Nasr, M. S., Canossa, A.: 'Game Analytics: Maximizing the Value of Player Data'; (M. Seif El-Nasr, A. Drachen & A. Canossa, Eds.). London: Springer London (2013). <https://doi.org/10.1007/978-1-4471-4769-5>
- [Hollins, Humphreys, Yuan, Sleightholme and Kickmeier-Rust 2017] Hollins, P., Humphreys, S., Yuan, L., Sleightholme, G., Kickmeier-Rust, M.: 'The 'Water Cooler' Game'; In *Proceedings of the 11th European Conference on Games Based Learning, ECGBL 2017* (2017).

- [Jivet, Scheffel, Drachler and Specht 2017] Jivet, I., Scheffel, M., Drachler, H., Specht, M.: 'Awareness Is Not Enough: Pitfalls of Learning Analytics Dashboards in the Educational Practice'; (2017), pp. 82–96. https://doi.org/10.1007/978-3-319-66610-5_7
- [Jivet, Scheffel, Drachler and Specht 2018] Jivet, I., Scheffel, M., Drachler, H., Specht, M.: 'License to evaluate: Preparing learning analytics dashboards for the educational practice'; In Learning Analytics & Knowledge Conference (2018).
- [Keim, Andrienko, Fekete, Görg, Kohlhammer and Melançon 2008] Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., Melançon, G.: 'Visual Analytics: Definition, Process and Challenges'; In Information Visualization (2008), pp. 154–175. <https://doi.org/10.1007/978-3-540-70956-5>
- [Loh, Sheng and Ifenthaler 2015] Loh, C. S., Sheng, Y., Ifenthaler, D.: 'Serious Games Analytics'; (C. S. Loh, Y. Sheng & D. Ifenthaler, Eds.). Cham: Springer International Publishing (2015). <https://doi.org/10.1007/978-3-319-05834-4>
- [I. Perez-Colado, Alonso-Fernandez, Freire, Martinez-Ortiz and Fernandez-Manjon 2018] Perez-Colado, I., Alonso-Fernandez, C., Freire, M., Martinez-Ortiz, I., Fernandez-Manjon, B.: 'Game learning analytics is not informagic!'; In 2018 IEEE Global Engineering Education Conference (EDUCON). IEEE (2018), pp. 1729–1737. <https://doi.org/10.1109/EDUCON.2018.8363443>
- [I. J. Perez-Colado, Rotaru, Freire-Moran, Martinez-Ortiz and Fernandez-Manjon 2018] Perez-Colado, I. J., Rotaru, D. C., Freire-Moran, M., Martinez-Ortiz, I., Fernandez-Manjon, B.: 'Multi-level game learning analytics for serious games'; 2018 10th International Conference on Virtual Worlds and Games for Serious Applications, VS-Games 2018 - Proceedings (2018). <https://doi.org/10.1109/VS-Games.2018.8493435>
- [Prasetya, Leek, Melkonian, Tusscher, van Bergen, Everink, et al. 2019] Prasetya, I. S. W. B., Leek, C. Q. H. D., Melkonian, O., Tusscher, J. ten, van Bergen, J., Everink, J. M., et al.: 'Having fun in learning formal specifications'; In Proceedings of the 41st International Conference on Software Engineering: Software Engineering Education and Training (2019), pp. 192–196.
- [Schwendimann, Rodriguez-Triana, Vozniuk, Prieto, Boroujeni, Holzer, et al. 2017] Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., et al.: 'Perceiving learning at a glance: A systematic literature review of learning dashboard research'; IEEE Transactions on Learning Technologies (2017), pp. 30–41. <https://doi.org/10.1109/TLT.2016.2599522>
- [Serrano-Laguna, Martinez-Ortiz, Haag, Regan, Johnson and Fernández-Manjón 2017] Serrano-Laguna, Á., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., Fernández-Manjón, B.: 'Applying standards to systematize learning analytics in serious games'; Computer Standards & Interfaces, Vol. 50, No. September 2016 (2017), pp. 116–123. <https://doi.org/10.1016/j.csi.2016.09.014>
- [Shoukry, Göbel and Steinmetz 2014] 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 - SeriousGames '14. New York, New York, USA: ACM Press (2014), pp. 21–26. <https://doi.org/10.1145/2656719.2656729>
- [Thomas and Kilmann 2008] Thomas, K. W., Kilmann, R. H.: 'Thomas-Kilmann Conflict Mode'; TKI Profile and Interpretive Report (2008), pp. 1–11.
- [TUGraz 2018] TUGraz: 'RAGE - Data Evaluation Watercooler pilot 1'; (2018). <https://doi.org/10.5281/zenodo.2577071>

[TUGraz 2019a] TUGraz: 'RAGE - Data Evaluation Watercooler pilot 2'; (2019a).
<https://doi.org/10.5281/ZENODO.2564932>

[TUGraz 2019b] TUGraz: 'RAGE - Data Final Evaluation RAGE Components'; (2019b).
<https://doi.org/10.5281/zenodo.2576902>

[Verbert, Duval, Klerkx, Govaerts and Santos 2013] Verbert, K., Duval, E., Klerkx, J., Govaerts, S., Santos, J. L.: 'Learning Analytics Dashboard Applications'; *American Behavioral Scientist*, Vol. 57, No. 10 (2013), pp. 1500–1509. <https://doi.org/10.1177/0002764213479363>

6.5. uAdventure: Simplifying Narrative Serious Games Development

6.5.1. Cita completa

Víctor Manuel Pérez Colado, Iván José Pérez Colado, Manuel Freire Morán, Iván Martínez Ortiz, Baltasar Fernández Manjón (2019). **uAdventure: Simplifying Narrative Serious Games Development**. In *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*. 2019. IEEE, 119-123 <https://doi.org/10.1109/icalt.2019.00030>

6.5.2. Resumen original de la publicación

Game engines and developer-friendly authoring tools have greatly simplified entertainment game development. However, this does not extend to serious games which, in particular, require involvement of non-developer stakeholders. We present the first evaluation of uAdventure, an easy-to-use game development environment for narrative point-and-click graphic adventure games. uAdventure is a re-implementation of the previously-validated eAdventure environment on top of the Unity game engine. The idea is to get most of the advantages of the Unity professional environment at a fraction of its complexity, and without requiring programming knowledge to use it. uAdventure include educational-oriented affordances, such as assessment and learning analytics; and has been formatively evaluated by heterogeneous users with different degrees of technical knowledge. The results of our evaluation show much simpler story creation for profiles that had no previous knowledge of the engine, and positive feedback from more technical profiles which would use the tool as a prototyping tool for complex projects.

uAdventure: Simplifying Narrative Serious Games Development

Pérez-Colado Víctor Manuel, Pérez-Colado Iván José, Freire-Morán Manuel, Martínez-Ortiz Iván,
Fernández-Manjón Baltasar

Software Engineering and Artificial Intelligence Department
Faculty of Informatics, Complutense University of Madrid
Madrid, Spain

e-mail: {victormp, ivanjper}@ucm.es, {manuel.freire, imartinez, balta}@fdi.ucm.es

Abstract—Game engines and developer-friendly authoring tools have greatly simplified entertainment game development. However, this does not extend to serious games which, in particular, require involvement of non-developer stakeholders. We present the first evaluation of uAdventure, an easy-to-use game development environment for narrative point-and-click graphic adventure games. uAdventure is a re-implementation of the previously-validated eAdventure environment on top of the Unity game engine. The idea is to get most of the advantages of the Unity professional environment at a fraction of its complexity, and without requiring programming knowledge to use it. uAdventure include educational-oriented affordances, such as assessment and learning analytics; and has been formatively evaluated by heterogeneous users with different degrees of technical knowledge. The results of our evaluation show much simpler story creation for profiles that had no previous knowledge of the engine, and positive feedback from more technical profiles which would use the tool as a prototyping tool for complex projects.

Keywords- *Serious game, authoring system, narrative games, learning analytics.*

I. INTRODUCTION

Serious Games (SGs) are games where the main purpose is not entertainment, being education one of the most extended. There are multiple benefits for students when it comes to learning through SGs and gamified experiences [1], most notably, a better engagement. However, despite of multiple initiatives having demonstrated positive impact in both educational institutions and companies, SGs are not yet widely used in education [2]. One of their biggest drawbacks are the high costs of using SGs for educational institutions, that must either purchase already validated games, or must develop their own games involving non-developer stakeholders in their creation, increasing development costs.

The popularity of free game engines with fully featured authoring tools has opened new possibilities in the SGs market, as it has happened with the indie game market [3]. However, the growth in freely available game authoring tools has not had as much impact for SGs, at least when compared to more mainstream types of video games. The use of SGATs (SG-specific authoring tools) can greatly simplify SG development [4], [5], making it accessible for non-programmers and non-technical SG stakeholders, most notably teachers, in the creation process. This paper describes uAdventure, an SGAT implemented on top of the Unity game engine, fully affordable for non-experts, and

with geolocalization and learning-oriented features such as learning analytics.

To simplify the development process and make it affordable to non-experts, SGATs should focus on facilitating high-level tasks, instead of overwhelming authors with unwanted flexibility in game mechanics. This can be achieved by specializing the tool in a single game genre. Specific genres support different types of learning [6]. Game-like simulators allow authentic learning through recreating real-life situations [7], but the complexity of their mechanics can be expensive to develop. Adventure games are adequate for creating educational games as they combine a strong narrative with puzzles and problem solving, promoting learning and reflection [8].

The rest of this introduction provides greater background on adventure games, how they can benefit SGs, and the general process of authoring adventure games. The rest of the paper is structured as follows: Section 2 explores uAdventure's (uA) features to simplify adventure SGs development for non-experts; Section 3 details the pilots performed with uA; Section 4 discusses the results; and finally, Section 5 comprises conclusions and future work.

A. Adventure genre for serious games

Adventures are story-driven games, and hence, they inherit elements from novels, including *characters*, which are the elements (normally persons) that populate the story and perform the different actions; *scenes*, where the action occurs; and finally, the use of *narrative* and *actions* to encompass narration, conversations and interactions of characters with their environment.

In adventure games, the player is placed inside of this narrative context, and must interact with the environment (scenarios) and the characters or items it contains, occasionally solving mini-game puzzles. These features are especially useful in educational games, as players need to



Figure 1. Broken Age (2014), a “point and click” game. The image shows two characters having a conversation and the conversation options given.

reflect on their knowledge to solve in-game puzzles and advance the story. According to [8], graphic adventures of the “point and click” (Figure 1) sub-genre, where the mouse or touch is used to select the protagonist actions, are one of the best genres for SGs.

Stories themselves can be non-linear, that is, allow access to a whole “sequence of possibilities considering changing world states.” [9]. Where a linear story follows a strict sequence, non-linear stories can result in different states (endings, scores, etc.) depending on the players. This is not without drawbacks, since, for example, in the “King’s Quest” game series, some interactions could lead to unresolvable game states named “dead ends” [10]. However, non-linearity is ideal for SGs, as it provides a greater sense of immersion for players, and the resulting game state can be used as an assessment measure, combined with the players’ responses to in-game questions and problem-solving history, among other sources of analytics.

B. eAdventure

The eAdventure (eA) SG engine (Figure 2) is a previous SGAT for the creation of “point and click” games [11]. The key goal of eA was to provide a multiplatform adventure game engine and reduce the complexity of using high-level concepts such as scenes, characters, interactions, and conversations and allowing to package games as educational learning objects using different educational standards like IMS Content Packaging and ADL SCORM [12]. For interaction, eA provided a high-level model based on *conditions* and *effects*, where the former had to be fulfilled to apply the changes in game-state described by the corresponding effects. At the time eA was created, game engine & authoring tool licenses were expensive. As a free and open-source project, eA was a good fit for educators and small teams.

eA games were mostly delivered via Java applets, which are now largely obsolete. At the time, applets allow to package and distribute a SG using a Learning Management Systems such as Moodle, but nowadays we have lost this advantage. In addition, the growth in the numbers of mainstream game platforms opened new opportunities, allowing us to focus on the distinctive eA’s features. Relying on game engines such as Unity to handle low-level tasks can greatly reduce development and maintenance costs and allow future development to focus on educationally-valuable features, that make uA unique.

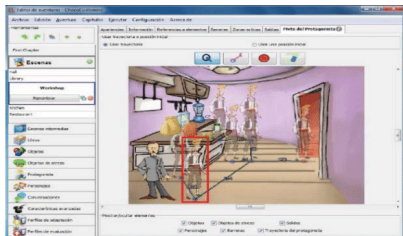


Figure 2. The scene editor in the eAdventure SGAT.

C. Unity

The Unity game engine is multiplatform, supporting the creation of games for all major operating systems, mobile devices (including iOS and Android) and consoles. It provides generic functionalities such as animations, physics, cameras and lightning systems, and support for almost any conceivable game mechanic. However, scripting is needed to create genre-specific mechanics, making it inaccessible to non-experts. Unity itself is designed to be highly customizable and supports a “Unity Store” where developers can share and plugins that add or modify its default behavior.

II. UADVENTURE FEATURES

uAdventure (uA) is a framework built on top of Unity with the goal of simplifying the creation of narrative “point and click” SGs by non-experts. By building on top of Unity, uA solves eA’s major problems: uA can always support the latest technologies and platforms, reusing most of Unity’s subsystems and focusing instead on providing features of a narrative and educational nature, making it more maintainable and less likely to become obsolete [13]. This is key to extending the lifecycle of the games created with uA and eA as it includes backwards compatibility.

This section provides an introduction of distinctive features and capabilities that uA offers. Sections II.A and II.B describe the high-level narrative model editor, which allows easy creation of narrative content to present to the players, promoting exploration and creativity through non-linearity, and allowing the addition of challenges through logic-puzzles and quizzes. Section II.C and II.D describe the simplified created to isolate non-experts from the full complexity of Unity (uA replaces most of the Unity interface with custom windows), designed to be used during narrative model creation. Hence, uA is not just a Unity tool, but a whole engine that redefines Unity’s purpose. Finally, subsections II.E and II.F describe additional features, such as learning analytics.

A. Narrative model editor

The model of uA is heavily based on narrative game elements [14]: scenes, interactive / non-interactive elements, characters, items and cutscenes. Scenes, where the game action takes place, contain all elements (interactive or not) that the players will explore and interact with. Moreover, scene definitions include areas of interest and walkable zone. Moreover, scenes can be connected through areas of interest, as the basic means to build the story. Characters (representing the player or other important characters of the game’s plot) may, among other actions, converse, move, and receive objects. Items are passive elements that can only be examined, and occasionally picked up and/or manipulated. Finally, Cutscenes can be used to introduce linear narrative sequences, such as to provide necessary exposition when a player reaches a milestone or enters a scene for the first time.

The uA narrative model includes a built-in set of actions, which, once triggered, can change the game state in both linear and non-linear stories. For example, an item in the scene can be examined, grabbed or used, and, depending on each action’s configuration, this can lead to different states,

endings or story branches. Instead of scripting, uA provides high-level effects that easily manipulate the game context, including the current scene, game-state variables, game elements, or feedback displayed to the player. Since effects have a high-level meaning and are documented within the interface, this makes it easier to trace the flow of the story, at least when compared with conventional scripting.

The conversation system provides a tree-based editor to manipulate all the different narrative branches that a non-linear story can require. Without the conversation system, dialogues could be created using effects, but the amount of effects would be just impossible to handle for large conversations with different branches. The conversation system also manages all the runtime tasks required to display the conversation, such as displaying different dialog bubbles from the character that is speaking, changing character animations, and displaying the different options in the screen for the player to select them.

B. Simplified scene editor

The uA scene editor allows the author to create scenes and manages their elements, including characters and items. Although Unity supports 3D, we have chosen to limit uA to 2D scenarios because they are more common in adventures and, due to their greater simplicity, are easier to author. Additionally, the capability of managing elements directly from the scene editor is a breaking change compared to Unity's scenario editor, where users must deal with different rendering components and behavior scripts. uA also sets scene sizes and boundaries automatically and handles the camera configuration and transitions. In addition, the scene itself may contain game-relevant elements that visually exist in the scene background. Hence, they can be handled by defining their region in the scene and defining their functionality such as exits or barriers.

Finally, the scene editor manages the properties of each element by displaying a Unity-like inspector, making it more affordable for native Unity users.

C. Integrated experience

uA provides different editors that simplify the content creation in an adventure. To provide a streamlined experience for non-experts, uA automatically switches Unity's layout, by closing unrelated windows such as the hierarchy window, and replacing them with SG-relevant views, including the model editor, context variables and flags, a gameplay window, and a project file explorer.

uA also inserts itself to respond to Unity's default functionalities. For example, when the "play" button is pressed to test the game, the corresponding uA scene is loaded transparently, without forcing users to deal with Unity scenes at all. Conversely, when the game stops, the uA window is opened back. Saving and loading projects works in a similar fashion.

Finally, uA uses the asset importing pipeline to automatically handle asset configurations when users copy assets into their project folders. For example, any image placed in the scene backgrounds folder will be automatically configured as a readable sprite, ready to be included into an



Figure 3. Simplified builder window.

actual uA scene. This allows authors to ignore the technicalities involved in configuring each asset type.

D. Simplified Builder

One of the most important features of Unity that helped it gain popularity is its multiplatform support (up to twenty-five platforms by 2019), reducing project costs by avoiding expensive platform ports, and maintenance costs by keeping a single code-base. Building a game is the process of generating executable distributions for each of the target platforms, and the builder is the Unity tool that allows it to be configured.

Supporting multiple platforms requires significant tuning to achieve good performance in each of them, for example in terms of quality or lighting. For this reason, uA provides a simplified builder that allows choosing only those platforms relevant to the educational field, and exporting the game in one step (see Figure 3), while automatically configuring all platform-dependent settings by assuming that the target audience (education) will either use the students' devices (according to "bring your own device" format [15]), or rely on institutional computers, which will frequently be running outdated software or hardware [16]. Therefore, the simplified builder enables Windows, Linux, MacOS, WebGL, Android and iOS platforms, discarding non-school-feasible platforms (e.g. consoles), and aiming for minimum requirements.

E. Learning analytics

Learning analytics (LA) are an important trend in learning understanding and assessment. For serious games, LA can allow game success (or failure) to be measured and help to understand its causes. LA applied to SG usually comprises the gathering of the main relevant game events (or, at worst, every interaction), in order to correlate these relevant events with educational objectives. Instead of defining a custom protocol, uA uses the xAPI standard, more precisely, uses the xAPI serious game's profile as interchange format for the events [17]. These events include element interaction, player movement, dialog choices and game progress.

uAd implements the xAPI profile out-of-the-box, without forcing authors to deal with analytics concepts. Since xAPI is event-based (although in xAPI terminology events are called statements), an xAPI-compliant server, such as an LRS, can gather the different events and states the non-linear stories pass through and later reconstruct the gameplay. Most major uA elements, such as scenes, can generate and send xAPI statements automatically; but to measure game progress and

success in so-called “completibles” (milestones such as quests or puzzles), uA provides an additional system to submit progress reports to the analytics component when the necessary conditions are met.

F. Extensions

Since game engines and SDKs are a very competitive environment, Unity adds novel functionalities pertaining to new technologies, such as Augmented Reality, in each new version. To leverage this evolution, uA is easily extensible, allowing the creation of new modules as trends emerge. Modules can extend the main editor, the effect system or the scene runner, among other entry points.

Two extensions have been developed so: GPS and QR. Both extensions are served in a pack of location-based game mechanics [18]. The GPS extension includes functionalities for outdoor positioning and a new type of map-based scenes, while the QR extension complements this by using QR codes and a scanner to determine player position indoors. The location-based extensions are best used in mobile platforms, allowing players to embark in geo-positioned adventures, a new genre of game that allows exploration and investigation in augmented real-world environments.

III. TESTING USER EXPERIENCE

To study the extent to which uA’s goals have been achieved we performed a preliminary evaluation with users from different profiles, including non-technical users such teachers; artists and programmers. The purpose of the evaluations was mainly finding out the difficulties and issues in uA but also to verify that with uA users are capable of: i) developing point and click game examples, ii) create non-linear stories with multiple endings, iv) develop their own stories autonomously and v) build their games.

A. Experimental design

To prove the objectives above, participants with no prior experience with uA were provided with the latest uA software and manual to teach them the narrative elements and uA concepts, which is going to be the average use case in uA developments. To make the users create a story and use the tool, the manual contains eleven guided examples, through which the user can create a simple non-linear story with multiple endings but excluding advanced tasks such as timers or macros. In the examples, the user creates a story based on existing emergency procedures in case of fire. Into the game, the player can either win or lose depending on how they choose to follow the protocol. Finally, participants were encouraged to develop their own stories to test tasks not included in the examples, and to build and test their games for the Windows and Android platforms.

After the intervention a survey was used to build a user’s profile and to measure the difficulty of each performed task in four levels: easy, normal, difficult or not-accomplished. In total, 110 simple tasks (e.g. “create a scene”) were identified to test the main uA functionality. The user profile consists in a self-classification into non-technical users, artists and programmers and a set of questions to create a narrative profile, regarding experience creating stories and

development of narrative games. Finally, users were also asked about their success (and doubts) in examples and goals and if the manual helped them through the learning process.

B. Results

The results were grouped by the different profiles and general overall success is calculated. Table I shows different sections regarding user’s background, manual and concept understanding, goals completed, and tasks completed. The last section analyzes task complexity and hence, uA success in terms of interface simplification.

Regarding the first goal, the results in Table 1 show that participants have finished 95% of the examples successfully, which is corroborated by the percentage of incomplete or “found hard” tasks, which shows that most of tasks are rated simple (~65%). Individually, artists have the lowest success in completing examples, with a 15% over the average value in tasks not tried. For goals two and three, regarding users’ creation of non-linear stories and multiple endings, the results show a 70% of success.

Regarding the manual, all users reported finding it helpful. However, all groups have also reported difficulties using uA. For the computer users, doubts with the model are related to the lack of narrative experience, as they are probably less familiarized with the concepts of characters, linear and non-linear stories and they haven’t authored stories before, which is also present in programmers. On the other hand, the doubts in effects and conditions are more visible in the groups with no programming background as expected. Finally, another relation is present between the doubts in users and the lower simplicity reported, being more evident in average computer users that only reported ~26% as simple.

Additionally, participants also reported 29 different issues, misconceptions and difficulties to be fixed in the software and its documentation for the final release.

TABLE I. SUMMARY OF RESULTS FOR DIFFERENT PROFILES (NON-EXPERTS, ARTIST, PROGRAMMER).

	User category			
	Non-tech.	Artist	Prog.	All
<i>User profile</i>				
Amount	2	2	6	10
Narrative experience (%) ^a	50	100	78	77
Experience with SGs (%) ^b	80	80	70	74
Manual				
Read entire	1	1	4	6
Found helpful	2	2	6	10
Doubts about model elements?	1	0	1	2
Doubts in effects and conditions?	1	1	2	4
Goals				
Examples completed (% n=11)	95	86	97	95
Non-linear story created	2	2	3	7
Multiple endings?	2	2	3	7
Have built the game	1	1	2	4
Individual use cases (110 in total)				
Completed (% n=110)	65,4	51,3	74,4	68,0
- Of which found simple	25,7	79,8	72,1	64,5
- Of which found normal	74,3	18,4	26,9	34,6
- Of which found hard	0,0	1,7	1,01	0,9
Not completed	0,0	3,6	0,90	1,2
Not needed / tried	34,5	45,0	24,6	30,6

a. 30% played narrative games, 40% created narrative content, 30% created narrative or SG
b. Avg. between users. Scale from 0% (I do not know what they are) to 100% (I know them, and I have used them)

IV. DISCUSSION

The software development industry has grown big and strong the last few years. However, the SG industry has not followed this trend, specially, due to costs. With uA we are trying to drastically break this trend providing a simple tool to easily create SGs ready for multiplatform support.

For this reason, uA takes Unity (a next-gen engine) and replaces all the functionalities to focus in high-level concepts without having to interact with native Unity concepts. Even so, it is impossible to hide all the Unity features and results show that users get distracted and confused by these options (e.g. the top menu in Unity). Although this may lead to some issues, results show that non-experts have been able to generate narrative stories in Unity.

Our development concluded with a pilot that reported good results on the users' objective consecution, and therefore, verified that even users with neither Unity nor programming skills can create at least simple examples. However, some results show the importance of narrative experience by users (prior to the development) to reduce the doubts. Also, some users have reported to complete all the examples but also denied having been able to create non-linear stories and multiple endings, which was the purpose of the examples. Therefore, we expect to generate a revised version of the manual, providing more background on narrative structure in the introduction, and improving uAdventure with additional in-software documentation.

V. CONCLUSIONS AND FUTURE WORK

This paper presents the preliminary evaluation of the game authoring environment uAdventure (uA) as an evolution of a pre-existing system called eAdventure (eA). With uA, it is now possible for non-experts to start developing adventure SGs without having to invest in expensive software and expert developers. uA simplifies game development by abstracting the main narrative elements of "point and click" adventure games (e.g. scenarios, characters, conversations) into a simplified, easy-to-understand model, allowing non-expert game developers to focus on the features they need.

uA substitutes the default Unity interface with genre-focused windows, simplifying Unity's editors and processes. Therefore, uA is not a Unity extension, nor a standalone editor, but an environment that transforms Unity into a "point and click" SG editor affordable to non-experts.

To improve the authoring tool, the uA environment has been formatively evaluated with different kinds of users. This evaluation has reported very promising results especially in the creation of simple narrative games by people with no previous experience in game development. This evaluation has also helped in creating a better user experience for different kinds of users.

As a next step, we would like to do a new iteration based on the analysis and observations gathered in the experiment presented in this paper and, later on, to use the updated uA version in a formal course on interactive narrative for graphical designers in a design school in Madrid. Finally, we are developing additional uA extensions to allow the creation

of location-based narrative adventure games and improving the game analytics obtained from those games. Location-based SGs creation can be greatly simplified with uA, producing games both connected to the real world and easier to integrate into the learning cycle through LA.

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).

REFERENCES

- [1] S. De Freitas, "Are Games Effective Learning Tools? A Review of Educational Games," *Educ. Technol. Soc.*, vol. 21, no. 2, pp. 74–84, 2018.
- [2] R. Dömer, S. Göbel, W. Effelsberg, and J. Wiemeyer, *Serious Games Foundations, Concepts and Practice*. Cham: Springer International Publishing, 2016.
- [3] A. Kovanto, "The Improvements for Indie Game Development," KARELIA UNIVERSITY OF APPLIED SCIENCES, 2013.
- [4] E. J. Marchiori, J. Torrente, Á. Del Blanco, P. Moreno-Ger, P. Sancho, and B. Fernández-Manjón, "A narrative metaphor to facilitate educational game authoring," *Comput. Educ.*, vol. 58, no. 1, pp. 590–599, 2012.
- [5] Á. del Blanco, J. Torrente, E. J. Marchiori, I. Martínez-Ortiz, P. Moreno-Ger, and B. Fernández-Manjón, "A framework for simplifying educator tasks related to the integration of games in the learning flow," *Educ. Tech. Soc.*, vol. 15, no. 4, pp. 305–318, 2012.
- [6] R. Aust, M. Nitsche, and J. Pelka, "Digital game-based learning and video games in teacher training. Conception, evaluation and results from Leipzig University," *Perspect. Innov. Econ. Bus.*, vol. 14, no. 3, pp. 113–131, Sep. 2014.
- [7] Center for Technology Implementation in Education, "Learning with Computer Games and Simulations," *Am. Institutes Res.*, 2014.
- [8] M. D. Dickey, "Game Design Narrative for Learning: Appropriating Adventure Game Design Narrative Devices of Interactive Learning Environment," *Educ. Technol. Res. Dev.*, vol. 54, no. 3, pp. 245–263, 2006.
- [9] U. Spierling, "Models for Interactive Narrative actions," in *Proceedings of the Sixth Australasian Conference on Interactive Entertainment - IE '09*, 2009, pp. 1–8.
- [10] K. Q. Omnipedia, "Dead end." [Online]. Available: http://kingsquest.wikia.com/wiki/Dead_end.
- [11] J. Torrente, A. del Blanco, E. J. Marchiori, P. Moreno-Ger, and B. Fernández-Manjón, "<e-Adventure>: Introducing educational games in the learning process," in *IEEE EDUCON 2010 Conference*, 2010, pp. 1121–1126.
- [12] Á. del Blanco, E. J. Marchiori, J. Torrente, I. Martínez-Ortiz, and B. Fernández-Manjón, "Using e-Learning standards in educational video games," *Comput. Stand. Interfaces*, vol. 36, no. 1, pp. 178–187, 2013.
- [13] I. J. Perez-Colado, V. M. Perez-Colado, I. Martínez-Ortiz, and B. Fernández-Manjón, "uAdventure: The eAdventure reboot," in *IEEE Education Engineering EDUCON 2017 Conference*, 2017.
- [14] A. Salter, *What Is Your Quest?: From Adventure Games to Interactive Books*. University of Iowa Press, 2014.
- [15] G. Disterer and C. Kleiner, "BYOD Bring Your Own Device," *Procedia Technol.*, vol. 9, pp. 43–53, 2013.
- [16] I. Škorić and M. Milić, "Computers in school: A student's perspective," in *MIPRO 2010 - 33rd International Convention on Information and Communication Technology, Electronics and Microelectronics, Proceedings*, 2010, pp. 1056–1061.
- [17] M. Freire, Á. Serrano-Laguna, and B. Iglesias, *Game learning analytics: Learning analytics for serious games*, 2016.
- [18] V. M. Pérez-colado, I. J. Freire-morán, I. Martínez-Ortiz, M. Freire-Morán, and B. Fernández-Manjón, "Simplifying location-based serious game authoring."

6.6. Simplifying Serious Games Authoring and Validation with uAdventure and SIMVA

6.6.1. Cita completa

Iván José Pérez Colado, Víctor Manuel Pérez Colado, Iván Martínez Ortiz, Manuel Freire Morán y Baltasar Fernández Manjón (2020). **Simplifying Serious Games Authoring and Validation with uAdventure and SIMVA**. In *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 106-108 <https://doi.org/10.1109/icalt49669.2020.00039>

6.6.2. Resumen original de la publicación

The e-UCM group is working to simplify the processes of both creation and scientific validation of educational adventure games, by making the entire process of game creation, data acquisition, and analysis more transparent and integrated. This paper presents the integration of the uAdventure authoring tool with SIMVA, a tool to simplify validation experiments, thus encompassing the full process of serious games authoring and validation.

Simplifying Serious Games Authoring and Validation with uAdventure and SIMVA

Pérez-Colado Iván José, Pérez-Colado Víctor Manuel, Martínez-Ortiz Iván, Freire-Morán Manuel, Fernández-Manjón Baltasar
Software Engineering and Artificial Intelligence Department
Faculty of Informatics, Complutense University of Madrid
Madrid, Spain
e-mail: {ivanjper, victormp}@ucm.es, {martinez, manuel.freire, balta}@fdi.ucm.es

Abstract—The e-UCM group is working to simplify the processes of both creation and scientific validation of educational adventure games, by making the entire process of game creation, data acquisition, and analysis more transparent and integrated. This paper presents the integration of the uAdventure authoring tool with SIMVA, a tool to simplify validation experiments, thus encompassing the full process of serious games authoring and validation.

Keywords- serious games, learning analytics, serious games validation, authoring tools, game-based learning

I. INTRODUCTION

For almost 12 years, the e-UCM research group has worked on different open code software tools that place teachers at the center of the development and validation process of serious games (SGs) [1], [2]. The goal is to provide teachers with easy-to-use tools that they can apply without in-depth technical knowledge. To simplify the authoring process, we first developed eAdventure [3], a game authoring platform that has recently evolved into uAdventure (uA). In parallel, we have worked on a learning analytics platform, RAGE Analytics [4], which allows the assessment of SGs, and of the students that play them. We facilitate this approach in uA using *game learning analytics* (GLA) [5], [6]. GLA enables richer game-based educational experiences. For example, by allowing interventions to help students that get stuck in the game or by assisting teachers to better control and understand the game application in the classroom. Teachers have access to per-user and per-group metrics such as student progress or in-game milestones reached.

However, applying GLA to serious games is complex, error-prone, and fragile: many technical details must be considered, including communication between components, and correct identification of games, gameplay sessions, and players. Any small glitch can cause the whole process to fail without teachers or researchers knowing very well why or being able to solve it. These issues hinder GLA adoption by non-technical users in uncontrolled, highly-diverse environments such as schools. Additional complications arise even before the games are deployed: the formal

validation of the actual educational impact of SGs is usually done by carrying out experiments where players are asked to answer pre-game and post-game questionnaires. While the key aspect of that validation is the definition of the tests, carrying out the experiments themselves is also difficult, error-prone, and cumbersome. To assist with SG validation experiments, we developed a supporting tool called SIMVA [4], [7], [8], which can partially automate the creation of online questionnaires for SG evaluation and its relation with the GLA data.

This work describes another step towards simplifying the whole process of SG authoring and validation: the integration between the uA authoring tool and SIMVA, an experimental validation support tool. Section II describes our previous workflow when authoring and validating SGs, while Section III describes how we are evolving these tools to address their limitations and create a unified authoring and validation experience. Finally, Section IV provides some conclusions and future lines of work.

II. AUTHORING AND VALIDATION OF SERIOUS GAMES USING E-UCM TOOLS

The design phase is the first step towards SG development. Stakeholders in charge of this SG design, such as teachers and educational designers (hereafter just authors) must consider not only educational aspects, such as how students will be assessed through the game but also the capabilities of the authoring tool, which dictate the set of available game mechanics. Then uA allows for the creation of “point and click” conversational games, and therefore supports the usual mechanics of this game genre: players can explore scenarios, participate in conversations by choosing responses, and ultimately solve logical puzzles involving objects and interactions with other characters.

Once a game has been designed and developed with uA, the next step is to carry out an evidence-based validation of its effectiveness. This validation is usually performed based on the comparison of players’ results in two formal external questionnaires, conducted before and after the gameplay, in what can be called pretest-game-posttest experiments [9]. To help with this process, we developed the SIMVA tool [7], [8], which allows the creation of one or more experimental

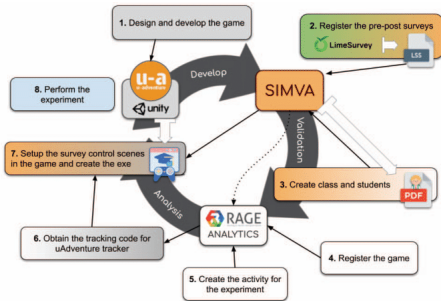


Fig. 1. uA + SIMVA + RAGE Analytics game validation process.

users' groups, managing and tracking the state of pre- and post- game tests, and collecting GLA data during and after the gameplay – a feature that is especially interesting when lacking an analytics server or experiencing unreliable internet connectivity.

SIMVA integrates a LimeSurvey (LS) [10] installation to manage questionnaires. Researchers must first author the pre and post surveys in LS, and then import them into SIMVA (step 2 of 8 in Fig. 1). Then, they set up an experimental group, with participants identified using randomly-generated alphanumeric tokens that will allow all tasks performed by the same participant to be labeled as such; and configure the LS surveys to only allow participation of players with valid tokens (step 3 in Fig. 1). Finally, to finish the setup of the experimental group, researchers can download a PDF document with the list of users' token codes, which can then be physically distributed among the participants for use during the experiments. SIMVA controls, for all participants, the specific tasks that they have performed, such as pretest, game, or posttest. Also, special template game scenes are available for authors to download, customize, and integrate into their uA games. These templates are a middleware that establishes communication between the uA game and the SIMVA server in a transparent way.

Analytics integration in uA is done using a parametrizable tracker component. When enabled in uA, this component automatically generates and sends xAPI-SG traces with relevant game events to the configured analytics server, typically an instance of RAGE analytics. The RAGE analytics dashboard then provides near real-time insights into players' actions and progress. SIMVA also automates class and student registration in RAGE Analytics (step 3 of Fig. 1). However, uA authors must still register the game, create a course, and set up the activity manually (steps 4 and 5 of Fig. 1). With the tracking code obtained in step 5, authors can proceed to step 6: configure uA's tracking component. Next, in step 7, the game authors provide developers with the survey tokens generated by SIMVA as part of step 2. With the identifiers, the developer configures the survey management scenes in uA, allowing the game to check whether its players have filled out the pre-test survey, played

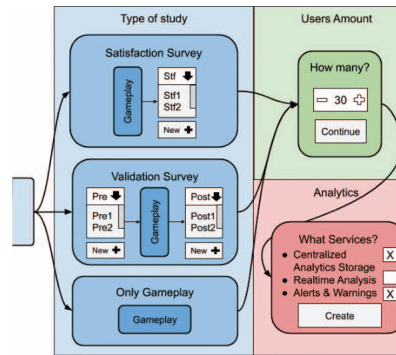


Fig. 2. Interface mockup of the three steps needed in order to setup a validation and/or assessment process in uA.

the game, and/or answered the post-test survey. Finally, as a backup to deal with the diversity of school environments, the game can also upload a backup copy of the GLA traces to SIMVA after the gameplay.

As a result, teachers can link their games to the corresponding validation experiments. Linked games will then show an initial screen where each player must enter their unique SIMVA-generated token to access the pretest, the game, and finally, the posttest. Additionally, during the experiment (step 8), the teacher will be able to access a dashboard to track the activity and observe the different decisions and progress of the players. After completing the experiment, the teacher can download all the responses as well as the GLA data from SIMVA for further analysis.

III. UNIFIED AUTHORING AND VALIDATION EXPERIENCE

We have used the configuration and setup process illustrated in Fig. 1 in several experiments, and consider it still challenging for less technically-oriented users, as there are several critical points where errors can compromise the entire validation and evaluation process. For example, failure to update a tracking-code can lead to misclassification or even loss of data. This section presents improvements to tackle both human and technological risk factors.

Ideally, the process of setting up the infrastructure for the validation and evaluation of an SG for the educator or educational researcher could be as simple as pressing a single "all-in-one button". By pressing it, any uA-authored SG would be ready to be validated. We are creating a wizard to guide the author during the three main steps of the development and validation process: i) choosing the type of study to be set up, ii) defining the number of users that will participate, and iii) selecting the analysis features to use. **¡Error! No se encuentra el origen de la referencia.** presents a mockup of this process where three types of studies are available, the number of users is established using a numerical selector, and multiple optional analysis services can be activated or deactivated depending on the teacher's

needs. Behind the scenes, uA requests SIMVA to configure and integrate all required tools. Based on wizard selections, the uA integration component: 1) creates a teacher account in SIMVA, or alternatively uses the uA author's login as an already-existing teacher; 2) generates students' users accounts using unique, random alphanumeric tokens; 3) creates a study group including both students and the teacher; 4) creates, within the study group, a study of the selected type; 5) if the study type includes surveys, the user may choose one of the survey templates or be redirected to create a new one; 6) depending on the chosen analytics capabilities, SIMVA will store the xAPI-SG traces and/or RAGE Analytics that will be configured to perform the near real-time analysis; and finally 7) creates intermediate game scenes, linked together with the current available game scenes, to orchestrate the user login and survey participation as required for the chosen study type. Additionally, a printable PDF with unique tokens to distribute to the users for the experiment is given. After the experiment is carried out, all gameplay traces and survey results (if applicable) will be downloadable from both uA and SIMVA.

Different stages of development of a serious game can benefit from different GLA uses. For instance, during the formative validation of a serious game, authors may wish to validate certain hypotheses regarding gameplay. To do so, it may be enough to simply collect some game metrics to analyze. This collection can be easily performed with uA. Once an SG is validated, it may be possible to apply machine learning techniques [11] to gain deeper insights on how students play and learn. At this point, data scientists usually require performing some exploratory analysis on the data before identifying the actual method that works best. Finally, teachers deploying the validated game in their classrooms may benefit from accessing other real-time metrics while their students are playing the game. To support these use-cases, we have begun to separate RAGE Analytics into several subcomponents, allowing us to move game analytics storage into SIMVA while keeping the near real-time capabilities inside RAGE Analytics. Additionally, by providing a file system interface (HDFS) for storage, analytics data is now easier to access from standard data science tools such as Jupyter Notebooks.

IV. CONCLUSIONS AND FUTURE WORK

The authoring and validation of serious games is a complex process. The integration of uAdventure, SIMVA, and RAGE Analytics seeks to provide a much simpler, robust and integrated flow. First, we have brought management and authoring of surveys directly into SIMVA, avoiding human error when manipulating survey files. Second, and regarding RAGE Analytics, game and activity creation, tracking-code retrieval, and linking are now automated by SIMVA, making them much easier to apply even by non-expert users. Thirdly, the integration of scenes that bind uA and SIMVA together has been automated, including all end-user access control. We are working on several further improvements. For example, we are working on partial automation of the analysis of pre-post surveys, allowing quick feedback on game effectiveness. For deeper

insights, we are testing integration of Jupyter Notebooks, which would allow exploratory analysis without leaving SIMVA. Finally, we are also working on a mechanism to streamline deployments so that students can complete a full evaluation by following a single URL, seamlessly jumping from one activity to the next. Options to achieve this include deploying games as HTML, or enabling a URI scheme which would automatically run the desired game with the specified user already authenticated in the system, similar to "uA://examplegame?authtoken=userauth".

SIMVA's latest version has great potential as an environment for creating games that can be validated following a simple, reliable, and easily configurable process, greatly increasing the value of the uAdventure tool and the games generated with it.

ACKNOWLEDGMENTS

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 (Erasmus+ IMPRESS 2017-1-NL01-KA203-035259) and by the Telefonica-Complutense Chair on Digital Education and Serious Games.

REFERENCES

- [1] Á. del Blanco, J. Torrente, E. J. Marchiori, I. Martínez-Ortiz, P. Moreno-Ger, and B. Fernández-Manjón, "A framework for simplifying educator tasks related to the integration of games in the learning flow," *Educ. Technol. Soc.*, 2012.
- [2] A. Calvo-Morata, C. Alonso-Fernandez, M. Freire-Moran, I. Martínez-Ortiz, and B. Fernández-Manjón, "Game Learning Analytics: Facilitating the Use of Serious Games in the Class," *IEEE Rev. Iberoam. Tecnol. del Aprendiz.*, vol. 14, no. 4, pp. 168–176, Nov. 2019.
- [3] I. Martínez-Ortiz et al., "Addressing Serious Games Interoperability: The eAdventure Journey," *J. Adv. Distrib. Learn. Technol.*, vol. 1, no. 1, pp. 60–76, 2013.
- [4] e-ucm, "e-UCM GitHub Page," 2017. [Online]. Available: <https://github.com/e-ucm/>.
- [5] M. Freire, A. Serrano-Laguna, B. M. Iglesias, I. Martínez-Ortiz, P. Moreno-Ger, and B. Fernández-Manjón, "Game Learning Analytics: Learning Analytics for Serious Games," in *Learning, Design, and Technology*, Cham: Springer International Publishing, 2016, pp. 1–29.
- [6] C. Alonso-Fernández, A. Calvo-Morata, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "Applications of data science to game learning analytics data: A systematic literature review," *Comput. Educ.*, vol. 141, 2019.
- [7] I. J. Perez-Colado, C. Alonso-Fernández, A. Calvo-Morata, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "Simva: Simplifying the scientific validation of serious games," in *9th IEEE International Conference on Advanced Learning Technologies (ICALT)*, 2019.
- [8] C. Alonso-Fernández, I. J. Perez-Colado, A. Calvo-Morata, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "Using Simva to evaluate serious games and collect game learning analytics data," in *LASI Spain 2019: Learning Analytics in Higher Education*, 2019.
- [9] 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.
- [10] The LimeSurvey Project Team, "LimeSurvey," 2013.
- [11] C. Alonso-Fernández, R. Caballero Roldán, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "Predicting students' knowledge after playing a serious game based on learning analytics data: A case study (in press)," *J. Comput. Assist. Learn.*, 2019.

6.7. Applications of Simva to Simplify Serious Games Validation and Deployment

6.7.1. Cita completa

Cristina Alonso Fernández, Iván José Pérez Colado, Antonio Calvo Morata, Manuel Freire Morán, Iván Martínez Ortiz y Baltasar Fernández Manjón (2020). **Applications of Simva to Simplify Serious Games Validation and Deployment**. In *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje* (Vol. 15, Issue 3, pp. 161–170). Institute of Electrical and Electronics Engineers (IEEE). <https://doi.org/10.1109/rita.2020.3008117>

6.7.2. Resumen original de la publicación

Serious games' evaluation and players' assessment is commonly done with experiments where the users play the game and fulfill one or more questionnaires. The tool Simva was designed to simplify these complex experiments, which commonly include the collection of game learning analytics data to provide further insight about players' progress and results. We present the latest updates on the tool Simva and three applications where it was used to validate serious games using pre-post experiments and collecting game learning analytics data of players' in-game interactions.

Applications of *Simva* to simplify serious games validation and deployment

Cristina Alonso-Fernández, Iván Pérez-Colado, Antonio Calvo-Morata,
Manuel Freire, Iván Martínez-Ortiz, Baltasar Fernández-Manjón

Title—Applications of *Simva* to simplify serious games validation and deployment

Abstract—Serious games' evaluation and players' assessment is commonly done with experiments where the users play the game and fulfill one or more questionnaires. The tool *Simva* was designed to simplify these complex experiments, which commonly include the collection of game learning analytics data to provide further insight about players' progress and results. We present the latest updates on the tool *Simva* and three applications where it was used to validate serious games using pre-post experiments and collecting game learning analytics data of players' in-game interactions.

Index Terms—Serious Games, Learning Analytics, Evaluation, Assessment

I. INTRODUCTION

This article is an extension of the conference paper presented at LASI-SPAIN 2019 [1] providing further details about the use of *Simva* and fully describing the updated version of the tool, which incorporates new features to simplify its use on more complex experiments.

Serious games (SGs) are applied with multiple purposes, including that of increasing knowledge, raising awareness, or changing attitudes or behaviors. Moreover, due to the interactive nature of SGs, it is possible to track players' actions during gameplay, including not only those related to the achievement of specific design goals, but also any other actions carried out by players that may provide further insight regarding their progress. This latter aspect opens the opportunity to analyze additional aspects of player behavior. In particular, analytics from tracking players has been applied in SGs for two main purposes: first, game design and validation; and second, students' assessment during deployment.

Before moving into their deployment phase, serious games need to be formally evaluated to ensure that their intended purposes are fulfilled. Their formal evaluation will ensure that games are indeed useful for their purposes and that any observed change on players characteristics (e.g. knowledge, attitude) can be attributed to their playing experience [2].

Once games have successfully undergone the evaluation process, it is usually necessary to measure how much of an effect they have had on the players that have used them. In other words, we want to be able to measure how much the serious games have changed whatever they set out to change,

such as the knowledge or awareness of its players.

The current approaches used on the literature to evaluate serious games and assess students playing them commonly rely on external paper-based questionnaires and have a correspondingly high cost, as there are no standard solutions that allow both the capture of data from in-games interactions and that gather responses to traditional paper-based evaluation questionnaires. When interactions and questionnaires are gathered separately, it takes additional effort from researchers and educators to link them back together, before attempting to extract useful information from the results.

We present the tool *Simva*, which aims to simplify these costly steps that are often part of serious games' complete evaluation and students' assessment: from interaction data collection and analysis, questionnaires management, and dealing with groups of students, including their anonymization and privacy.

The rest of this paper is structured as follows: Section 2 reviews some of the state of the art regarding evaluation of serious games data collection. Section 3 provides an overview of the tool *Simva* to simplify the validation of serious games, and also describes its latest features. The three following sections exemplify how *Simva* has been used with different serious games to simplify validation and deployment experiments. Section 4 describes the experience of using it with *Conectado*, a serious game to raise awareness about bullying and cyberbullying. In these experiments, *Simva* was used to conduct the experiments to evaluate the game while also collecting interaction data. Section 5 describes the experience with the 15 Objects test, a visual discrimination task. In these experiments, *Simva* was used to evaluate and compare two different versions of the game and two formats (paper-based and computer-based). Section 6 describes the experience with First Aid Game, a game to teach first aid techniques, where *Simva* was used to collect questionnaires and interaction data over both an initial experiment and a later recall experiment. Section 7 discusses the lessons learned from these experiences. Finally, Section 8 summarizes the conclusions of our work.

II. STATE OF THE ART

The validation of serious games and the measurement of the change they cause on their players has been performed through various methods. Researchers have carried out studies assessing games using methods such as

Manuscrito recibido el día de mes de año; revisado día de mes de año; aceptado día de mes de año.

English version received Month, day-th, year. Revised Month, day-th, year. Accepted Month, day-th, year.

Nombres de los autores, Lugares actuales de trabajo, ciudad, país (email ejemplo@ejemplo.es).
(https://orcid.org/...)

questionnaires, interviews, discussions or logs [3]. These methods can be combined in different ways according to the needs of research and validation. Assessment of the players using these games has also been done with different methods, although there is a trend towards capturing interaction data from gameplays for assessment purposes.

The first step towards evaluating games or assessing the students that play them using their interaction data is to define how interaction data is going to be collected, stored and managed. A common option is to use Learning Record Stores (LRSs), data collection repositories that allow the storage and retrieval of data following the xAPI standard [4]. However, these repositories offer no support to manage important aspects of experiments beyond interaction traces, such as student groupings, privacy, or ancillary data that is part of the experiment but not formatted as xAPI statements.

On the following subsections, we briefly describe the two most common methods to evaluate serious games: pre-post experiments and game learning analytics. Both methods are essential to understand the tool *Simva*, which allows using both together while also managing groups of students.

A. Pre-post experiments

The most common method to evaluate serious games and assess the students who play them is the use of pre-post experiments [3]. These experiments usually comprise three steps:

1. A pre-test: an initial questionnaire which assesses students' characteristic (e.g. knowledge) before the intervention. It can be paper-based or computer-based, and it should be a valid measure of the characteristics that the game aims to change. Therefore, the questionnaire itself must also be formally validated.
2. An intervention: the activity that is intended to change the students' characteristics. In the case of serious games, it will be the gameplay itself, usually from beginning to end. No time should elapse between the pre-test and the intervention, nor between the intervention and the post-test.
3. A post-test: a post-game questionnaire used to assess students' characteristics (e.g. knowledge) after the intervention. This questionnaire has the same requirements as the pre-test, and as such should include the same questionnaire to measure the characteristics, possibly in addition to questions regarding the experience.

The change in the students' characteristics is then measured by comparing the results of the pre-test and the post-test. As the only intervention between the pre-test and the post-test is the gameplay, any significant change (usually an increase in knowledge or awareness) in the characteristic measured by the questionnaires can be attributed to the intervention, which can then be considered as formally validated. Once this process is completed for a serious game, it is ready for deployment in real settings, as it has been proven as useful for its intended goals.

During the deployment phase of games, the educators, managers or researchers applying them will typically want to know how much effect the game is having on their players. For this purpose, it is possible and common to make use of

additional pre-post experiments. This, however, may not be the most precise solution, since it is an external measure. In comparison, analytics are internal to the game, and provide time and causality dimensions. However, pre-post experiments can be convenient, and even cost-effective, if the specified educational scenario or research experiment measures a characteristic for which there is no analytics available, such as cases where analytics cannot be provided or there is a lack of resources to include them.

These post-validation questionnaires again provide a measure of the characteristic before the intervention (pre-test) and after the intervention (post-test). For instance, for a game attempting to teach something, the pre-test will show how much players know about the topic before playing, the post-test will report on their knowledge on the topic after playing, and the comparison of both measures will yield how much players have learned with the game. This information will be relevant not only to teachers and educators that wish to measure the effect of games, but also to institutions which can use it to make evidence-based decisions about the application of serious games in their courses.

B. Game Learning Analytics

Besides the external measures provided by the questionnaires, another option to effectively measure the changes on students' characteristics or to obtain insight of students' gameplays on serious games is to analyze their in-game interactions. In the field of Game Analytics for entertainment games, information on player interactions has been collected transparently (in a process called tracking) for many years, primarily with profitability purposes [5]. For serious games, the combination of these Game Analytics techniques with the purposes of Learning Analytics [6] (applied in all kind of learning environments) can be termed Game Learning Analytics (GLA) [7].

GLA data collected from serious games provides information regarding in-game interactions both from an educational perspective and a gaming perspective. This information can therefore be used both to evaluate and improve the game itself, and to gain insights on the progress and results of students, and even to assess them. The information captured from players' interactions can therefore provide a rich insight for a wide set of stakeholders (teachers, managers, educational authorities, researchers, students) and for a variety of purposes (validate game design, students' assessment, improve the game, display real-time feedback, provide overview metrics) [8]–[11].

III. SIMVA

The current state-of-the-art regarding the evaluation of serious games and assessment of the students that play them, and our previous experiences in this area, led us to create a tool to simplify these complex and costly processes for any serious game after a short configuration step. For the creation of this tool we first identified several important requirements:

- Data storage and management: collection of questionnaires and game interaction data should be performed in the tool, where results should be available to be downloaded. This avoids dealing with paper-based questionnaires and simplifies linking together all information gathered from each

- Student.
- Student management and anonymization: in-tool creation of groups of students, and assignment of groups of students to questionnaires greatly simplifies keeping track of which questionnaires must be completed by which students. In addition, by issuing students with unique tokens that contain no personal information, and linking all student interaction to those tokens, students can remain fully anonymous without compromising data storage and management.
- Access control and level of completion: access to each questionnaire, and if conducting experiments with serious games, the games themselves should be managed by the tool. This ensures that, for example, students are can only play the game after filling in the pre-test questionnaire. The tool should also provide information regarding player progress and completion of pre- and post-questionnaires, allowing experimenters to gain a quick overview of the progress of the experiment, and to address problems as soon as possible.

By building a tool that fulfills these requirements, we have created *Simva*, which greatly simplifies the validation of serious games, as well as the deployment of games in educational scenarios where additional features such as questionnaires or interaction data collection are desired. *Simva* is capable of managing most of the common requirements for conducting experiments which are guided by questionnaires and where the players' interactions are collected in log files or through a GLA platform. These items include: management of any number of questionnaires, for example the two used for pre-post experiments; of groups of students and users; and of interaction data. Additionally, *Simva* also ensures both privacy and anonymity. With these features, *Simva* can, once suitably configured, be applied to any serious game's evaluation or process of student assessment.

Simva is tightly integrated with LimeSurvey [12], a software that manages questionnaires. *Simva* keeps track of the questionnaires that must be filled by each student, and stores responses for later retrieval. Students are grouped into *clases*, and each class can be assigned a different sequence of activities, such as filling in particular questionnaires or playing specific game versions. For each student in a class, *Simva* generates anonymous 4-letter tokens to be used as their username, instead of any personal identifier that could compromise privacy requirements. *Simva* can connect to a game learning analytics framework, using these tokens to link together all data pertaining to each individual user, regardless of source. Currently supported sources include questionnaires, GLA data, and arbitrary external experimental data uploaded using *Simva*'s web interface.

After generating new tokens for students, *Simva* generates a document that teachers can download and print, to then cut out tokens and hand them out to each student (see Fig. 1, bottom image). Teachers have the choice of using this document to write in the names of students next to their tokens for later reference; this allows them to track each student's learning process. However, since no copy of this correspondence between students and tokens is available in the system itself, this still ensures that privacy requirements and regulations are not infringed (e.g. helping to comply with GDPR). This technique also supports performance of recall experiments and longitudinal studies.

Serious games can then be configured to access the specific questionnaires created in *Simva*. A simple configuration file identifies the questionnaires (commonly pre-test and post-test) that the game should access. This is the only configuration required to link the game with *Simva*, therefore, simplifying the use of the tool with any serious game. These questionnaires are linked in *Simva* to one or more classes that should use them. Each class is defined by the set of tokens of its students, and the specific sequence of questionnaires and games that these students will be expected to fill or play. Data collected from each student is stored, and remains linked

Clase ElCaton 1A

<input type="checkbox"/>	Code	Conectado (Pre/Post/Other)					
<input type="checkbox"/>	GYRJ	FINISHED	FINISHED	FINISHED		TRACES	
<input type="checkbox"/>	WEFF	FINISHED	FINISHED	STARTED		TRACES	
<input type="checkbox"/>	YEYT	FINISHED	NOT FOUND	NOT FOUND		TRACES	
<input type="checkbox"/>	ZMBL	FINISHED	STARTED	NOT FOUND		TRACES	

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

Fig. 1. The *Simva* tool provides access to students' questionnaires and collected interaction traces (top figure) and a ready-to-print list of anonymous tokens that can be handed out to students (bottom figure). Figures retrieved from [22].

together by their anonymous token, including: pre-test, post-test and game interaction data (see Fig. 1, top image). All the information is then available from *Simva* to download by authorized users, typically teachers and/or researchers.

An overview of the *Simva* architecture can be seen in Fig. 2. *Simva* has been designed to be modular and extensible, and currently has two main modules: surveys (using LimeSurvey) and integration with an analytics framework (the H2020 Rage Analytics platform [13]). *Simva* allows authorized users (teachers or experimenters) to create and manage classes, surveys, games, and link them together to set up experiments.

Classes are groups of students designed to unify student management. When adding a class to *Simva*, teachers or managers set the number of students, and a corresponding amount of four-letter anonymous identifiers (4-letter random tokens such as “FJCD” or “PWNB”) are created to hand them out to students, granting them access both to surveys and to the game that is going to be used. When a class is created, *Simva* handles the setup of the corresponding users in both the survey system and the analytics framework, allowing students to send authenticated data with their *Simva*-generated tokens.

A typical experiment has 2 or 3 surveys: a pre-gameplay survey, post-gameplay survey and, if needed, an auxiliary survey. When surveys are created, *Simva* creates the necessary LimeSurvey schema files and uploads them using the corresponding API. Assigning a survey to a class adds the anonymous token identifiers from the class as participants of as valid respondents. Also using the LimeSurvey API, *Simva* can later retrieve survey completion status and use it to allow or deny access to the game itself, preventing the students from playing without answering the survey with a simple status request. When the game ends, a *Simva* API endpoint is available for game-results upload, where logs, scores, or statistics can be appended for later analysis. This can be used to add arbitrary experimental data or metadata for later

analysis.

A. Latest *Simva* version

The *Simva* tool has recently undergone an update process to add features and re-structure its internal model of experiments and surveys. The new implementation provides features for both teachers, researchers and students to simplify both the validation and deployment stages of using games in classrooms.

After these updates, users of *Simva* with a role of either teacher or researcher can manage four different entities: groups, studies, tests and activities. Groups of students are managed by the teacher/researcher, who can create them, generate printable anonymous tokens for each group, and assign them to studies. The hierarchy of studies-tests-activities is the main new feature of the new implementation of *Simva*. A study is defined as a full research work to be carried out, and it can involve different stakeholders. For example, a study to validate a serious game may be composed of a validation of the game with students, and another validation with teachers. Studies have one or more associated owners which manage the study, and contain a list of tests. A test can be used to carry out a specific part of the study or a variation of it, for instance, a test may be “Serious game validation with teachers” or “Serious game validation with students”, or even “Variant A” vs. “Variant B” for subgroup testing. Tests, in turn, contain a list of activities. Each activity corresponds to a specific task carried out during the study: for instance, an activity may be a questionnaire (e.g. pre-test), a gameplay, etc. Fig. 3 provides an example hierarchy of a study to validate a serious game with two different tests, each of them containing three and two activities, respectively.

Teachers, like researchers, can manage the creation and assignment of studies, tests and activities. As owners of the studies, they have access to the tests that belong to their studies and, in turn, the activities associated to those tests.

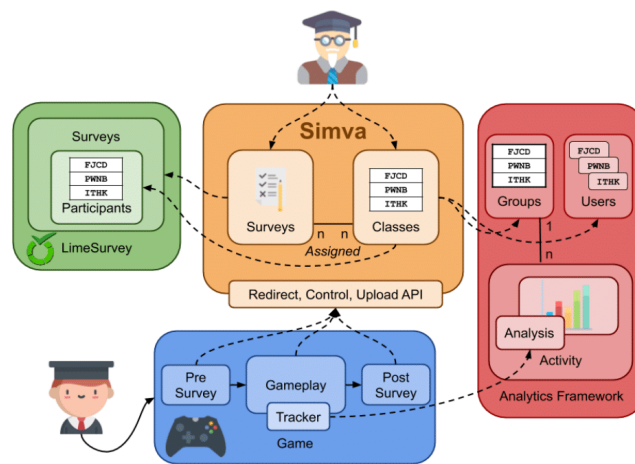


Fig. 2. Overview of the *Simva* architecture managing surveys, classes, and the Analytics Framework. Teachers can access surveys and classes in *Simva*; players typically complete a pre-test, a gameplay and a post-test. All data is sent to *Simva*, linked together using pseudonymous user tokens, and available for later retrieval.

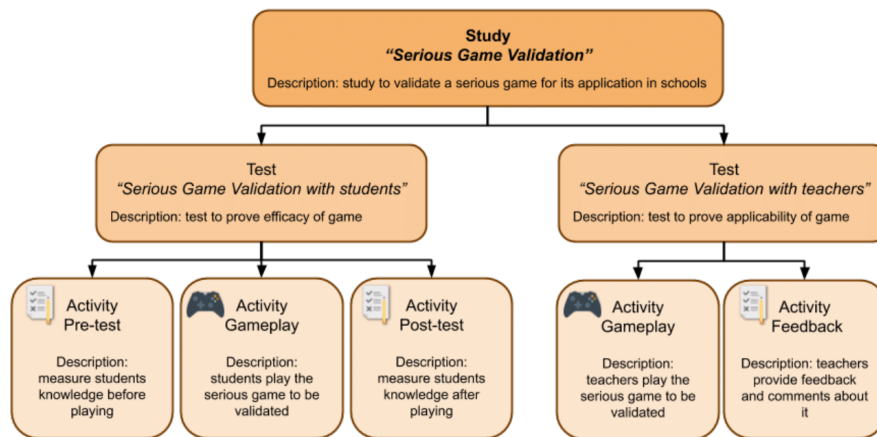


Fig. 3. Example hierarchy of a study for “Serious Game Validation”, containing two tests that, in turn, contain three and two activities, respectively. A study with a hierarchy such as this is now possible to be carried out in *Simva*.

This way, teachers can deploy games with associated features such as questionnaires to measure students/players’ learning or to provide additional activities together with the gameplay.

Students, on the other hand, simply request information via *Simva* on the studies that they have been added to (as list of activities), and can also request information on each specific activity that they have been assigned to, to know whether it is open or completed, to retrieve a list of links to currently-open activities, or simply to jump to the next activity in a given study. Student activity schedules can be obtained and updated through an API, allowing the user experience to be automated and simplified. For example, the game can connect to this API to make sure that students fill a survey before playing games; or that students complete simple tests between chapters; or even to jump through a set of scheduled mini-games.

Simva has been tested in three case studies, using different serious games, described in the following sections. In the first case study, described in Section 4, *Simva* was used to manage questionnaires and students, and to collect game interaction data, to evaluate a serious game that raises awareness about bullying and cyberbullying. Section 5 describes the experience of using *Simva* to compare two different game versions of a test for active aging. *Simva* managed all questionnaires information and its metadata feature was used to link additional information for each participant. The third experience, described in Section 6, deals with conducting a recall experiment with a game that teaches first aid techniques, and which required collection both of questionnaires and interaction data and using the anonymous identifiers to link both experiments together. *Simva* itself is open-source and available for download and use on GitHub (<https://github.com/e-ucm/simva>).

IV. EVALUATING A SERIOUS GAME TO RAISE AWARENESS: CONECTADO

Conectado is a graphic adventure video game that aims to

raise awareness about bullying and cyberbullying. The game is designed as a tool that teachers can use to kick-start discussions or debriefing sessions on the topics covered in the game with their students. The game places players in the shoes of a transfer student that suffers (in-game) bullying in a new school, and creates a strong shared experience that has proven successful both to increase awareness and to spark debate. So far, the game has been validated through several experiments in high schools with more than 1000 students between 12 and 17 years old, as well as with more than 200 teachers and educational science students [14], [15].

The game validation consisted of pre-post experiments using a formal questionnaire which assesses the players’ awareness about bullying and cyberbullying. This questionnaire was used as pre-test and post-test in the experiments to compare players’ awareness before and after playing *Conectado*. Additionally, the relevant interactions of the players with the game were collected to further analyze the players’ progress, their interactions with other game characters, and their in-game choices and attitudes. This information can provide further insight on how students have used the game and their behavior on bullying and cyberbullying situations such as the ones depicted on the game.

Before conducting the experiments, the main researcher who managed the experience prepared the pre-post questionnaires in *Simva*. This included the following steps: the surveys were registered in *Simva*, and the groups which would be using them were created, with 30 students per group, reserving a few empty student slots as a safety margin. All student slots were identified by their tokens, unique sets of 4 random letters. The list of tokens provided by *Simva* was printed in advance and carried to participating schools. The interaction data was captured by the game and sent to *Simva*, and to the Analytics Server for processing. Therefore, the users created with *Simva* were linked to the ones created in the Analytics System that collected and analyzed user

interaction data during the experiments

During the different sessions of the experiments, the main researcher only had to distribute the different printed tokens, one per student. Students then used their tokens to access the game, which had a welcome screen that requested the token to use as identifier for the data which would be sent to the analytics system. At this stage, the game checks in the configuration file for the questionnaires to be used. Then, the game accesses *Simva*, and checks that the assigned pre-test questionnaire is available for the token supplied by the student. If it is available, and not yet submitted, it automatically opens a browser with the initial survey that players must fill in. *Simva* checks that the surveys are correctly configured for the user, as identified by the unique token used to enter the game. If *Simva* indicates that the survey does not exist, or that it is not available for the indicated user, the game will not continue. When the pre-test is completed, the results are sent to *Simva* and users can access the game. After the gameplay is finished, the interaction data is sent to *Simva*. For the post-test, the same checks and process for the pre-test are repeated. If everything is correctly configured, the post-test survey is opened and, when completed, the results are sent to *Simva*.

When the experiments were completed, the main researcher could download the answers to both questionnaires as well as the interaction data from the corresponding *Simva* screen. The different data sources captured from each user (pre-test, post-test and game interactions) were linked together by the unique identifier of each player, facilitating the next analysis step.

With the information gathered in these experiments using *Simva*, the evaluation of the game *Conectado* could be performed, concluding that the game indeed increases the awareness about bullying and cyberbullying, as measured by the pre-post questionnaires. Additionally, analysis of the interaction data captured allowed researchers to extract further information such as times taken to complete the game, progress-over-time, or the different in-game choices and interactions with game characters.

V. COMPARING TWO VERSIONS OF A SERIOUS GAME FOR ACTIVE AGING: 15-OBJECTS TEST

The 15-Objects Test (15-OT) is a visual task that presents 15 overlapping objects that users need to identify as fast as possible. The aim of this test of visual discrimination is to evaluate the slowing of cognitive processing in players that may have Parkinson's disease [16]. The test is carried out with two figures of superimposed images, each with 15 objects, traditionally provided to participants on paper.

For these experiments, in addition to the traditional paper-based version of the test, we developed a new computer-based version of the 15-Objects Test with the same structure and characteristics. This new version was tested with 18 adults [17]. For this test, two different configurations of the 15-OT were created (A and B), each with a different placing of the 15 superimposed objects. To further compare the paper and computerized scores of each participant, as well as the two versions of the game (A and B), participants were randomly assigned to four experimental conditions, balanced by age and gender (see Fig. 4). These experiments were a proof-of-concept to test whether the computerized version of

this traditional test could be considered equivalent to the paper-based versions when used to research active aging.

For this experiment, the required groups of participants were first created in *Simva*. Then, both questionnaires (pre-test and post-test) were created and managed using *Simva*. These questionnaires were then linked to the groups of participants that were going to use them. All participants needed to complete both questionnaires in different moments according to the conditions shown on Fig. 4. Paper-based versions of the test were prepared in advance and handled to participants either at the beginning of the experiment (for participants assigned to experimental conditions II and IV) or at the end (participants assigned to conditions I and III). All participants were provided with their anonymous 4-letter identifiers at the beginning of the experiment. They were asked to write down their unique identifiers on both the pre-test questionnaires, and in the paper-based 15-OT tests as well as to introduce them on their computer-based 15-OT tests.

The experiment was carried out with two different game versions (A and B) depending on the objects used. As seen on Fig. 4, some participants (those assigned to conditions II and III) completed version A of the 15-OT test on paper, and version B on a computer. The rest of participants (assigned to conditions I and IV) completed version B on paper and version A on a computer. For the analysis, it was important to know which version of the game each participant was performing on the computer (and therefore which one they were completing on paper). We uploaded this information to *Simva* as additional experimental metadata, linked to each to participants' answers and questionnaires, as this was an early version that did not yet support studies composed of multiple tests. After the additional data was uploaded, the class view clearly showed which version of the game each participant was completing on the computer (A or B). This information was displayed in *Simva* as shown in Fig. 5.

After the experiment, participants' responses to both questionnaires were stored in *Simva*, linked together with the information of the game version used in each condition by the unique identifier token provided for students. Researchers could then analyze the questionnaire responses together with the game version to compare both the paper-based and the computer-based versions of the test, as well as to study the

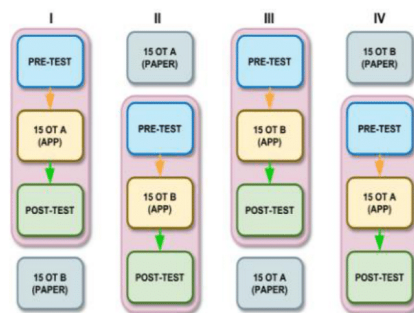


Fig. 4. Four experimental conditions in the 15-Object test experiments. Participants were randomly assigned to one of these four conditions. Picture adapted from [17].

<input type="checkbox"/>	Code	AB
<input type="checkbox"/>	YNHW	A ▼
<input type="checkbox"/>	PWPI	B ▼
<input type="checkbox"/>	DVVB	A ▼

Fig. 5. A group of participants in the *15-Object test* experiment, showing the game version each one used (A/B). This example shows a possible use of the metadata feature in *Simva*.

equivalence between the two versions of the game used.

The comparison of both game versions was easily carried out, with the aid of the information gathered in these experiments using and *Simva*, also comparing the paper-based and the computer-based versions of the test. Results yielded no significant differences between both game versions, proving their equivalence. Since no significant differences were found in results between the paper-based and the computer-based version of the test, we concluded that the computerized version of the test is a valid and equivalent alternative to the traditional paper-based version.

VI. COLLECTING GLA DATA AND CONDUCTING A RECALL EXPERIMENT: FIRST AID GAME

The First Aid Game is a videogame to teach young players life-saving maneuvers in different emergency situations. The game had already been formally validated using a traditional paper-based pre-post setup in a previous experiment. In that validation, the game was even compared with a control group in a theoretical-practical demonstration of the same topics covered in the game [18]. This was, in itself, a good example of the problems that can arise with these types of experiments: after completing the experiments, researchers had to deal with a large number of questionnaires on paper, which were often hard to read, and which had to be painstakingly transcribed to a computer for their analysis.

A new set of experiments was carried out using *Simva*, where data were collected for more than 300 students from 12 to 17 years old [19], [20]. Students completed the pre-post questionnaires assessing their knowledge about first aid techniques, adapted from the questionnaires used on the original validation experiment [18]. For this experience, the game had been rebuilt using a different game technology while maintaining the same design and game mechanics. Nevertheless, these experiments were used to validate that the updated version of the game was still effective at increasing players' knowledge. Additionally, the tracking of in-game interaction data was incorporated to this new version of the game. The interaction data collected while students played the game included game scores, in-game choices and responses, and their interactions with the different game elements. This information helped to follow students' progress and results while they were playing.

In these experiments, the pre-post questionnaires and the groups of students were handled using *Simva*. At the

beginning of each session, teachers provided tokens that they had previously downloaded and printed from *Simva* to their students. During the session, teachers wrote down the name of each student next to their token in the printed copies. In this way, teachers were the only stakeholder who had the correspondence between anonymous 4-letter tokens and the students they corresponded to – in a physical medium, safe from online misuse. Tokens could be reused by the same student in the future. No personal information was ever entered in the game. Teachers were encouraged to keep the token print-outs for possible future activities.

A few weeks after completing the training with the game in the school, researchers returned to perform an additional experiment to measure recall of knowledge learned with the game. For this experiment, teachers provided the same token to each student from the paper files they kept (where they had manually written the name of each student next to their assigned token). *Simva* tokens allowed all information from students to be grouped by student while preserving anonymity (from anyone except their teachers), both for the original experiment and the subsequent recall experiment. This greatly simplified the process of analyzing whether students recalled what they had learned, as all the information from their questionnaires and in-game interactions from both set of experiments could be linked together with the anonymous token. **Fig. 6** depicts the experimental setting of these two consecutive experiments using the First Aid Game.

The combination of both sets of experiments helped to measure not only how much students learned while playing, but also how much they remembered a few weeks after the original validation experiment. From their initial knowledge (measured in the pre-test of the original experiment) to their final knowledge (measured in the post-test of the recall experiment), we could determine how much their knowledge had improved with the experience, and in the time in-between (where they could have had other interventions related to the topics covered in the game). In a more fine-grained analysis, recall from the first experience to the second one could be measured by comparing their final knowledge (post-test) on the original experiment with the knowledge they had a few weeks later, before any other intervention (pre-test in the recall experiment). This shows not only that players learn while playing, but also that they are able to recall the things they have learned with the game.

VII. DISCUSSION

Simva aims to reduce the complexity of experiments used when validating and deploying serious games. The three experiences described in this paper (Sections 4, 5 and 6) exemplify how *Simva* has helped to simplify the validation of different serious games as well as the assessment of the students playing them. The evaluation of games has been performed by combining both traditional pre-post experiments with the information collected from in-game interactions in the form of GLA data, after a simple configuration step to link the serious games with *Simva*. Additionally, some features included in *Simva* have also simplified the execution of experiments with complex requirements, such as comparing two versions of a game or conducting a recall experiment.

In the experiments described in this paper, all the

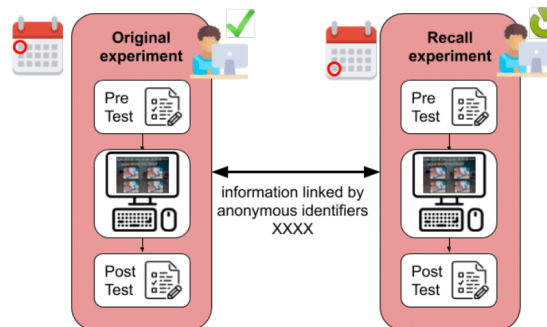


Fig. 6. Experimental setting for *First Aid Game* original and recall experiments. *Simva* simplified linking information between both experiments, and between the different data sources (pre-test, game analytics, post-test) on each experiment.

information gathered from the different sources (questionnaires and game interactions) was collected in *Simva* and linked to each user by their unique anonymous identifier tokens. These identifiers are provided to players by the managers of the activity, who obtained them from the lists of tokens returned by *Simva* when creating the required classes of students. For each participant, researchers have then been able to extract all the information of the experiment: pre-test, post-test, GLA interaction data, and, where applicable, the version of the game played – or the pre-test, post-test, and GLA interaction data from the following recall experiment. All information collected can be used to simplify games validation and deployment by teachers on their own.

From these experiences, we have learnt several lessons that that we consider are key when conducting experiments in real settings to validate serious games, or to deploy them and assess their players:

- **Ensuring user privacy:** to adequately conduct the pre-post experiments, it was essential that the tool we were using to manage students and questionnaires, in this case *Simva*, automatically dealt with and ensured privacy. To effectively ensure privacy, the safest option is to collect and handle no personal information from the experiments. In our experiences, neither *Simva* nor the Analytics System, where interaction data was also being sent to, collected any personal information. Despite ensuring privacy, it still linked together all required information from each student (pre-test, post-test, game interactions, any additional metadata), which was necessary for later analysis. For this purpose, pseudo-anonymization, via the 4-random-letters unique identifier tokens, automatically provided by *Simva* when classes of students are created, has been an effective solution, as privacy is ensured while keeping all student information linked together. For other researchers in similar scenarios, we encourage the use of such a simple anonymization system, which effectively links all the information gathered from each user, simplifying later analysis, while ensuring privacy as the user identifier does not provide any personal information and is the only

identification input into the system.

- **Collecting different data sources:** the collection of questionnaires done with *Simva*, as automatically launched from the game and performed online collecting the results without any effort, reduces the time and cost of carrying out pre-post experiments, as well as the use of paper and printer supplies. An additional option was also available during the experiments to collect the information offline. With this option, all data was stored in the computers were students were playing, to be later manually collected by researchers in case of network connection problems. All the interaction data was also stored and linked with the questionnaires online and offline. The option to upload additional metadata to *Simva* as well as the possibility of linking the information from several experiments together has also been very useful, as it simplifies the later analysis of each users' data, regardless of source. We encourage researchers to consider similar options to link together their different data sources as it greatly simplifies the later steps of an experiment's analysis.
- **Moving from pre-post experiments to GLA:** although, in the three cases presented, we have used the traditional pre-post experiments to evaluate serious games' efficacy and assess students playing the game, we consider that the information extracted from in-game interactions is also essential, and research should move towards always including this type of information. In our case, we have used the xAPI-SG Profile [21] as the data collection standard for in-game interactions. This Profile describes a common set of interactions that describe the main events found in serious games, and a specific set of verbs and activity types to use when storing them as xAPI statements. We used different trackers for each game, but since they all sent interaction data using the same xAPI-SG Profile, all of them could be analyzed and displayed using the same tools. We also recommend researchers to use this or other standard when collecting interaction data as it simplifies data reuse and integration in larger systems, including the

collection process itself [11].

We consider that the combination of these lessons learned from our work, collecting different data sources from traditional pre-post experiments to more informative GLA data while ensuring privacy, can benefit the execution of experiments to evaluate serious games and assess the students who play them. The tool *Simva* has been effective in the described experiments as it has greatly simplified the execution of the three case studies. New features added after these experiments can simplify the execution of even more complex, long-term experiments and deployments by using its new hierarchy of studies, tests and activities.

VIII. CONCLUSIONS

Increased adoption of serious games in education builds upon their previous formal validation. Despite their drawbacks, pre-post experiments are still one of the most common evaluation methods for serious games and for the assessment of students who play them. Together with these classical experiments, the use of game interaction tracking has also increased, providing richer information based on the in-game actions of students/players. The combination of pre-post experiments and collection of game learning analytics data make the formal validation of games complex and costly. Simplifying these experiments and making them more user-friendly and reducing their costs both in time and effort can greatly improve the application of games in real settings, simplifying their evaluation and deployment, and increasing their application including assessment of players.

The tool *Simva* was created for this purpose to simplify all the issues that need to be dealt with when conducting experiments to validate serious games. *Simva* manages both questionnaires and groups of players, addresses privacy concerns, allows the collection and linking of information from different data sources (both questionnaires and in-game interactions), and includes additional features for requirements such as adding arbitrary experimental metadata or carrying out recall experiments. Additionally, *Simva* has been recently updated to extend and improve management of complex or multi-part experiments. By adopting a three-layer hierarchy of studies, tests and activities, experiments that include complex setups, span multiple stages, or involve several stakeholders are much easier to carry out. As future work, we plan to validate this latest *Simva* extension on new experiments with these characteristics.

The three specific applications of *Simva* described in this paper show how the tool has been used in real settings and for different goals related to the evaluation of serious games and the assessment of the students playing them. We have also presented several lessons learned from these experiences, which we hope will contribute to further research on this area.

ACKNOWLEDGMENT

This work has been partially funded by the Regional Government of Madrid (eMadrid P2018/TCS4307), 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] C. Alonso-Fernández, I. J. Perez-Colado, A. Calvo-Morata, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "Using Simva to evaluate serious games and collect learning analytics data," in *LASI-SPAIN*, 2019.
- [2] C. S. Loh, Y. Sheng, and D. Ifenthaler, *Serious Games Analytics*. Cham: Springer International Publishing, 2015.
- [3] 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.
- [4] ADL, "xAPI Lab." [Online]. Available: <http://adlnet.github.io/xapi-lab/>. [Accessed: 03-Jun-2016].
- [5] M. El-Nasr, A. Drachen, and A. Canossa, *Game Analytics: Maximizing the Value of Player Data*. London: Springer London, 2013.
- [6] P. Long and G. Siemens, "Penetrating the Fog: Analytics in Learning and Education," *Educ. Rev.*, pp. 31–40, 2011.
- [7] M. Freire, Á. Serrano-Laguna, B. M. Iglesias, I. Martínez-Ortiz, P. Moreno-Ger, and B. Fernández-Manjón, "Game Learning Analytics: Learning Analytics for Serious Games," in *Learning, Design, and Technology*, Cham: Springer International Publishing, 2016, pp. 1–29.
- [8] V. Shute and M. Ventura, "Stealth Assessment," in *The SAGE Encyclopedia of Educational Technology*, 2455 Teller Road, Thousand Oaks, California 91320: SAGE Publications, Inc., 2013, p. 91.
- [9] A. R. Cano, B. Fernández-Manjón, and Á. J. García-Tejedor, "Using game learning analytics for validating the design of a learning game for adults with intellectual disabilities," *Br. J. Educ. Technol.*, vol. 49, no. 4, pp. 659–672, Jul. 2018.
- [10] M. Muratet, A. Yessad, and T. Carron, "Understanding Learners' Behaviors in Serious Games," in *Advances in Web-Based Learning - ICWL 2015*, vol. 9412, F. W. B. Li, R. Klamma, M. Laanpere, J. Zhang, B. F. Manjón, and R. W. H. Lau, Eds. Cham: Springer International Publishing, 2016, pp. 195–205.
- [11] C. Alonso-Fernández, A. Calvo-Morata, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "Applications of data science to game learning analytics data: a systematic literature review," *Comput. Educ.*, 2019.
- [12] The LimeSurvey Project Team, "LimeSurvey," 2013. [Online]. Available: <https://www.limesurvey.org/>.
- [13] eUCM Research Group, "RAGE Analytics," 2016. [Online]. Available: <https://github.com/e-ucm/rage-analytics>. [Accessed: 13-Nov-2016].
- [14] A. Calvo-Morata, D. C. Rotaru, C. Alonso-Fernández, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "Validation of a Cyberbullying Serious Game Using Game Analytics," *IEEE Trans. Learn. Technol.*, pp. 1–1, 2018.
- [15] A. Calvo-Morata, M. Freire-Moran, I. Martínez-Ortiz, and B. Fernández-Manjón, "Applicability of a Cyberbullying Videogame as a Teacher Tool: Comparing Teachers and Educational Sciences Students," *IEEE Access*, vol. 7, pp. 55841–55850, 2019.
- [16] B. Pillon *et al.*, "Cognitive slowing in Parkinson's disease fails to respond to levodopa treatment: The 15-objects test," *Neurology*, vol. 39, no. 6, pp. 762–762, Jun. 1989.
- [17] M. C. D.-M. Dan-Cristian Rotaru, Sara García-Herranz, Manuel Freire, Iván Martínez-Ortiz, Baltasar Fernández-Manjón, "Using Game Technology to Automate Neuropsychological Tests and Research in Active Aging," *GOODTECHS 2018 - 4th EAI Int. Conf. Smart Objects Technol. Soc. Good*, 2018.
- [18] E. J. Marchiori, G. Ferrer, B. Fernández-Manjón, J. Povar-Marco, J. F. Suberviola, and A. Gimenez-Valverde, "Video-game instruction in basic life support maneuvers," *Emergencias*, vol. 24, no. 6, pp. 433–437, 2012.
- [19] C. Alonso-Fernández, A. R. Cano, A. Calvo-Morata, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "Lessons learned applying learning analytics to assess serious games," *Comput. Human Behav.*, vol. 99, pp. 301–309, Oct. 2019.
- [20] C. Alonso-Fernández, R. Caballero Roldán, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "Predicting students' knowledge after playing a serious game based on learning analytics data: A case study (in press)," *J. Comput. Assist. Learn.*, 2019.
- [21] Á. Serrano-Laguna, I. Martínez-Ortiz, J. Haag, D. Regan, A. Johnson, and B. Fernández-Manjón, "Applying standards to systematize learning analytics in serious games," *Comput. Stand. Interfaces*, vol. 50, pp. 116–123, 2017.
- [22] I. J. Perez-Colado, C. Alonso-Fernández, A. Calvo-Morata, M.

Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "Simva: Simplifying the scientific validation of serious games," in *9th IEEE International Conference on Advanced Learning Technologies (ICALT)*, 2019.



Cristina Alonso Fernández obtained her Bachelor in Computer Science and her Bachelor in Mathematics for the Complutense University of Madrid in 2016. A year later, she finished the Master in Data Mining and Business Intelligence, also for the UCM. Since September 2016, she is part of eUCM, where she has worked for the H2020 Beacons Project. She is currently doing her PhD in Computer Science. Her research interests include educational videogames and application of data analysis and data mining for their improvement.



Iván José Pérez Colado received a Bachelor in Software Engineering for the Complutense University of Madrid in 2014. Two years later, he completed the Master in Computer Science, also at the UCM. As his Master Thesis was uAdventure, a new implementation of eAdventure in Unity 3D, Dr. Baltasar Fernández Manjón offered Iván the chance to continue the development of uAdventure being a researcher of the eUCM Group. He also started his PhD in Computer Science that same year. Six months after that, Iván became part of the H2020 Beacons Project. His research interests include educational videogames and the authoring tools used to create them. Due to his work in Beacons, the study of Learning Analytics techniques and standards is one of his main research areas.



Antonio Calvo-Morata obtained his bachelor in Computer Science for the Complutense University of Madrid in 2014. In 2017, he completed the Master in Computer Science, also in the Complutense University. He is currently doing his PhD in Computer Science. Antonio has been part of the research group e-UCM since 2014, as a contract researcher for projects eMadrid and H2020 RAGE. His research interests include the study of educational videogames and their application in schools, as well as the use of Learning Analytics techniques to improve their efficacy and their validation as an educational tool.



Manuel Freire has a PhD in Computer Science from the Universidad Autónoma de Madrid (UAM). He is interested in Information Visualization, Human-Computer Interaction, Online Learning, Serious Games, and Plagiarism Detection. He performed a 2008 post-graduate Fulbright scholarship in the Human-Computer Interaction Lab (HCIL-UMD), working with Ben Shneiderman and Catherine Plaisant. In 2010, he became a member of the e-UCM group in the Universidad Complutense de Madrid (UCM), where since 2013 he is an Associate Professor.



Iván Martínez-Ortiz works as Associate Professor in the Department of Software Engineering and Artificial Intelligence (DSIA) at the Complutense University of Madrid (UCM). He has been assistant to the Vice-Rector of Technology at UCM and Vice-Dean for Innovation in the Computer Science Studies. He has been Lecturer in the Computer Science School at UCM in the Computer Science School at the Centro de Estudios Superiores Felipe II. He received a Bachelor in Computer Science (first in the Dean's List "premio extraordinario") and a Master and PhD in Computer Science from the UCM. His research interests include e-learning technologies and the integration of educational modeling languages, serious games and e-learning standardization.



Baltasar Fernández-Manjón received the PhD degree in physics from the Universidad Complutense de Madrid in 1996. He is a Full Professor of computer science at UCM and director of the e-UCM e-learning research group. He has the Honorary Complutense-Telefonica Chair on Digital Education and Serious Games. His research interest is focused on the applications of ICT in education and in serious games and educational simulations applied to different domains (e.g. medicine, education). He is also working in game learning analytics and the application of e-learning standards to the integration of those technologies in e-learning systems. He is a senior member of the IEEE.

6.8. e-learning Standards in Game-Based Learning?

6.8.1. Cita completa

Iván José Pérez Colado, Víctor Manuel Pérez Colado, Manuel Freire Morán, Iván Martínez Ortiz y Baltasar Fernández Manjón (2021). **e-learning Standards in Game-Based Learning?** In *2021 International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 81-82 <https://doi.org/10.1109/icalt52272.2021.00032>

6.8.2. Resumen original de la publicación

Despite the proven benefits of Serious Games when compared to traditional e-learning, uptake of game-based learning in mainstream education is still very low. Increasing uptake requires making suitable games easier to deploy as an activity type in existing e-learning platforms, without compromising game quality or increase development cost; and leveraging data generated by players to improve their learning outcomes. We briefly discuss available e-learning standards to address these issues and describe our current standards-based approach to both game deployment and evaluation (including learning analytics).

e-learning Standards in Game-Based Learning?

Pérez-Colado, Iván José; Pérez-Colado, Victor Manuel
Dept. of Software Engineering and Artificial Intelligence
Complutense University of Madrid
Madrid, Spain
{ivanjper,victormp}@ucm.es

Freire, Manuel; Martínez-Ortiz, Iván; Fernández-Manjón, Baltasar
Dept. of Software Engineering and Artificial Intelligence
Complutense University of Madrid
Madrid, Spain
{manuel.freire, imartinez, balta}@fdi.ucm.es

Abstract—Despite the proven benefits of Serious Games when compared to traditional e-learning, uptake of game-based learning in mainstream education is still very low. Increasing uptake requires making suitable games easier to deploy as an activity type in existing e-learning platforms, without compromising game quality or increase development cost; and leveraging data generated by players to improve their learning outcomes. We briefly discuss available e-learning standards to address these issues and describe our current standards-based approach to both game deployment and evaluation (including learning analytics).

Keywords—serious games, e-learning standards, learning analytics, learning management systems

I. INTRODUCTION

Teachers that want to include game-based learning in their courses confront multiple challenges, including: i) finding suitable games, ii) providing access to students as activities integrated into their institutional Learning Management System (LMS), including grading and evaluation; and iii) leveraging the high degree of interactivity of games to enjoy the benefits of Learning Analytics. Ad-hoc solutions for these challenges are expensive to develop and maintain, and lead to vendor lock-in, compromising the institution's freedom to change platforms or manage the resulting data. The logical answer is to rely on open standards, which allow disparate systems to interoperate and fuel a thriving ecosystem of e-learning systems. This paper describes how current and upcoming e-learning standards can facilitate deployment of serious games in educational contexts while providing access to learning analytics.

II. E-LEARNING STANDARDS AND EDUCATIONAL RESOURCES

Table I provides an overview of some of the main e-learning standards and their features, with columns grouped into metadata, to describe and later locate relevant games; game packaging and access; and communication/analytics. Several standards cover multiple concerns at once.

A. Accessing activities: IMS CP, LTI and CC

IMS Content Package has been widely adopted by the most popular LMSs and is actively used for content/game packaging in learning environments. When packaging games or other highly interactive content, IMS CP is limited to web-based distribution, which is suitable only for technologies such as WebGL. IMS CP packages are self-contained, placing a significant load on LMS hosting infrastructure when delivering large games and complicating updates for deployed games.

The IMS Learning Tools Interoperability standard focuses on communications between LMS platforms and tools. It provides a mechanism to forward user authentication and authorization from the LMS to an external tool for the duration

of a session. During gameplay, tools can report back completion and grades for tasks carried out by students. LTI is simple to use for teachers, which only need to provide a link to the LMS that can then automatically configure the corresponding activity. Nevertheless, on the developer side, it is a complex standard to implement and it cannot handle native applications by itself. LTI is currently adopted by most major LMSs.

Finally, IMS Common Cartridge (CC) is a specification by IMS similar to ADL's SCORM, using IMS CP as a packaging base, but with a more flexible approach. It provides better mechanisms to build packages from modules, greatly facilitating reuse of content within a package. It can also include LTI links as content, providing web-distribution and communicating activity with LMSs. However, there is a low adoption rate among the main open-source LMSs such as SAKAI and Moodle; additionally, supported versions tend to be outdated, and may lack support for LTI.

TABLE I. MAIN E-LEARNING STANDARDS AND THEIR FEATURES

E-L standard	Feature					
	Metadata	Packaging	Delivery	Protocol	Report	LA
LOM	X	-	-	-	-	-
CP	LOM	X	-	-	-	-
SCORM	LOM	CP	-	X	x	-
LTI	-	-	X	-	X	-
CC	LOM	X	LTI	-	LTI	-
xAPI	-	-	-	-	-	X
CMI-5	Basic	X	X	X	X	X

B. Communication & LA.: SCORM, xAPI, Caliper and CMI5

The Shareable Content Object Reference Model (SCORM) is a web-based collection of standards and specifications for Learning Objects developed by ADL. It builds on IMS CP for packaging and IMS LOM for metadata, and then adds a simple communication layer. SCORM versions 1.4 and 2004 are well-supported and used for game-based learning. However, SCORM inherits several problems from IMS CP, provides a very constrained information stream which is mostly limited to result-score pairs, and only supports web-based platforms.

The xAPI specification is an activity stream data model for Learning Analytics that requires an external communication protocol to be implemented and relies on an external Learning Record Store (LRS) to store the information. Its focus is to provide a flexible format to trace all meaningful interactions,

using simple statements that include an actor, an object and a verb, and can optionally include results, score and a set of customizable extensions. The xAPI model can be extended with *application profiles* that add more verbs, objects and extensions to describe specific domains, such as Serious Games when using xAPI-SG [1]. IMS Caliper is a related standard that also supports activity streams, using events instead of statements, each with an actor, action and object which align with xAPI's actor, verb and object. These events can also use specific vocabularies. However, adoption of Caliper is currently lagging that of xAPI in e-learning systems.

The ADL/AICC CMI-5 protocol is a packaging and communication standard for xAPI-based courses. It is the spiritual successor of SCORM, addressing most of its weaknesses. CMI-5 uses a specific xAPI profile to communicate with the LRS and describe the state of activities, using extensions for its outcomes. Launching a CMI-5 activity from an LMS transfers the user to the activity through a URL-based protocol. This allows native applications to be launched by using URI Schemes, supported in most operating systems. When launching activities, LMSs can provide authorization and user information as actors to the corresponding application, solving the task of identifying the user and gathering user information for activities. While CMI-5 was launched in 2016, it remains, as of this writing, unsupported by mainstream LMSs.

III. A STANDARDS-BASED PROPOSAL

Our proposal (Fig.) focuses on the task of delivering games using standards that go beyond the limits of IMS CP and SCORM, providing support for native applications and better game distribution, while integrating grading, evaluation, and rich analytics with xAPI. Since CMI-5 is not yet widely supported, we use an alternative approach for hosting xAPI-based CMI-5 packages, using LTI to delegate the CMI-5 package launch to an intermediate xAPI-compliant tool that also contains an LRS to store the traces. With LTI, the LMS creates a bidirectional communication channel with the mediator tool. In one direction, it delivers the user's authorization session to the mediator platform, and in the other direction results from the activity are reported back to the LMS. Once activities finish, xAPI reports become accessible via LTI Submission Review.

User experience is also an important factor in this approach. Students enjoy a seamless experience with a one-click launch of the game by following a link presented by their LMS, since the CMI-5 launch is seen as a redirection. This gives developers freedom to choose where the game is hosted, and can boost the speed with which the game can be launched and simplify maintenance of deployed games. The teachers' task is also streamlined when game packages are managed by developers, teachers just have to include the appropriate LTI links in their LMS courses, instead of having to upload packages to the mediator tool to obtain the corresponding LTI links.

We have implemented most of this proposal as a proof of concept, using uAdventure (an authoring tool built on top of Unity 3D) for game development, packaging and xAPI support; and Simva (a multi-purpose validation tool that includes user management, experiment design, cloud storage, and data-science [2]) for LTI connection, a CMI-5 inspired launch, and to

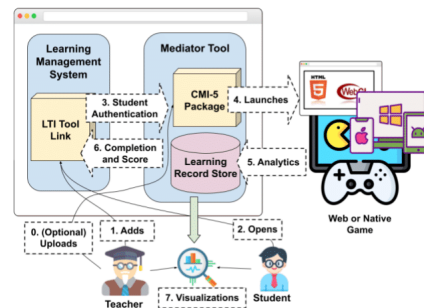


Fig. 1. Using games in the class with LTI and CMI-5 through a mediator tool that supports CMI-5 Launch for xAPI-Based analytics.

provide basic LRS-like functionality. This “proxy” approach can launch both online and native games and transparently handle player authorization during the activity. Our vision is for game developers to automatically package and export games to Simva, generating LTI links that allow teachers to include the resulting game in their courses, resulting in both interoperable games and simple-to-integrate game-based activities.

IV. CONCLUSIONS AND FUTURE WORK

Game-based learning offers educators significant benefits compared to traditional methods. But to generalize its use, serious games should be easier to deploy in real setting. Upcoming e-learning standards such as CMI-5 could be a solution but are not yet supported by any major LMSs.

We propose an approach to integrate games in LMS courses that relies on the widespread LMS support of LTI to integrate a mediator tool that supports launching xAPI activities, packaged using CMI-5. We consider that our approach would result in more interoperable games that include rich learning analytics and are simple-to-integrate with other e-learning activities.

ACKNOWLEDGMENT

This work has been partially funded by the Regional Government of Madrid (eMadrid S2018/TCS-4307, co-funded by the European Structural Funds FSE and FEDER), by the Ministry of Education (TIN2017-89238-R) and by the Telefónica-Complutense Chair on Digital Education and Serious Games.

REFERENCES

- [1] Á. Serrano-Laguna, I. Martínez-Ortiz, J. Haag, D. Regan, A. Johnson, and B. Fernández-Manjón, “Applying standards to systematize learning analytics in serious games,” *Comp. Std. Interfaces*, vol. 50, pp. 116–123, 2017.
- [2] I. J. Pérez-Colado, V. M. Pérez-Colado, I. Martínez-Ortiz, M. Freire, and B. Fernández-Manjón, “A Scalable Architecture for One-Stop Evaluation of Serious Games,” *LNCS*, vol. 12517 LNCS, pp. 69–78, 2020.

Referencias

- Ablegamers Foundation. (2012). Game accessibility guidelines | A straightforward reference for inclusive game design. Retrieved November 1, 2022, from <https://gameaccessibilityguidelines.com/>
- Abt, C. (1970). *Serious games*. New York: The Viking Press.
- Afreen, R. (2014). Bring Your Own Device (BYOD) in Higher Education: Opportunities and Challenges. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, 3(1), 133–164.
- Alonso-Fernandez, C. (2017). *Applying Data Mining Techniques to Game Learning Analytics*. UCM.
- Alonso-Fernández, C., Calvo-Morata, A., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2019). Applications of data science to game learning analytics data: A systematic literature review. *Computers & Education*, 141, 103612. <https://doi.org/10.1016/j.compedu.2019.103612>
- Alonso-Fernandez, C., Calvo, A., Freire, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2017). Systematizing game learning analytics for serious games. In *2017 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1111–1118). IEEE. <https://doi.org/10.1109/EDUCON.2017.7942988>
- Alonso-Fernández, C., Martínez-Ortiz, I., Caballero, R., Freire, M., & Fernández-Manjón, B. (2020). Predicting students' knowledge after playing a serious game based on learning analytics data: A case study. *Journal of Computer Assisted Learning*, 36(3), 350–358. <https://doi.org/10.1111/jcal.12405>
- Anastasiadis, T., Lampropoulos, G., & Siakas, K. (2018). Digital Game-based Learning and Serious Games in Education. *International Journal of Advances in Scientific Research and Engineering*, 4(12), 139–144. <https://doi.org/10.31695/IJASRE.2018.33016>
- Aslan, S. (2016). *Digital Educational Games: Methodologies for Development and Software Quality*.
- Baker, R. S. J. D., & Yacef, K. (2009). The State of Educational Data Mining in 2009 : A Review and Future Visions. *Journal of Educational Data Mining*, 1(1), 3–16.
- Bakharia, A., Corrin, L., De Barba, P., Kennedy, G., Gašević, D., Mulder, R., ... Lockyer, L. (2016). A conceptual framework linking learning design with learning analytics. *ACM International Conference Proceeding Series*, 25-29-April, 329–338. <https://doi.org/10.1145/2883851.2883944>
- Bakhouyi, A., Dehbi, R., Talea, M., & Hajoui, O. (2017). Evolution of standardization and interoperability on E-learning systems: An overview. In *2017 16th International Conference on Information Technology Based Higher Education and Training (ITHET)* (pp. 1–8). IEEE. <https://doi.org/10.1109/ITHET.2017.8067789>
- Barko, T., & Sadler, T. D. (2013). Practicality in Virtuality: Finding Student Meaning in Video Game Education. *Journal of Science Education and Technology*, 22(2), 124–132. <https://doi.org/10.1007/s10956-012-9381-0>

- Barra, E., Gordillo, A., Gallego, D., & Quemada, J. (2013). Integration of SCORM packages into web games. In *Proceedings - Frontiers in Education Conference, FIE*. <https://doi.org/10.1109/FIE.2013.6684913>
- Bellotti, F., Berta, R., De Gloria, A., & Primavera, L. (2009). Adaptive Experience Engine for Serious Games. *IEEE Transactions on Computational Intelligence and AI in Games*, 1(4), 264–280. <https://doi.org/10.1109/TCIAIG.2009.2035923>
- Bellotti, F., Kapralos, B., Lee, K., Moreno-Ger, P., & Berta, R. (2013). Assessment in and of Serious Games: An Overview. *Advances in Human-Computer Interaction*, 2013(i), 1–11. <https://doi.org/10.1155/2013/136864>
- Blow, J. (2004). Game Development: Harder Than You Think. *Queue*, 1(10), 28–37. <https://doi.org/10.1145/971564.971590>
- Blumberg, F. C., Altschuler, E. A., Almonte, D. E., & Mileaf, M. I. (2013). The Impact of Recreational Video Game Play on Children's and Adolescents' Cognition. *New Directions for Child and Adolescent Development*, 2013(139), 41–50. <https://doi.org/10.1002/cad.20030>
- Calderón, A., & Ruiz, M. (2015). A systematic literature review on serious games evaluation: An application to software project management. *Computers & Education*, 87, 396–422. <https://doi.org/10.1016/j.compedu.2015.07.011>
- Callaghan, M., McShane, N., Eguíluz, A. G., & Savin-Baden, M. (2018). Extending the Activity Theory Based Model for Serious Games Design in Engineering to Integrate Analytics. *International Journal of Engineering Pedagogy (IJEP)*, 8(1), 109. <https://doi.org/10.3991/ijep.v8i1.8087>
- Calvo-Morata, A. (2020). USO DE TÉCNICAS DE LEARNING ANALYTICS PARA LA VALIDACIÓN, MEJORA Y APLICACIÓN DE JUEGOS SERIOS EN LA CLASE APLICADO AL CIBERBULLYING, (Anexo I), 0–7.
- Calvo-Morata, A., Alonso-Fernández, C., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2021). Creating awareness on bullying and cyberbullying among young people: Validating the effectiveness and design of the serious game Conectado. *Telematics and Informatics*, 60(3), 101568. <https://doi.org/10.1016/j.tele.2021.101568>
- Calvo-Morata, A., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2018). Conectado: A Serious Game to Raise Awareness of Bullying and Cyberbullying in High Schools. *Educational Communications and Technology*, (3), 59–64.
- Calvo-Morata, A., Rotaru, D. C., Alonso-Fernandez, C., Freire, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2018). Validation of a Cyberbullying Serious Game Using Game Analytics. *IEEE Transactions on Learning Technologies*. <https://doi.org/10.1109/TLT.2018.2879354>
- CAST. (2018). Universal Design for Learning Guidelines version 2.2. Retrieved from <http://udlguidelines.cast.org>
- Chatti, M. A., Muslim, A., & Schroeder, U. (2017). Toward an Open Learning Analytics Ecosystem. In B. Kei Daniel (Ed.), *Big Data and Learning Analytics in Higher Education* (Vol. 9, pp. 195–219). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-06520-5_12

- Chaudy, Y. (2016). An Assessment and Learning Analytics Engine for Games-based Learning An Assessment and Learning Analytics Engine for Games-based Learning Yaëlle Chaudy Borghini Thesis submitted in partial fulfilment of the requirements of the University of the West of Sc, (December 2015). <https://doi.org/10.13140/RG.2.1.4932.5040>
- Chittaro, L. (2006). Visualizing information on mobile devices. *Computer*, 39(3), 40–45. <https://doi.org/10.1109/MC.2006.109>
- Corti, K. (2006). Games-based Learning: a serious business application. *Informe de PixelLearning*, 34(6), 1–20.
- Creutzfeldt, J., Hedman, L., & Felländer-Tsai, L. (2012). Effects of pre-training using serious game technology on CPR performance – an exploratory quasi-experimental transfer study. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, 20(1), 79. <https://doi.org/10.1186/1757-7241-20-79>
- De Freitas, S. (2018). International Forum of Educational Technology & Society Are Games Effective Learning Tools? A Review of Educational Games. *Source: Journal of Educational Technology & Society*, 21(2), 74–84.
- de Souza e Silva, A. (2017). Pokémon Go as an HRG: Mobility, sociability, and surveillance in hybrid spaces. *Mobile Media & Communication*, 5(1), 20–23. <https://doi.org/10.1177/2050157916676232>
- DeKanter, N. (2005). Gaming redefines interactivity for learning. *TechTrends*, 49(3), 26–31. <https://doi.org/10.1007/BF02763644>
- del Blanco, Á., Serrano-Laguna, Á., Martínez-Ortiz, I., Fernández-Manjón, B., & Stanescu, I. A. (2013). Integrating Serious Games into e-Learning Platforms: Present and Future. In *9th International Scientific Conference eLearning and software for Education, Bucharest, April 25-26, 2013*. Bucharest.
- Di Mitri, D., Börner, D., Scheffel, M., Ternier, S., Drachslar, H., & Specht, M. (2017). Learning pulse: A machine Learning approach for predicting performance in self-regulated learning using multimodal data. *ACM International Conference Proceeding Series*, 188–197. <https://doi.org/10.1145/3027385.3027447>
- Dickey, M. D. (2006). Game Design Narrative for Learning: Appropriating Adventure Game Design Narrative Devices and Techniques for the Design of Interactive Learning Environments. *Educational Technology Research and Development*, 54(3), 245–263. <https://doi.org/10.1007/s11423-006-8806-y>
- Djaouti, D., Alvarez, J., Jessel, J., & Rampnoux, O. (2011). *Origins of Serious Games*. (M. Ma, A. Oikonomou, & L. C. Jain, Eds.), *Serious Games and Edutainment Applications*. London: Springer London. <https://doi.org/10.1007/978-1-4471-2161-9>
- Egenfeldt-Nielsen, S. (2006). Overview of research on the educational use of video games. *Nordic Journal of Digital Literacy*, 1(03), 184–214. <https://doi.org/10.18261/issn1891-943x-2006-03-03>
- Faber, T., Dankbaar, M., & van Merriënboer, J. (2018). Applying an Instructional Design Method to Serious Games - Experiences and Lessons Learned. In *2018 9th International Conference on Information, Intelligence, Systems and Applications*

- (IISA) (pp. 1–3). IEEE. <https://doi.org/10.1109/IISA.2018.8633666>
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. <https://doi.org/doi.org/10.1609/aimag.v17i3.1230>
- Ferguson, R., Hoel, T., Scheffel, M., & Drachsler, H. (2016). Guest Editorial: Ethics and Privacy in Learning Analytics. *Journal of Learning Analytics*, 3(1), 5–15. <https://doi.org/10.18608/jla.2016.31.2>
- Freire, M., Martinez-Ortiz, I., Alonso-Fernandez, C., Perez-Colado, I., & Fernandez-Manjon, B. (2015). D4 . 7 – Game analytics and adaptation components reference implementation Appendix with First Aid Game, (687676), 1–67.
- Freire, M., Serrano-Laguna, Á., Manero-Iglesias, B., & Martínez-Ortiz, I. (2016). *Game learning analytics: Learning analytics for serious games. Learning, design, and Technology*. <https://doi.org/10.1007/978-3-319-17727-4>
- Fundación Telefónica. (2020). *Sociedad Digital en España 2019*.
- Hauge, J. B., Berta, R., Fiucci, G., Manjon, B. F., Padron-Napoles, C., Westra, W., & Nadolski, R. (2014). Implications of Learning Analytics for Serious Game Design. In *2014 IEEE 14th International Conference on Advanced Learning Technologies* (pp. 230–232). IEEE. <https://doi.org/10.1109/ICALT.2014.73>
- Hitchcock, C., & Stahl, S. (2003). Assistive Technology, Universal Design, Universal Design for Learning: Improved Learning Opportunities. *Journal of Special Education Technology*, 18(4), 45–52. <https://doi.org/10.1177/016264340301800404>
- Hollins, P., Humphreys, S., Yuan, L., Sleightholme, G., & Kickmeier-Rust, M. (2017). The “Water Cooler” Game. In *Proceedings of the 11th European Conference on Games Based Learning, ECGBL 2017*.
- Horn, B., Hoover, A. K., Barnes, J., Folajimi, Y., Smith, G., & Hartevelde, C. (2016). Opening the Black Box of Play, 142–153. <https://doi.org/10.1145/2967934.2968109>
- Hsu, H. Y., & Wang, S. K. (2010). Using Gaming Literacies to Cultivate New Literacies. *Simulation and Gaming*, 41(3), 400–417. <https://doi.org/10.1177/1046878109355361>
- IMS Global Learning Consortium. (2007). 1EdTech Content Packaging v1.1.2 Information Model | IMS Global Learning Consortium. Retrieved November 3, 2022, from http://www.imsglobal.org/content/packaging/cpv1p1p2/imscp_infov1p1p2.html
- Jean Justice, L., & Ritzhaupt, A. D. (2015). Identifying the Barriers to Games and Simulations in Education. *Journal of Educational Technology Systems*, 44(1), 86–125. <https://doi.org/10.1177/0047239515588161>
- Johnson, L., Becker, S. A., Cummins, M., Estrada, V., & Meira, A. (2012). *Technology Outlook for Brazilian Primary and Secondary Education 2012-2017*. Austin, Texas.
- Kafai, Y. (2001). The educational potential of electronic games: From games-to-teach to games-to-learn. *Playing By The Rules, Cultural Policy Center, University of Chicago, Chicago, IL, 2001*.

- Kato, P. M. (2010). Video Games in Health Care: Closing the Gap. *Review of General Psychology, 14*(2), 113–121. <https://doi.org/10.1037/a0019441>
- Kato, P. M. (2012). Evaluating Efficacy and Validating Games for Health. *Games for Health Journal, 1*(1), 74–76. <https://doi.org/10.1089/g4h.2012.1017>
- Kazimoglu, C., Kiernan, M., Bacon, L., & Mackinnon, L. (2012). A Serious Game for Developing Computational Thinking and Learning Introductory Computer Programming. *Procedia - Social and Behavioral Sciences, 47*, 1991–1999. <https://doi.org/10.1016/j.sbspro.2012.06.938>
- Larson, K. (2020). Serious Games and Gamification in the Corporate Training Environment: a Literature Review. *TechTrends, 64*(2), 319–328. <https://doi.org/10.1007/s11528-019-00446-7>
- Learning Technology Standards Committee of the IEEE. (2002). *1484.12.1-2002 - IEEE Standard for Learning Object Metadata*. <https://doi.org/10.1109/IEEESTD.2002.94128>
- Liarakou, G., Sakka, E., Gavrilakis, C., & Tsolakidis, C. (2012). Evaluation of serious games, as a tool for education for sustainable development. *European Journal of Open, Distance and E-Learning, 15*(2).
- Lim, T., Carvalho, M. B., Bellotti, F., de Freitas, S., Louchart, S., Suttie, N., ... De Gloria, A. (2015). The LM-GM framework for Serious Games Analysis. *British Journal of Educational Technology, 46*(2), 391–411.
- Liñán, L. C., Alejandro, Á., & Pérez, J. (2015). Educational Data Mining and Learning Analytics: differences, similarities, and time evolution Learning Analytics: Intelligent Decision Support Systems for Learning Environments. *RUSC. Universities and Knowledge Society Journal, 12*(3), 98–112.
- Loh, C. S., Sheng, Y., & Ifenthaler, D. (2015). Serious games analytics: Methodologies for performance measurement, assessment, and improvement. *Serious Games Analytics: Methodologies for Performance Measurement, Assessment, and Improvement*, (November), 1–477. <https://doi.org/10.1007/978-3-319-05834-4>
- Long, P., & Siemens, G. (2011). Penetrating the Fog: Analytics in Learning and Education.
- Mangaroska, K., Vesin, B., & Giannakos, M. (2019). Cross-Platform Analytics. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge - LAK19*, 71–75. <https://doi.org/10.1145/3303772.3303825>
- Marchiori, E. J., Ferrer, G., Fernández-Manjón, B., Povar-Marco, J., Suberviola, J. F., & Giménez-Valverde, A. (2012). Instrucción en maniobras de soporte vital básico mediante videojuegos a escolares: Comparación de resultados frente a un grupo control. *Emergencias, 24*(6), 433–437.
- Marchiori, E. J., Torrente, J., del Blanco, Á., Moreno-Ger, P., Sancho, P., & Fernández-Manjón, B. (2012). A narrative metaphor to facilitate educational game authoring. *Computers & Education, 58*(1), 590–599. <https://doi.org/10.1016/j.compedu.2011.09.017>
- Matcha, W., Uzir, N. A., Gasevic, D., & Pardo, A. (2020). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning

- Perspective. *IEEE Transactions on Learning Technologies*, 13(2), 226–245. <https://doi.org/10.1109/TLT.2019.2916802>
- Michael, D. R., & Chen, S. L. (2005). Serious Games: Games That Educate, Train, and Inform. *Education*, October 31, 1–95.
- Mislevy, R. (2005). Evidence-Centered Assessment Design: Layers, Structures, and Terminology. *Principled Assessment Designs for Inquiry Technical Report 9*, 9(July), 46.
- Moreno-Ger, P., Burgos, D., Martínez-Ortiz, I., Sierra, J. L., & Fernández-Manjón, B. (2008). Educational game design for online education. *Computers in Human Behavior*, 24(6), 2530–2540. <https://doi.org/10.1016/j.chb.2008.03.012>
- Moreno-Ger, P., Sierra, J. L., & Fernández-Manjón, B. (2008). Game-Based Learning in e-Learning Environments. *UPGRADE-Novática*, IX(3), 15–20.
- Muñoz-Cristóbal, J. A., Rodríguez-Triana, M. J., Gallego-Lema, V., Arribas-Cubero, H. F., Asensio-Pérez, J. I., & Martínez-Monés, A. (2018). Monitoring for Awareness and Reflection in Ubiquitous Learning Environments. *International Journal of Human-Computer Interaction*, 34(2), 146–165. <https://doi.org/10.1080/10447318.2017.1331536>
- Muslim, A., Chatti, M. A., & Guesmi, M. (2020). Open Learning Analytics: A Systematic Literature Review and Future Perspectives (pp. 3–29). Springer, Cham. https://doi.org/10.1007/978-3-030-41099-5_1
- Nickols, F. W. (2005). Why a Stakeholder Approach to Evaluating Training. *Advances in Developing Human Resources*, 7(1), 121–134. <https://doi.org/10.1177/1523422304272175>
- Nishikawa, K. A., & Jaeger, J. (2011). A computer simulation comparing the incentive structures of dictatorships and democracies. *Journal of Political Science Education*, 7(2), 135–142. <https://doi.org/10.1080/15512169.2011.564915>
- Oblinger, D. G. (2004). The Next Generation of Educational Engagement. *Journal of Interactive Media in Education*, 2004(1), 10. <https://doi.org/10.5334/2004-8-oblinger>
- Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L., & Fuller, T. (2003). User-centered design in games. In *The human-computer interaction handbook* (pp. 883–906). L. Erlbaum Associates Inc. Hillsdale, NJ, USA.
- Palavitsinis, N., Manouselis, N., & Sanchez-Alonso, S. (2014). Metadata quality in learning object repositories: A case study. *Electronic Library*, 32(1), 62–82. <https://doi.org/10.1108/EL-12-2011-0175>
- Papazoglou Papazoglakis, P. (2013). The Past, Present and Future of SCORM. *Economy Informatics*, 13(1), 16–26.
- Perez-Colado, I. J., Perez-Colado, V. M., Freire-Moran, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2017). Integrating Learning Analytics into a Game Authoring Tool. In *International Conference on Web-Based Learning (ICWL)* (pp. 51–61). https://doi.org/10.1007/978-3-319-66733-1_6
- Perez-Colado, I. J., Perez-Colado, V. M., Martinez-Ortiz, I., Freire-Moran, M., &

- Fernandez-Manjon, B. (2017). uAdventure: The eAdventure reboot: Combining the experience of commercial gaming tools and tailored educational tools. *IEEE Global Engineering Education Conference, EDUCON*, (April), 1755–1762. <https://doi.org/10.1109/EDUCON.2017.7943087>
- Pérez-Colado, V. M., Rotaru, D. C., Freire-Morán, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2018). Learning analytics for location-based serious games. In *2018 IEEE Global Engineering Education Conference (EDUCON)* (Vol. 2018-April, pp. 1192–1200). Santa Cruz de Tenerife, Spain: IEEE. <https://doi.org/10.1109/EDUCON.2018.8363365>
- Pérez Colado, I. J. (2016). *uAdventure : desarrollo del intérprete y de un emulador de videojuegos de eAdventure sobre Unity3D*.
- Pérez Colado, V. M. (2022). *Contribuciones a la creación y aplicabilidad de los juegos serios (serious games) en el ámbito del geoposicionamiento*.
- Pishtari, G., Rodríguez-Triana, M. J., Sarmiento-Márquez, E. M., Pérez-Sanagustín, M., Ruiz-Calleja, A., Santos, P., ... Våljataga, T. (2020). Learning design and learning analytics in mobile and ubiquitous learning: A systematic review. *British Journal of Educational Technology*, *51*(4), 1078–1100. <https://doi.org/10.1111/bjet.12944>
- Polsani, P. R. (2005). Use and Abuse of Reusable Learning Objects 1 Movements in the Learning Object. *Journal of Digital Information*, *3*(4), 1–10.
- Pons Betrián, D., Hilera González, J. R., & Pagés Arévalo, C. (2011). ISO/IEC 19788 MLR: Un nuevo estándar de metadatos para recursos educativos. *Revista Iberoamericana de Tecnologías Del Aprendizaje*, *6*(3), 140–145.
- Prasetya, W., Leek, C., Melkonian, O., ten Tusscher, J., van Bergen, J., Everink, J., ... van Zon, W. (2019). Having Fun in Learning Formal Specifications. In *2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET)* (pp. 192–196). IEEE. <https://doi.org/10.1109/ICSE-SEET.2019.00028>
- Prensky, M. (2001). *Digital game-based learning*. New York: McGraw-Hill.
- Quinn, C. N. (1994). Designing educational computer games. In K. Beattie, C. McNaught, & S. Wills (Eds.), *Interactive multimedia in university education: Designing for change in teaching and learning* (pp. 45–57). Amsterdam: Elsevier.
- Randel, J. M., Morris, B. A., Wetzel, C. D., & Whitehill, B. V. (1992). The Effectiveness of Games for Educational Purposes: A Review of Recent Research. *Simulation & Gaming*, *23*(3), 261–276. <https://doi.org/10.1177/1046878192233001>
- Ravenscroft, A., Lindstaedt, S., Delgado Kloos, C., & Hernández-Leo, D. (2012). *21st Century Learning for 21st Century Skills* (Vol. 7563). <https://doi.org/10.1007/978-3-642-33263-0>
- Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, *40*(6), 601–618. <https://doi.org/10.1109/TSMCC.2010.2053532>
- Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(3), 1–21. <https://doi.org/10.1002/widm.1355>

- Rus Cano, A. (2019). *Aplicación de analíticas en la sistematización del diseño y validación de juegos serios para usuarios con discapacidad intelectual*. Universidad Complutense de Madrid.
- Samuelsen, J., Chen, W., & Wasson, B. (2019). Integrating multiple data sources for learning analytics—review of literature. *Research and Practice in Technology Enhanced Learning*, 14(1), 11. <https://doi.org/10.1186/s41039-019-0105-4>
- Sawyer, B. (2007). The “Serious Games” Landscape. Camden, USA.
- Serrano-Laguna, A., & Fernandez-Manjon, B. (2014). Applying learning analytics to simplify serious games deployment in the classroom. In *2014 IEEE Global Engineering Education Conference (EDUCON)* (pp. 872–877). IEEE. <https://doi.org/10.1109/EDUCON.2014.6826199>
- Serrano-Laguna, Á., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., & Fernández-Manjón, B. (2017). Applying standards to systematize learning analytics in serious games. *Computer Standards & Interfaces*, 50, 116–123. <https://doi.org/10.1016/j.csi.2016.09.014>
- Serrano Laguna, Á. (2012). GLAS: framework to improve assesment in educational videogames.
- Setya Murti, H. A., Dicky Hastjarjo, T., & Ferdiana, R. (2019). Platform and Genre Identification for Designing Serious Games. In *5th International Conference on Science and Technology (ICST)* (pp. 1–6). IEEE. <https://doi.org/10.1109/ICST47872.2019.9166177>
- Siemens, G., & Baker, R. S. J. D. (2012). Learning analytics and educational data mining: Towards communication and collaboration. *ACM International Conference Proceeding Series*, 252–254. <https://doi.org/10.1145/2330601.2330661>
- Sliney, A., & Murphy, D. (2008). JDoc: A Serious Game for Medical Learning. In *First International Conference on Advances in Computer-Human Interaction* (pp. 131–136). IEEE. <https://doi.org/10.1109/ACHI.2008.50>
- Squire, K., & Jenkins, H. (2003). *Harnessing the Power of Games in Education*. InSight (Vol. 3).
- Suh, S., Kim, S. W., & Kim, N. J. (2010). Effectiveness of MMORPG-based instruction in elementary English education in Korea. *Journal of Computer Assisted Learning*, 26(5), 370–378. <https://doi.org/10.1111/j.1365-2729.2010.00353.x>
- Szczurek, M. (1982). *Meta-analysis of simulation games effectiveness for cognitive learning*. Bloomington: Indiana University.
- Thomas, K. W., & Kilmann, H. (1976). Thomas-Kilmann Conflict Mode Instrument. *Group & Organization Studies*, 1(2), 249–251. <https://doi.org/10.1037/t02326-000>
- Torrente, J., Del Blanco, Á., Marchiori, E. J., Moreno-Ger, P., & Fernández-Manjón, B. (2010). e-Adventure: Introducing Educational Games in the Learning Process. In *IEEE Education Engineering (EDUCON) 2010 Conference* (pp. 1121–1126). <https://doi.org/10.1109/EDUCON.2010.5493056>
- Torrente, J., Del Blanco, Á., Moreno-Ger, P., & Fernández-Manjón, B. (2012). Designing serious games for adult students with cognitive disabilities. *Lecture Notes in*

Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 7666 LNCS(PART 4), 603–610. https://doi.org/10.1007/978-3-642-34478-7_73

- Torrente, J., Moreno-Ger, P., Martínez-Ortiz, I., & Fernandez-Manjon, B. (2009). Integration and deployment of educational games in e-learning environments: The learning object model meets educational gaming. *Educational Technology and Society*.
- Van der Kooij, K., Hoogendoorn, E., Spijkerman, R., & Visch, V. (2015). Validation of Serious Games. *International Journal of Serious Games*, 2(3). <https://doi.org/10.17083/ijsg.v2i3.75>
- van Eeden, E., & Chow, W. (2018). Perspectives from the Global Entertainment & Media Outlook 2018–2022: Trending now: Convergence, Connections and Trust, 34.
- van Merriënboer, J. J. G., & Kirschner, P. A. (2017). *Ten Steps to Complex Learning. Ten Steps to Complex Learning*. Third edition. | New York : Routledge, 2018. | “First edition published by Routledge 2007”—T.p. verso. | “Sixth edition published by Routledge 2013”—T.p. verso.: Routledge. <https://doi.org/10.4324/9781315113210>
- Verduin, M. L., LaRowe, S. D., Myrick, H., Cannon-Bowers, J., & Bowers, C. (2013). Computer simulation games as an adjunct for treatment in male veterans with alcohol use disorder. *Journal of Substance Abuse Treatment*, 44(3), 316–322. <https://doi.org/10.1016/j.jsat.2012.08.006>
- Westera, W., Nadolski, R. J., Hummel, H. G. K., & Wopereis, I. G. J. H. (2008). Serious games for higher education: A framework for reducing design complexity. *Journal of Computer Assisted Learning*, 24(5), 420–432. <https://doi.org/10.1111/j.1365-2729.2008.00279.x>
- Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2), 249–265. <https://doi.org/10.1037/a0031311>
- Xu, F., Buhalis, D., & Weber, J. (2017). Serious games and the gamification of tourism. *Tourism Management*, 60, 244–256. <https://doi.org/10.1016/j.tourman.2016.11.020>
- Yusoff, A. (2010). A Conceptual Framework for Serious Games and its Validation. *Science, PhD*(October).