UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

TESIS DOCTORAL

An Evidence-Based Framework for Optimizing Serious Game Design in Software Engineering Education

Un Marco Basado en Evidencia para Optimizar el Diseño de Juegos Serios en la Educación de Ingeniería de Software

D.ª Manal Kharbouch

2024

UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

TESIS DOCTORAL

An Evidence-Based Framework for Optimizing Serious Game Design in Software Engineering Education

Un Marco Basado en Evidencia para Optimizar el Diseño de Juegos Serios en la Educación de Ingeniería de Software

Autor: D.ª Manal Kharbouch

Director/es: D. José Luis Fernández Alemán

UNIVERSIDAD DE
MURCIA

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD
DE LA TESIS PRESENTADA EN MODALIDAD DE COMPENDIO O ARTÍCULOS PARA OBTENER EL TÍTULO DE DOCTOR
Aprobado por la Comisión General de Doctorado el 19-10-2022

D./Dña. Manal Kharbouch

doctorando del Programa de Doctorado en

Informática

de la Escuela Internacional de Doctorado de la Universidad Murcia, como autor/a de la tesis presentada para la obtención del título de Doctor y titulada:

An Evidence-Based Framework for Optimizing Serious Game Design in Software Engineering Education / Un Marco Basado en Evidencia para Optimizar el Diseño de Juegos Serios en la Educación de Ingeniería de Software

y dirigida por,

D./Dña. José Luis Fernández Alemán
D./Dña.
D./Dña.

DECLARO QUE:

La tesis es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, de acuerdo con el ordenamiento jurídico vigente, en particular, la Ley de Propiedad Intelectual (R.D. legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, modificado por la Ley 2/2019, de 1 de marzo, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), en particular, las disposiciones referidas al derecho de cita, cuando se han utilizado sus resultados o publicaciones.

Además, al haber sido autorizada como compendio de publicaciones o, tal y como prevé el artículo 29.8 del reglamento, cuenta con:

- *La aceptación por escrito de los coautores de las publicaciones de que el doctorando las presente como parte de la tesis.*
- *En su caso, la renuncia por escrito de los coautores no doctores de dichos trabajos a presentarlos como parte de otras tesis doctorales en la Universidad de Murcia o en cualquier otra universidad.*

Del mismo modo, asumo ante la Universidad cualquier responsabilidad que pudiera derivarse de la autoría o falta de originalidad del contenido de la tesis presentada, en caso de plagio, de conformidad con el ordenamiento jurídico vigente.

En Murcia, a 3 de octubre de 2024

Fdo.: Manal Kharbouch

Información básica sobre protección de sus datos personales aportados	
Responsable:	Universidad de Murcia. Avenida teniente Flomesta, 5. Edificio de la Convalecencia. 30003; Murcia. Delegado de Protección de Datos: dpd@um.es
Legitimación:	La Universidad de Murcia se encuentra legitimada para el tratamiento de sus datos por ser necesario para el cumplimiento de una obligación legal aplicable al responsable del tratamiento. art. 6.1.c) del Reglamento General de Protección de Datos
Finalidad:	Gestionar su declaración de autoría y originalidad
Destinatarios:	No se prevén comunicaciones de datos
Derechos:	Los interesados pueden ejercer sus derechos de acceso, rectificación, cancelación, oposición, limitación del tratamiento, olvido y portabilidad a través del procedimiento establecido a tal efecto en el Registro Electrónico o mediante la presentación de la correspondiente solicitud en las Oficinas de Asistencia en Materia de Registro de la Universidad de Murcia

For my unwavering support system—forever grateful!

And to the naysayers—well, here we are!

Acknowledgment

First and foremost, I would like to express my deepest gratitude to my PhD supervisor, Pr. Jose Luis Fernandez-Aleman. His distinctive approach to supervision pushed me to my limits, challenging me in ways that led to a profound understanding of both my research and academia. Your expertise and mentorship have been pivotal throughout this journey, and I am sincerely thankful for the opportunity to work under your guidance.

I would also like to extend my heartfelt thanks to Pr. Aurora Vizcaino, who hosted me during my research stay at University Castilla La Mancha (Spain) and continued to collaborate with me, providing invaluable insights and guidance alongside my supervisor. Your generosity, patience, and support have had a profound impact on my work, and I am deeply appreciative of the time and effort you have dedicated to helping me succeed.

To my family—my parents and my brother—your unwavering support has been my anchor throughout the highs and lows of this journey. Your steadfast belief in my abilities, combined with your boundless love and encouragement, has provided me with the strength and resilience needed to navigate the most challenging moments. This achievement is as much yours as it is mine, and I truly could not have reached this milestone without your enduring support and encouragement.

I would also like to extend my heartfelt thanks to my friends. Your unwavering support and constant check-ins on both my academic progress and mental well-being have been a pillar of strength for me. Your patience in listening to my complaints and your companionship, even in moments of complete disinterest, have been a true source of comfort and joy throughout this process.

I would also like to take a moment to acknowledge myself—for enduring this journey with an often questionable level of sanity. To my own mental endurance, which faced an ongoing battle as sleepless nights, fleeting moments of peace, and the occasional panic attack prompted a re-evaluation of every life choice that led me here. To the ever-elusive sleep, my long-time adversary, and to the health complications that accompanied this demanding journey—your challenges were a constant reminder of the physical and emotional toll of this PhD. And to my once-thriving social life, which slowly faded into a distant memory as I navigated the trials of anxiety and self-doubt—This thesis stands more as a testament to survival through the challenges than as a mark of academic achievement.

I would also like to acknowledge my ever-present imposter syndrome, which, in conjunction with those who doubted my ability to succeed, conspired to keep me on edge. Paradoxically, I am grateful for the motivation this created, as it drove me to persevere and prove myself.

Additionally, coffee deserves a mention for aiding me through the busiest days of this journey. Despite caffeine eventually losing its edge, it remained a steadfast companion during crucial

networking moments and intense periods of productivity, adding a bittersweet flavor to this challenging experience.

Finally, to everyone and everything that played a part in this beautifully chaotic journey: thank you. This thesis reflects not just perseverance and hard work but also the value of embracing chaos and surviving it with humor intact. Without your presence, this process might have been more straightforward—but far less memorable. Here’s to the dark humor, sarcasm, and resilience that carried me through!

Glossary of terms and abbreviations

SG	Serious Game
GBL	Game-Based Learning
SE	Software Engineering
DSR	Design Science Research
TAM	Technology Acceptance Model
SWEBOK	Software Engineering Body of Knowledge
GDBL	Game Development-Based Learning
SLR	Systematic Literature Reviews
SMS	Systematic mapping study
AI	artificial intelligence
LM-GM	Learning Mechanics - Game Mechanics
PICO	Population, Intervention, Comparison, and Outcomes
RQ	Research Questions
NPC	Non Player Character
PEM	Player Experience Modelling
NLP	Natural language processing
PBL	problem-based learning
RE	Requirements Engineering
MEEGA	Model for Evaluation of Educational Games
MOOC	Massive Online Open Course
BPMN	Business Process Modeling Notation
MDA	Mechanics-Dynamics-Aesthetics
DPE	Design, Play, and Experience
EE	Effectiveness and Efficiency
QR	Quality and Resources
IO	Interoperability
PU	Perceived Usefulness
PEU	Perceived Ease Of Use
ATU	Attitude Towards Using
BIU	Behavioral Intention To Use

Content

List of Figures

List of Tables

ABSTRACT

This thesis explores the design and impact of serious games (SGs) in software engineering (SE) education using a Design Science Research methodology. Chapter 1 provides a comprehensive mapping study of SGs in SE, identifying key characteristics, challenges, and success factors, while offering practical recommendations for educators and developers. Chapter 2 details a pedagogical experiment that demonstrates the positive impact of SGs on students' academic performance, highlighting the importance of instructor experience in successful SG integration. Chapter 3 introduces the RAF-SGD, a novel framework aimed at improving SG design through collaboration, reuse, and automation. This framework, validated through expert review and a comparative analysis with other frameworks, highlights its comprehensiveness and potential. While the framework's practical application is still being validated through the design of SGs.

Keywords: Serious Games, Software Engineering, Education, RAF-SGD, Design Science Research, Pedagogical Experiment, Game Design Framework.

RESUMEN

Introducción

Los Juegos Serios (JS) se diseñan con un propósito más allá del entretenimiento, a menudo con fines educativos o de formación profesional. A través de la integración de elementos lúdicos, los JS pueden hacer que el proceso de aprendizaje sea más interactivo y motivador. Se diferencian de la gamificación, que aplica elementos de juego en entornos no lúdicos con el objetivo de aumentar la participación y el compromiso. En la educación, especialmente en Ingeniería de Software (IS), los JS tienen un gran potencial para enfrentar algunos de los problemas más difíciles en la enseñanza de temas complejos, como la abstracción de conceptos y la aplicación práctica de teorías en entornos del mundo real. El uso de juegos serios puede proporcionar entornos ricos en retroalimentación, en los que los estudiantes pueden practicar habilidades técnicas y de toma de decisiones en contextos controlados pero inmersivos.

Esta tesis investiga el diseño de JS en la educación de IS, su efectividad como herramientas pedagógicas y cómo se pueden mejorar los enfoques actuales de diseño de JS para este campo. Para ello, se desarrolló un nuevo marco para el diseño de JS, que fue posteriormente validado a través de retroalimentación de expertos en IS y diseño de juegos. El objetivo es cubrir un vacío importante en la literatura: la falta de un marco estandarizado que guíe el diseño de JS en la educación de IS. Esta investigación, por tanto, no solo busca ampliar la comprensión del papel de los JS en la enseñanza de IS, sino también proporcionar un marco práctico que pueda ser utilizado por educadores y desarrolladores de juegos para crear herramientas educativas más efectivas.

Hipótesis

La hipótesis principal de la tesis es que un marco de diseño basado en evidencia, y respaldado por la retroalimentación de expertos, puede mejorar significativamente la calidad y la efectividad de los JS en la educación de IS. Específicamente, el marco puede ayudar a crear JS que se alineen mejor con los objetivos educativos, optimicen los recursos disponibles y aumenten los resultados de aprendizaje.

Objetivos

Los principales objetivos de esta investigación son los siguientes:

1. Comprender los juegos serios en la educación de IS: Desarrollar una comprensión integral del estado actual de los JS en la educación de IS, analizando los ejemplos existentes y categorizándolos según los dominios de IS y la Taxonomía de Bloom para evaluar su alineación con los objetivos educativos.

2. Evaluar el impacto de los JS en los resultados de los estudiantes, considerando la experiencia de los profesores: Realizar estudios empíricos que no solo midan la efectividad de los JS en términos

de rendimiento académico, satisfacción y compromiso de los estudiantes en la educación de IS, sino que también evalúen cómo la experiencia previa de los profesores en la implementación de estos juegos influye en dichos resultados. Es importante analizar si los profesores con mayor familiaridad y práctica con los JS logran un mayor impacto positivo en los estudiantes, comparado con aquellos que tienen menos experiencia o formación en el uso de estas herramientas.

3. Aportar nuevo conocimiento a la educación en IS: Identificar las mejores prácticas para el diseño de JS y desarrollar pautas que puedan ayudar a los educadores a integrar estos juegos en los planes de estudio de IS.

4. Desarrollar un nuevo marco de diseño de juegos serios: Proponer un marco que no solo estandarice el proceso de diseño de juegos serios, sino que también promueva la reutilización de ciertos componentes de los juegos de manera fluida. Además, este marco debe automatizar ciertas tareas dentro del proceso de diseño, lo que lo hará más sostenible y ahorrará tiempo y esfuerzo a los interesados, facilitando así la colaboración entre educadores de IS y diseñadores de juegos.

5. Validar el marco a través de retroalimentación de expertos y análisis comparativo: Asegurar que el marco propuesto sea práctico y efectivo mediante su validación a través de revisiones y retroalimentación de expertos en diseño de JS y en educación de IS. Además, se llevó a cabo un análisis comparativo con marcos novedosos existentes para evaluar su efectividad. Actualmente, el marco está siendo validado en la práctica real mediante su aplicación en el diseño de un juego serio, lo que permitirá obtener resultados concretos sobre su funcionalidad y adaptabilidad en escenarios de uso real.

Metodología

La tesis siguió la metodología de*Investigación en Ciencias del Diseño (Design Science Research - DSR), que se organiza en cuatro fases fundamentales:

1. Identificación del Problema: Una revisión exhaustiva de la literatura reveló varias deficiencias en el diseño e implementación de JS en la educación de IS. Se realizó un estudio de mapeo sistemático de los JS existentes, que clasificó estos juegos según su alineación con la Taxonomía de Bloom y los dominios de IS definidos en el Cuerpo de Conocimientos de Ingeniería de Software (SWEBOK). Además, se llevaron a cabo experimentos controlados para evaluar el impacto de los JS en los resultados de los estudiantes.

2. Desarrollo del Artefacto: Basado en los hallazgos de la fase de identificación del problema, se desarrolló un nuevo marco de diseño de JS que busca estandarizar y mejorar el proceso de creación de estos juegos en la educación de IS. Este marco proporciona una guía práctica tanto para educadores como para desarrolladores de juegos, con un enfoque en la alineación con los objetivos educativos y la optimización de recursos.

3. Validación del Artefacto: El marco fue validado mediante un proceso de revisión por expertos en el campo de la IS y del diseño de juegos. También se realizó un análisis comparativo entre el marco propuesto y otros marcos de diseño de JS existentes.

4. Contribución al Conocimiento: Los resultados de la investigación se consolidaron en un conjunto de mejores prácticas para el diseño de JS, y el marco propuesto fue refinado en función de la retroalimentación de los expertos y los estudios comparativos.

Conclusiones

En capítulo 1, se realizó un estudio de mapeo sistemático de los JS existentes en la educación de IS, con el fin de comprender mejor el estado del arte en este campo. Los juegos serios identificados fueron categorizados según los dominios de IS y los niveles de la Taxonomía de Bloom, lo que permitió evaluar qué tan bien estos juegos apoyaban los objetivos educativos de diferentes áreas de IS, como la gestión de proyectos, el diseño de software y las pruebas. El análisis reveló que muchos JS se concentraban en las etapas inferiores de la Taxonomía de Bloom, como la comprensión y la aplicación, mientras que había pocos juegos diseñados para las etapas superiores, como el análisis y la evaluación. Esto señala una brecha importante en el diseño de JS en IS, ya que las habilidades cognitivas de nivel superior son cruciales para los profesionales de IS. El estudio también destacó la subrepresentación de ciertos dominios de IS en los JS existentes, en particular las áreas de seguridad de software y garantía de calidad, lo que indica la necesidad de desarrollar más JS que cubran estas áreas. Además, se identificó que muchos juegos fallaban en tener en cuenta los diferentes perfiles de los jugadores, lo que podría afectar negativamente la experiencia de aprendizaje. Los hallazgos de este estudio sirvieron de base para el desarrollo de un marco de diseño de JS que abordara estas deficiencias, proponiendo un enfoque más equilibrado y estructurado para el diseño de juegos educativos en IS.

El segundo capítulo describió un experimento controlado que se llevó a cabo en el curso de gestión de proyectos de desarollo software en varios años académicos. El propósito del experimento fue evaluar el impacto de los JS en el rendimiento de los estudiantes en comparación con los métodos tradicionales de enseñanza. Se midieron tres variables clave: calificaciones finales, asistencia y rendimiento en el examen final. Además, el estudio consideró el papel de la experiencia del instructor en la implementación de los JS. Para ello, se compararon cursos impartidos por instructores con diferentes niveles de experiencia en el uso de JS como herramienta pedagógica. Los resultados del experimento mostraron que los estudiantes que participaron en cursos donde se utilizaron JS tuvieron un mejor rendimiento que aquellos en los cursos tradicionales. Esto fue particularmente cierto para los cursos impartidos por instructores con una experiencia considerable en el uso de JS. Se observó que estos instructores no solo eran más eficaces en la implementación de los juegos en sus cursos, sino que también pudieron integrar mejor los elementos educativos clave dentro del juego. Esto sugiere que la formación de los instructores es un factor crítico para maximizar los beneficios de los JS en la educación de IS.

El capítulo 3 es el corazón de la tesis y se centra en el desarrollo del nuevo marco de diseño de JS, así como en su validación. El marco propuesto se basa en las deficiencias identificadas en el estudio de mapeo y los hallazgos del experimento controlado, ofreciendo un enfoque más estructurado y accesible para el diseño de JS en la educación de IS. El marco tiene tres componentes principales:

1. Alineación con los objetivos educativos: El marco asegura que los JS se diseñen para alinearse con los objetivos educativos específicos de IS, basándose en la Taxonomía de Bloom y en los dominios definidos por SWEBOK.

2. Mecánicas y dinámicas de juego adaptadas: El marco proporciona pautas claras sobre cómo seleccionar e implementar las mecánicas de juego que mejor se adapten a los objetivos educativos y a las características de los estudiantes. Esto incluye recomendaciones sobre la personalización de la experiencia de juego para diferentes perfiles de jugadores.

3. Colaboración entre educadores y desarrolladores: Dado que el diseño de JS puede ser un proceso complejo, el marco facilita la colaboración entre educadores y diseñadores de juegos, proporcionando un lenguaje común y herramientas que simplifiquen la comunicación entre ambas partes.

Además, el capítulo incluye una descripción detallada del proceso de validación del marco. Se realizaron entrevistas con expertos en el campo de la Ingeniería de Software y el diseño de juegos serios. Estos expertos evaluaron el marco propuesto, proporcionando retroalimentación crítica para su refinamiento. Los resultados de la validación demostraron que el marco era considerado efectivo y flexible por los expertos, destacando especialmente su capacidad para guiar el diseño de juegos serios que promuevan habilidades cognitivas de nivel superior. Los expertos también señalaron que el marco podía ayudar a reducir la brecha de comunicación entre educadores y desarrolladores de juegos, lo que facilitaba un diseño más coherente y alineado con los objetivos educativos.

También se llevó a cabo una comparación entre el marco propuesto y otros marcos de diseño de JS, utilizando criterios como la facilidad de uso, la alineación con los objetivos pedagógicos y la capacidad de adaptación a diferentes contextos educativos. Se examinaron marcos utilizados en diversas disciplinas, incluyendo la educación general y la capacitación profesional, para identificar sus fortalezas y debilidades en comparación con el nuevo marco desarrollado para IS. El análisis mostró que, aunque varios marcos existentes son eficaces en ciertos contextos, pocos proporcionan una alineación clara con los objetivos educativos específicos de la Ingeniería de Software. Muchos marcos tienden a centrarse en áreas como la motivación y el compromiso, lo cual es importante, pero no abordan de manera adecuada los desafíos específicos de enseñar habilidades técnicas y analíticas complejas como las que se requieren en IS. Además, el marco propuesto en esta tesis sobresale en términos de flexibilidad y adaptabilidad a diferentes entornos de aprendizaje. Mientras que algunos marcos anteriores se enfocaban en un solo tipo de mecánica de juego o estilo de enseñanza, el nuevo marco permite una personalización más profunda, ajustándose tanto a las necesidades del estudiante como a las del instructor.

Trabajo Futuro

La investigación realizada en esta tesis ha abordado varias áreas importantes en la intersección de los JS y la educación en IS. A través de un análisis riguroso del estado actual de los JS en IS, la tesis identifica áreas de mejora y propone un nuevo marco de diseño que busca superar las limitaciones de los enfoques anteriores. El nuevo marco no solo alinea los juegos serios con los objetivos educativos, sino que también proporciona una estructura flexible y adaptativa que puede ser utilizada por educadores y desarrolladores de juegos. La retroalimentación de los expertos ha confirmado su validez, y los estudios comparativos han mostrado que tiene ventajas claras sobre otros marcos existentes. A pesar de los avances logrados con esta investigación, hay varias áreas que podrían beneficiarse de estudios adicionales en el futuro:

1. Evaluación longitudinal: Aunque los estudios controlados mostraron que los juegos serios pueden tener un impacto positivo en el corto plazo, sería valioso realizar estudios longitudinales para evaluar el impacto de estos juegos en el aprendizaje a largo plazo y en la retención de habilidades técnicas.

2. Expansión a otros dominios de IS: Aunque esta tesis se centró en ciertos dominios clave de IS, como la gestión de proyectos y el diseño de software, el marco podría expandirse y adaptarse a otras áreas críticas como la seguridad del software y la garantía de calidad. El desarrollo de JS específicos para estos dominios podría tener un impacto significativo en la educación y formación en IS.

3. Aplicaciones en la industria: Si bien esta tesis se centró en el uso de juegos serios en un entorno educativo formal, existe un gran potencial para aplicar el marco en entornos de capacitación profesional en la industria de software. Las empresas podrían beneficiarse del uso de juegos serios para entrenar a sus empleados en nuevas tecnologías, metodologías y mejores prácticas.

4. Integración con tecnologías emergentes: A medida que las tecnologías emergentes como la realidad virtual (VR) y la inteligencia artificial (IA) continúan avanzando, existe una oportunidad para integrar estos avances en el diseño de JS. El marco podría ser adaptado para aprovechar estas tecnologías, creando juegos más inmersivos y personalizados que respondan a las necesidades individuales de los estudiantes.

5. Diversificación de perfiles de estudiantes: Los juegos serios diseñados para educación en IS deberían tener en cuenta la diversidad de los estudiantes, tanto en términos de habilidades previas como de estilos de aprendizaje. En el futuro, sería interesante explorar cómo adaptar los juegos serios para grupos más amplios y diversos de estudiantes, incluidos aquellos con necesidades educativas especiales o estudiantes de diferentes niveles de experiencia en IS.

1

INTRODUCTION

1.1 Introduction

Serious Games (SG) are games used for purposes other than mere entertainment in accordance with specific rules, which use leisure as a tool to promote educational settings of various kinds [1]. In several studies, the terms SGs and gamification are used interchangeably to refer to Game-Based Learning (GBL) approaches for adults [2]. which is an umbrella term encompassing any approach that uses game components as a medium for learning [3]. Although the use of game components in both SGs and gamification is what leads to this confusion, the difference lies in the way in which they were used. While gamification is a task to which game components were added in non-game contexts [4], a SG is a purpose-built game that uses these components like any other game [5].

According to a report on the game-based learning market revenue worldwide in 2018 and 2024 [6], the SGs market is expected to grow from 3.5 billion U.S. dollars in 2018 to 24 billion in 2024. This does not only comply with the global interest that SGs gained during the last few years, but also suggests that SGs are a fast-growing trend that is worth exploring. According to Laamarti et al. [7], SGs are becoming ever more important in the domain of education. In the same line, statistics confirm that in the year 2020 the main category of SGs developed by the Spanish videogame industry was 'education' with a percentage of 79%, significantly higher than any other category of SGs [8]. Facing the very rich taxonomy of SGs identified in the literature [9,10], serious gaming appears to be a growing market worldwide, and an interesting area for interdisciplinary academic research [11]. This market keeps on expanding to other sectors such as education. This surge has captured the attention of investors, researchers, and developers, emphasizing a collective commitment to improving the design quality of these educational tools [12]. Notably, serious games are on the rise in higher education, reflecting a broader acknowledgment of their potential to enhance traditional learning methods [13] as an adjunct to teaching and learning materials [14], as the latter produce more tangible educational results and are more inspiring than other learning approaches such as classroom learning, eLearning, and hands-on-learning [15]. A comparative analysis led by Tahir and Wang [16] highlighted that learning/pedagogy and game factors are the most essential attributes for the design of educational games, linking these attributes was considered an efficient way to facilitate affective reactions such as flow, enjoyment, and immersion towards educational games.

In the context of Software Engineering (SE) education, SGs offer a promising avenue to address the challenges associated with traditional teaching methods. However, despite their growing popularity, there remains a need for comprehensive understanding and systematic evaluation of SGs' effectiveness and design principles. This thesis addresses these needs by exploring the impact of SGs on SE education and developing practical guidelines for their implementation. Existing literature underscores the benefits of SGs in in education and training, , as they align with modern theories of effective learning [17]. These theories emphasize that learning is most effective when it is active, experimental, situated, problem-based, and provides immediate feedback—all features commonly found in SGs [18]. However, there is a noticeable gap in empirical studies specifically addressing the use of SGs in Software Engineering (SE) education. Current research has demonstrated that SGs can improve student motivation and engagement [19], but there is limited evidence on how these benefits translate into tangible improvements in academic performance and learning outcomes within SE courses. Moreover, the design and implementation of SGs in SE contexts often lack a standardized approach, leading to varied effectiveness and integration challenges [20].

The primary objective of this thesis is to bridge the gap between theoretical benefits and practical applications of SGs in SE education. This involves evaluating the impact of SGs on student performance, engagement, and satisfaction, as well as developing actionable guidelines for SE educators and SG developers. This study is significant for several reasons. First, it contributes to the academic literature by providing empirical evidence on the impact of SGs in SE education, filling a critical gap in current research. The findings will offer valuable insights for educators seeking to enhance their teaching methods with game-based learning strategies. Second, the practical guidelines developed through this research will support SG developers in creating more effective and contextually relevant educational games. By addressing both theoretical and practical aspects, this thesis aims to advance the field of educational technology and improve learning outcomes in SE education. This research employs a Design Science Research (DSR) methodology to address the identified problems and objectives. The DSR approach will guide the development and evaluation of artifacts designed to improve SG integration in SE education. The methodology includes conducting controlled experiments to assess the effectiveness of SGs, performing a systematic literature review to identify design challenges and opportunities, and developing practical guidelines based on empirical findings and theoretical insights.

The thesis is structured as follows: Section 2 presents a comprehensive review of the state of the art, focusing on the use of SGs in educational contexts, particularly within SE education. Section 3 outlines the research methodology employed in this study, detailing the DSR approach, including the design, development, and evaluation of the study's artifacts. Section 4 provides an overview of the systematic mapping study conducted to explore the current landscape of SGs in SE. Section 5 focuses on the influence of instructor expertise with SGs on student performance in SE courses. Section 6 introduces the proposed solution, describing the development of a novel SG design framework, followed by its validation through expert consultations and ongoing practical application in SG design. Section 7 concludes the thesis, discussing the broader implications of the research, identifying limitations, and offering suggestions for future research directions.

1.2 Hypothesis:

Guided by an evidence-based framework and validated through expert feedback, the design of SGs in SE education can lead to more efficient and effective outcomes, optimizing both resource use and educational impact.

1.3 Goals and Tasks

Based on the hypothesis, we can outline the following goals for the thesis and break them down into specific tasks:

1.3.1 Goal 1: To Develop a Comprehensive Understanding of SGs in SE Education

- Task 1.1: Perform a systematic mapping study to identify existing SGs in SE education.

- Task 1.2: Categorize SGs used in SE education according to their focus on specific SE domains (e.g., Software Engineering Body of Knowledge (SWEBOK) areas) and Bloom's Taxonomy levels.

- Task 1.3: Identify key game elements, mechanics, and dynamics that contribute to the effectiveness of SGs in SE education.

- Task 1.4: Investigate the diversity of player profiles targeted by SGs in SE education and identify potential gaps.

1.3.2 Goal 2: To Evaluate the Impact of Serious Games on Student Outcomes in SE Education

- Task 2.1: Conduct a controlled experiment to assess the impact of SGs on student performance, exam attendance, and passing rates in SE courses.

- Task 2.2: Analyze the role of instructor expertise in the effectiveness of SG implementation in SE education.

- Task 2.3: Compare student satisfaction and engagement levels between traditional teaching methods and SG-enhanced courses.

1.3.3 Goal 3: To Synthesize and Contribute New Knowledge to the Field of SE Education

- Task 3.1: Synthesize findings from the controlled experiments, literature reviews, and theoretical proposals to formulate best practices for SG design and implementation in SE education.

- Task 3.2: Develop guidelines for SE educators on effectively integrating SGs into their curricula, considering factors such as instructor expertise and student diversity.

- Task 3.3: Identify and discuss future research directions, focusing on addressing identified gaps and exploring emerging trends such as the use of AI in SGs for SE education.

1.3.4 Goal 4: To Develop a Novel Framework for Serious Games Design

- Task 4.1: Curate and integrate the expertise gained from controlled experiments, literature reviews, and practical experiences into a comprehensive framework for SG design in SE education.

- Task 4.2: Aligning the framework with established frameworks and incorporating validated proposals to ensure it is evidence-based, making it adaptable to various educational contexts.

- Task 4.3: Create a structured process that simplifies the development of SGs to address the current lack of standardization in SG design, ensuring that the design process is more straightforward and accessible for educators.

- Task 4.4: Establish clear guidelines and protocols within the framework to facilitate more efficient communication between educators and game design experts, enhancing collaboration and ensuring that educational objectives are effectively translated into game mechanics.

1.3.5 Goal 5: To Validate the Framework Through Expert Review

- Task 5.1: Validate the framework's efficiency and effectiveness by gathering and analyzing expert feedback on its potential to optimize resource utilization and enhance educational quality in SG design.

- Task 5.2: Conduct a Technology Acceptance Model (TAM) assessment to evaluate expert perceptions of the framework's usability, practicality, and acceptance in educational contexts.

- Task 5.3: Assess the framework's interoperability by obtaining expert insights on its applicability across diverse educational and development settings, ensuring it facilitates effective collaboration between educators and game designers.

- Task 5.4: Refine the framework based on the validation feedback, ensuring it meets high standards for quality, resource efficiency, and applicability in the design and implementation of serious games for SE education.

2

STATE OF THE ART

2.1 Uncharted Dimensions in Serious Games for Software Engineering Education

2.1.1 Game-Based Approaches in Software Engineering Education

Through investigating the aspect of how gaming can affect the player's real world, Frank et al. [21] have found that gaming has a positive and measurable effect not only on players' skills but also on their knowledge, while players' real-life state of mind is not necessarily affected. In this respect, and since teaching SE is a difficult task due to its excessive reliance on theoretical courses that may lead to unenthusiasm and a lack of interest among learners [22], SE educators are particularly interested in supporting and motivating SE students [23] and thus, enhancing learning processes and outcomes through gaming [24]. Several studies reported a positive tendency toward the adoption of game-based approaches in SE education [2,25,26]. While SGs are receiving increasing interest in the field of SE [27], Barreto and França [28] conducted a mixed-method literature review, covering 130 studies, re-questioning gamification's definition in SE studies and coming to the conclusion that researchers in this field tend towards a strict view of gamification, , closely aligning it with Deterding et al. [4] definition of SGs. In another systematic survey, Caulfield et al. [29], examined games or simulations used in the SWEBOK areas of SE education and/or training has shown that, as pedagogical tools, these games have become more common and that students enjoy playing them and feel that this experience has an added value. The author also emphasized that although most studies lacked experimental design, there is sufficient evidence to say that SGs are beneficial and interesting complements to other teaching methods that educators can enrich their courses with. Moreover, Souza et al. [30], performed a systematic mapping of game-related approaches in SE education and their support of the Software Engineering Education Knowledge (SEEK) areas., analyzing 106 primary studies and identifying GBL and GDBL (Game Development-Based Learning) as the most prevalent approaches.

2.1.2 *Serious Games in Software Engineering Education*

Given that empirical studies in SE education with students as subjects have been beneficial for researchers for a long time [31], many studies in SE education validate SGs solutions or put a combination of SGs into test using students in their SE courses as subjects of experiments. This subsection examines various studies that investigate the use of SGs in software engineering courses, shedding light on their potential to enhance learning outcomes and student satisfaction.

In an empirical study conducted by Flores et al. [32], an experiment regarding teaching SE topics through pedagogical game design patterns highlights that students who played SGs not only achieved most of the expected learning objectives but did so in a fun and enjoyable way. Similarly, in a quasi-experimental study by Vizcaíno et al. [33], students showed significant knowledge improvement when exposed to a Global Software Development (GSD) SG in a SE course. Satisfaction in terms of game quality, especially in terms of usability, was high. Furthermore, Students found the game fun and enjoyable while some even reported being fully immersed in the game. In a similar vein, an empirical study conducted by Ghanbari et al. [34] demonstrated the effectiveness of utilizing SGs for requirement gathering in a SE course, incorporating key contextual challenges of GSD. This approach proved successful in enhancing the ability of less-experienced individuals to identify a greater number of requirements. Furthermore, it streamlined the process of requirements elicitation, making it more accessible and enjoyable for most participants. Additionally, it fostered improved collaboration and communication among students who utilized the SGs in contrast to those adhering to the traditional approach. In contrast, in the experiment by Kemell et al. [35], although participants enjoyed the learning experience, the SG in question was perceived as predictable and less useful when played for a long time. Just like the latter, no significant learning outcome linked to the SG was reported in their experiment results.

An experiment conducted by García et al [22] assessing a SG for teaching the fundamentals of the ISO/IEC/IEEE 29148 and SE at the undergraduate level showed that the SG did not only help to strengthen students' understanding of the theoretical concepts learned in the classroom but was also perceived by students to be an efficient and motivating teaching/learning tool. Along the same line, the study by Gulec et al. [36] investigating programming knowledge levels of computer engineering students using a web-based SG revealed that the students who often used the SG increased significantly their knowledge levels, unlike students who were in the control group for this experiment. In another context, in an experiment by Von Wangenheim et al. [37] regarding a SG they developed for teaching SCRUM in computer courses, results indicate that the SG has real potential to contribute to SCRUM learning engagingly, suggesting that the SG offers a low-budget alternative to complement traditional instructional strategies for teaching SCRUM in the classroom. Additionally, in another experiment conducted by Sánchez-Gordón et al. [38] assessing a SG they developed for the ISO/IEC 29110 Standard understanding, it was reported that participant involvement has the highest score, with 94% of the participants agreeing on playing the game again. A rate of 79% of participants reported that the SG is not only engaging but also a good alternative to traditional classroom activities. The lowest scores were reported regarding the SG's potential to improve their knowledge of the standard and to encourage them to know more. In contrast, the study by Drappa et al [39] found no learning effect caused by their designed SG in their experimental results.

2.1.1 *Gaps and Opportunities in Serious Games for Software Engineering Education*

Many studies were investigating how to learn/teach SE through playing SGs [27]. Nevertheless, to the best of our knowledge, there are no Systematic Literature Reviews (SLR), Systematic Mapping Studies (SMS), or other secondary studies in literature, that address the topic of SGs in SE in the same depth or with the same aims as our thesis, namely by: (1) Examining SWEBOK areas that have been targeted for SGs in SE in the literature and map the aforementioned SGs accordingly, enhancing the relevance of SGs to their aligned SWEBOK areas in SE education. This way both educators and students could access easily SGs addressing their targeted knowledge areas when learning SE. (2) Extracting general characteristics and components addressed in these SGs in SE studies. This could provide educators with a comprehensive view of current trends in SGs for SE education. Simultaneously, it would offer insights to SG designers regarding the various gamification elements and Game Mechanics - rules and systems designed to enhance the player's gaming experience - implemented within these SGs. (3) Studying the potential importance of having a deep knowledge of the player profiles and their characteristics, addressing a concept previously overlooked in SG studies. It could provide insights for educators and students on personalized gaming experiences and guide SG designers in tailoring games to specific player preferences and characteristics. (4) Identifying Game Dynamics which are interactive mechanisms and elements integrated into SGs to facilitate the accomplishment of each level of Bloom's taxonomy in the learning process. Likewise, Learning Mechanics referring to design elements that facilitate effective learning processes and knowledge retention were identified. Both game dynamics and learning mechanics were addressed by the SGs in these studies and the different levels of the Bloom's taxonomy incorporated in these SGs were investigated as the latter were never explicitly cited in any of the former SGs in SE studies. Thus, educators and students could be enlightened about the effectiveness of these elements in the learning process. Simultaneously, would guide SG designers in incorporating these elements for enhanced educational impact.

On a different note, while there were experiments that have shown a positive impact of SGs on learning outcomes and students performance and satisfaction, it's worth noting that not all SGs produce similar results. In contrast, some studies have reported mixed or even negative outcomes. Moreover, these studies did not address the potential impact of a professor's experience in integrating SGs into their course on the overall student experience. To the best of our knowledge, no comprehensive study has been conducted on the influence of SGs on students' academic achievements in the field of software project management. This gap is particularly notable given the recognized need for innovative strategies to enhance the teaching and perception of software project management concepts during theoretical sessions [40]. The literature lacks detailed investigations into the effects of SGs on key academic metrics such as students' final scores, exam attendance, and overall course success. Additionally, the impact of teachers' experience with SG implementation on its effectiveness as a learning tool has not been thoroughly explored. Finally, there is a scarcity of practical recommendations for SG stakeholders and SE educators to maximize the benefits of SGs in software engineering education.

2.2 Advances and Challenges in Serious Game Design

2.2.1 Distinctive Features of Serious Game Design Compared to Traditional Game Design

The design process of serious games differs significantly from that of regular games. Serious game design inherits many of its traditions from entertainment-oriented game design [41]. However, serious game design is a multidisciplinary process that requires a collaborative approach, as it aims to have a purposeful impact on the players' lives beyond the self-contained aim of the game itself [42]. This is in contrast to regular game design, which primarily focuses on entertainment value and gameplay experience. The design of serious games also involves the use of standardized frameworks and a systematic process, including phases such as context identification, user requirements, planning, design, construction of interaction devices and video game, and evaluation [43]. Additionally, the complexity of serious game design may hamper the games' effectiveness [44]. Furthermore, the design of serious business process games requires game designers to have business process modeling skills and instructions on how to represent business process elements in the game context [45]. Moreover, the design of serious games for specific sectors involves a multidisciplinary process that aims to streamline the design and facilitate the reuse of domain knowledge and personalization algorithms [46]. In contrast, regular games do not necessarily require the same level of multidisciplinary collaboration and standardized frameworks. They are primarily focused on entertainment value and gameplay experience and may not have the same level of systematic and iterative design processes as serious games.

2.2.2 The Role and Benefits of Serious Game Design Frameworks and Methodologies

Serious game design frameworks and methodologies have played a pivotal role in the development of serious games elevated both their development process and effectiveness. These frameworks provide a structured approach to the design and development of serious games, ensuring a balance between entertainment and education while supporting the design process [47]. By providing a clear methodology for design and development, these frameworks have facilitated the creation of serious games that can be effectively utilized in formal education, training, and societal scales, dealing with the skepticism of instructors and addressing various real-world challenges [48]. Additionally, these methodologies contribute to the purposeful design of serious games by analyzing a game's formal conceptual design and its elements based on the game's intended purpose [49]. Despite the initial investment, serious game design frameworks prove to be a cost-effective game design method in the long run as they can be reused repeatedly simplifying the design process of personalized serious games without incurring additional costs [46].

2.2.3 Gaps and Challenges in Serious Game Design

Designing SGs is not an easy task as it is often perceived as a challenging task by educators, which necessitates the nurturing of diverse ideas, fostering academic objectivity towards SGs, and the adoption of creative design and development methodologies [50]. One major issue is the lack of standards and guidelines for SGs development [51]. This challenge impedes the consistent and reliable design of serious games. Carrión et al. [52] conducted a systematic literature review to collect a set of documents related to the origins, definitions, classifications, evaluations, and frameworks of SGs design. Although some research work in the field of serious gaming frameworks was found, none of the approaches studied by Carrión et al. [52] use techniques to promote creativity and most of them are not generalizable and non-integrable with agile approaches. Another significant challenge is the limited integration of pedagogical principles into game design. Many serious games focus primarily on gameplay mechanics, often at the expense of incorporating well-established educational strategies, which can restrict their educational benefits. User experience is also frequently insufficiently considered. Although engagement and motivation are crucial, the design process often overlooks these elements, which are essential for effective learning experiences. Accessibility and inclusivity are additional areas of concern. Serious games should be designed to accommodate diverse learners, including those with disabilities or varying learning styles. However, many games fail to include necessary accessibility features. Interdisciplinary collaboration, a vital component of effective serious game design, is often lacking. Effective design requires input from various fields, including game design, education, psychology, and technology, but this collaboration is not always adequately realized. Scalability and adaptability represent further gaps. Serious games are often tailored for specific contexts, limiting their flexibility and broader applicability. Lastly, ethical and social considerations are frequently overlooked in the design process. Issues such as privacy, data security, and cultural sensitivity need to be addressed to ensure that serious games are developed responsibly and inclusively. Addressing these gaps is crucial for advancing serious game design and maximizing its impact on education and learning outcomes.

3

RESEARCH METHOD

This thesis employs the DSR methodology, which is particularly well-suited for research that involves the creation and evaluation of artifacts to solve identified problems. The DSR methodology provides a structured approach to developing, refining, and validating artifacts that contribute to both academic knowledge and practical applications. The research conducted in this thesis follows the DSR process through four main stages: Problem Identification, Artifact Development, Artifact Evaluation, and Contribution to Knowledge.

3.1 Problem Identification

The research begins with a thorough analysis of the current landscape of SGs in SE education. The lack of a standardized approach to designing SGs for SE, combined with the need to understand their effectiveness in the classroom, was identified as a significant gap in the literature. The problem was further articulated through the following activities:

- Systematic Mapping Study: A systematic mapping study was conducted to identify existing SGs in SE education. This study classified the identified SGs according to their focus on specific SE domains (e.g., SWEBOK areas) and their alignment with Bloom's Taxonomy levels. The mapping also included an analysis of the game elements, mechanics, dynamics, gamification elements, and player profiles targeted by these SGs. Additionally, a comparison was made between the success factors of SGs in SE and those of general SGs across various fields.

- Gap Identification: Through the mapping study, gaps were identified in the current use of SGs in SE education. These gaps included a lack of standardized design frameworks, underrepresentation of certain SWEBOK areas, and insufficient consideration of player profiles and game dynamics tailored to SE contexts.

- Controlled Experiments: A controlled experiment across multiple academic years was conducted to assess the impact of incorporating SGs into SE education compared to traditional teaching methods. This evaluation compares the performance, attendance, and satisfaction of students who participated in SG-enhanced courses with those who followed traditional teaching methods. The experiment also examines the role of instructor expertise in the successful deployment of SGs.

3.2 Artifact Development

Building on the insights from the problem identification stage, the research focused on developing new artifacts to address the identified gaps. The artifacts include a novel framework for SG design in SE education and a set of guidelines for SE educators.

- Framework Development: The findings from the mapping study, combined with the results from the subsequent experiments, informed the development of a comprehensive framework for SG design in SE education. This framework was designed to be evidence-based, incorporating best practices from existing frameworks and proposals. It aimed to standardize the SG design process, making it more straightforward for educators and facilitating more effective communication between educators and game design experts.

- Guidelines for Educators: Practical guidelines for SE educators were developed to help them effectively incorporate SGs and advanced gamification techniques into their teaching practices. These guidelines were based on the findings of controlled experiments and the analysis of the role of instructor expertise in SG implementation.

3.3 Artifact Evaluation

The developed artifacts were evaluated through a series of controlled experiments and expert validation processes to ensure their effectiveness and practical applicability.

- Expert Validation: The proposed SG design framework was further validated through consultations with experts in both SE education and game design. This validation process ensured that the framework was both theoretically sound and practically relevant, capable of addressing the needs of SE educators and game designers alike.

- Application Validation: To further validate the proposed SG design framework, it is currently being applied in the design and development of new serious games tailored for Software Engineering education. This ongoing process involves using the framework to guide the creation of SGs, ensuring that the design process aligns with the evidence-based practices and standards set out in the framework. By applying the framework in real-world game design scenarios, we aim to assess its practical effectiveness and adaptability across various educational contexts.

The results of this validation process, including feedback from educators and game designers, will be used to refine the framework further. While this phase of validation is still in progress, the preliminary outcomes suggest that the framework is facilitating a more structured and efficient SG design process, ultimately contributing to the creation of more effective educational tools.

3.4 Contribution to Knowledge

The final stage of the DSR process involved synthesizing the findings and contributing new knowledge to the field of SE education.

- Synthesis of Findings: The results from the mapping study, controlled experiments, and expert consultations were synthesized to formulate best practices for SG design and implementation in SE education. These best practices were incorporated into the developed framework and guidelines.

- Identification of Future Research Directions: Based on the gaps identified and the findings from the research, several future research directions were proposed. These include exploring the use of Artificial Intelligence (AI) in SGs for SE education, further refining the proposed framework, and investigating the long-term impact of SGs on students' professional development.

- Framework and Guidelines Publication: The developed framework and guidelines were published/to-be-published to provide SE educators and game designers with a standardized, evidence-based approach to SG design and implementation. This contribution aims to improve the quality and effectiveness of SGs in SE education and to foster greater collaboration between educators and game designers.

This methodology ensures that the research not only addresses a significant gap in the literature but also provides practical tools and frameworks that can be used to enhance SE education through the effective design and implementation of SGs. The iterative nature of the DSR process, involving continuous refinement and validation of artifacts, ensures that the final contributions are both rigorous and relevant to the field.

This thesis employs the DSR methodology to systematically investigate, design, implement, and evaluate the integration of SGs and advanced gamification techniques into SE education. The research is structured to align with the core principles of DSR, encompassing problem identification, artifact creation, evaluation, and knowledge contribution. The methodology section is organized to reflect these stages, ensuring a coherent approach that links the research objectives and contributions into a unified framework.

4

CHAPTER 1: Systematic Mapping Study on Serious Games in Software Engineering

4.1 Introduction

The application of the SG approach is particularly promising in the SE field, where one of the essential competencies, as in all engineering, is the practical application of knowledge, i.e. "learning by doing". Therefore, there is a growing interest in the use of games in SE education, which suggests a smooth transition to a more game-like environment blended with traditional teaching methods as a leading strategy to greatly affect future software engineers [53]. Along the same line, educational SE was advocated as an emerging area of research in which game technologies often play an important role along with SE technologies for teaching and learning SE [54]. Given the limitations that SE education has when provided in its traditional teaching approach, and which are related to the applied nature of SE that requires learners to experience real-world issues to acquire an appreciation for SE concepts and best practices, game-related approaches came in handy to overcome some of these limitations [55]. One such approach is GDBL, where students learn through the process of creating games themselves, gaining hands-on experience in designing, programming, and testing games, which can offer valuable insights into subjects in the realm of SE [56].

However, although the use of games is littered throughout the history of SE, existing research in the field can be considered quite preliminary [53,57]. Furthermore, among other aspects, to solidify knowledge in this field and to facilitate secondary studies in summarizing results, the research community is advised to seek a common vocabulary for game-related approaches in SE education [30]. Because of this trend, we considered it convenient to differentiate between SGs, gamification, GBL, and GDBL afore performing a mapping study to determine and characterize the state of the art of SG in SE, analyzing existing initiatives, depicting the SWEBOK areas [58] that have been the most addressed, the different levels of Bloom's taxonomy that have been covered by the found SGs, as well as the covered game related approaches, game dynamics, learning mechanics, and player profiles. In the realm of SE, these comprehensive dimensions of SGs have hitherto remained unexplored in earlier work. Our study attempts to address this crucial gap in the literature, offering a comprehensive investigation into the uncharted facets of SGs within the context of SE. As such, our research is meant to serve as a pivotal resource, providing educators and developers with exclusive and insightful information by shedding light on these unexplored dimensions, empowering stakeholders to make informed decisions, and fostering a deeper understanding of the complex dynamics at play.

4.2 Methodology

In this section, the study's goal, methodology, the research questions are explained, and so is the motivation behind them.

4.2.1 *Study Process*

We carried out the mapping from August 2020 to October 2021 and the search for literature was held between December 2020 and May 2021. Three researchers participated in the planning and execution of this study following five process steps adapted from [59]: identifying research questions, defining data sources and research strategy, as well as the selection, classification, and evaluation criteria of literature. Each step of the process has an outcome, converging on the systematic map as the end result of the process. We offer a best practices checklist, as shown in Table 10 in the Appendix A to ensure the research process's rigor. The steps and sub-steps of our Research Process are aligned with these guidelines, as illustrated in the checklist.

4.2.2 *Goal and Motivation*

The aim of this study is to determine and characterize the state of the art of SGs in SE, analyze existing initiatives, and identify challenges for future research. This leads to the main research question that prompted this study, which is: How did SGs get adopted into SE education? To achieve this systematic mapping, we have followed the guidelines by Petersen et al. [59,60] which are consistent with the guidelines by Kitchenham and Charters [61]. Thus, we present the planning of the activities conducted in each stage of the mapping study.

4.2.3 *Research Questions*

The following Table 1 describes the research questions addressed in this paper and the motivations behind them.

Questions	**Motivation**
RQ1. What SWEBOK software areas have been targets for SGs in SE?	In this question, we study SGs in SE and the SWEBOK areas that the latter have targeted
RQ2. How do Bloom's Taxonomy levels and game dynamics align in SGs for SE?	In this question, we aim to identify the game dynamics implemented in SGs for SE through the game design metric proposed by Pendleton and Okolica [62], according to which the achievement of every level of Bloom's taxonomy requires that SGs

	implement a set of these dynamics. We also assess particular mastery levels in SGs for SE by identifying the levels of Bloom's Taxonomy [63] as applicable to these SGs through the Learning Mechanics - Game Mechanics (LM-GM) framework [64].
RQ3. How are game-related approaches and learning mechanics implemented in SGs for SE?	In this question, we aim to identify both the gamification elements alongside the Game mechanics implemented in SGs for SE. Gamification elements were identified following the periodic table of gamification elements proposed by Marczewski [65] while Game mechanics were extracted from the LM-GM framework [64]. Also, we aim to derive the Learning mechanics implemented in SGs for SE through the LM-GM framework [64].
RQ4. To what extent do SGs in SE take into consideration player profiles in their gameplay?	In this question, we aim to distinguish what kind of player profiles have been addressed in SGs for SE, enumerate the classifications adopted for the player profiles in SGs for SE, highlight the characteristics of the player profiles identified, and investigate to what extent SGs in SE adapt the gameplay to match the player profiles.

Table 1: Research Questions

4.2.4 Search Strategy

Search Scope

The PICO (Population, Intervention, Comparison, and Outcomes) [66] was developed to identify keywords and formulate search strings from research questions. In this study, we will be breaking our question into key concepts using the evidence-based practice search formula PICO.

- Population: In SE, the population can refer to a specific role in SE, to a category of software engineer, to an application area, or to an industrial group [61]. In our context, the population is SE studies.

- Intervention: In SE, intervention refers to a methodology, a tool, a technology, or a software procedure. Within the framework of this study, we will focus on SGs.

- Comparison: In this study, no empirical comparison is made. Thus, no term was selected for this aspect.

- Outcomes: All existing outcomes regarding SGs in the field of SE are of interest in this mapping study. Therefore, no term was selected for the outcome aspect in order not to restrict the result set.

Search Terms

The search strategy is illustrated in Figure 1. According to Kitchenham and Charters guidelines [61], PICO criteria were used for developing the search string. However, one constraint when defining a search string is that the result set should be of manageable size, but still has the maximum possible coverage. Therefore, some additional synonyms and the most common relevant terms driven by the research questions for each attribute were selected and added to the search string in accordance with the recommendations of Petersen et al. [59].

The identified keywords are SGs, SE, and the SWEBOK areas which were grouped into sets, and their synonyms were considered to formulate the search string. Search terms were constructed using the steps described in [67], where Boolean OR is used to incorporate alternate spellings, synonyms, or related terms, and Boolean AND is used to link main terms. Then the search was mainly focused on the Title, Abstract, and Keywords of the studies. This search string was reviewed and agreed on among all authors.

Search Resources

Concerning the search resources, we considered using mainly: Science Direct, IEEE Xplore, ACM, Scopus, as well as Wiley. These databases were selected considering the number of relevant conferences and journals indexed by them. Being aware that the number of papers addressing the topic would be small, the authors tried to undertake a comprehensive search, which has been used on all fields in databases that do not allow search in specific fields such as Title, Abstract, and Keywords. In order to ensure that no important articles were overlooked, a cross-check was performed using the same search process on all selected databases. Unfortunately, some databases did not allow using the full search string as defined. Therefore, slightly adapted, and simplified search strings were used instead to suit the specific requirements of the different search interfaces.

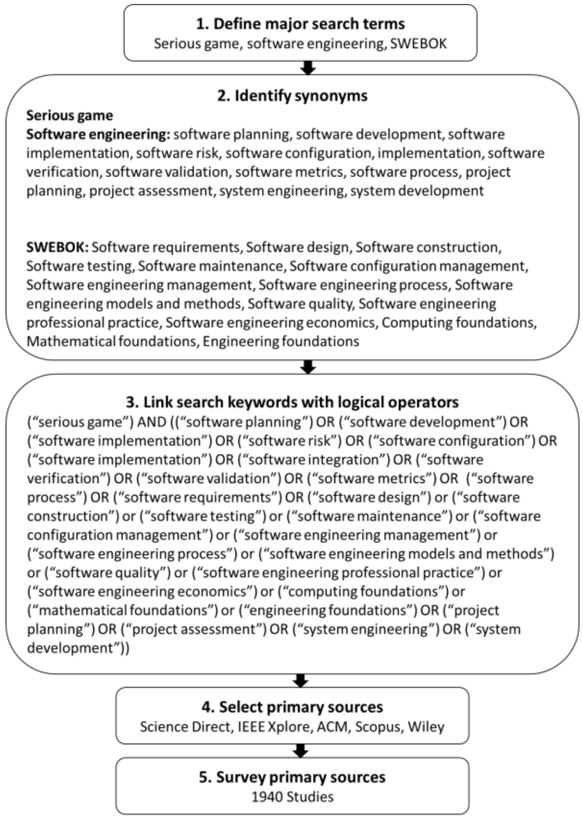


Figure 1: Search Strategy

4.2.5 Study Selection

After obtaining the potentially relevant studies, these should be evaluated for their real relevance to the scope of our study.

Search Criteria

The following Table 2 displays the inclusion and exclusion criteria that were applied to Titles, Abstracts, and Keywords and based on which the papers were kept or discarded in our selection process.

Inclusion Criteria	IC1: Studies that investigate SGs. IC2: Studies are in the field of SE. IC3: Studies written in English.
Exclusion Criteria	EC1: Papers not accessible in full text. EC2: Studies that are duplicates of other studies or duplicate papers of the same research in different databases. EC3: Papers available only in the form of abstracts or papers which present workshop abstract submissions, posters... EC4: Studies that are focused on the use of SE to build SGs for other fields.

Table 2: Inclusion and Exclusion Criteria

Selection Process

The search for primary studies was recursive. In other words, once relevant studies had been identified, the references from the latter were routinely searched using the same search criteria. Our study only considered articles that met all the inclusion criteria and discarded those that met any of the exclusion criteria identifying 125 studies as detailed in Table 3. Among these 125 studies, borderline ones were deemed relevant during the selection phase based on Title, Keywords, and Abstract. After being reviewed by the first author by reading the full text, irrelevant studies were discussed among the rest of the authors and thus excluded later in this process. The remaining 46 papers were used to perform snowball sampling [68], which led to the addition of 28 studies to bring the total number of selected studies to 74 papers.

	Scopus	Science Direct	IEEE	Wiley	ACM	Springer	TOTAL
Search Results	607	13	46	63	246	1631	1940
IC1	607	13	46	17	121	236	1040
IC2	563	11	13	13	25	153	778
IC3	545	11	13	13	25	150	757
EC1	108	11	1	13	9	150	292
EC2	288						
EC3	285						
EC4	125						
Full Text	46						
Snowballing	74						

Table 3: Selection Process Summary

4.2.6 Data extraction

In order to extract the data from the identified primary studies, we developed a comprehensive Excel template. Each data extraction field has a data element and a value. The extraction was carried out by the first author and revised by the rest of the authors by tracking the information contained in the extraction form to the declarations of each article and checking their accuracy. Details regarding the data items in question are described in Table 9 in the Appendix A.

4.2.7 Studies Classification

The information for each item extracted has been visually illustrated. The extracted items were grouped per Research Questions (RQ) by the first author during the analysis. Subsequently, the papers belonging to each RQ were counted. The classification categories covered by each research question were based on the thematic analysis performed on extracted data. Details regarding the classifications in question and the different SGs studies belonging to each classification are described in Appendix C.

4.3 Results and Discussion

In this section, we present and analyze the results obtained after the classification of the extracted data from the set of selected papers for this mapping. Thus, this section aims at showcasing the demographic data and answering the RQs defined in Table 1.

4.3.1 Demographic Data

Before addressing the RQs, various demographic attributes of the selected papers will be presented in the subsections below.

Distribution of Papers

Aiming to track the publication evolution of SGs in SE studies and identify any raise of interest toward these studies, Figure 2 displays the number of SGs in SE studies identified from the early 2000s to May 2021. The first study was published by Drappa et al. [69], and it was the only study in the year 2000. While the interest in SGs in SE studies started to gain consistent interest in 2010, it was until 2015 that a significant increase of 166.66% was observed followed by a growing interest throughout the last 5 years. Although SGs are not a recent topic, and nor is their use in educational contexts [23], this increase in the number of SGs in SE studies in recent years indicates that these games are considered highly relevant by the SE research community for SE education and/or training.

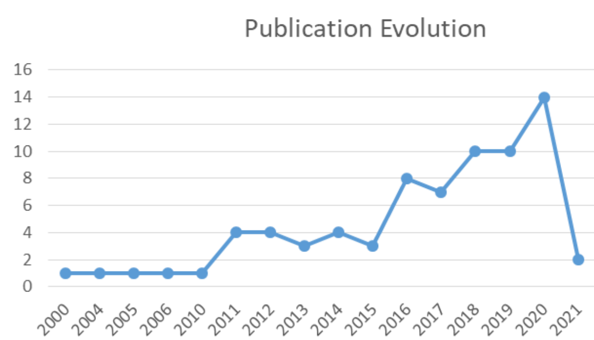


Figure 2: Yearly Publication Evolution of SGs in SE

Research Types & Methods

As Figure 3 shows, nearly 60% of selected studies are Validation Research, followed by Solution Proposals and Evaluation Research which make up 17.56% and 13.51% respectively of the selected studies. Experience Papers come 4th with a rate of 3.05% followed by both Opinion Papers and Philosophical Papers which make each 2.70% of the selected set. According to this ranking, the top three research approaches for SGs in SE studies are novel solutions or significant extensions of an existing technique. The latter applicability is shown by a small example or a good line of argumentation (Solution Proposal) such as studies by Aydan et al. [70], Chung et al. [71], and Miljanovic et al. [72], an experiment within an academic setting prior its implementation in practice (Validation Research) namely studies by Xenos et al. [73], Viscaino et al. [33], and Caserman et al.

[74], or a truthful evaluation in an industrial setting when implemented in practice (Evaluation Research) citing for instance studies by Alexandrova et al. [75], Przebylek et al. [76], and Rodriguez et al. [77].

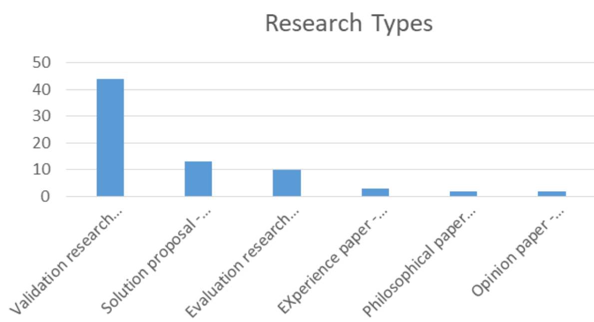


Figure 3: Research Types of SGs in SE

Aiming to identify the main trends in research methods that have been adopted in the evaluation of SGs in SE we look further to Figure 3, according to the latter 28.37% of selected papers are either Solution Proposals, Opinion Papers, or Philosophical Papers that do not tackle by their nature the evaluation of the solutions that these papers discuss. Thus, no research methodology for SGs in SE evaluation was reported in these studies. As for papers that are Validation Research, Evaluation Research, and Experiment Papers, 72% of the latter do report information about their research methodology in the evaluation of the presented SGs in SE. The distribution of these papers between Case Studies, Surveys, and Controlled Experiments is displayed in Figure 4. Notably, a significant percentage of 28% of these papers did not report any specific research methodology. This omission can be attributed to several factors, including the exploratory nature of some studies, a lack of rigorous methodological frameworks in certain research, or simply an oversight in reporting. The presence of a substantial N/A percentage highlights a gap in the field where methodological transparency and rigor could be improved. Addressing this gap in future research is crucial for advancing the robustness and reliability of findings in the evaluation of SGs in SE.

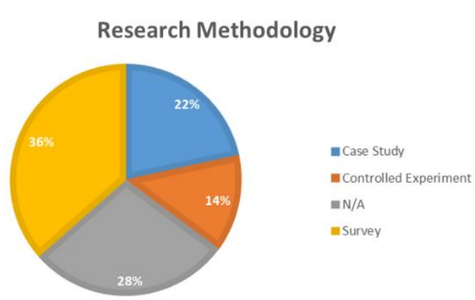


Figure 4: Research Methodologies of SGs in SE

Publication Channels

Seeking to classify SGs in SE studies by publication channels, peer-reviewed venues including Journals, Conferences, and Workshops were considered in this classification. Figure 5 gives an overview of the distribution of the selected studies between these venues.

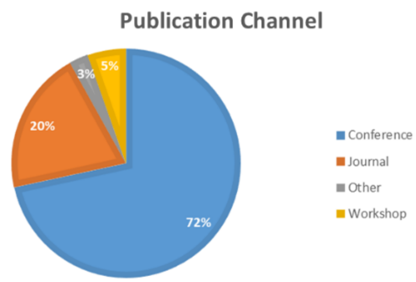


Figure 5: Publication Channels of SGs in SE

4.3.2 *Quantitative Summary*

After this study's analysis, the first point to be emphasized is that SGs in SE are still at a very early stage. As can be seen from the results presented in the previous section, most of the selected studies were published at conferences, and only 20.27% of the studies were published as journal articles, This parallels the situation observed in studies focused on gamification within SE, where only 7% made it to journal publications while 47% of the studies observed were published at conferences [78], Clearly, both SGs and gamification in SE are still navigating preliminary phases of research. However, SGs in SE research diverges from its gamification counterpart since only 5.4% of the studies on SGs in SE were published in workshops, while 39% of studies on gamification in SE are published in these same venues [78]. This variance may be attributed to the promising nature of SGs within SE. It suggests that SGs in SE is an emerging field with considerable potential compared to gamification in SE, and that the SGs in SE studies published to date have already reached a level of maturity that enables them to offer comprehensive and substantial findings.

Another interesting remark regarding SGs in SE studies revealed in Figure 6 is that the few papers that were published in the early 2000s were either solution proposals or validation research presenting preliminary results. In 2010, more validation papers were published than solution proposals and only one opinion paper was published. It was until 5 years later that SGs in SE gained great interest, and more various studies were published from all distinct research types. However, validation papers remained the leading research type followed by evaluation papers till the late 2020. This leads to an interesting conclusion that studies on SGs in SE are getting more mature and that this area presents a real trend that is about to soar.

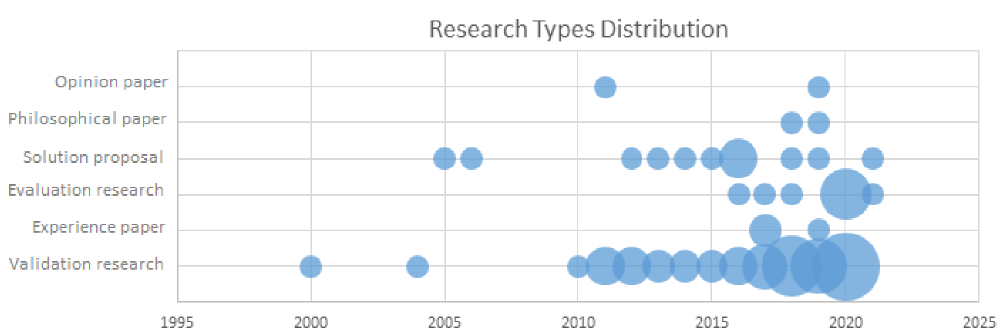


Figure 6: Research Type Yearly Distribution

In addition, most invested research types regarding SGs in SE studies are validation and evaluation research, this reflects the growing interest that this topic gained not only in the academic world but also in the industrial one. Moreover, and as shown in Figure 7, starting in the year 2011 and every other year, there were new SGs in SE focused on professionals rather than the solely sweep of academic-focused ones. It wasn't until 2016 that this new opening into the professional world became steady and increasing throughout the years. Although academia is highly research and discovery-focused. In contrast, the industry seeks a sense of immediate impact on scoped users under real-life constraints. This stresses the fact that SGs in SE have proven their impact in both educational settings with students, and training settings with practitioners and professionals in the industry. However, a total of 20.02% of analyzed studies did not provide information about the scope of their presented SGs. Further attention regarding this metric will help better identify the public market for SGs in SE and better encounter SGs that presumed users' needs.

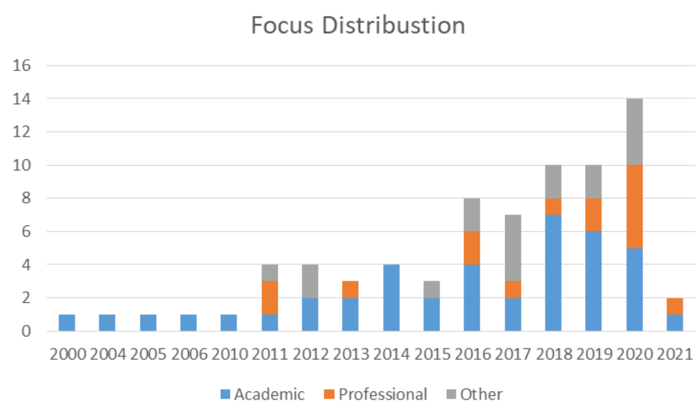


Figure 7: SGs' Focus Yearly Distribution

Despite the expansion of both the focus and scope of SGs in SE, only a few studies reported explicit information about the pricing strategy of these SGs, to whom belong their intellectual property and their license nature. Although the studies reporting intellectual rights properties were each assigned to an academic institution [37] and a recognized company [79]. However, both were academic-focused and targeted students as their main scope users. In addition, both studies were in the Software Engineering Management knowledge area. The same applies to studies reporting their pricing strategy [79,80]. Given the fact that the very first publications regarding SGs in SE were in Software Engineering Management, and that the latter is the knowledge area with the most publications throughout the years and the one with the most comprehensive information regarding SGs in its SWEBOK area. This leads to the conclusion that SGs tackling this SWEBOK area are not only the first ones to be investigated among the other SGs in SE field, but also the ones with most popularity and the most mature ones.

4.3.3 Research Questions

In this study, our research questions are structured to comprehensively explore the landscape of SGs within the realm of SE. The Figure 8 below provides a detailed overview of the dimensions addressed by these questions and how they were methodically extracted from the literature or derived from established taxonomies.

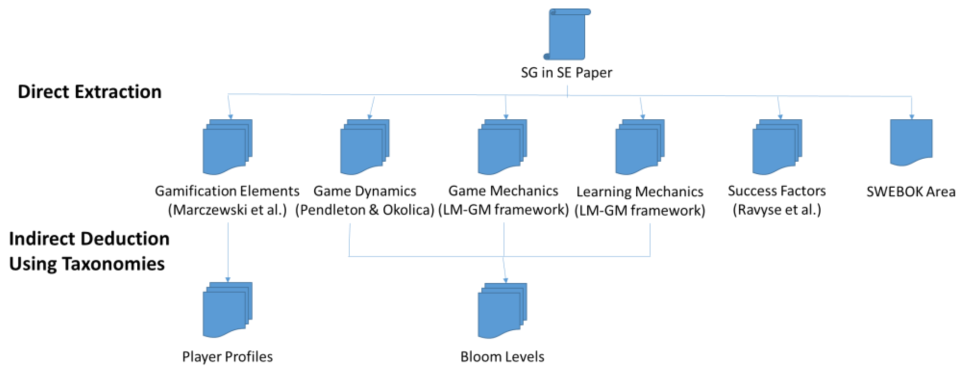


Figure 8: Visual Representation of Dimensions Addressed by Research Questions

RQ1. What SWEBOK software areas have been targets for SGs in SE?

This research question aims to track the distribution of SGs in SE studies on the different SWEBOK areas and identify the main targeted areas by the identified SGs.

As illustrated by Figure 9, Software Engineering Management was found to be the SWEBOK area with the most reported SGs in SE, which makes a significant 33.78% of the total selected SGs for this study, based on the number of papers displayed in the figure. This followed by SE Professional Practice which represents 13.51%, and 12.16% of SGs in SE studies tackling Computer Foundations. Moreover, SGs in the SE Process rank 4th with a rate of 10.81%, and studies addressing SGs in the Software Requirements area come 5th with a 9.45% ratio. In contrast, SGs studies in SE Models and Methods, Mathematical Foundations, and Software Maintenance, rank last with a minimum share of 1.35% each. According to these findings, a significant gap rises in SGs in SE studies that approach the latter's targeted knowledge areas.

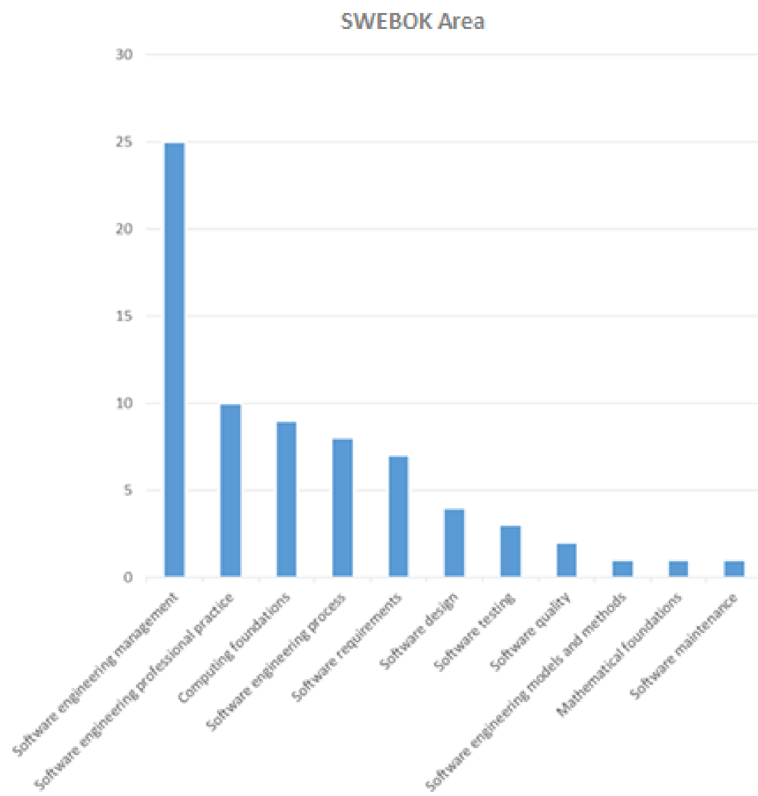


Figure 9: SWEBOK Areas of SGs in SE

RQ2. How do Bloom's Taxonomy levels and game dynamics align in SGs for SE?

This mapping was done using the game dynamics classification based on the game design metric proposed by Pendleton and Okolica [62]. According to this metric the achievement of every level of Bloom's taxonomy requires that the SGs implement a set of these dynamics while the assessment of the tendency of particular mastery levels in SGs for SE by identifying the levels of Bloom's Taxonomy as applicable to these SGs through the LM-GM framework [64]. The latter is a classification based on Bloom's ordered thinking skills which maps the corresponding Game mechanics and Learning mechanics to each mastery level.

As shown in Figure 10, barely 44.82% of the total Game dynamics were reported in the selected SGs. Among this set, Teams rank first with a rate of 16.21%, followed by Realism with a rate of 13.51%, and third come Competition and Limited Actions with an equal share of 6.75%. Given those results, we can conclude that a lot of Game dynamics are underused in SGs in SE and that further investigation should study their importance and impact on the learning/training of the players if implemented in SGs in SE.

Figure 10: Game Dynamics of SGs in SE

As shown in Figure 11, all levels of Bloom's taxonomy were tackled in SGs for SE with different variations, depending on the difficulty of each level. The level that was most addressed by the selected studies is Applying, with a rate of 58.10%. Understanding comes second, being popular among 52.70% of the analyzed SGs. Evaluating ranked 3rd, with a representation of 44.59%. Next comes Analyzing, which was addressed in 39.18% of the studies. Fifth comes Remembering, with a rate of 37.83%. Creating ranked last with a share of 10.81%.

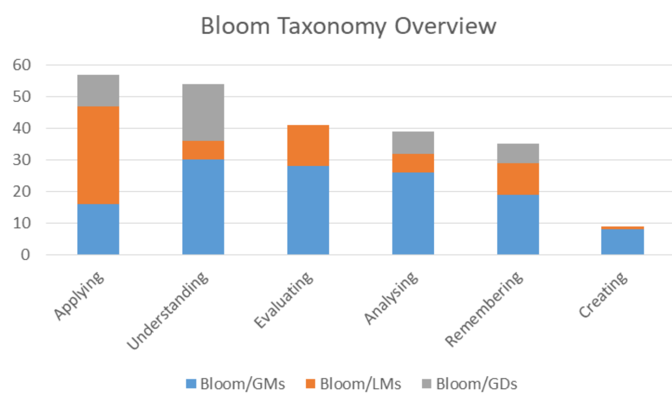


Figure 11: Bloom's Taxonomy of SGs in SE

We can conclude that the contents of the SE subjects require higher Bloom's Taxonomy levels. The reason why SGs in SE implemented even advanced levels of Bloom's taxonomy, which are difficult to attain. However, it is crucial to foster higher-order thinking in SE players by building from lower-level cognitive skills. Along the same line, the study by Wankhede and Kiwelekar [81] suggests that the qualitative assessment of SE with Bloom's Taxonomy should focus on Higher Order Cognitive Skills (HOCS).

RQ3. How are game-related approaches and learning mechanics implemented in SGs for SE?

This question is going to be divided into three other sub-questions where we will be identifying in each the gamification elements, the Game mechanics, and the Learning mechanics derived from the former using the LM-GM framework [64].

What are the gamification elements implemented in SGs for SE?

In this question, we aim to assess the popularity rate of gamification elements among SGs for SE by grouping the latter by their implemented gamification elements. As illustrated in Figure 12, only 26 out of the 51 gamification elements of Marczewski's periodic table [65], were reported in our set of selected SGs in SE. The most common gamification element was Points with a rate of 41.89% among selected studies, followed by Progress/Feedback which is implemented in 35.13% of the SGs in SE, Time-Dependent ranked 3rd with a representation of 28.37%, followed by Time Pressure with a share of 25.67%, and 5th comes Level/Progression with a representation of 24.32% over selected studies. The fact that SGs in SE implement only 50.98% of the available gamification elements, represents a significant gap, limiting the comprehensive leverage from these technologies.

In alignment with the findings of the mapping study by Pedreira et al. [78], our research indicates that within the context of SGs in SE, gamification elements exhibit notable patterns of adoption. Notably, the simplest gamification elements have prominently featured among the most frequently considered elements within SGs in SE, mirroring the trends observed in gamification contexts in SE. It is essential to acknowledge that the said gamification in SE systematic mapping identifies this particular tendency as potentially hazardous, as it may hinder the completion of the Gartner prediction. However, our study unveils a more nuanced landscape. Intriguingly, we observe a distinct prominence of time-dependent elements, which have emerged as leading components among gamification elements implemented in SGs in SE. This temporal focus suggests a conscientious integration of time-sensitive mechanics, potentially enhancing engagement and user experiences within the SE educational domain. Moreover, our investigation reveals that SGs in SE implement more comprehensive and diverse gamification elements compared to gamification in SE. While the implementation percentages of these diverse elements may be relatively modest, their presence signifies a concerted effort to optimize SGs in SE through well-informed and diversified gamification practices.

According to Werbach and Hunter [82], gamification elements can be classified into Dynamics, Mechanics, and Components. Following this taxonomy, Narrative, Virtual economy, and Level/Progression fall under Dynamics. Moreover, Level/Progression, Challenges, Competition, Progress/feedback, Fixed Rewards, and Prizes are considered Mechanics. Last, Points, Badges, Leaderboard, Time-dependent, Time pressure, Unlockables, Level/Progression, and Quests are Components. This leaves Strategy, Learning, Customization, Development Tools, Theme, Flow, Innovative Platform, Investments, Anonymity, Voting, and Consequences unclassified. Given that the list of the gamification elements proposed by Werbach and Hunter [82] is roughly different from the one presented in Marczewski's periodic table [65], and so is the way some common elements were formulated in both taxonomies. Level/Progression is categorized as a gamification Dynamic, Mechanic, and Component all at once, which contradicts the purpose of a taxonomy.

Figure 12: Gamification Elements of SGs in SE

What are the game mechanics implemented in SGs for SE?

In this question, we aim to assess the popularity rate of game mechanics among SGs for SE by grouping the latter by their implemented mechanics. As can be seen in Figure 13, only half of the game mechanics were reported in the analyzed SGs. The most common game mechanics within selected SGs in SE is Role Play, with a rate of 27.02%, followed by Realism, which is implemented in 25.67% of the SGs in SE, Cut Scene/Story ranks 3rd, with a representation of 21.62%, followed by both Rewards/Penalties and Resource Management, with a share of 18.91% each, and 5th comes Tutorial, with a representation of 14.86% over selected studies. Just like gamification elements, game mechanics seem to be under-implemented in SGs for SE, which highlights a significant gap obstructing the full advantage of these technologies in improving SGs in SE.

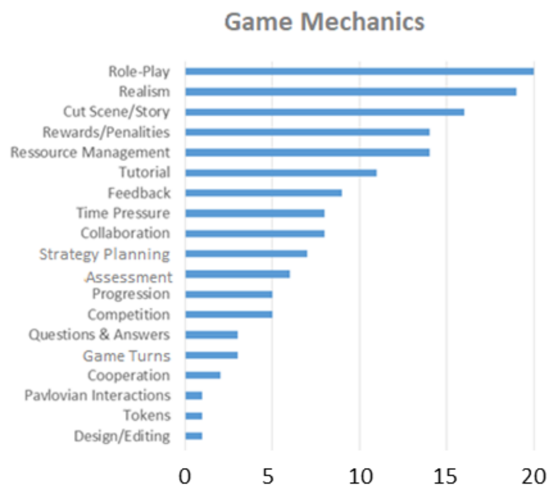


Figure 13. Game Mechanics of SGs in SE

In the forthcoming paragraphs, we present illustrative examples of game mechanics observed within our mapped SGs in SE, while underlying the motivations and benefits associated with their implementation, elucidating the reasons behind their selection to enhance learning experiences among learners using these SGs.

Kemell et al. [35] proposed an innovative board game teaching future project managers and software engineers the Essence Theory. This SG was considered reasonably realistic in simulating a SE project as it relied on role-playing. This feature was not only used to simulate the gameplay but also to make the SG more challenging. The players assume the roles of the project team members with one player acting as a project manager. However, to add an element of competition into the game, the role of "the son of the boss" was introduced. Thus, the player taking this role will be part of the project team, but will also seek to disrupt the project for his own benefit.

Maarek et al. [83] presented the combination of SG design approaches they used in the co-creation of a SG to investigate developer-centered security. According to this approach, an instructional tutorial was made in place for players to familiarize themselves with the gameplay. While so, the layout of the game elements taught players how to play, and directional heuristics helped direct players to specific game elements as they played.

Bell et al. [84] proposed a Halo-based SG called Secret Ninja Testing for software testing. They concluded that by hiding testing behind a short story and series of quests, Halo shields students from discovering that they are learning testing practices, making it easier to expose students to software testing at an early stage. They also stressed the fact that story designers could help improve the implementation of Halo for a course of SE and expressed their interest in further studying game design for this purpose.

Longstreet and Cooper [85] proposed a meta-model for developing simulation games in higher education and professional development training. The player character takes on the function of the protagonist in the role of a SE student, who progresses through the game challenges acquiring rewards and/or penalties. There are also some game mechanics that can limit and/or motivate the player in their engagement with the challenges. These mechanics are timed and limited resources and competition with Non-Player Character (NPC) antagonist characters. The former teaches players to manage resources efficiently while the latter adds a layer of realism to this simulation.

What are the Learning mechanics implemented in SGs for SE?

In this question, we aim to investigate trends regarding the implementation of Learning mechanics within SGs for SE following the LM-GM framework [64]. Figure 14 shows that 40.62% of the adopted Learning mechanics in this study are not implemented in SGs in SE, and only 19 Learning mechanics were reported in analyzed SGs. Simulation takes the lead with a rate of 65.62%, Action/Task comes second, with a rate of 37.5%, followed by Instruction, which is adopted in 18.75% of the SGs in SE, Feedback and Reflect/Discuss rank 4th, with a share of 15.62% each, and 5th come Competition, Motivation, and Assessment, with a representation of 12.5% each over the selected studies. Similarly, to game-related approaches, SGs for SE overlooked Learning mechanics in their implementation, which highlights another significant gap obstructing the full advantage of these technologies in improving SGs in SE, especially in education and/or training settings.

Figure 14. Learning Mechanics of SGs in SE

Rodriguez et al. [77] measured the impact of ScrumGame on the learning experience and motivation of users. ScrumGame aims to assess Scrum concepts according to the knowledge level gained by the user along the game. It provides the player with a way to validate the concepts learned in an attractive and motivating way. The motivation here consists of all the internal determinants that stimulate actions and tasks to be carried out considering Scrum learning as the main aspect in addition to other soft skills such as presentation, communication, and teamwork. The authors also concluded that improving key concepts such as competition and anxiety management can provide better in-game psychological preparation thus improving furthermore the motivation value after playing the game.

Calderón et al. [86] have proposed a SG for project management training named ProDec. The latter was developed to provide an automatic assessment of the player's performance after gameplay. Simulating a software project execution and given that this SG is not a competitive one, it fostered reflection and communication among game players so as to work collaboratively to win the game.

RQ4. To what extent do SGs in SE take into consideration player profiles in their gameplay?

In this question, we aim to distinguish what kind of player profiles have been addressed in SGs for SE, through implemented game-related approaches.

We adopted the player profiles from Marczewski's taxonomy [87] according to which there are 6 player profiles. First, come Players who are the most willing to play among the rest of the player profiles. They will do what is needed for them to collect rewards from a system. They are in it for themselves. Second, come Socializers who are motivated by relatedness. They want to interact with others and create social connections. Then Free Spirits are motivated by autonomy and self-expression. They want to create and explore. In the fourth place are Achievers who are motivated by mastery. They are looking to learn new things and improve themselves, but also want challenges to overcome. Then Philanthropists who are motivated by purpose and meaning. This group is altruistic, wanting to give to other people and enrich the lives of others in some way with no expectation of reward. While these last 4 profiles are less willing to play than the Player one, the player profile who is not willing to play is the Disruptor. In general, Disruptors want to disrupt your system, either directly or through other users to force positive or negative change.

As can be seen in Figure 15, the Achiever profile ranks first with a rate of 48.64%, followed by Player by a slight difference of 1.35%, while Socializer comes 3rd, with a rate of 10.81%, followed by both Free Spirit and Disruptor with an equal share of 9.45%, and last comes Philanthropist, with no representation among selected SGs. Based on the Klock et al. literature review [88], just like player profiles from Bartle's taxonomy, player profiles from the Hexad typology are not mutually exclusive and therefore each user is a composition of player types, and one or some of them are usually predominant. Moreover, according to the study by Echeverría and Jurado [89], the topmost three reasonable player profiles within an educational context would be Achiever, Explorer, and Socializer, while the least likely player profile in this same context is the Killer. Looking into the characteristics of the Killer player profile in Bartle's Taxonomy [90], it corresponds to the Disruptor player profile in our study, which explains the low coverage of this player profile among SGs in SE. The Free Spirit player profile in our study can be seen as the same as Bartle's Explorer [90], however, Marczewski [87] amplifies that players with Free Spirit profile both like to explore and create. It explains the low coverage of this profile among SGs in SE despite the fact that the Explorer profile is one of the predominant profiles in educational contexts and highlights the lack of induced creativity in SGs in SE that tended to preclude such low coverage. Also, regarding the Player profile in our study, the fact that the later ranked 2nd in our study just behind the Achiever profile may have to do with Marczewski [87] loosening up the players' profiles to the point that a Player might at the same time be an Achiever if that helps him get rewarded.

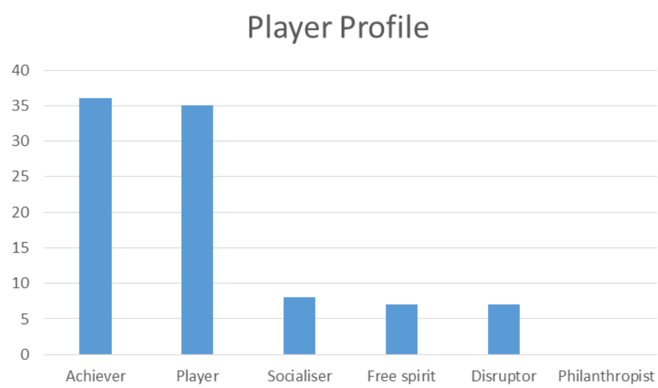


Figure 15: Player Profiles of SGs in SE

According to these results, and although SGs can support multiple players' profiles at once, we can still conclude that most SGs in SE heavily target players with the profile of Achiever, and Player, while other profiles can be underrepresented or not represented at all among these games. This is a limitation of SGs in SE as developed SGs will not be comprehensive when it comes to the variety of player profiles that they support. Given that these users' predominant player profiles may not always be among most aforementioned targeted player profiles, many players can find the game uninteresting or fail in the game for the simple reason that their predominant player profile is not supported in that game. Considering that the predominant player profile is indeed one of the targeted profiles in these SGs, as stated by Klock et al. [88], suggesting game elements based solely on a dominant player type, is likely to induce misunderstandings and inaccuracy.

4.3.4 Main findings

In this section, we provide more in-depth analysis results, and we describe the main gaps in the literature alongside highlighting our contribution and the limitations of our study. We also prepared Table 11 in Appendix B, which provides a comprehensive list of the SGs identified in our study with insightful information that can help SE students, practitioners, and teachers identify the SGs that may serve best their interests.

SWEBOK Areas

Our mapping study revealed that there is a growing trend to use SGs in SE. This trend extends to covering more SWEBOK areas throughout the years. The Bubble chart in Figure 16 shows the growth of these different knowledge areas covered by SGs in SE over time. The first SWEBOK area tackled by SGs in SE is Software Engineering Management which goes back to the year 2000. From the mid-2000s to 2010, only three studies were invested in SGs in SE, and each of them covered one distinct SWEBOK area. These covered areas were Software Engineering Process, Software Engineering Management, and Software Design respectively. It wasn't until early 2010 that a growing interest in SGs in SE occurred and this time, in addition to Software Engineering Process and Software Engineering Management, more SWEBOK areas such as Software Quality and

Software Testing were approached. However, until 2015, Software Engineering Process and Software Engineering Management kept being the areas most covered by SGs in SE. From 2015 upward, more attention was brought to SGs in SE, and far more SWEBOK areas were covered by these trending studies. To this date, Software Engineering Management, alongside Computing Foundations and Software Engineering Professional Practice keep being the most invested in SWEBOK areas by studies in SGs in SE.

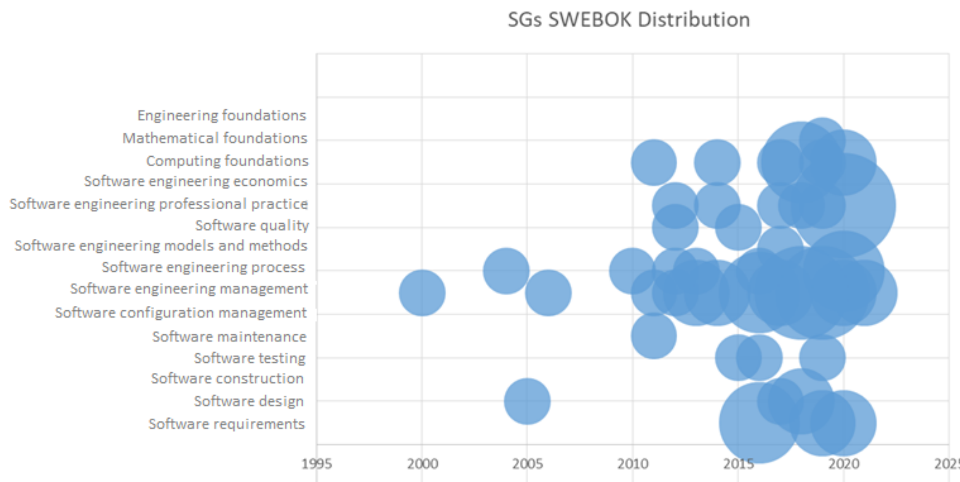


Figure 16: SGs SWEBOK Areas Yearly Distribution

Bloom's Taxonomy

Knowing how to use Bloom's taxonomy to compare the knowledge and skills acquired in SE courses to the breakdown of SWEBOK's areas is highly needed [91]. In this light, although games can help students to strengthen the theoretical concepts learned by traditional teaching, according to an SLR by Garcia et al. [92], most games establish learning objectives by considering the first rung of Bloom's taxonomy. However, according to our study, and as previously illustrated in Figure 11, Remembering was the penultimate learning levels among studied SGs in SE, while the higher rungs of Bloom's taxonomy were more addressed by these studies such as Applying and Understanding followed by evaluating and analyzing.

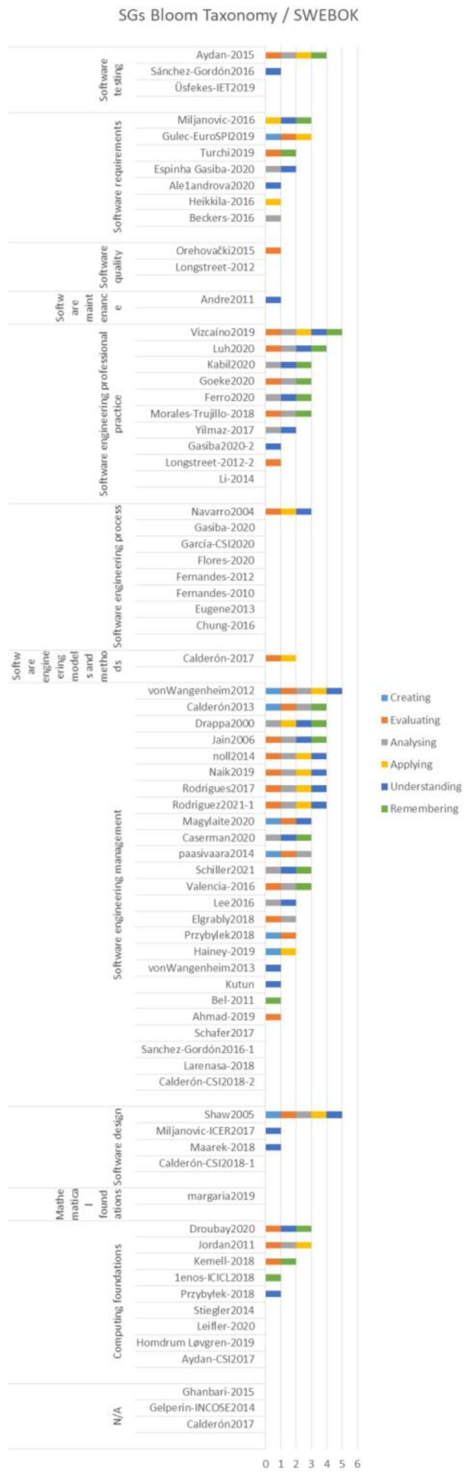


Figure 17: SGs Bloom Taxonomy' SWEBOK Distribution

As illustrated in Figure 17 above, no SG combines all learning levels of Bloom's taxonomy. Only three SGs in three distinct SWEBOK areas tackled five out of the six learning levels of Bloom's taxonomy. The corresponding SWEBOK areas of these SGs are Software Engineering Professional Practice, Software Engineering Management, and Software Design, which happen to be knowledge areas in which over 75% of found SGs address at least one learning level. Moreover, those knowledge areas rank right after Software Requirement which is the only one in which all SGs within its scope cover at least one learning level each. Furthermore, Software Quality appears to be the SWEBOK area with the least coverage of Bloom's taxonomy. The Survey results by Alarcón et al. [93] show that the latter is the most difficult and complex knowledge area among SWEBOK areas. This can explain the low number of SGs found under this knowledge area and the fact that only one of the two reported SGs tackled one Bloom's taxonomy only, and that the Bloom level in question was Evaluating, which is a high level of Bloom's taxonomy and thereby conforms to the complexity of this knowledge area. Therefore, it is a critical and urgent gap within SGs in SE, which requires further attention.

SGs' Key Elements, Mechanics, and Dynamics

Since investigating and modeling methods and tools for an effective and thorough infusion of learning pedagogy inside SGs through pedagogical theories and approaches is highly encouraged [94], Lim et al. [64] designed the LM-GM framework, which provides insights into the pedagogical and gaming patterns of a game. This framework addresses the process of identifying Game mechanics that pair well with particular Learning mechanics in the creation of digital SGs based on their common Bloom's classification [95]. Since SGs can be considered a continuous assessment of gained knowledge as the player proceeds from level to level, the use of such taxonomy to link commonly found Game mechanics and Learning mechanics is logical. Following the same logic, Pendleton and Okolica [62] mapped ideal Game dynamics and Game mechanics, targeting new game designers to empower them to have a much greater level of effectiveness in designing SGs with playable mechanics and well-integrated learning objectives. From a different perspective, Werbach and Hunter [82] proposed in the six steps to gamification (6D) taxonomy a hierarchy of game elements using Dynamics, Mechanics, and Components. In this classification, the top of the hierarchy is composed of dynamics, which are the abstractions related to the task being gamified. These dynamics are used to create the motivation to perform the task and manifest through mechanics. Mechanics are the processes used to drive user actions and are presented through components. Finally, these components are extrinsic rewards and feedback features. This presented taxonomy, however, does not provide clear strategies on how to properly combine these elements [96]. At the same time, we have the GAME framework by Marczewski [65], which provides an extensive periodic table of 52 gamification elements divided by 6 player profiles. Within this taxonomy, these gamification elements may work better depending on the users' player profile. However, gamification frameworks nowadays are no longer a novelty.

Our findings cast a new light on gamification elements, game mechanics, learning mechanics, game dynamics, and the different roles they play in the SGs sample we analyzed. There are studies focused on reimagining the classroom or lesson layout, where game elements promote engagement and fun to keep students motivated to learn. While other studies focus on gamifying the learning topic itself, inducing students to apply or develop new skills to achieve better performance at the meta-game or to move the narrative forward, therefore promoting a change in the learning.

As shown in Figure 18 bellow, the gamification elements and Game mechanics were more recurrent in our sample, while the Game dynamics and Learning mechanics were under-investigated which makes them good opportunities for future research on SGs in SE education and/or training.

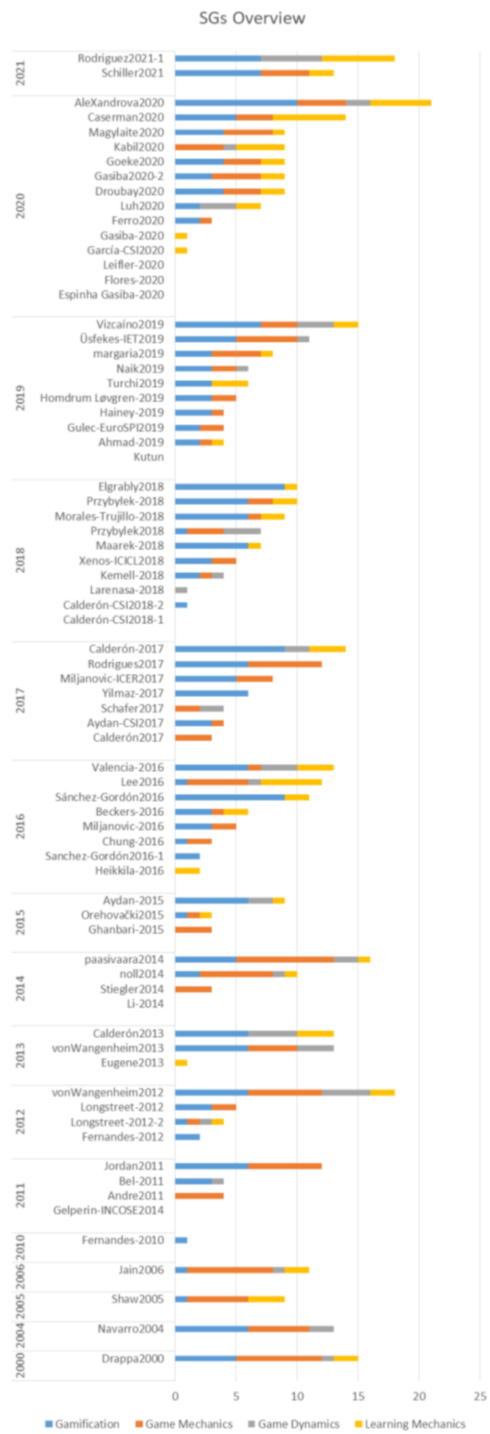


Figure 18: SGs Elements' Yearly Distribution

Player Profiles

In accordance with the results showcased in the previous section, and although SGs can support multiple players' profiles at once, Figure 19 illustrates how most SGs in SE heavily target players with the profiles of Achiever and Player both combined or separately. While only 9.45% of SGs tackle 3 player profiles at once, and nearly 5.40% target 4 player profiles at once. Still, the Philanthropist player profile was absent in all SGs studied.

In previous work, Gustavo et al. [97] analyzed the correlations between the participants' scores in each of the Hexad user [87] with their scores on each Big Five personality trait [98], and they found positive correlations between all Hexad user types with the expected game design elements, except Philanthropist. In light of these results, this study stressed the need for further studies regarding the Philanthropist user type not only for SGs in SE but in game design in general. The same data by Gustavo et al. [97] suggest that the four user types based on intrinsic motivation – Philanthropist, Socializer, Free Spirit, and Achiever – are similarly common as the main user types with a distribution percentage of 24%, 19%, 22%, and 23% respectively. The Player type was reported to be half as common as any of the intrinsic types, making a percentage of 10% among this study's participants. Finally, the Disruptor type was reported to be the least common if not absent as the main user type with a poor distribution percentage of 1%. However, according to our study, Achiever and Player are the most common player profiles in SGs in SE, followed by Socializer, Distributor, and Free Spirit.

The contrast between our study results and Gustavo et al. [97] suggests that the user's types distribution for players in SE field is different than the one of general players proposed by the former study. Such difference in player's type distribution between these studies can be due to some unique traits of players in SE field that distinguish them from the general players. In this regard, further research should be carried out to investigate in-depth player profiles for SGs in SE.

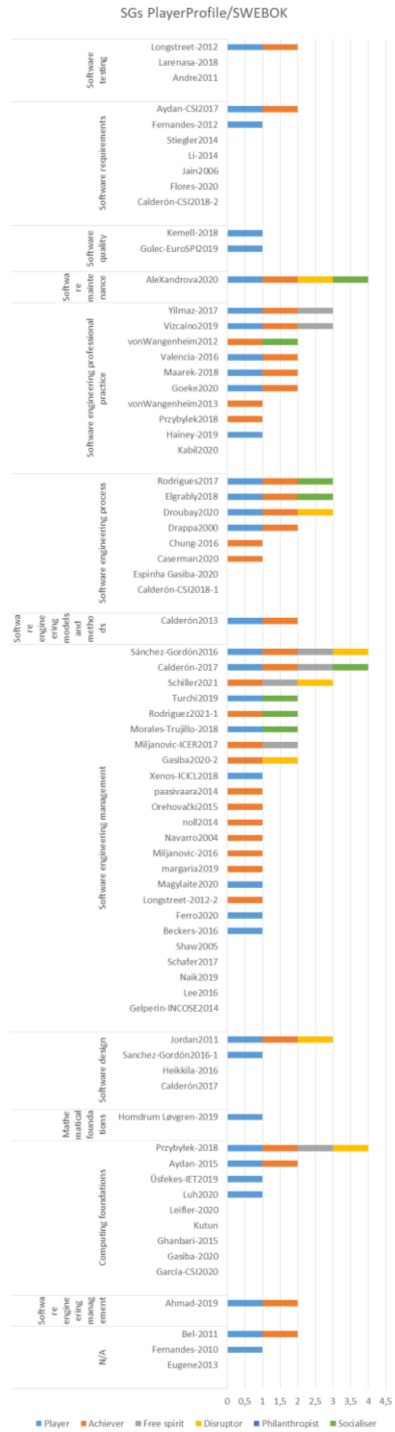


Figure 19: SGs Player Profiles' SWEBOK Distribution

4.3.5 *Success factors and challenges for future research*

This section isolates the success factors of SGs that had an encouraging impact on the gameful SE learning experience and compare the finding with the shared SGs success factors previously discussed by Ravyse et al. [99]. Additionally, we discuss challenges observed in the selected studies and address the open issues highlighted in the previous section, proposing valuable areas for future research.

There is an abundance of factors that make learning with SGs successful [100]. However, research papers dealing with these factors tend to focus on selected elements of SGs and do not combine all of the salient factors for successful learning with SGs. The study by Ravyse et al. [99], analyzed existing academic literature from 2000 to 2015, and extracted shared SG success factors that have had an encouraging impact on gameful learning experiences to reveal 5 core SG themes, that require thoughtful deliberate intertwining with pedagogical content to ensure the successful learning of SGs players. (1) Back story and production, refers to the storyline and game-world players encounter and expectantly immerse themselves in as they play the game. (2) Realism, stipulates how close a game under scrutiny replicates or resembles real life. (3) Feedback & Debriefing, which refer to both: in-game feedback experience through a variety of in-game rewards and NPC interactions, and post-game debriefing and reflection sessions which ultimately elucidate the learning material and place the game-learning experience into a greater context. (4) AI & Adaptability, refers to what is broadly considered as an unscripted game response to player activity, and they influence SGs on two fronts: (a) adjustment within the game through agents; and (b) adjustment of the game itself by means of adaptivity. (5) Interaction, which sets SGs apart from other forms of edutainment by requiring user input and responding accordingly within the gameplay, in turn instigating the next player action and continuing in a repeated player game feedback loop.

As illustrated in Figure 20, Backstory & Production is the top-rated success factor among SGs in SE with a rate of 45.94%, second come both Realism and Feedback & Debriefing, with an equal share of 28.37%, followed by AI & Adaptability, with a rate of 22.97%, and last Interaction that was reported in 14.86% of SGs in SE studies.

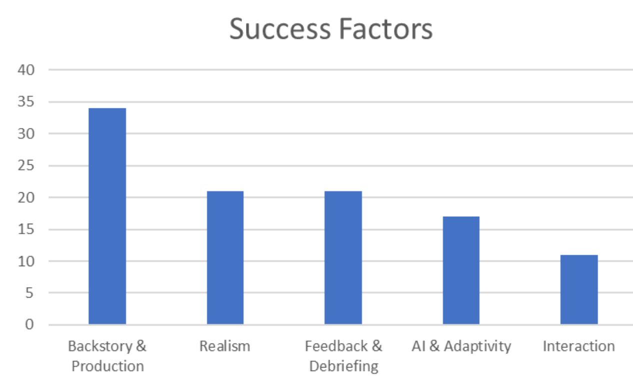


Figure 20: Success Factors of SGs in SE

In the study by Schiller et al. [116], ten user experience tests were conducted during the implementation of the user-centered game called ENC#YPTED. The early involvement of potential users during the development process was deemed an important factor in the successful implementation of the SG due to user feedback. Aside from the early involvement of potential users, and the call for further research to determine and possibly optimize the actual learning success of the SG, several SG's success factors were identified according to Ravyse et al. [99] adopted taxonomy. First comes Realism on which the design concept is based to enhance immersion and control the attention of users during the game. To make the SG even more immersive, AI was used to induce a psychological conflict in the SG. In this vein, two game characters were developed. One of them is an NPC impersonating the pirate and the other one embodied the player. The application was iteratively improved and refined according to user-centered development through user experience testing, proving that Adaptation creates a high level of usability. Last, Backstory & Production were also stated in this study, and all respondents to the UX tests enthusiastically stated that they found the application to be an exciting simulation due to its dystopian story.

Through our analysis, we revealed that no SG in SE combines all five success factors by Ravyse et al. [100]. Moreover, we reported that the most implemented success factors among SGs in SE are Backstory and Production, followed by both Realism and Feedback & Debriefing. Third rank AI & Adaptability, followed by Interaction.

Although the Interaction factor is what sets SGs apart from other forms of edutainment, it is the success factor with the least representation over SGs in SE according to our results, which makes it a real gap that needs further investigation, especially since a recent study by Dimitriadou et al. [101] stress the need of SGs to be interactive in order to be engaging and provide "individualized feedback for learners". In the same vein, the study by Iten and Petko [102] concludes that the key to success when teaching and learning with SGs, is the combination of 'fun' and 'engagement', which covers not only emotional engagement but also behavioral and cognitive engagement. However, reference was made to the need for strong data to support success when it comes to educational SGs [101]. In this vein, future studies should address closely the success factors of SGs in SE, validate them with support data, and investigate if there is any particularity about the success factors of SGs in SE that distinguish them from the ones of other SGs. In the light of this statement by Dimitriadou et al. [101], Daoudi et al. [103] proposed a success-oriented model based on the emotional states of learners and different features of the SG, which consists of Success Indicators and Success Factors. These indicators are supposed to indicate the success degree of the game-based learning environment and are impacted by success factors. In the proposed simulator, the Success Factors will represent the input while the Success Indicators will represent the output. A multi-agent-based simulator, which would be able to predict the impact of operating a SG on classroom teaching was developed and tested. Findings show that the simulator gives results close to real feedback which helps teachers to predict the impact of adopting SGs in a particular learning process by analyzing the resulting learner's emotions. Further extension of the proposed simulator by considering other success indicators and factors was proposed for future work in this paper. Empirical studies supported by the simulator could be performed to determine the most impactful success factors, which could drive the design of the new generation of SG in SE. Given the strong data support of the former study, the Success Indicators and Success Factors proposed may thereby be more interesting than the success factors proposed by Ravyse et al. [100] for future work.

Aside from being one of the success factors of SGs, AI was reported to be an important component of game technology if not the core of SGs [104]. Still, despite the potential of AI, our study found that only 17.56% of SGs in SE support AI within their gameplay. This AI integration mainly covered three key components: Player Experience Modelling (PEM), Natural Language

Processing (NLP), and advanced NPC modeling. These areas are considered to be among the finest and biggest game AI research according to the study by Westera et al. [105]. This same study suggests that the relevance of the aforementioned AI components for SGs is easily explained by the pedagogical frame of teaching, which assumes a teaching agent (cf. NPC) that frequently probes and assesses the learner's mental states (cf. PEM) and, when needed, engages in supportive dialogue with the player (cf. NLP) to provide guidance or feedback.

According to our study, among the games that reported the use of AI, 76.92% do implement NPC within their gameplay to challenge players, assist game storytelling, and be one solution in multiplayer games while live opponents are not available. NLP was under-investigated among studies SGs as only 15.38% of these games do support NLP, given that pedagogical SGs still replicate the framework of traditional games, in which the trainer usually does not exist, these algorithms are generally used to support players' interactions with the game [106].. Last, none of the studied SGs in SE tackled PEM. This may be due to the ethical issues arising by Westera et al. [105], regarding creating detailed user profiles and having them analyzed with advanced algorithms that may uncover sensitive personal traits, behaviors, preferences, capabilities, and opinions.

However, aside from the aforementioned AI fields, from the wider domain of AI, various additional high-potential game AI concepts and technologies will continue to foster and innovate the domain of serious gaming, opening up many new possibilities. The Artificial Intelligence in Education (AIED) for instance, can make use of AI to promote instructive gameplay, manage the level of challenge of the user experience, provide scaffolding selectively where needed, and support learners in their efforts to reflect on their play and improve their skills [107]. By the same token, the Dynamic Difficulty Adjustment (DDA) seeks to adapt the challenge a game poses to a human player and has shown that when implemented within SGs, can substantially decrease the time a human player has to wait for suitable game levels [108] which affect positively the user experience and th(e learning curve of the player. Furthermore, various additional high-potential game AI areas can upscale SGs in SE and deserve further attention. Yannakakis and Togelius [109] identified a few which include procedural content generation that allows creating design-tailored game content, the computational narrative that helps optimize procedures for game storytelling, event generation, and generating sequences of game events, and AI-assisted game design to support creative game design and the development of SGs.

4.3.6 Recommendations Based on Findings

In light of our comprehensive analysis of existing SGs in SE, we present practical recommendations to guide the development of future games. These insights aim to enhance the effectiveness and engagement of SGs tailored to both educational and industrial settings in SE.

1. Contextual Adaptation: While many SGs are designed for SE students within a university context, there is a need for games tailored to SE practitioners in the industry. These games should consider the unique constraints and interests of professionals, such as time limitations and specific industry needs.

2. Assessment of Importance and Impact: Not all underrepresented SWEBOK areas necessarily require more games. According to the same logic, even if certain SWEBOK areas appear saturated with games, they may not adequately address the industry's requirements. Therefore, it is crucial to evaluate the significance and impact of each SWEBOK area within both educational and industrial settings before developing new games. Our mapping study highlighted that areas such as Software Quality, Software Testing, and Software Design, despite some coverage, still present significant opportunities for further development of SGs. Based on the growing trend and our findings, developers should focus on creating SGs for these underrepresented yet vital SWEBOK areas. This effort could help bridge specific gaps in both educational curricula and industry requirements. Additionally, developers should consider which levels of Bloom's taxonomy are most appropriate for each setting to avoid creating overly complex games that might overwhelm users. For example, foundational knowledge might be targeted with simpler SGs, while advanced topics could leverage more complex game mechanics.

3. Generational Differences and Game Familiarity: When designing SGs for industry professionals, it is important to account for generational differences in gaming experience. Many senior practitioners might not be as familiar with video games as younger students. Therefore, games targeting industry professionals should have simple, intuitive gameplay to ensure accessibility and avoid intimidating less experienced players.

4. Accurate Representation of Player Profiles: Our study suggests certain trends in targeted player profiles, but these may not accurately reflect the actual demographics within SE. Most SGs in this field were designed intuitively, without a formal framework incorporating player profiles. We recommend conducting player profile tests with a sample of the target audience to better tailor game mechanics to their preferences and needs. When reflecting the results of personality tests in SE from the study conducted by Gulati et al. [110] on Marczewski's taxonomy, it becomes apparent that specific player profiles align well with certain roles within the SE field. Based on the personality traits and player profiles, we recommend the following alignments for various SE jobs: (1) For designers and developers, who predominantly exhibit ISTJ (Introverted, Sensing, Thinking, Judging) and INTJ (Introverted, Intuitive, Thinking, Judging) personality traits, the recommended player profiles are Achievers and Free Spirits. Designers and developers often demonstrate a strong motivation for mastering their craft, overcoming challenges, and continuous self-improvement, which aligns closely with the Achiever profile. Additionally, these roles frequently require a high degree of creativity and autonomy, particularly among developers who value the freedom to explore and innovate. Thus, the Free Spirit profile is also highly suitable. (2) Analysts, characterized by Extrovert NT (Intuition and Thinking), NF (Intuition and Feeling), and INTJ personality types, should align with the Socializer and Achiever player profiles. Analysts, especially those with Extrovert NT and NF traits, thrive on interactions and collaboration, fitting well with the Socializer profile. The INTJ analysts, being detail-oriented and driven by a desire for mastery, align with the Achiever profile, as they strive to enhance their analytical skills and tackle complex problems. (3) For architects and testers, who exhibit high extroversion and conscientiousness, the recommended player profiles are Socializers and Achievers. Architects often engage with various stakeholders and work collaboratively, making the Socializer profile appropriate. Testers, who need to be meticulous and aim for high-quality standards, align well

with the Achiever's motivation for mastery, as their roles require detailed attention and a commitment to excellence. (4) Presenters, who display high agreeableness, are best suited to the Philanthropist player profile. Presenters are often motivated by helping others and creating positive experiences, which matches the altruistic nature of the Philanthropist profile. Their cooperative and friendly disposition makes them well-suited to roles that involve enriching others' lives without the expectation of personal gain. (5) Pair programmers, characterized by open-mindedness, should align with the Free Spirit player profile. The nature of pair programming benefits from individuals who are open to new ideas, flexible, and innovative in their approach. This aligns with the Free Spirit profile, which values autonomy, creativity, and self-expression. In addition to the specific roles, certain player profiles can be beneficial across various SE roles. For instance, Disruptors, although not explicitly linked to specific roles, can drive innovation and positive change in research and development or roles focused on improving existing systems. Additionally, roles incorporating gamified elements or direct rewards might attract Player profiles, even though these roles were not specifically highlighted in the personality test results.

5. Utilizing Formal Frameworks: We advocate for the use of formal frameworks in SG design. In our recent work [111], we proposed a framework that guides educators and game experts through a collaborative process to design SGs. This framework ensures the incorporation of appropriate learning and game mechanics for the intended Bloom levels and aligns the game with the target player profiles, thereby enhancing both engagement and effectiveness.

4.4 Validation

Our study applies the Resonance Scheme [112] to systematically analyze the impacts and contributions across several dimensions. This methodological approach, as detailed in Table 4, categorizes our study's contributions distinct categories. These include Synthesizing (SYN), which involves summarizing and organizing existing literature by time trends (SYN-T), venues (SYN-V), and classification of existing frameworks (SYN-C) to establish an order and transparently relate research contributions to each other, our study also employs Aggregating Evidence (AE) to extract and evaluate empirical evidence, focusing on the soundness of empirical study conduct (AE-I). Additionally, our study contributes to Theory Building (TB) by introducing categorization frameworks (TB-C), and proposing new model constructs (TB-M). We also highlight Identifying Research Gaps (RG), emphasizing forward-oriented knowledge development by pinpointing domain voids (RG-D) and deficiencies in empirical evidence (RG-E) that warrant further investigation. Moreover, our Research Agenda (RA) offers actionable recommendations for closing research gaps (RA-S), supporting strategic decisions in SE, and suggesting specific empirical settings (RA-E) to guide future research. Finally, we engage in Criticizing (CRI) by identifying logical or conceptual (CRI-C) issues, and problematizing assumptions (CRI-A) to foster rigorous advancement within the SGs in SE research domain. Each category aligns with components of the Resonance Scheme, validated through consultation with the schema's authors to ensure accuracy and reliability in characterizing our study's contributions.

SYN	AE	TB	RG	CRI	RA
SYN-T, SYN-C, SYN-V	AE-I	TB-C,TB-M	RG-E, RG-D	CRI-A, CRI-C	RA-S, RA-E

Table 4: The Expected Resonance

4.5 Our Contribution

In this section, we highlight the distinctive contributions our study brings to the understanding of SGs in the realm of SE. Our analysis has revealed several key aspects that hold significant implications for educators, students, and SG designers alike. The essence of our findings and their potential impacts on these crucial stakeholders are encapsulated as follow:

- Crossed Classification of SG in SE based on Bloom's Taxonomy Levels and Learning Mechanics: Educators gain insights into the educational levels targeted by each SG, helping them match games with specific learning goals. Students benefit from a more tailored educational experience, and SG designers enhance their ability to align game elements with desired learning outcomes.
- Derived player profiles classification of SG in SE based on game-related approaches: Educators can identify SGs that cater to diverse player profiles. Students can identify SGs that incorporate their player profile for a better personalized and more immersive gaming experiences, and SG designers can receive guidance on player preferences and characteristics for future game development.
- Proposal of Success Factors for SGs in SE: Educators gain insights into the factors contributing to the effectiveness of SGs in SE education, enabling them to optimize teaching strategies. SG designers learn about domain-specific success factors, guiding them in creating more impactful games.
- SGs in SE catalog as a Resource: Educators receive a practical resource in the form of detailed tables, facilitating the selection of SGs based on specific learning objectives in SE courses.

Overall, the study addresses a gap in the existing literature and presents novel insights that even the authors of the mapped SGs may not have been previously aware of, ensuring that SGs in SE are systematically examined and categorized for the benefit of educators, students, and SG designers. The insights from our study not only categorize existing SGs but also lay a foundation for future innovations. The section preceding offers practical guidance for SG designers to adapt games to industry/education needs, assess SWEBOK area relevance, consider generational differences in gaming experience, accurately represent player profiles, and utilize formal frameworks for effective game design. This approach ensures that our study’s contributions are comprehensive and forward-looking, bridging the gap between current understanding and future advancements in SGs tailored for SE education and industry applications.

4.6 Threats of Validity

This section presents the measures taken to mitigate the threats to validity. We mitigated these threats concerning the construct, internal and external, and conclusion threats [113].

Construct validity. This threat concerns the measures that ensure we followed proper operational instructions that support the statements in which research questions are investigated. It consists of mitigating the risks that could prevent our SMS from investigating what it claims to be investigating. The inclusion of improper terms in the search string we used in the automatic search may cause the exclusion of some studies. We followed the guidelines described in [59,61,67] for constructing the appropriate Search terms This measure also attenuates a gap between the keywords of potential studies and our search string. We also consider an extensive list of search engines related to the field of research, to mitigate the risk of venue inaccessibility. Furthermore, non-peer-reviewed materials, or the so-called gray literature—such as patent specifications and technical reports, which are typically absent from scientific review processes but sometimes included in multivocal systematic literature reviews—were excluded from the selection process. We followed a rigorous search protocol for systematic mapping studies to maximize the inclusion of relevant papers and avoid including studies [61]. These guidelines define the creation of inclusion and exclusion criteria and a multi-step filtering process for selecting primary studies. We added to this process the snowballing methodology to increase the coverage of our search and minimize the possibility of a relevant paper not being indexed in the search engines employed in our SMS. According to Petersen et al. [59] and Kitchenham and Charters [61]the combination of manual with automatic searches based on keywords helps cover possible gaps present in the search string terms. The usage of this rigorous and systematic methodology also mitigates any threat concerning the possibility of the definition of improper inclusion and exclusion criteria and the adoption of an incorrect search method. However, we acknowledge a potential threat to construct validity in our search strategy: the decision not to include broader synonyms for "serious game" such as "game" was made to avoid an unmanageable number of results. Our validation process showed that including the term "game" resulted in a substantial number of irrelevant papers focused on game design, game development, game theory, or general gaming—topics far removed from our research objectives. This significant noise (98.96% irrelevant papers) demonstrated that including "game" in the search string did not provide a meaningful benefit but instead introduced substantial inefficiencies. Therefore, we concluded that sticking with "serious games" alone is a more efficient and effective strategy for our study, ensuring a focused and manageable dataset while still capturing the most relevant research. Nonetheless, it is possible that some relevant papers were missed. This is a common threat in systematic reviews and has been considered in our validity assessment. Although we have not reviewed all the additional articles, we believe that the conclusions would likely remain the same or similar. Certainly, while not all additional articles have been reviewed, it is anticipated that the conclusions would likely align or maintain similarity. It should be noted that a replication of this secondary study, feasible only within a specialized domain with expert-led investigations, did not yield identical papers. However, the overarching conclusions drawn remained consistent [114].

Internal validity. This category of threat mitigates whether the data used to ascertain conclusions are internally valid. Thus, reducing the risks of deriving results from inadequate data. As a measure, only one author conducted the selection and filtering process, and two other authors assisted in decision-making when doubts arose about the application of the inclusion and exclusion criteria. During the studies' selection, in the case neither the title nor the abstract present enough details to consider the study for inclusion in the next step, the author responsible for the filtering process also read the integrity of the articles. This conduction can threaten the validity of the inclusion of studies due to the authors' subjectivity in the primary studies' selection. To mitigate this problem, and standardize the application of the criteria, a round of cross-validation was carried out with two other authors. Moreover, the selection was guided by rigorously defined inclusion and exclusion criteria. To reduce the errors produced by the fatigue effect, two hours was established as the time limit for holding the review sessions. Another threat to internal validity is the presence of duplicate studies. The author in charge of the filtering process identified duplicated studies using the JabRef tool. This tool matches studies in duplicate by title. In the case of matching results, the author removed the duplicated ones after checking the abstracts of both studies. Subjectivity could be another threat happening when planning and executing due to the length of this study. As we were so immersed in the work that objectivity could be thought of as an issue. To mitigate this, we used Petersen et al. best practices checklist [59] to check our work as illustrated in Table 10 from Appendix A.

External validity. This category of threat deals with risks concerning the generalization of the findings of this study [113]. For instance, the inaccessibility of papers could jeopardize the representability of our sample. The university guaranteed access to the articles used in this research, turning the paper inaccessibility a reduced risk. Another threat to external validity is the range of years that the inclusion criteria considered to select potential study candidates. The upper and lower bounds could lead to the exclusion of relevant studies. We did not define any time constraint to retrieve potentially relevant study candidates in search engines.

Conclusion validity. The main conclusion validity threat is the data collection. Since we do not know how the digital libraries' search engines work, we run the risk of getting different results for each search (even because libraries can index new articles daily). Therefore, we ran the search string and, to eliminate the possibility of changes to the list of papers returned by the digital libraries, stored the returned studies in a bibliography management tool for later analysis and data extraction. To mitigate the issue of data extraction, we decomposed the research question and structured a spreadsheet workbook as a form to receive the data necessary to answer the research question, as recommended in [59]. In this way, we know precisely what we want to extract from the papers and how to store the extracted data in an organized way. Classifying and analyzing the selected studies correctly was a constant concern during this research. For this, we dedicated a significant effort to examine the relevant studies. After the analysis, we focused on avoiding the fishing problem as recommended in [115]. Consequently, we did not assert conclusions before analyzing the results. For this, we formulated the study conclusions only after examining the results we collected.

4.7 Conclusion

This chapter presents a comprehensive mapping study focusing on SGs in the realm of SE. Notably, no prior mapping study has specifically explored the same concerning SGs in the field of SE. Through a systematic analysis of the included papers, this study presents novel insights that even the authors of the mapped SGs may not have been previously aware of. For each SG, we deduced the levels of Bloom's Taxonomy based on the covered gamification elements and game mechanics and identified the learning mechanics by analyzing the incorporated game dynamics. Furthermore, the study provides a comprehensive list of targeted player profiles for each SG, based on the gamification elements and game mechanics utilized. Notably, the success factors for SGs in SE were found to demonstrate distinct tendencies compared to general SGs, indicating unique characteristics that contribute to their effectiveness as educational tools within the SE domain. In addition to the research findings, this chapter offers detailed tables listing all the identified SGs, which can serve as a valuable resource for educators seeking to utilize SGs for specific learning objectives in their courses. The challenges identified in this study provide valuable insights for developers aiming to create effective SGs tailored for SE education. Overall, this mapping study contributes to the understanding of SGs in SE, shedding light on their design, alignment with educational objectives, and potential impact on learners. The findings open avenues for further research and innovation in the integration of SGs in SE education, fostering a more effective and engaging learning experience for students. For future research on SGs in SE, a greater presence in journals should be considered and more attention should be paid to AI, player profiles, and Bloom’s taxonomy. More evaluation research should be carried out in order to validate the perceived benefits of these SGs, and evaluate the success factors of SGs in SE. Moreover, given the lack of a systematic approach to gamifying SE education, we believe that further research should investigate a systematic methodology for incorporating gamification and SGs into SE education. Furthermore, since no statistics were reported in the identified papers regarding the popularity of SGs in SE among different genders, further research should investigate whether or not women are less interested in SE’s SGs.

5

CHAPTER 2: Impact of Instructor Expertise with Serious Games on Software Engineering Student Performance

5.1 Introduction

Investigating the benefits of SGs in education, a previous study identified Problem-Based Learning (PBL), cooperative learning, realism and immersion, motivation and competition, and interaction and feedback as their main benefits [14]. Another study claimed that SGs have a positive impact on players' skills development [116]. A more niche-down study in SE education highlighted that the use of SGs had benefits for both students and educators. It helped in increasing students' scores and incorporating novelty in educators' teaching approaches [25]. Furthermore, in the context of project management within SE, SGs provide a risk-free environment for learning, allowing students and practitioners to develop crucial skills, solve problems, and acquire new abilities [117]. As the SG market continues to expand, with the educational category gaining prominence, it is worth exploring the realm of SGs in the context of SE education, investigating their impact on both students and educators offering valuable insights into the potential of this teaching approach.

5.2 Methodology

This section describes the methodology used to design and carry out an experiment in SE education with undergraduate students as subjects. The aim of this study was to implement SGs in a SE course during its theoretical sessions and to determine the impact of this teaching strategy on students' final scores, exam attendance, chance of passing, and overall perception regarding SGs.

This study follows the principles used to perform experiments described in [115]. Moreover, it was planned through the GQM (Goal, Question, Metric) paradigm [118], which is composed of four steps: Planning; Definition; Data Collection; and Interpretation.

- Goal: to investigate the SGs' impact on students' academic results.
- Question: when implemented in a SE undergraduate course, does SGs usage improve students' academic achievement in a software project management course?
- Metric: students' academic data.
- Object of study: undergraduate SE students.
- Purpose: to evaluate the improvement of students' academic results.
- Focus: investigate the introduction of SGs in SE undergraduate course.
- Perspective: from a neutral and unbiased standpoint
- Context: at the undergraduate SE course level

5.2.1 Hypothesis

A set of SGs within a SE course was implemented to evaluate the impact that SGs have on students' academic achievement. The following hypotheses were investigated:

- H1. Students achieve better final scores when SGs are used.
- H2. The chance of passing is higher among students enrolled in courses that employed SGs compared to those enrolled in courses with a traditional approach.
- H3. The chance of attendance at the final exam is higher among students enrolled in courses that employed SGs compared to those enrolled in courses with a traditional approach.
- H4. Students express satisfaction with the use of SGs, irrespective of their teachers' experience in incorporating SGs into their lectures.

Hence, one dependent continuous variable (Test Score) for H1, one dependent dichotomous variable (Passing) for H2, one dependent dichotomous variable (Attending) for H3, and three independent dichotomous variables (SGs2013_14nSGs, SGs2021_22nSGs, and SGs2013_14SGs2021_22) for H1, H2 and H3 were defined.

Student satisfaction (H4) was assessed through teaching diaries. We asked the teaching group to write a diary at the end of each semester on their students' perceived satisfaction regarding the SGs used. In particular, professors asked students to critically discuss in the last 30 minutes of the last lecture whether the SGs helped to motivate the students, improved their learning or enhanced their learning experience. No standardized questions were elaborated. Professors asked students the same question: "what is your opinion on the use of SGs in lectures?" Students openly expressed their thoughts and experiences, while also providing professors with quality feedback. The teaching diaries were analyzed, and the feedback was assorted with respect to the model by Savi et al. [119].

5.2.2 *Participants*

Students enrolled in a SE course called "Project Management for Software Development" in a Bachelor of Computer Science Degree at the University of Murcia participated in the experiment. Their ages ranged from 21 to 23 years. This third-year undergraduate course describes and synthesizes practical aspects of software requirements and software development project management: creating project plans that address real-world management challenges, reaching a consensus on project goals and deliverables, assessing a project for strategic risks, choosing the Software Development Life Cycle that best meets an organization's needs, among others. The course consists of 14 weeks of lectures in the second semester of the academic year. Participants were the students who enrolled in this course and attended the theoretical classes. No specific training session was held to familiarize the students with the SGs that were used in the course, as the rules and motivation behind each of them were explained when introducing the SG in question in class. All participants were involved in the use of the SGs that were mandatory in the course.

5.2.3 *Teaching Intervention*

Our analysis is based on the results obtained in a SE course through the integration of SGs as an innovative educational tool. Two time points were selected to carry out the study with SGs. The academic year 2013-2014 was selected to analyze the impact of the SGs on the student's learning with inexperienced professors in teaching with SGs. The second time point was the academic year 2021-2022, once professors have gained increasing experience and expertise in the implementation of SGs throughout the years, which is a more stable scenery to mitigate the threats to validity. The use of these SGs was proactively adapted to the learning objectives of the course based on the perceived students' response and interaction with these SGs, and the observed results and limitations from previous courses. In addition, some SGs were added, removed, or replaced in the progress of this subject during the last 8 years to better match the students' needs, the class dynamic, and the learning objectives when changed. The activities with SGs used in this intervention practically filled the entire course dates. Some of the SGs were embedded in the course as in-class activities while others were flagged as extra activities. It should be noted that the latter were presented according to the course agenda, and remained open for students to choose whether or not to do them and decide when to do them throughout the duration of the course. Initially, the course coordinator incorporated different SGs in the course, each had different difficulty levels. The search for these SGs along with the analysis of their applicability in the subject and class environment was conducted in collaboration with students from the academic year 2012/2013, the year in which SGs were not yet used in this subject. Re-O-Poly, ReGo, and Guess What We Want were the SGs selected for in-class activities while Requirements Island and Software Quantum were selected to be conducted online from home as extra activities.

5.2.4 *Selected Serious Games*

The serious games used in the lectures are described below. It is worth mentioning that all were free of charge, and there was no need to pay for a subscription license to use them.

Re-O-Poly [120] is a board game based on the famous MonoPoly. It was designed to introduce and reinforce good practices in Requirements Engineering (RE). The goal of the game is to help students acquire the most important knowledge of RE, covering four clearly identifiable phases in RE: elicitation, analysis and verification, documentation, and change management. To play the game a minimum of 2 players and a maximum of 8 players are required. Each game will last approximately one hour, although if all the players agree, the game period could be extended. This SG is on-site, so materials such as notes, scenarios, and task cards must be printed in advance (Figure 21).

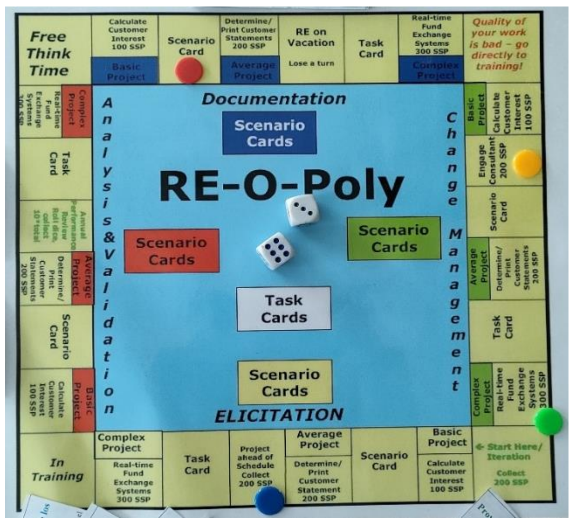


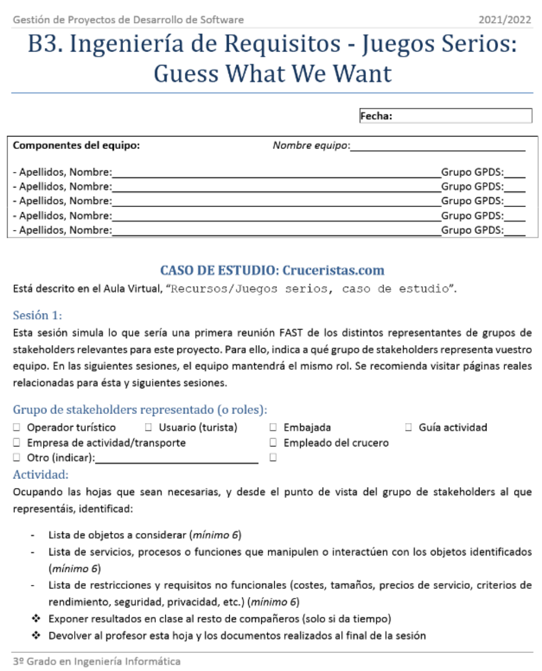
Figure 21: Materials of SG Re-O-Poly

ReGo [121] is a SG that could clearly be classified as a knowledge reinforcement group activity that can be applied to all areas of RE. Although this game was originally used in RE courses, it can be used to test students' knowledge of any discipline. Its objective is to evaluate students' knowledge at the beginning or the end of the academic year, carrying out a set of games using the theory topics to be taught throughout the course as the object of study. Thus, ReGo would be used on two occasions: (1) at the beginning of the academic year before acquiring any theoretical knowledge from the course, to assess their knowledge; (2) the game could also be used at the end of the course once all knowledge is given as a review of knowledge. As in the previous one, the SG is on-site and resources such as bingo cards must be delivered before the lecture (*Figure 22*). The students note down the answers to the questions in the test. Correct answers are in green, wrong answers are in red, and pending corrections are in black. The first student who has a column, row, or diagonal of correct answers signs “line!” The first student to get all the answers right signs “Bingo!” and wins.

1-a	2-d	9-b	7-b
4-a	16-b	6-d	11-a
3-c	5-d	15-a	13-d
8-b	14-a	12-b	10-b

Figure 22: A bingo card of ReGo SG

Guess What We Want [121] is an single-player game to consolidate knowledge of RE. It has been designed to help students understand the different levels of the hierarchy of requirements. In addition, the game tries to make students aware of the importance of having quality requirements specifications with a good level of detail to obtain satisfactory results in the software projects addressed. This game demonstrates that without having detailed and good-quality requirements, we cannot achieve a system that meets all the needs of the stakeholders. As in the two previous ones, the SG is played on-site. Therefore, wordings must be ready before starting the session (*Figure 23*).



Gestión de Proyectos de Desarrollo de Software 2021/2022

B3. Ingeniería de Requisitos - Juegos Serios: Guess What We Want

Fecha:

Componentes del equipo: *Nombre equipo:*__________

- Apellidos, Nombre:__________ Grupo GPDS:____
- Apellidos, Nombre:__________ Grupo GPDS:____
- Apellidos, Nombre:__________ Grupo GPDS:____
- Apellidos, Nombre:__________ Grupo GPDS:____
- Apellidos, Nombre:__________ Grupo GPDS:____

CASO DE ESTUDIO: Cruceristas.com

Está descrito en el Aula Virtual, "Recursos/Juegos serios, caso de estudio".

Sesión 1:

Esta sesión simula lo que sería una primera reunión FAST de los distintos representantes de grupos de stakeholders relevantes para este proyecto. Para ello, indica a qué grupo de stakeholders representa vuestro equipo. En las siguientes sesiones, el equipo mantendrá el mismo rol. Se recomienda visitar páginas reales relacionadas para ésta y siguientes sesiones.

Grupo de stakeholders representado (o roles):

☐ Operador turístico ☐ Usuario (turista) ☐ Embajada ☐ Guía actividad
☐ Empresa de actividad/transporte ☐ Empleado del crucero
☐ Otro (indicar):__________ ☐

Actividad:

Ocupando las hojas que sean necesarias, y desde el punto de vista del grupo de stakeholders al que representáis, identificad:

- Lista de objetos a considerar (*mínimo 6*)
- Lista de servicios, procesos o funciones que manipulen o interactúen con los objetos identificados (*mínimo 6*)
- Lista de restricciones y requisitos no funcionales (costes, tamaños, precios de servicio, criterios de rendimiento, seguridad, privacidad, etc.) (*mínimo 6*)

❖ Exponer resultados en clase al resto de compañeros (solo si da tiempo)
❖ Devolver al profesor esta hoja y los documentos realizados al final de la sesión

3º Grado en Ingeniería Informática

Figure 23: A wording of the Guess What We Want SG

Requirement Island [122] is an online game designed for single players to reinforce their knowledge of RE. Therefore, only one person will play, who will be in charge of solving the challenges proposed by the game in the role of Jack, a Requirements Engineer who has fallen on an island after a plane crash. On this island, the protagonist meets two local tribes that protect a Magic Stone with which they can only communicate through the use of RE techniques and tools. Jack must help the tribes to leave the island before the explosion of the island volcano. The game mainly consists of moving around a map clicking on the possible destinations and answering questions about the RE in order to advance in the game.

To access the game the following website must be visited: http://qgsoft.com.br/requirement_island. A screenshot of the game is shown in *Figure 24*.

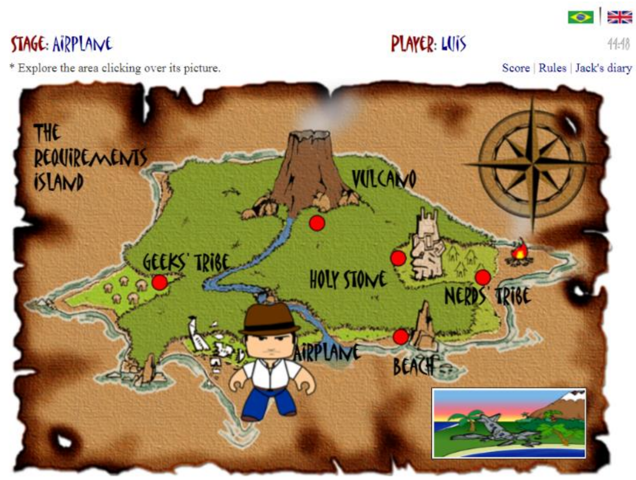


Figure 24: Map of the SG The Island of Requirements

Software Quantum [123] is an online SG that can only be accessed via the Web navigator as it does not have a web app for mobile devices nor a desktop one for computers. The game aims to make the player aware of the importance of planning the time spent in each of the development phases of a software development project. The game indicates that it is very important to dedicate more time to the early stages of the software project, that is, to the requirements analysis and software system design stages in order to obtain great customer satisfaction. Making the player see the importance of a rigorous requirements analysis with the subsequent agreement of the stakeholders is another objective of this SG. The requirements cannot be constantly modified nor can they be frozen in the early stages of development, but there must be an agreement on the main requirements of the application, which will be the success factors of the system. A screenshot of the SG is depicted in Figure 25.

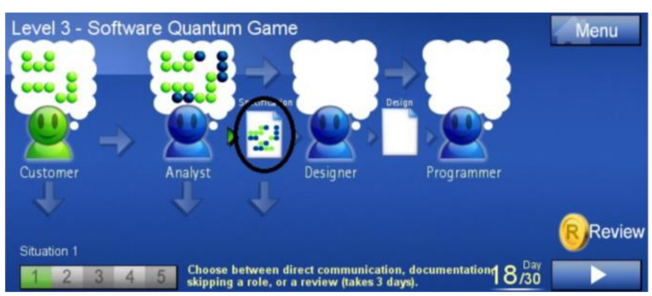


Figure 25: SG Software Quantum

5.3 *Results and Discussion*

This section displays an overview of the influence of the use of SGs in a SE course on students' achievement. Both the descriptive and the hypothesis results of the study are presented.

5.3.1 *Descriptive Results*

Table 5 summarizes and organizes the characteristics of the data collected from 101 students that were enrolled in a SE course during the years 2012/2013-2013/2014 and 2021/2022. The performance rate is calculated by dividing the number of students who passed the course by the number of students enrolled, while the success rate is calculated by dividing the number of students who passed the course by the number of students who took the exam. It was noticed that the average score from the courses in which SGs were implemented (7.090) is significantly higher than the average scores from the course in which no SG was implemented nor used (6.680).

Also, findings suggest that the overall success rate of students from all courses analyzed is slightly affected by the use of SGs and is on average 94.500% during the analyzed courses in which SGs were used compared to 92% when no SG was involved. This can be argued to be due to the fact that students who decide to take an exam, regardless of the examination call and the course in which it was taken, are highly likely well prepared to pass the exam.

Another promising finding was that the use of SGs was proven to improve the performance rate of students in the last year of its implementation in the subject, although it had a sudden drop in the first year when SGs were first implemented. The noticed drop in the course 2013-2014 is highly likely because the course was overwhelming with all the additional activities in which SGs were taking part so some students were not confident enough to take the exam in that course.

					ALL EXAMINATION CALLS					**1st CALL**		**2nd CALL**		**3rd CALL**	
Year	**Enrolled**	**Presented**	**Abandoned**	**Passed**	**Performance Rate**	**Success Rate**	**Scores' Standard**	**Scores' Median**	**Average Score**	**Performance Rate**	**Success Rate**	**Performance Rate**	**Success Rate**	**Performance Rate**	**Success Rate**
12-13	29	26	3	24	83%	92%	1.520	6.600	6.680	52%	65%	50%	70%	29%	67%
13-14	27	19	9	17	63%	89%	1.760	7.300	7.010	37%	63%	41%	78%	0%	0%
21-22	45	40	0	40	89%	100%	1.060	7.000	7.180	71%	91%	25%	100%		

Table 5: Descriptive Statistics for Students' Scores, Performance, and Success Rates With Regards to The Usage of SGs In a SE Undergraduate Course

While the improvement seen in the course 2021-2022 is likely due to teachers becoming better at introducing SGs in a way that matches the courses' schedules and learning objectives, and better tailoring these games to students' needs and learning dynamics. The present findings confirm the positive impact that the use of SGs in a SE course can have on both the students' success rate and performance rate when implemented adaptably to the course's terms and the students' needs. If we focus on students' academic results in each examination call in the course 2012-2013, in which no SGs were implemented, the course 2013-2014, in which SGs were first introduced, and the last course from the academic year 2021-2022, the improvement that the success rate had from the first examination call to the second during the years 2013-2014 and 2021-2022 in which SGs were used was twice as the one from the year 2012-2013, where no SGs were implemented. From the short review above, key findings imply that the implementation of SGs in the course provided an interactive ludic learning material that the students can still use after the end of the course to revise and prepare for their different examination calls. This helps students significantly increase their chances to pass the 2nd and 3rd examination calls when failing on the 1st attempt.

5.3.2 Hypothesis Testing Results

H1. According to our analysis, students' average scores in the course in which SGs were first introduced as a learning approach is 7.010, and the one in the last course after 8 years of expertise in SGs usage in this course is 7.180. Whilst the average score in the course in which no SGs were used is 6.680. We can notice that the students' average scores from the courses in which SGs are used are slightly higher than the one in which SGs were absent. A Mann-Whitney test indicated that this difference was statistically significant between the course in which no SGs were used and the one when they were first introduced as a learning approach ($U= 160$, $p=0.045$). A T-Test showed that this difference was not significant comparing the course in which no SGs were used with the latest course from the year 2021-2022 ($T\ (64)= 1.583$; $p=0.059$) although the p value was close to 0.05. Last, another Mann-Whitney test revealed that this difference was not statistically significant between the courses in which SGs are used ($U= 331$, $p=0.426$).

We can conclude that the courses in which SGs were used as a learning approach had a significant positive influence on students' average scores compared to the course in which only traditional teaching approaches were used. Also, given that no statistically significant difference was found comparing students' average scores from the courses in which SGs were used, this leads to the conclusion that regardless of the teaching group's experience with SGs implementation as a learning approach, courses using SGs still have a positive influence on students' average scores.

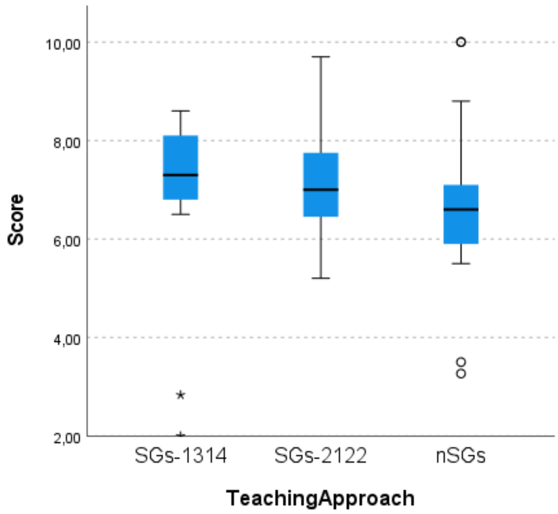


Figure 26: Boxplot of students' scores when and when not using SGs

Looking at Figure 26, this diagram shows that the median of the scores from the course with no SGs is low compared to the ones from the course in which SGs are used. Notice that the scores of the 2nd and 3rd quartiles from the SGs-based course are higher compared with the ones from the non-SG-based course. This leads to the conclusion that the distribution of 50% of the scores in the SG-based courses is higher than the one of the non-SG-based course.

H2. The odds ratio for passing the final exam comparing the years 2012-2013 and 2013-2014 (OR = 0.354; 95% CI 0.102-1.224) indicated that students from the course following a traditional approach and not exposed to SGs have a 2.824 times higher chance of passing the exam than students from the course in which SGs were introduced for the first time. While the odds ratio comparing the years 2012-2013 and 2021-2022 (OR = 1.667; 95% CI 0.437-6.358) points out that students from the course in which SGs were implemented by 8-year-experienced teachers have 1.667 times higher chance of passing the exam than students from the course in which no SGs were implemented. From another perspective, this hypothesis testing when applied to data from the years 2013-2014 and 2021-2022 (OR = 4.706; 95% CI 1.397-15.849) highlights that students from the course in which SGs were implemented by 8-year-experienced teachers have a 4.706 times higher chance of passing the exam than students from the course in which SGs were implemented with no prior experience. This leads to the conclusion that SGs when introduced properly to SE courses and when adapted to both the students' needs and the course learning objectives have a positive influence on students' chance to pass their final exam.

H3. The odds ratio for attending the final exam comparing the years 2012-2013 and 2013-2014 (OR = 0.274; 95% CI 0.064-1.172) indicated that students from the course which solely relied on traditional teaching approaches have a 3.649 times higher chance of attending the exam than students from the course in which SGs were introduced for the first time. While The odds ratio comparing the years 2012-2013 and 2021-2022 (OR = 0.923; 95% CI 0.203-4.196) points out that students from the course in which SGs were implemented by 8-year-experienced teachers have a 1.083 times higher chance of attending the exam than students from the course in which no SGs were implemented. From another perspective, this hypothesis testing when applied to data from the years 2013-2014 and 2021-2022 (OR = 3.368; 95% CI 0.971-11.683) highlights that students from the course in which SGs were implemented by 8-year-experienced teachers have a 3.368 times higher chance of attending the exam than students from the course in which SGs were implemented with no prior experience. This stress the fact that SGs when implemented in SE courses with prior experience in the matter from the side of the teaching group have a positive influence on students' attendance chance to their final exam compared to when implemented for the first time with no such experience. However, the findings suggest that traditional teaching approaches still have a better influence on students' attendance chance to their final exam compared to SGs.

H4. On the one hand, according to the learning diaries, the overall reaction was very positive. Teachers explained, based on the students' opinions, that the incorporation of SGs was found motivating, useful, and an attractive resource to compensate for the tediousness of some of the subject's contents. Some of the comments from students stated that "Serious games in class motivated us to study". It was also reported that with the use of SGs the number of students attending the class increased as students received feedback in class that helped them to better understand the subject and to have a better idea of how to apply its theoretical concepts. The following comment was captured from students "With the serious games in class we receive feedback that helps us to understand the subject better". It was also mentioned on several occasions that students recommended promoting the use of SGs in other subjects, especially the ones with heavy theoretical content. Some of the students' comments claimed that "Serious games should be used in other subjects", and "Serious games should be promoted and used more in class to get a better idea of how to apply theoretical concepts". On the other hand, our investigation heavily emphasizes the teachers' experiences during the experiment and their long-term exposure to incorporating SGs into the classroom. We compared the initial and final SG utilization over an 8-year period, with professors engaging in in-depth discussions with students at the conclusion of each course. These dialogues, lasting around 30 minutes after the final class, provided direct insights into student experiences with SGs. Student feedback consistently leaned toward the highly favorable end of the spectrum, underlining the profound impact of SGs on learning and motivation. Teaching strategies evolved over sessions, initially facing challenges like abrupt endings due to time constraints, highlighting the need for longer gaming sessions. Additionally, the discussions highlighted certain limitations, particularly concerning language. For instance, "Software Quantum," one of the SGs employed, was exclusively available in English, posing difficulties for some students. Similarly, "The Island of Requirements" was only accessible in English and Portuguese, and thus language barriers remained an issue for certain learners. While English proficiency should not have been a hindrance, it posed challenges for specific students. Teachers took note of these language-related challenges and recognized the need to consider language barriers when searching for new SGs or explaining those that were in foreign languages.

5.3.3 Results Interpretation and Hypothesis Evaluation

H1. Given that SGs were found to have a positive influence on students' academic results in our study. Previous studies seem to have more conclusive results regarding our hypothesis. Although there are previous studies validating SGs in SE management subjects among master students, such as Play-Scrum at the University of UMinho [124] that relied solely on satisfaction forms, only a few experiments were conducted to assess the students' academic achievements. This is the case of a controlled experiment with a SG entitled E4—Expert and Efficient Estimators Enterprise at the University of Porto [32], to assess the game's effectiveness to teach software estimation through a baseline group and an experimental group. This effectiveness was measured with pre and post-game knowledge questionnaires about the subject. A statistically high significant difference regarding knowledge intake ($T(34) = 3.134$, $p < 0.004$) was found between both groups, in favor of the experimental group who acquired more correct information about the subject than the baseline group, confirming our results in H1. In that same vein, another experiment with a set of SGs entitled Innovation Games® was conducted among international master's students with varying levels of work experience in software development at the University of Oulu [34]. To assess players' performance in software requirement elicitation through the number of requirements that one can identify, two experimental groups were formed with different work experiences. No statistically significant difference was found between the two groups ($p = 0.089$). Although the hypothesis of equality of the two groups' performance cannot be rejected at the significance level $\alpha = 0.05$, the paper argues that there is still a possibility that using SGs for requirements elicitation could result in higher performance compared to non-game-based approaches. Last, according to the study conducted by Subhash and Cudney [125], although a few studies regarding gamified learning in higher education did not observe an improvement in final exam scores, an improvement in learning was observed in the majority of students compared to non-gamified groups, and higher average scores were also observed in these gamified learning groups.

H2. In our study, SGs were found to have a positive influence on students' chance to pass the exam, in the same line, previous studies tackling gamified learning in higher education have backed up this claim. When using gamification and game-based learning in class, reduced failure rates were observed in a systematic literature review on gamified learning in higher education [125]. According to three of the studies covered in this review [126] [127] [128], having the freedom to fail in in-class gamified activities was not only seen as one of the benefits of gamification in education but also identified as an element of gamification suitable for use in education systems. Game-based learning has also proven to be as effective as traditional teaching in terms of knowledge acquisition but is much preferred by students over the latter [129].

H3. To the best of our knowledge, there are no studies investigating SGs' effect on students' participation in the final exam. However, previous studies analyzing game-based learning approaches' effect on students' behavior and participation in the course proposed activities were encountered in our literature search and are in the same line with our findings. Dias et al. [130] have studied class attendance when gamification is implemented in a controlled experiment among first-year students in a modeling course for the Bachelor's degree in Management, results reported regular attendance rate raise of 20% among the gamified group compared to the non-gamified group. In a controlled experiment by Caton et al. [131], evaluating a gamified framework implemented in a game production course to improve attendance and participation among undergraduate computing module students, findings suggest that the average attendance for a gamified year (83%) was also consistently higher than a control year (76%). Likewise, in a similar study conducted by Fotaris et al. [132], attendance in a gamified Python programming class using "Kahoot" among fundamentals of software development students (78%) was found to be higher compared to the control group (65%). The findings also revealed that the gamification approach was observed to have a positive impact on students' punctuality when attending class. In addition, the results revealed that traditional teaching methods continue to have a greater impact on students' likelihood of attending their final exam than SGs. Several factors could explain this result. It is worth noting that SGs must be created as an integrated idea of play and learning in order to fully reap the potential of their benefits [133]. SGs are quick and simple self-assessment tools that allow students to quickly assess their grasp of course contents. In this vein, by integrating knowledge and skill acquisition, SGs are an efficient teaching tool for learners to make smart and realistic decisions [134]. The use of SG in the course may have led students to decide not to take the final exam when they realized their poor mastery of the subject. On the other hand, the use of SGs in academia is still in its infancy, particularly in fields like SE. Finding an appropriate SGs in SE topics can be challenging for teachers [135]. As a result, undesired events may occur when employing SGs in lectures. In this study, the Software Quantum SG generated technical difficulties with updates of Java Runtime Environment. This could have led to less impact of the expected in the learning pathways of the students. The Requirement Island, another SG investigated in this paper, was only available in English and Portuguese. Language barriers may have arisen if students did not have a strong enough knowledge skills, which may have led to complaints and unhappiness among them.

H4. While there are only a few experiments assessing SGs influence on students' academic achievement, there is a significant number of studies assessing the quality of educational games for teaching SE through surveys and satisfaction forms. The findings in the literature are assorted in accordance with the Model for Evaluation of Educational GAmes (MEEGA) [119] used to review the aspects of (1) motivation, (2) user experience, and (3) learning of educational SE games. (1) Concerning motivation, previous works regarding teaching Software Project Management [136], SCRUM [137] [77], fundamentals of ISO/IEC/IEEE 29148 systems and SE [22], and requirement elicitation [138] have reported that students perceived that they become more motivated when using SGs in the course, reporting that students found a sense of exploration, experimentation, curiosity, and manipulation within the game. (2) As regards the user experience, different studies such as the ones at the Federal University of Santa Catarina [37], two Mexican universities [136], Aalto University [139], and Singapore Management University [137] have found that SGs can provide a fun and engaging learning experience as they were proven to have a great potential to impact positively students' process to achieve the learning objectives and to allow students to immediately apply their learning in practice rather than just consuming lectures. Other satisfaction surveys in using SGs to teach SCRUM [124], modeling language Essence and SE project work [35], software project management [32], and requirement elicitation [138] have found SGs to be perceived as enjoyable by students when implemented in SE management subjects. (3) Finally, concerning learning through educational SE games, previous studies [124] [139] [37] have reported through satisfaction questionnaires that students perceived that their knowledge acquisition regarding SE management subjects had increased when using SGs in the course. From a different perspective as regards students' previous knowledge necessity for SGs to be an efficient teaching/learning tool, in a study by Lee et al. [137], students reported that SGs provided a good overview of the topic and that no prior knowledge was necessary to play the SG. In contrast, in another survey by Kemell et al. [35], students expressed that SGs did a poor job in increasing their knowledge acquisition and were sometimes misleading to students with little prior knowledge of the topic. However, SGs helped pave the way for future mastering of the topic among participants. Thus, this matter can be subject to argument and amenable to further investigations regarding the characteristics of the SGs in question, the topics of application, and the way in which SGs were implemented in these topics.

In any case, the meetings with the students at the end of the course, in which about half an hour of the last class is dedicated for them to express directly what they thought of the use of the SGs, were important. All years, they valued the use of the SGs very positively both for the learning and for the motivation to attend the classes. In summary, the main factor that has been improved by the use of the SGs in the lectures has been attendance. By entering the students' grades and providing them with entertainment, the students' attendance in class was observed to continue until the last week of the semester.

Traditionally, research on educational games has often overlooked the pivotal role of teachers in the selection, preparation, instruction, and assessment of educational games [140]. While the studies we draw upon are not specifically focused on assessing teachers' experience in delivering classes with SGs and their direct influence on student performance, they collectively underscore the broader positive impact of experienced teachers when SGs are incorporated into the classroom setting. The study by Romero and Usart. [141] explored SGs integration within an entrepreneurship Massive Online Open Course (MOOC), accentuating the pivotal role of facilitators and teaching presence. Participants in the MOOC consistently expressed high levels of satisfaction, with a notable emphasis on the active involvement of facilitators. This example highlights that experienced teachers familiar with the SGs in question can significantly enhance students' satisfaction and overall class experience. Moreover, Klemetti et al. [142] delve into the role of teaching presence and experienced educators in SG-enhanced learning, and considered it essential to study the practical implications for teachers as they often encounter challenges when incorporating new games into their classrooms. Therefore, in his study conducted among Finnish school teachers, he sheds light on these challenges. In their experiences, teachers reported initial hurdles when familiarizing themselves with new digital learning materials, including SGs. These initial interactions sometimes led to increased workload and apprehension. However, it became evident that with repeated use and growing familiarity, teachers found it more manageable to integrate SGs into their teaching methods. The findings suggest that the first few interactions with new SGs may pose challenges, but as teachers gain experience with the same game, it becomes more seamless and less taxing on their workload. Furthermore, In a study by Jäskä and Aaltonen [143], teachers discovered that the integration of GBL into their courses yielded valuable insights. While GBL elevated student experiences, it also served as a catalyst for teacher growth. The process of sharing and reflecting on these experiences during and after game sessions not only deepened instructors' subject understanding but also refined their teaching skills. Particularly noteworthy were the debriefing sessions, which emerged as pivotal moments. These sessions not only aided students but also played a vital role in fostering teachers' professional development. Consequently, teachers leveraged this growth to refine their approaches to incorporating games in future courses, ultimately enhancing the overall student learning experience. It is essential to recognize that while specific performance metrics may not be assessed in the aforementioned studies, they gave irrefutable evidence the presence of experienced educators remains a crucial factor in optimizing the SG-based learning environment and fostering an enriching educational experience.

5.3.4 Recommendations Based on Findings

Based on the findings of the experiment and their subsequent validation, several practical recommendations have been developed to enhance the integration of serious games in SE education. These recommendations not only serve as direct implications of the experimental results but also provide actionable insights for educators and game designers to optimize learning outcomes.

To ensure the relevance and applicability of these recommendations, they have been categorized according to their source of validation: those based on direct observations made by professors during the experiment are marked as (PO) for 'Professor Observations', and those supported by existing literature are marked as (RL) for 'Relevant Literature'. This dual approach ensures that the recommendations are both empirically grounded and theoretically sound,

For Teachers:

- (RL) SGs should be used as a course complement along with the course learning material rather than a stand-alone resource. The understanding of the game and its dynamic can be perceived as difficult for students with a poor background in the subject [124]. When used as a stand-alone resource, the effectiveness of SGs should be investigated in a self-paced setting and compared to results from professor-directed settings [129].
- (RL) SGs should be played for a restricted period only until their pedagogical goals are reached. Otherwise, when played for lengthened periods for entertainment purposes it may negatively affect players’ intention to play the game and the enjoyment[35].
- (PO) Implement gradually SGs so that teachers can gain experience in better adapting the SGs to their courses.
- (RL) Adapt the integration of SGs gameplay into students’ coursework according to the students’ backgrounds. When students are already familiar with digital games development in earlier courses, the implementation of SGs in coursework was proved smooth [32].

For SGs Developers:

- (PO) Design patterns such as the one proposed by Bjork and Holopainen [144] can be useful for SGs design.

- (RL) A methodology should be used to support the SGs' usage in specific contexts such as software engineering. Good practices for relating game design patterns to teaching techniques are deemed to be necessary but not enough [32].
- (RL) Design the SG in a way no previous knowledge of the adopted topic is necessary to play the game; otherwise, it may affect negatively the SG-based learning efficiency, and the students' knowledge acquisition [35].
- (RL) Select appropriate gamification elements to address the course learning objectives and players' profiles. For instance, introduce rewards, teams play, and bonus rounds to SGs design to motivate students, and make the game more interesting [138].
- (PO) Consider the environment and conditions in which the SGs are implemented when selecting the game mechanics. For example, introducing anonymity and competition to SGs design when the game is used in a professional setting was found to boost players' creativity and engagement, as its organizational hierarchies can hinder honest arguments/opinions [75].

5.4 Threat of validity

This study has some limitations that may present a threat to its results validity [115].

Internal validity issues primarily deal with causality issues related to the presented results. Therefore, we made sure that the students had no prior knowledge of this experiential assessment so that they would not alter their behavior while performing any of the class activities and cause any bias in the results. Threats from other potentially confounding variables related to prior knowledge and experience regarding the course topic did not apply, as subjects had the same academic background and were taking this course for the first time. However, we cannot confirm that the students who participated in this experiment were enthusiastic about the use of SGs, or that they were smarter than other students who did not participate in this experiment. Finally, considering that three teachers monitored the different play sessions, we cannot confirm that students' perceptions of SGs were not influenced by the teachers' different levels of enthusiasm.

External validity issues may arise given the experiment's characteristics: the nature of the subjects of our experiment (students), the number of participants (101 students and 3 teachers), the complexity of the activities involved in the course, and the field of application of the implemented SGs. Nevertheless, Salman et al. [145] and Falessi et al. [146] demonstrated that a small group of students can be considered sufficient to obtain reliable results as they represent a valid simplification of reality needed in laboratory settings. Moreover, the results of this study may therefore be of interest to teachers of other fields outside the case studied, since the effects of the SGs analyzed may be relatively common among students of different majors. Finally, and as Carver [31] suggests, although we cannot overemphasize the general significance of the results obtained via empirical studies with students, these results have relevance for advances in the SE field.

Construct validity issues can arise from errors in the assessment. Therefore, to address these threats, we focused on collecting quantitative data after each course was completed to ensure that the results were the same regardless of who analyzed the data. To reduce the problems of coherence for all the different teachers participating in this experiment, a preparatory meeting was organized at the beginning of each course to deal with the way in which the SGs will be implemented to reproduce the same experimental conditions. Another possible threat could be related to evaluation apprehension. This threat was mitigated by ensuring that the subjects were unaware of the study's objectives and studied hypothesis.

The threat of conclusion validity is a factor that can lead to an incorrect conclusion about an observed relationship. A statistical interpretation method was used to mitigate this threat using parametric/non-parametric statistical tests to investigate three of our hypotheses (H1, H2, and H3).

5.5 Conclusion

This chapter presented a comprehensive pedagogical experiment aimed at measuring the impact of SGs within a SE course. We assessed various aspects, including student scores, attendance rates, and performance in the final exam of a software project management course. One key innovation of this experiment was the evaluation of an SG-based learning approach, led by instructors with differing levels of experience in SG implementation, as compared to the traditional learning method in an SE course context. We investigated their overall influence on students' academic achievement. Our findings demonstrate that the integration of SGs into the course environment resulted in an improvement in students' scores. However, a noteworthy insight emerged when analyzing the chances of students passing the final exam. Significantly higher pass rates were observed when SGs were implemented by instructors with extensive 8-year experience in SGs. This outcome underscores the pivotal role of teachers' experience and expertise in the successful deployment of SGs within SE education. Furthermore, this study has yielded a valuable set of recommendations that can guide SG stakeholders and SE educators in fully harnessing the potential of SGs for SE education. These recommendations are rooted in the insights gained from our research.

In future research endeavors, we plan to delve deeper into understanding potential gender differences in SG usage among SE students, building upon prior research in this area [129]. We also intend to conduct larger-scale experiments, which will explore the impact of SGs on the achievements of both students and industry practitioners, considering their varying levels of enthusiasm for gaming. In-depth case studies involving SE students and experts will provide additional insights. To consolidate and expand our knowledge base, we aim to conduct a higher number of controlled experiments and industry case studies, ultimately enabling a meta-analysis that will contribute to a broader understanding of the effectiveness of SGs versus traditional approaches in SE education and training. Additionally, we have plans to design a comprehensive methodology, complemented by game design patterns, tailored to support the design of effective SGs in specific contexts, such as SE education and training.

6

CHAPTER 3: Development and Expert Validation of an Evidence-Based Framework for Serious Games Design in Software Engineering

6.1 Introduction

In the ever-changing realm of educational technology, it becomes increasingly crucial to recognize the importance of a properly designed serious game, ensuring high levels of student satisfaction and game usability [147]. However, amid this enthusiasm, the field of serious game design encounters numerous challenges. First and foremost, the lack of standardization in serious game development and evaluation [20]. Additionally, the need to balance game and pedagogical elements [148], the call for an interdisciplinary design, and the demand for effective communication between pedagogy and game experts [149] are all deemed crucial for the successful design of serious games. Other challenges include the use of counterproductive game preferences that prioritize player experiences over learning efficacy [150]. Finally, The development process of serious games also faces challenges in terms of involving the target group and conducting iterative testing and improvement [151].

Creating a framework for serious game design is motivated by the need to address the aforementioned challenges and complexities associated with developing effective serious games. Therefore, frameworks are proposed to streamline the design process, facilitate knowledge transfer, enhance coding behavior, and improve the effectiveness of serious games. These frameworks aim to provide clear instructions, guidelines, and methodologies for developers, educators, and practitioners to create effective serious games. Thus these frameworks consider several factors. Firstly, serious games have been recognized as effective tools for learning and training in various fields, such as education, healthcare, and business. However, it is crucial to address the needs and preferences of stakeholders involved in the design and implementation of serious games. Additionally, the quality of serious games has become a focus for investors, researchers, and developers, leading to the development of evaluation frameworks and questionnaires [152]. Furthermore, the use of personalized interventions in serious games is gaining importance, but there is a lack of standardized approaches in this area [12]. Finally, the complexity and cost of developing serious games necessitate the need for frameworks that restructure the design process and facilitate the reuse of domain knowledge and personalization algorithms [46].

6.2 Methodology

To construct a robust framework for serious game design, we adhered to a DSR methodology [153].This methodology involves the development of artifacts to solve real-life business problems following a comprehensive phased approach, with iterations of problem identification, artifact design, construction, and evaluation, contributing to the understanding and practical development of solutions [154].

Research strategies such as semi-structured interviews and questionnaire distribution are used in the process of problem identification and artifact evaluation to gather data and insights from stakeholders, users, and experts, which is crucial for understanding the problem context, identifying requirements, and evaluating the effectiveness of the developed artifacts [155]. Semi-structured interviews allow researchers to have a flexible conversation with participants, enabling them to explore relevant topics while maintaining a certain level of structure. Questionnaires, on the other hand, provide a standardized set of questions that can be distributed to a larger sample size, allowing for quantitative analysis [156].

Design models like Business Process Modeling Notation (BPMN) are employed to depict data, user interface, and processes involved in the development of artifacts by providing visual representations and analytical tools that enhance the effectiveness and efficiency of DSR [157]. These design models do not only serve as communication tools, facilitating collaboration and understanding among researchers, stakeholders, and practitioners involved in the DSR process, but can also aid in the analysis and evaluation of systems and processes, allowing researchers to identify potential improvements and design solutions [158].

Systematically, in this study we adopt the novel approach of DSR by De Sordi [159], incorporating the following steps by analyzing prominent frameworks and methodologies as :

- Problem identification and motivation.: Conducted a meticulous literature review to gain a comprehensive understanding of existing frameworks, methodologies, and their applications in serious game design [42][160]. This step allowed us to define how can analyzing existing frameworks and methodologies guide the development of a novel and comprehensive approach to serious game design and grasp the foundational concepts and practices in the field, providing a solid basis for our framework development.

- Objectives of a solution: Delved into the identified frameworks to discern gaps, limitations, and areas for improvement within current methodologies for serious game design [42][160]. This critical analysis served as a cornerstone for our framework's development, enabling us to pinpoint specific areas where innovation and enhancements were warranted.

- Design and Development: Constructed the designed solution by distilling crucial elements from existing frameworks while addressing identified limitations, establishing the groundwork for a comprehensive and innovative serious game design framework.

- Demonstration: Showcased the effectiveness of the framework in facilitating serious game design. This involved presenting a prototype of the framework to provide a tangible proof of concept.

- Evaluation: Examined and measured the predefined objectives of the designed framework by comparing anticipated outcomes to the observed results from the utilization of the prototype during the demonstration.

- Communication: Articulated the significance of the identified problem, proposed the serious game design framework, highlighted its utility and novelty, underscored the rigor of its design, and communicated its effectiveness to researchers and other pertinent audiences.

By employing a Design Science Research methodology, our approach to framework development was methodically structured, integrating insights from existing literature, addressing the collaborative approach, customization, adaptability, integration of pedagogy, and game design along with relevant gaps identified from the literature. This approach aims not only to facilitate the design process of serious games but also to enhance the effectiveness and efficiency of serious game development.

6.3 Solution Proposal

Our proposed Reusable Automation Framework for Serious Game Design (RAF-SGD) can improve serious games by streamlining the design process, facilitating the transferability of knowledge among stakeholders and reusability of some game components, and simplifying the design process of personalized serious games. It identifies the responsibilities of the involved stakeholders throughout the different stages of the design process. The framework also emphasizes the importance of iterative design methodologies, which have proven valuable in other areas of engineering and commercial game design. By incorporating iterative design, the framework helps ensure that serious games are engaging, effective, and aligned with its overall objectives.

Importantly, the entire framework process is conceptualized as one sprint within the serious game project. This sprint-based approach allows for a structured, iterative cycle that aligns with agile project management methodologies, facilitating rapid development and continuous improvement.

Figure 27 presents an innovative framework designed for the development of serious games. The framework unfolds in incremental steps: starting with an Initialization Phase, followed by a Selection Phase in which components deemed key for the serious game are selected by both educators and game experts. Followed by a Mapping Phase in which game mechanics and learning mechanics are generated automatically. Next, we move on to the Aggregation Phase. Here, game experts develop gameplay content incorporating curated game mechanics, while educators craft the learning content based on the learning mechanics identified earlier and other essential models if chosen in the Initialization Phase. Only after this process are both contents aggregated into the game content. Last, Validation occurs prior to the development of the serious game. Educators carefully review the game content to ensure it complies with quality indicators' thresholds defined in the Initialization Phase, offering feedback for refinement. This iterative feedback loop ensures continuous improvement and refinement of the game content.

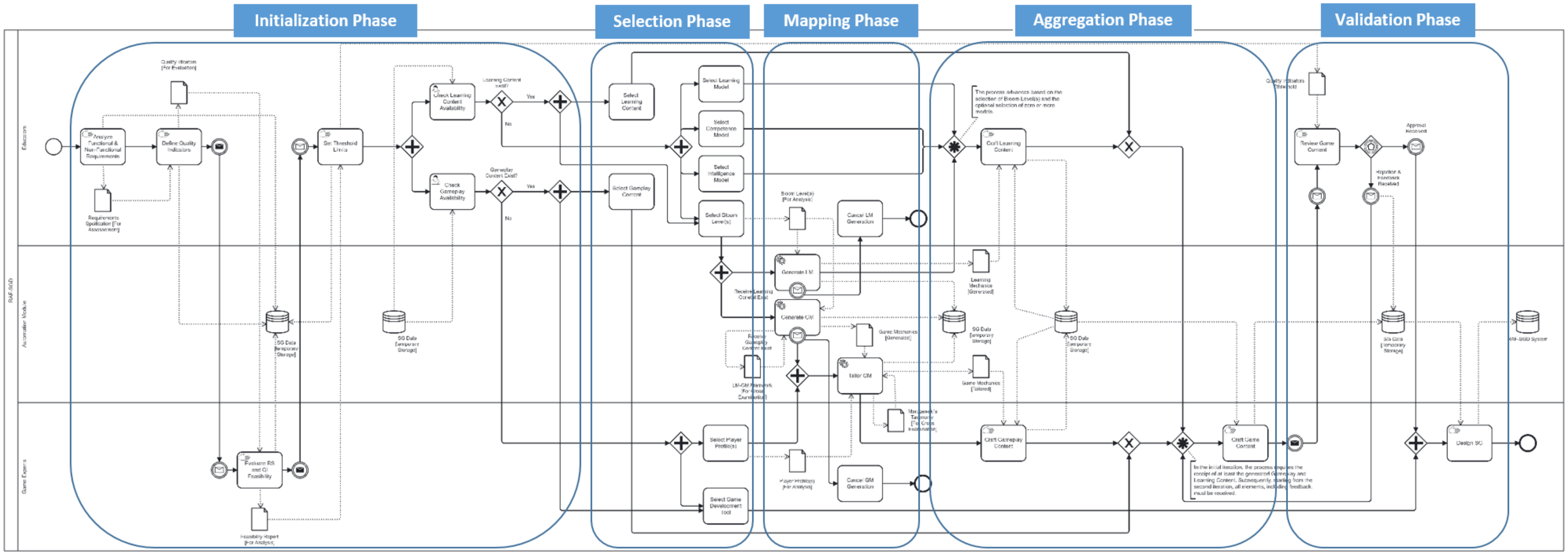


Figure 27: The RAF-SGD Process

6.3.1 Phases explained

Initialization Phase

In this phase, educators analyze the serious game's functional and non-functional specifications and establish quality indicators, considering the quality measurement metric proposed by Suryapranata et al. [161]. Next, Stakeholders collaboratively establish thresholds for these indicators, which will be used to validate the game content in the final stage. This phase also involves educators deciding whether to reuse existing gameplay and/or learning content stored in the system or to craft new content. If they opt to reuse any content, they will select the desired elements during the Selection Phase, thereby skipping the corresponding tasks in the subsequent phases to save time and effort.

Selection Phase

In the proposed framework, a fundamental component involves a targeted phase where educators and game experts make selections based on specific pre-established criteria depending on whether or not they chose to craft new learning and/or gameplay contents in the previous phase. If educators decide upfront to reuse existing content, they select the desired learning or gameplay content during this phase, thereby skipping redundant tasks in the Selection Phase itself and in the Mapping and Aggregation Phases. Otherwise, during this Selection Phase, on the one hand, the educators select (1) the appropriate learning model from a list of evidence-based models presented in the literature review conducted by Wantu et al. [162], providing educators and researchers with guidance on models tailored to suit various learning environments; (2) the competence model - vital measurement tools that cultivate a common understanding of outstanding performance among individuals, synchronizing internal behaviors and skills with educational objectives, and transforming goals into tangible actions [163] - that they want to implement within the serious game among a comprehensive list of the most used and well-cited competence models in the realm of education [164] regardless of the discipline or subject for which the serious game is conceived; (3) the intelligence model that they want to incorporate based on their own understanding of the various models of intelligence proposed by experts in the field [165], such as the Cattell-Horn-Carroll (CHC) model, and their implications for educational practice; (4) the bloom level of knowledge [166] based on the intended purpose behind the use of the serious game so to serve best its aim. On the other hand, the game experts decide on the intended player profile or profiles that the serious game should target based on forms they can pass on to their target students [167] so to have a realistic idea about the player profiles that the gameplay content should be shaped for by understanding what game mechanics are the best to get their attention and ensure their engagement during the gameplay. They also select the game development tool upfront from a predefined list that we provide, which includes recommendations drawn from Barczak and Woźniak comparative analysis [168] of Unity, Unreal Engine, and CryEngine, highlighting their strengths and pitfalls.

Mapping Phase

The selections made in the previous phase of the framework serve as the foundation for a meticulous mapping process, facilitated and cross-referenced through a set of taxonomies. This phased approach represents a departure from attempting to incorporate all conceivable game mechanics. Instead, it streamlines the design process by customizing the serious game. This targeted approach not only optimizes the design process but also enhances the relevance and effectiveness of the serious game, aligning it more closely with the desired educational outcomes.

The framework integrates learning mechanics and game mechanics through a taxonomy that involves Bloom levels. This Learning Mechanics - Game Mechanics (LM-GM) framework [64], serves as a structured guide for aligning the learning and gaming components with the predefined Bloom levels identified by educators. This meticulous mapping ensures that the serious game adheres to the educational objectives set forth by educators. By linking the mechanics to Bloom levels, the framework systematically addresses cognitive domains, enhancing the educational value of the serious game. While the generated learning mechanics are final, the generated game mechanics are mapped next against Marczewski's taxonomy [65] elements corresponding to the selected player profile(s) before generating the final curated set. This strategic alignment ensures that the serious game is tailored to engage and resonate with the intended audience, optimizing the potential gameplay for maximum impact.

In cases where educators have opted to reuse pre-existing content from the previous phase, the subsequent generation of new content is adjusted accordingly. If pre-existing learning content has been selected, the task of generating new learning mechanics is omitted. Similarly, if pre-existing gameplay content is chosen, the task of generating new game mechanics is skipped. As a result, the framework proceeds directly to the Aggregation Phase, bypassing the intermediate tasks.

Aggregation Phase

Building upon the insights gathered from the Mapping Phase, and the Selection Phase if content reuse was involved, the Aggregation Phase is a pivotal stage where educators and game experts converge to synthesize the elements generated in the prior stages into a cohesive and purposeful whole.

Educators take charge of incorporating the learning mechanics generated in the Mapping Phase. This involves a seamless integration of these mechanics with the preselected learning model, intelligence model, and competence model. The synthesis of these elements forms the bedrock for crafting well-informed learning content for the serious game. In this process, educators leverage their pedagogical expertise to ensure that the learning objectives are not only met but also align harmoniously with the selected bloom levels.

Simultaneously, game experts harness the generated game mechanics from the Mapping Phase. These elements are carefully integrated into the gameplay design, forming the structural framework for the serious that aligns with the identified player profiles and enhances the overall gaming and learning experience. In the final stage of the Aggregation Phase, game experts converge the crafted gameplay and learning content into a cohesive game content. This content extends beyond mere mechanics and learning objectives, encompassing elements such as storytelling, scene generation, dialog generation, and asset creation.

The collaboration between educators and game experts during this phase ensures that the educational and gaming aspects seamlessly intertwine resulting in a purposeful and engaging serious game.

Validation Phase

The Validation Phase serves as a pivotal step in ensuring the alignment of the serious game content with the educational vision envisioned by educators. During this phase, educators meticulously examine the game content crafted by game experts to verify its congruence with the predefined quality indicators' thresholds.

Educators play a crucial role in the validation process, examining the game content to ensure that it aligns seamlessly with their educational vision for the serious game. This involves a comprehensive assessment of the storytelling, scene generation, dialog generation, and asset creation, ensuring that these elements harmonize with the predefined models and objectives.

Should educators detect any inconsistencies or deviations from their envisioned educational framework, they are encouraged to engage in constructive feedback sessions with the game experts. This collaborative dialogue serves as the foundation for iterative refinement, whereby developers revisit the Aggregation Phase to address identified issues. Subsequent iterations entail the creation of new game content, which undergoes further validation by educators. Importantly, all generated artifacts during the game design process are stored, facilitating potential reuse of components from earlier versions or revisions based on educator feedback. Additionally, educators' feedback is archived for enhanced traceability, and elements from other games designed within the same framework may be repurposed if beneficial, reducing time and effort in the design process.

This iterative process continues until the educators are satisfied that the game content not only meets the educational criteria but also reflects a harmonious fusion of educational and gaming elements. Each iteration allows for a nuanced refinement of the serious game content, addressing any concerns raised by educators. The collaborative nature of this Validation Phase ensures that both educators and game experts contribute to the iterative enhancement of the serious game.

Upon successful validation and alignment with the educational vision, educators grant approval for the implementation process to commence. This signals the transition from the Validation Phase to the practical implementation of the serious game, where the meticulously crafted game content comes to life contributing to the overall effectiveness and coherence of the serious game.

6.3.2 Stakeholders

Different stakeholders in the process of serious game design interact with each other through dialogue, collaboration, and understanding of their roles and interests. Co-design workshops for serious games involve rapid knowledge exchange between educators and game experts, ensuring interdisciplinary communication and inclusivity in the design process [169]. Figure 27 showcases our framework, emphasizing a collaborative approach between educators and game experts facilitated by an integrated automated mapping module. The framework orchestrates a structured process that incorporates learning design, driven by educators, and game design, led by game experts.

- Educators: The learning design front is enriched by four pivotal components: the learning model, competence model, intelligence model, and Bloom's taxonomy. These components collectively guide the creation of the learning content, while selecting appropriate Bloom levels, directs an automated generation of corresponding game mechanics aligned with the established learning mechanics.
- Game Experts: On the game design front, the framework dictates the generation of both the game content and gameplay through the selection of player profiles and game development tools. The chosen player profile(s) guide the automation module in crafting game mechanics tailored to the specific player profiles. Thus, game experts contribute by creating storytelling, scene generation, dialog generation, and asset creation, aligning with the previously generated gameplay and learning content.

6.4 Validation

In adherence to the DSR methodology, the evaluation of the proposed serious game design framework RAF-SGD is designed to systematically measure its predefined objectives. This process involves a comparative analysis of the various components within the framework against existing related work, thereby emphasizing the distinctive contributions of the proposed framework. Furthermore, to validate the effectiveness, efficiency, and impact of the framework on the serious game design process, a practical application is undertaken. The framework is actively utilized in the design of a specific serious game, providing a real-world context for evaluation. To gather comprehensive insights and data on the framework's performance, research strategies are employed. Stakeholders, users, and experts in the field are actively engaged to contribute their perspectives. Through semi-structured interviews, in-depth discussions are conducted to delve into nuanced aspects of the framework's utility, effectiveness, and efficiency. Simultaneously, questionnaires are distributed to collect quantitative data, ensuring a well-rounded understanding of the framework's impact on the serious game design process.

This multifaceted evaluation approach is crucial in ascertaining not only the theoretical underpinnings and comparative advantages of the proposed framework but also its practical implications and acceptance within the user community. By triangulating insights from both qualitative and quantitative research strategies, this evaluation aims to provide a robust and comprehensive assessment of the serious game design framework's efficacy, thereby contributing to the advancement of design science in the domain.

6.4.1 Comparative Analysis

Having established the pivotal role played by methods and frameworks in the realm of serious game design, this section undertakes a comprehensive literature review covering the latest publications from 2020 to 2024 as the most recent and robust examination of existing serious game design frameworks available at the time of our study was the one conducted by Carrión et al. [52]. Although some research work in the field found, none of the approaches studied by Carrión et al. [52] use techniques to promote creativity and most of them are not generalizable and non-integrable with agile approaches. To address these limitations, Carrión et al. [52] proposed a new methodology called iPlus built upon found gaps outperforming previous frameworks. this assessment was pivotal in guiding our research focus, as it suggested that a review extending beyond iPlus would likely yield diminishing returns in terms of new insights or advancements.The methodology employed involves a meticulous search of the Scopus database, focusing on advancements in the field during this timeframe. The search strategy employs various keywords and their combinations, specifically tailored to highlight relevant content related to serious game design, covering approaches, methods, and frameworks. The selection criteria ensure a concentration on the domain of Computer Science and closely related fields. The subsequent step involves an inspection and elimination process to refine the collection filtering out works with limited information, followed by the classification of articles into distinct thematic sections within this survey in order to cross-compare our proposed solution against the latest methods and frameworks for serious game design found in the literature. By leveraging the most recent and comprehensive review of serious game design frameworks, we aimed to ensure that our proposed framework is evaluated against the most relevant and up-to-date benchmarks. This approach not only ensures the relevance and applicability of our findings but also reinforces the contribution of our research to advancing serious game design methodologies.

Ref	Iterative Design	Collaborative Approach	Player Profiles	LM-GM	Bloom's Taxonomy	Pedagogy	Metrics	Reusability	Automation	Adaptability	MDA or DPE Support
[153]	Yes	No	Yes	Yes	No	No	No	No	No	Yes	Yes
[170]	Yes	No	Yes	No	No	Yes	Yes	Yes	Yes	Yes	No
[160]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes
[171]	No	Yes	No	No	Yes	Yes	No	Yes	No	Yes	No
[172]	Yes	No	Yes	Yes	Yes	Yes	No	Yes	No	Yes	No
[47]	Yes	Yes	No	No	No	No	No	No	Yes	No	No
[173]	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes
[149]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
[174]	Yes	No	No	No	No	Yes	No	No	No	Yes	Yes
[175]	No	No	No	Yes	Yes	No	No	Yes	No	Yes	Yes
[176]	Yes	No	No	No	No	No	No	No	Yes	Yes	No
[177]	No	Yes	Yes	Yes	Yes	No	No	Yes	No	Yes	Yes
[178]	Yes	Yes	Yes	No	No	Yes	No	Yes	Yes	No	No
[179]	Yes	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	No
[180]	No	Yes	Yes	Yes	No	Yes	No	No	No	No	Yes
[181]	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
[182]	Yes	No	Yes	No	No	Yes	No	Yes	Yes	Yes	Yes
[183]	Yes	No	No	No	No	Yes	No	Yes	No	Yes	No
[184]	No	No	No	Yes	No	Yes	No	No	No	No	No
[185]	No	No	No	No	No	No	No	No	No	No	No
[186]	No	Yes	No	No	No	Yes	Yes	No	No	Yes	Yes
[187]	Yes	Yes	Yes	Yes	No	No	No	No	No	Yes	Yes
[188]	No	No	Yes	No	No	No	No	No	No	Yes	Yes
[189]	Yes	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
[190]	No	No	No	No	No$	No	No	No	No	No	No
[191]	No	No	No	No	Yes	Yes	No	Yes	No	No	No
[192]	No	No	No	No	No	Yes	No	Yes	Yes	Yes	No
[193]	Yes	Yes	Yes	No	No	Yes	No	Yes	Yes	Yes	No
RAF-SGD	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 6: Comparative Analysis Summary

In examining found game design methods and frameworks for SGs, several commonalities emerged, underscoring key principles that contribute to their effectiveness. The comparison of our proposed framework with existing frameworks, highlighting its potential advantages, is presented in Table 6. This table outlines key features such as automation, reusability, and iterative design, providing a clear contrast with other frameworks to emphasize the unique strengths of our approach.

Iterative Design

Serious game design frameworks often include processes for the iterative evaluation of a game's effectiveness in achieving learning objectives. This involves collecting data, analyzing player performance, and implementing improvements to enhance the educational value of the game. The concept of "Iterative Design" is supported in various ways across frameworks. [172] emphasizes iterative processes through the DIJS model, which involves cycles of action and reflection within a prototyping process. Similarly, [178] highlights the iterative nature of the IPEOF model, which facilitates continuous refinement and adaptation. [187] also supports iterative design, though it does not provide specific details beyond a general endorsement of iterative methods. In contrast to existing frameworks, our proposed framework, The RAF-SGD emphasizes iterative methodologies by incorporating feedback loops in the Validation Phase. The process ensures continuous refinement of game content before and during development, with repeated iterations until educators are satisfied.

Collaborative Approach

Many serious game design frameworks advocate for collaboration and an interdisciplinary approach. This entails engaging experts from diverse fields, including game design, education, psychology, and technology, to ensure a comprehensive and well-rounded design that meets the needs of both educators and learners. However, the "Collaborative Approach" is addressed differently among frameworks. [171] does not explicitly mention a collaborative approach but supports decision-making by educators and game designers based on learning objectives, indirectly suggesting a collaborative process. In contrast, [179] actively incorporates a collaborative approach through multiplayer SGs, which blend serious games with collaborative learning techniques to enhance connection, cooperation, and engagement among participants. When compared to other approaches, the RAF-SGD relies heavily on collaboration between educators and game experts throughout each phase. Co-design workshops and interdisciplinary communication ensure that both educational and gaming elements are developed in tandem, ensuring a holistic approach.

Player Profiles

Serious game design frameworks consistently prioritize the creation of immersive and engaging player experiences. This involves leveraging various elements tailored to target player profiles in order to enhance motivation and engagement. Different frameworks consider player profiles in diverse ways. To illustrate, [160] takes into account different player profiles by integrating learning styles, game genres, and environments. [172] classifies player profiles using Kolb's topology and player types, while [173] incorporates Bloom's Taxonomy and the Theory of Multiple Intelligences. [177] emphasizes multidisciplinary collaboration and tailored support materials for player profiles. [178] addresses the characteristics of digital natives, combining game design with learning theories and participatory design methods. [179] uses HCI techniques like personas to consider player profiles, and [182] focuses on individual task setups and in-game progress tracking. [187] acknowledges diverse player interactions and choices, including AI-driven agents, and [188] incorporates cognitive styles and personality traits for personalized learning experiences. Unlike found methods and frameworks, when it comes to RAF-SGD framework, player profiles are determined in the Selection Phase, where game experts select appropriate player profiles based on surveys and understanding of target students. These profiles help shape gameplay and game mechanics tailored to engage specific audiences.

LM-GM Framework

A focal point for serious game design frameworks is the intentional design and implementation of game mechanics. These mechanics are crafted to engage players effectively, incorporating elements such as challenges, rewards, feedback systems, and progression mechanisms. In the same line, SGs design frameworks consistently emphasize the alignment of clear learning objectives with game mechanics and activities. These objectives delineate the specific knowledge, skills, or behaviors players should acquire through gameplay. This can be exemplified by, but is not limited to [172], were the LM-GM framework's integration was notably addressed and which employs this methodology to link game mechanics with learning outcomes, thereby engaging users through effective game design and mechanics. Where other models fall short, our framework integrates the LM-GM framework in the Mapping Phase, aligning game mechanics with Bloom's Taxonomy. Thus, learning mechanics are directly mapped to educational objectives, ensuring that game design is purposefully educational.

Bloom's Taxonomy

Serious game design frameworks incorporate principles of instructional design, ensuring that the learning experience is effective. This involves employing instructional strategies, scaffolding techniques, and thoughtful sequencing of content to support a structured learning progression. Bloom's Taxonomy can be considered part of this as it provides a hierarchical framework for structuring learning objectives, which helps align game mechanics with desired cognitive outcomes, ensuring a systematic progression of learning. Frameworks vary in their treatment of Bloom’s Taxonomy. [173] does not explicitly address Bloom's Taxonomy in relation to its framework. However, [149] incorporates Bloom’s extended taxonomy alongside a Learning Qualities Framework to structure educational objectives and activities. Regarding our framework, Bloom’s taxonomy is used to align learning mechanics with specific educational goals. During the Mapping Phase, learning mechanics are tied to Bloom’s cognitive domains, ensuring that educational content matches the desired learning outcomes.

Pedagogy

Pedagogical considerations are addressed with varying emphases across frameworks. [170] supports pedagogical principles indirectly through discussions on interactive and engaging learning experiences. [160] balances pedagogy with entertainment value and considers game genre relative to Bloom's educational objectives. [171] uses Kolb’s Experiential Learning Cycle and the Learning Objective-Game Design framework to align pedagogical intent with game design. [173] supports pedagogical concepts through the DPE framework and iPlus metamodel. [149] emphasizes Bloom’s extended taxonomy and the Learning Qualities Framework. [174] incorporates constructivism, humanism, and cognitivist theories, along with strategies such as challenges and scaffolding. [180] supports situated learning and the educational product lifecycle, while [182] discusses adaptation and user data for enhancing learning activities. [183] lists 22 learning and teaching functions to align with pedagogical concepts. [184] supports GBL concepts, and [186] emphasizes Game-Based Learning Units (GBLUs) and knowledge transfer. [191] proposes ludo-interpretation to connect ludification with learning, and [192] aligns game goals with learning outcomes. [193] supports various pedagogical concepts related to educational game design. With respect to other frameworks, ours supports pedagogy by incorporating educational models such as competence models, intelligence models, and Bloom’s Taxonomy in the design process. Educators guide the development of learning content, ensuring pedagogical alignment.

Metrics

Frameworks also vary in their approach to metrics. [170] supports metrics related to player knowledge, character evolution, narrative progression, and game flow. [173]'s iPlus supports metrics such as process life cycle, story design, and creativity promotion according to ISO standards. [186] provides support for metrics related to game data and learning analytics, while [189] emphasizes metrics for large-scale projects and various delivery platforms. What sets our framework apart is its ability to set quality indicators in the Initialization Phase as metrics to measure the success and quality of the serious game. These metrics are used in the Validation Phase to evaluate game content against predefined educational and functional goals.

Reusability

Reuse refers to the practice of utilizing game and learning components across multiple serious games. This approach can streamline the development of new games by leveraging existing components and knowledge, thereby saving time and resources while promoting standardization and interoperability. Different frameworks address reusability in varying ways. By way of example, [170] supports reusability by enabling the application of its framework in programmable environments for future game development. [160] proposes a methodology that identifies game components for educational and gameplay purposes, facilitating designer-educator communication. [171] introduces the Learning Objective-Game Design framework as a reusable tool for connecting learning objectives with game design elements. [173]'s iPlus supports reusability through an agile approach, formalization via a metamodel, and integration with other methodologies. [179] supports reusability through design patterns, frameworks, models, and personas. Unlike previously suggested methods, Reusability is a key feature of our framework, allowing educators to reuse existing learning or gameplay content from SGs previously designed using the same system. This reduces redundant tasks, making the process more efficient and scalable for future game development projects.

Automation

In this context, automation refers to the use of algorithms and computer programs to automate tasks and processes involved in SGs design. While automation aims to enhance efficiency, reduce development time, and improve the quality and consistency of serious games, its implementation varies across different frameworks. For instance, [170] facilitates automation by streamlining game design processes, character evolution, and narrative progression. [176] supports automation through a structured approach for analytical games, while [193] emphasizes automation in data collection and game design considerations. [192] provides automation through modifiable prototypes and component interoperability. What differentiates our framework from others is that RAF-SGD includes automation in the Mapping Phase, where learning mechanics and game mechanics are generated automatically based on preselected educational models and player profiles. Automation streamlines content creation and reduces manual effort.

Adaptability

Some serious game design frameworks highlight the importance of adaptability and customization. This flexibility allows games to be tailored to different contexts, learner profiles, or specific learning objectives, facilitating personalized and relevant learning experiences. Found frameworks address adaptability in various ways. [153] supports adaptability through a framework with multiple levels of analysis for development and evaluation. [177] integrates domain experts and is flexible in game jam delivery. [179] supports adaptability through design patterns and careful consideration of game activities for different personas, and [187] emphasizes adaptability through ongoing interpretation of circumstances and collaborative learning. In terms of innovation, our framework excels by being adaptable to different educational environments and player needs. Through the Selection Phase, educators and game experts can choose different models, mechanics, and content depending on the game's objectives and target audience.

MDA or DPE Support

Two widely recognized frameworks that are the MDA framework [194] and the DPE framework [195] were used as references. We examined existing serious game design frameworks to determine whether they incorporate or support either or both of these frameworks, as they are indicators of rigor and have been proven to enhance the effectiveness of serious game design. Among found listed frameworks, Support for MDA and DPE varied. [153] supports MDA, while [160] and [189] support both MDA and DPE. [187] supports both MDA and DPE by promoting complex dynamics simulation and integrative knowledge development. [188] supports DPE. Relative to existing solutions, our framework aligns with DPE principles: (1) the Initialization, Selection, and Mapping Phases focus on structured design based on educational objectives. (2) iterative feedback process ensures that gameplay is tested and refined even at the prototype stage. (3) the Validation Phase emphasizes evaluation against predefined metrics, ensuring continuous improvement.

6.4.2 Experts Evaluation

Expert Profiles and Selection

The evaluation process involved two distinct groups of experts, each playing a critical role in the iterative development and validation of the proposed framework. These participants consisted of professionals actively engaged in serious game design and development, including subject matter experts, game designers, and developers. The first group, referred to as the 'Pilot' group, comprised experts with 5 to 12 years of experience in their respective fields. This group was selected for its depth of experience and its multicultural perspectives, which were instrumental in providing comprehensive feedback on the initial version of the framework. The input gathered from the Pilot group was pivotal in identifying areas for refinement and improvement, leading to a more robust and adaptable framework. The second group, referred to as the 'Validation' group, included experts with a broader range of experience, spanning from 1 to 15 years. Like the Pilot group, the Validation group brought together a diverse set of multicultural experiences, which was essential for ensuring the framework's applicability across different contexts. This group was tasked with evaluating the final, refined version of the framework, providing critical insights into its effectiveness, efficiency, and overall quality.

Evaluation Procedure

Both groups followed a similar evaluation procedure. Prior to completing their evaluation forms, participants attended a presentation of the framework, which was followed by a practical application example. This session was designed to ensure that all experts had a clear and thorough understanding of the framework's components and its intended application. During the presentation, the coauthors of the framework conducted a question-and-answer session, addressing any uncertainties and clarifying key aspects of the framework to ensure that the experts' evaluations were informed and accurate. The structured engagement with both the Pilot and Validation groups ensured that the framework was rigorously assessed, drawing on the diverse expertise and cultural perspectives of the participants.

Evaluation Instruments

For the evaluation process, we designed a questionnaire consisting of 7 items gathering sociodemographic information about the experts participating in our study. Additionally, there were 30 items aiming to measure how these experts perceived the value of the RAF-SGD framework as presented in Table 21 of Appendix D. The Likert scale used ranged from strongly disagree (1) to strongly agree (5). Our evaluation focused on specific performance criteria including efficiency and effectiveness, impact on the quality of serious game design, considerations of time and cost. Additionally, stakeholder involvement and collaboration were examined while concurrent and predictive validity were assessed separately through the TAM [196].

Findings and Analysis of Expert Feedback

The feedback questionnaire aimed to uncover impressions and perceptions regarding the value of the RAF-SGD framework. Quantitative data analysis was performed. To complement our quantitative findings, we conducted qualitative interviews with the participants. The interviews were recorded and/or transcribed for further analysis. This approach allowed us to delve deeper into their perceptions of our framework, creating a triangulation of both qualitative and quantitative evidence to enhance the overall evaluation of the RAF-SGD framework.

H1. Enhancements made to the serious game design framework based on feedback from the 'Pilot' expert group will lead to improved perceptions by the 'Validation' expert group.

The evaluation of the framework's components across both the 'Pilot' and 'Validation' groups yielded the following insights:

- Effectiveness and Efficiency (EE): Both groups rated the framework positively in terms of effectiveness and efficiency, with scores ranging from 3 to 5. The 'Validation' group showed slightly higher scores in some areas, though no statistically significant differences were observed between the groups. Specifically, components EE5 and EE6 showed a trend towards improvement in the 'Validation' group ($t = 1.7424$, $p = 0.1250$ for both), while EE8 showed a slight trend towards a decrease ($t = -1.8359$, $p = 0.1090$), though neither reached statistical significance.
- Quality and Resources (QR): The 'Pilot' group rated the quality and resources aspects between 3 and 4, while the 'Validation' group provided slightly higher ratings, ranging from 3 to 5. Notably, QR6 demonstrated a statistically significant improvement in the 'Validation' group ($t = 2.9250$, $p = 0.0222$), suggesting progress in this area. However, other QR components did not show significant differences between the groups.
- Interoperability (IO): Both groups consistently rated the interoperability aspects highly, with scores between 4 and 5. The data indicated strong positive feedback in this area, with no statistically significant differences observed between the groups.
- Technology Acceptance Model (TAM) Components:
    - Perceived Usefulness (PU): Both groups provided positive assessments, with scores between 3 and 5, but, no significant difference was found between the groups ($t = 0.0000$, $p = 1.0000$).
    - Perceived Ease of Use (PEU): The 'Validation' group showed a significant improvement in perceived ease of use ($t = 3.2404$, $p = 0.0142$), marking the most notable finding among all TAM components.
    - Attitude Toward Using (ATU): Although there was a trend towards improvement in the 'Validation' group ($t = 1.7838$, $p = 0.1176$), it did not reach statistical significance.
    - Behavioral Intention to Use (BIU): No significant difference was observed between the groups ($t = -1.0502$, $p = 0.3285$).

Overall, the framework was well-received by both expert groups, with the 'Validation' group generally showing slightly higher scores, indicating improvements based on feedback from the 'Pilot' group. The most significant improvement was observed in Perceived Ease of Use (PEU), suggesting that the framework has become easier to use over time. Additionally, there was a statistically significant improvement in one aspect of Quality and Resources (QR6). However, most other components did not exhibit significant changes, implying that the framework's effectiveness, efficiency, and interoperability have remained relatively stable between the pilot and validation groups.

The lack of significant changes in many components might be attributed to the small sample size, which limits the ability to detect significant differences. A larger sample size could potentially reveal more substantial improvements. The trend towards improvement in Attitude Toward Using (ATU), although not statistically significant, indicates a positive shift in users' attitudes towards the framework. These findings suggest that while there have been some notable improvements, particularly in perceived ease of use and quality, further enhancements could be made in areas that showed trends towards improvement but did not reach statistical significance. Future iterations of the framework should focus on these areas to achieve greater effectiveness and user satisfaction.

H2. The level of experience and multicultural background of the experts may influence their ratings of the framework's components.

To explore the relationships between experience, multicultural background, and the TAM components more deeply, we used correlation analysis and t-tests to examine these relationships.

	Correlation
PU_Mean	-0.10992625200912384
PEU_Mean	0.4816727030991568
ATU_Mean	-0.08111071056538063
BIU_Mean	0.2567271307599146

Table 7: Correlation Between Experience Years and TAM Components

As showcased in Table 7, Perceived Ease of Use (PEU) demonstrates a moderate positive correlation with years of experience, indicating that more experienced individuals tend to find the framework easier to use. Similarly, Behavioral Intention to Use (BIU) exhibits a weak positive correlation with experience years. In contrast, Perceived Usefulness (PU) and Attitude Toward Using (ATU) both show weak negative correlations with experience years, suggesting that as experience increases, the perceived usefulness and favorable attitude toward using the framework slightly decrease.

	T-Test	**p-Value**
PU_Mean	-1.2396	0.2551
PEU_Mean	-1.1114	0.3031
ATU_Mean	-5.2915	0.0011
BIU_Mean	-0.7228	0.4933

Table 8: Influence of Multicultural Experience on TAM Components

Table 8 presents the results of a t-test analysis assessing the influence of multicultural experience on the components of the TAM. Attitude Toward Using (ATU) is the only component showing a statistically significant difference, with a t-statistic of -5.2915 and a p-value of 0.0011, indicating that multicultural experience significantly influences attitudes toward using the framework. In contrast, the other components—Perceived Usefulness (PU), Perceived Ease of Use (PEU), and Behavioral Intention to Use (BIU)—show no statistically significant differences, as reflected by their p-values of 0.2551, 0.3031, and 0.4933, respectively.

In short, data suggest that while experience years alone may not strongly influence the TAM component ratings and only have a moderate influence on Perceived Ease of Use (PEU), multicultural experience could play a role in shaping perceptions as it significantly affects Attitudes Toward Using (ATU) the framework.

H3. Specific components of the framework (e.g., Perceived Ease of Use) have a more significant impact on the overall acceptance and intention to use the framework than others.

The correlation analysis between the TAM components revealed the following relationships:

- Perceived Ease of Use (PEU) and Perceived Usefulness (PU): A weak positive correlation (0.09) was found, indicating a slight relationship between these two components, though not strong.
- PEU and Attitude Toward Using (ATU): A moderate positive correlation (0.58) was observed, suggesting that ease of use likely influences the attitude toward using the framework.
- PU and ATU: A moderate positive correlation (0.32) was also noted, indicating that perceived usefulness contributes to forming a positive attitude toward using the framework.
- ATU and Behavioral Intention to Use (BIU): A moderate positive correlation (0.47) was found, implying that a positive attitude toward using the framework tends to lead to a stronger intention to use it.

Further regression analyses provided additional insights regarding the statistical significance of these revealed relationships.

- PEU and PU: The relationship was very weak (R-squared = 0.008) and not statistically significant (p-value = 0.815), indicating no meaningful link between ease of use and perceived usefulness.
- PEU and ATU: Although the relationship was moderate (R-squared = 0.333), it was not statistically significant at the 0.05 level (p-value = 0.101), suggesting that while ease of use influences attitude toward using, this effect was not confirmed by the data.

- PU and ATU: The relationship was weak (R-squared = 0.100) and not statistically significant (p-value = 0.406), indicating only a minor influence of perceived usefulness on attitude toward using.
- ATU and BIU: The relationship was moderate (R-squared = 0.222) but also not statistically significant (p-value = 0.200), suggesting some influence of attitude on behavioral intention, though not confirmed by the sample.

Overall, the relationships between the TAM components were generally positive, consistent with TAM theory. However, none of the relationships reached statistical significance at the 0.05 level, likely due to the small sample size (n=9), which limits the power of the statistical tests. The strongest observed relationship was between PEU and ATU (R-squared = 0.333), indicating that Perceived Ease of Use has the most substantial impact on Attitude Toward Using. The moderate relationship between ATU and BIU (R-squared = 0.222) suggests some influence of attitude on behavioral intention, though not statistically confirmed in this sample. The weak relationship between PEU and PU suggests that ease of use does not necessarily imply perceived usefulness within this context.

It is important to acknowledge the significant limitation posed by the small sample size. With only nine observations, the ability to detect statistically significant relationships is limited, even if such relationships exist in the broader population. A larger sample size would yield more reliable results and could potentially reveal significant relationships not apparent in this analysis.

6.5 Threats to validity

Internal Validity: As we're comparing pilot and validation groups, selection bias is a significant concern. The differences we observe might be due to inherent differences between the groups rather than the intervention itself. To mitigate this concern we made sure to have coherent groups with similar experience ranges. If the validation group had more time to familiarize themselves with the framework or related concepts, this could affect the results, particularly the improvement in Perceived Ease of Use (PEU). To mitigate this threat we made sure to follow the same evaluation protocol with both the pilot and validation group including session duration.

External Validity: The sample size appears to be relatively small, which limits our ability to generalize the findings to larger populations. We don't have information about the specific context in which the framework was tested, so it's unclear how well these results would generalize to real-world settings. Thus, as part of the ongoing work part of this thesis we are continuing to conduct a border evaluation of this framework with more experts in the field and have another evaluation to be conducted by serious game stockholders reporting their in-action feedback using the framework to design a serious game.

Construct Validity: The use of multiple items for each construct (EE1-EE8, QR1-QR7, IO1-IO3, etc.) suggests an attempt to fully capture each construct. However, the need to examine the specific questions to ensure they adequately represent the constructs is a must. To have this threat covered we made sure to have the questionnaire revised by fellow experts in this context prior to conducting the evaluation. In addition, since all measures were self-reported on a similar scale, which could introduce bias. Mixed methods (qualitative data alongside quantitative) were used to triangulate findings and reduce mono-method bias and strengthen construct validity.

Conclusion Validity: The small sample size (evident from the degrees of freedom in the t-tests) limits the statistical power. This could explain why we see trends towards significance in some measures but not full statistical significance. To address this threat we resumed the evaluation with a larger pool of experts to improve statistical power and reduce the impact of random variance.

6.6 Conclusion

In conclusion, the ongoing development of the RAF-SGD represents a significant leap toward addressing critical challenges in the serious game design landscape. The framework, emphasizing collaboration, adaptability, and a holistic approach to integration, holds promise in transforming the way serious games are conceptualized and developed. The iterative nature of the RAF-SGD not only minimizes redundancy but also paves the way for sustained and efficient serious game development. The RAF-SGD has significant implications for educators and game developers. Educators benefit from a guided and easy-to-follow co-design process, ensuring alignment with educational objectives and efficient integration of pedagogical models and mechanics. Game developers gain automation support, improved collaboration, and effective stakeholder communication. The main takeaway is that the RAF-SGD stands out as a game-changing approach, surpassing existing frameworks in its comprehensive coverage, reuse, and automation aspects. Moving forward, an imperative aspect of our work intend to conduct a comparative study involving the development of a serious game using the RAF-SGD against other methodologies and spontaneous development. This comparative analysis will offer valuable insights into the unique contributions and advantages of the RAF-SGD in the realm of serious game design.

7

CONCLUSION

This thesis has presented a comprehensive investigation into serious game design through the application of a design science methodology, culminating in the development and validation of an evidence-based framework. The research journey began with a mapping study of serious games in the realm of software engineering, revealing critical gaps and opportunities for improvement. Along with this, a long-term controlled experiment was conducted to evaluate the impact of serious games integrated into classroom settings alongside traditional teaching methods. This experiment provided valuable insights into the effects of serious games on students' academic performance and the influence of educators' familiarity with these games on their effectiveness. The findings from both studies underscored the need for a structured approach to serious game design and implementation. To address this, a framework was proposed that strives to balance the serious and enjoyable aspects of game design, ensuring efficacy, efficiency, and quality while optimizing resource use through reuse and automation. The framework was rigorously validated through expert reviews, cross-comparison with existing frameworks, and preliminary application in a real-world serious game design project. These validation steps have confirmed the framework's utility and its alignment with the core principles of effective serious game design. Ongoing efforts will focus on further validation of the framework by engaging additional experts and obtaining feedback from stakeholders actively using the framework in serious game design. Future research will also involve a comparative evaluation of serious games designed using the framework against those developed intuitively or with alternative frameworks. This comparative analysis aims to provide a deeper understanding of the framework's contributions and advantages in game design, ultimately enhancing the development and implementation of serious games. In conclusion, this research contributes significantly to the field of serious game design by providing a robust framework that addresses existing challenges and promotes best practices. The evidence-based approach adopted throughout this study ensures that the framework is not only theoretically sound but also practically applicable, paving the way for more effective and engaging serious games in educational contexts.

8

REFERENCES

[1] D. Djaouti, J. Alvarez, J.-P. Jessel, O. Rampnoux, Origins of Serious Games, in: Serious Games Edutainment Appl., 2011: pp. 25–43. doi:10.1007/978-1-4471-2161-9_3.

[2] G. Ivanova, V. Kozov, P. Zlatarov, Gamification in software engineering education, in: 2019 42nd Int. Conv. Inf. Commun. Technol. Electron. Microelectron. MIPRO 2019 - Proc., Croatian Society MIPRO, 2019: pp. 1445–1450. doi:10.23919/MIPRO.2019.8757200.

[3] J.L. Plass, B.D. Homer, C.K. Kinzer, Foundations of Game-Based Learning, Educ. Psychol. 50 (2015) 258–283. doi:10.1080/00461520.2015.1122533.

[4] S. Deterding, D. Dixon, R. Khaled, L. Nacke, From game design elements to gamefulness: Defining "gamification," in: Proc. 15th Int. Acad. MindTrek Conf. Envisioning Futur. Media Environ. MindTrek 2011, 2011: pp. 9–15. doi:10.1145/2181037.2181040.

[5] N. Wenk, S. Gobron, Reinforcing the difference between Simulation, Gamification, and Serious Game, Gamification Serious Game. (2017) 1–3.

[6] Serious games revenues worldwide 2024 | Statista, (n.d.). https://www.statista.com/statistics/733616/game-based-learning-industry-revenue-world/ (accessed July 1, 2022).

[7] F. Laamarti, M. Eid, A. El Saddik, An overview of serious games, Int. J. Comput. Games Technol. 2014 (2014) 1–16. doi:10.1155/2014/358152.

[8] Serious games: national demand by category Spain | Statista, (n.d.). https://www.statista.com/statistics/1027194/most-demanded-categories-of-serious-games-developed-in-spain/ (accessed July 1, 2022).

[9] A. De Gloria, F. Bellotti, R. Berta, Serious Games for education and training, Int. J. Serious Games. 1 (2014) 1–15. doi:10.17083/ijsg.v1i1.11.

[10] R.P. De Lope, N. Medina-Medina, A Comprehensive Taxonomy for Serious

Games, J. Educ. Comput. Res. 55 (2017) 629–672. doi:10.1177/0735633116681301.

[11] J.S. Breuer, G. Bente, Why so serious? On the relation of serious games and learning, J. Comput. Game Cult. 4 (2010) 7–24. doi:10.7557/23.6111.

[12] R. Bakhtiari, Z. Habibzadeh, Designing a framework and validating a tool for evaluating the educational quality of serious games: a meta-synthesis, Int. J. Serious Games. 10 (2023) 61–83. doi:10.17083/ijsg.v10i2.576.

[13] F. Almeida, Z. Buzady, Exploring the Impact of a Serious Game in the Academic Success of Entrepreneurship Students, J. Educ. Technol. Syst. 51 (2023) 436–454. doi:10.1177/00472395231153187.

[14] E. Tsekleves, J. Cosmas, A. Aggoun, Benefits, barriers and guideline recommendations for the implementation of serious games in education for stakeholders and policymakers, Br. J. Educ. Technol. 47 (2016) 164–183. doi:10.1111/bjet.12223.

[15] K. Borna, H.M. Rad, Serious Games in Computer Science Learning Goals, 2018 2nd Natl. 1st Int. Digit. Games Res. Conf. Trends, Technol. Appl. DGRC 2018. (2018) 161–166. doi:10.1109/DGRC.2018.8712030.

[16] R. Tahir, A.I. Wang, Insights into design of educational games: Comparative analysis of design models, in: Adv. Intell. Syst. Comput., Springer, Cham, 2019: pp. 1041–1061. doi:10.1007/978-3-030-02686-8_78.

[17] O. Pedreira, F. Garcia, M. Piattini, A. Cortinas, A. Cerdeira-Pena, An Architecture for Software Engineering Gamification, Tsinghua Sci. Technol. 25 (2020) 776–797. doi:10.26599/TST.2020.9010004.

[18] T.M. Connolly, E.A. Boyle, E. MacArthur, T. Hainey, J.M. Boyle, A systematic literature review of empirical evidence on computer games and serious games, Comput. Educ. 59 (2012) 661–686. doi:10.1016/j.compedu.2012.03.004.

[19] I. Zadeja, J. Bushati, GAMIFICATION AND SERIOUS GAMES METHODOLOGIES IN EDUCATION, in: Int. Symp. Graph. Eng. Des., 2022: pp. 599–605. doi:10.24867/GRID-2022-p66.

[20] R. Damaševičius, R. Maskeliūnas, T. Blažauskas, Serious Games and Gamification in Healthcare: A Meta-Review, Inf. 14 (2023). doi:10.3390/info14020105.

[21] I. Frank, N. Sanbou, K. Terashima, Some positive effects of online gaming, in: Int. Conf. Adv. Comput. Entertain. Technol. 2006, 2006: pp. 73-es. doi:10.1145/1178823.1178909.

[22] I. García, C. Pacheco, A. León, J.A. Calvo-Manzano, A serious game for teaching the fundamentals of ISO/IEC/IEEE 29148 systems and software engineering – Lifecycle processes – Requirements engineering at undergraduate level,

Comput. Stand. Interfaces. 67 (2020) 103377. doi:10.1016/j.csi.2019.103377.

[23] M.M. Alhammad, A.M. Moreno, Gamification in software engineering education: A systematic mapping, J. Syst. Softw. 141 (2018) 131–150. doi:10.1016/j.jss.2018.03.065.

[24] J. Rice, The Gamification of Learning and Instruction: Game-Based Methods and Strategies for Training and Education, Int. J. Gaming Comput. Simulations. 4 (2012) 81–83. doi:10.4018/jgcms.2012100106.

[25] P. Rodrigues, M. Souza, E. Figueiredo, Games and gamification in software engineering education: A survey with educators, in: IEEE Front. Educ. Conf. FIE, Institute of Electrical and Electronics Engineers Inc., 2019: pp. 1–9. doi:10.1109/FIE.2018.8658524.

[26] M.R. d. A. Souza, L. Veado, R.T. Moreira, E. Figueiredo, H. Costa, A systematic mapping study on game-related methods for software engineering education, Inf. Softw. Technol. 95 (2018) 201–218. doi:10.1016/j.infsof.2017.09.014.

[27] A. Calderón, M. Ruiz, R. V. O'Connor, A multivocal literature review on serious games for software process standards education, Comput. Stand. Interfaces. 57 (2018) 36–48. doi:10.1016/j.csi.2017.11.003.

[28] C.F. Barreto, C. Franca, Gamification in Software Engineering: A literature Review, in: Proc. - 2021 IEEE/ACM 13th Int. Work. Coop. Hum. Asp. Softw. Eng. CHASE 2021, IEEE Computer Society, 2021: pp. 105–108. doi:10.1109/CHASE52884.2021.00020.

[29] C. Caulfield, J.C. Xia, D. Veal, S.P. Maj, A systematic survey of games used for software engineering education, Mod. Appl. Sci. 5 (2011) 28–43. doi:10.5539/mas.v5n6p28.

[30] M.R.D.A. Souza, L.F. Veado, R.T. Moreira, E.M.L. Figueiredo, H.A.X. Costa, Games for learning: Bridging game-related education methods to software engineering knowledge areas, in: Proc. - 2017 IEEE/ACM 39th Int. Conf. Softw. Eng. Softw. Eng. Educ. Track, ICSE-SEET 2017, 2017: pp. 170–179. doi:10.1109/ICSE-SEET.2017.17.

[31] J. Carver, L. Jaccheri, S. Morasca, F. Shull, Issues in using students in empirical studies in software engineering education, in: Proc. - Int. Softw. Metrics Symp., IEEE Computer Society, 2003: pp. 239–249. doi:10.1109/METRIC.2003.1232471.

[32] N. Flores, A.C.R. Paiva, N. Cruz, Teaching software engineering topics through pedagogical game design patterns: An empirical study, Inf. 11 (2020) 153. doi:10.3390/info11030153.

[33] A. Vizcaíno, F. García, I.G.R. De Guzmán, M.Á. Moraga, Evaluating GSD-aware: A serious game for discovering global software development challenges, ACM

Trans. Comput. Educ. 19 (2019). doi:10.1145/3218279.

[34] H. Ghanbari, J. Similä, J. Markkula, Utilizing online serious games to facilitate distributed requirements elicitation, J. Syst. Softw. 109 (2015) 32–49. doi:10.1016/j.jss.2015.07.017.

[35] K.K. Kemell, J. Risku, A. Evensen, P. Abraharnsson, A.M. Dahl, L.H. Grytten, A. Jcdryszek, P. Rostrup, A. Nguyen-Duc, Gamifying the Escape from the Engineering Method Prison, in: 2018 IEEE Int. Conf. Eng. Technol. Innov. ICE/ITMC 2018 - Proc., Institute of Electrical and Electronics Engineers Inc., 2018. doi:10.1109/ICE.2018.8436340.

[36] U. Gulec, M. Yilmaz, A.D. Yalcin, R. V. O'Connor, P.M. Clarke, CENGO: A Web-Based Serious Game to Increase the Programming Knowledge Levels of Computer Engineering Students, in: Commun. Comput. Inf. Sci., Springer, Cham, 2019: pp. 237–248. doi:10.1007/978-3-030-28005-5_18.

[37] C.G. Von Wangenheim, R. Savi, A.F. Borgatto, SCRUMIA - An educational game for teaching SCRUM in computing courses, J. Syst. Softw. 86 (2013) 2675–2687. doi:10.1016/j.jss.2013.05.030.

[38] M.L. Sánchez-Gordón, R. V. O'Connor, R. Colomo-Palacios, S. Sanchez-Gordon, A learning tool for the ISO/IEC 29110 standard: Understanding the project management of basic profile, in: Commun. Comput. Inf. Sci., Springer, Cham, 2016: pp. 270–283. doi:10.1007/978-3-319-38980-6_20.

[39] A. Drappa, J. Ludewig, Simulation in software engineering training, in: Proc. - Int. Conf. Softw. Eng., IEEE, 2000: pp. 199–208. doi:10.1145/337180.337203.

[40] A. Mas, A.L. Mesquida, R. Colomo-Palacio, Enhancing the Student Perception on Software Project Management in Computer Science, IEEE Trans. Educ. 64 (2021) 1–11. doi:10.1109/TE.2020.2998429.

[41] R. Khaled, A. Vasalou, Bridging serious games and participatory design, Int. J. Child-Computer Interact. 2 (2014) 93–100. doi:10.1016/j.ijcci.2014.03.001.

[42] D. Jaccard, L. Suppan, E. Sanchez, A. Huguenin, M. Laurent, The co.lab generic framework for collaborative design of serious games: Development study, JMIR Serious Games. 9 (2021). doi:10.2196/28674.

[43] M. Del Pilar Beristain-Colorado, J.F. Ambros-Antemate, M. Vargas-Treviño, J. Gutiérrez-Gutiérrez, A. Moreno-Rodriguez, P.A. Hernández-Cruz, I.B. Gallegos-Velasco, R. Torres-Rosas, Standardizing the development of serious games for physical rehabilitation: Conceptual framework proposal, JMIR Serious Games. 9 (2021). doi:10.2196/25854.

[44] W. Van Der Vegt, W. Westera, E. Nyamsuren, A. Georgiev, I.M. Ortiz, RAGE Architecture for Reusable Serious Gaming Technology Components, Int. J.

Comput. Games Technol. 2016 (2016). doi:10.1155/2016/5680526.

[45] T.M. de Classe, R.M. de Araujo, G.B. Xexéo, S.W. Matsui Siqueira, The Play Your Process Method for Business Process-Based Digital Game Design, Int. J. Serious Games. 6 (2019) 27–49. doi:10.17083/ijsg.v6i1.269.

[46] S. Carlier, V. Naessens, F. De Backere, F. De Turck, A Software Engineering Framework for Reusable Design of Personalized Serious Games for Health: Development Study, JMIR Serious Games. 11 (2023). doi:10.2196/40054.

[47] A. Gauthier, K. Porayska-Pomsta, S. Mayer, I. Dumontheil, E.K. Farran, D. Bell, D. Mareschal, Redesigning learning games for different learning contexts: Applying a serious game design framework to redesign Stop & Think, Int. J. Child-Computer Interact. 33 (2022). doi:10.1016/j.ijcci.2022.100503.

[48] A.F.S. Barbosa, P.N.M. Pereira, J.A.F.F. Dias, F.G.M. Silva, A new methodology of design and development of serious games, Int. J. Comput. Games Technol. (2014). doi:10.1155/2014/817167.

[49] K. Mitgutsch, N. Alvarado, Purposeful by design?: A serious game design assessment framework, in: Found. Digit. Games 2012, FDG 2012 - Conf. Progr., 2012: pp. 121–128. doi:10.1145/2282338.2282364.

[50] J.A.V.M.K. Jayakody, W.M.J.I. Wijayanayake, Challenges for adopting DevOps in information technology projects, in: Int. Res. Conf. Smart Comput. Syst. Eng. Proc., Institute of Electrical and Electronics Engineers Inc., 2021: pp. 203–210. doi:10.1109/SCSE53661.2021.9568348.

[51] M. Carrion, M. Santorum, M. Perez, J. Aguilar, A participatory methodology for the design of serious games in the educational environment, in: 2017 Congr. Int. Innov. y Tendencias En Ing. CONIITI 2017 - Conf. Proc., Institute of Electrical and Electronics Engineers Inc., 2018: pp. 1–6. doi:10.1109/CONIITI.2017.8273363.

[52] M. Carrión-Toro, M. Santorum, P. Acosta-Vargas, J. Aguilar, M. Pérez, iPlus a user-centered methodology for serious games design, Appl. Sci. 10 (2020) 1–33. doi:10.3390/app10249007.

[53] M. Kosa, M. Yilmaz, R. V. O'Connor, P.M. Clarke, Software engineering education and games: A systematic literature review, J. Univers. Comput. Sci. 22 (2016) 1558–1574. http://www.jucs.org/jucs_22_12/software_engineering_education_and/jucs_22_12_1558_1574_kosa.pdf (accessed October 10, 2022).

[54] T. Xie, N. Tillmann, J. de Halleux, J. Bishop, Educational software engineering: Where software engineering, education, and gaming meet, in: Comput. Games Softw. Eng., 2015: pp. 113–132. doi:10.1201/b18453.

[55] M.R. d. A. Souza, L. Veado, R.T. Moreira, E. Figueiredo, H. Costa, A systematic

mapping study on game-related methods for software engineering education, Inf. Softw. Technol. 95 (2018) 201–218. doi:10.1016/j.infsof.2017.09.014.

[56] B. Spieler, W. Slany, Game development-based learning experience: gender differences in game design, in: Proc. Eur. Conf. Games-Based Learn., Dechema e.V., 2018: pp. 616–625. https://arxiv.org/abs/1805.04457v3 (accessed July 19, 2023).

[57] P.-W. Ng, Making Software Engineering Education Structured, Relevant and Engaging through Gaming and Simulation, J. Commun. Comput. 10 (2013) 1365–1373.

[58] A. Abran, P. Bourque, R. Dupuis, J.W. Moore, L. Tripp, Software Engineers: What does industry expect them to know & master?, (2004) 1–68. www.swebok.org (accessed July 31, 2023).

[59] K. Petersen, S. Vakkalanka, L. Kuzniarz, Guidelines for conducting systematic mapping studies in software engineering: An update, Inf. Softw. Technol. 64 (2015) 1–18. doi:10.1016/j.infsof.2015.03.007.

[60] K. Petersen, R. Feldt, S. Mujtaba, M. Mattsson, Systematic mapping studies in software engineering, in: 12th Int. Conf. Eval. Assess. Softw. Eng. EASE 2008, 2008. doi:10.14236/ewic/ease2008.8.

[61] S. Kitchenham, B. and Charters, Guidelines for performing systematic literature reviews in software engineering, Tech. Report, Ver. 2.3 EBSE Tech. Report. EBSE. 5 (2007) 1–65.

[62] A. Pendleton, J. Okolica, Creating serious games with the game design matrix, Springer International Publishing, 2019. doi:10.1007/978-3-030-34350-7_51.

[63] L. Buchanan, F. Wolanczyk, F. Zinghini, Blending Bloom's Taxonomy and Serious Game Design, Proc. 2011 Int. Conf. Secur. Manag. H.R.Arabnia, M.R.Grimaila, G. Markowsky, S. Aissi, Eds. CSREA Press. V (2011) 518–521. http://securedecisions.com/wp-content/uploads/pubs/SAM5062.pdf.

[64] T. Lim, M.B. Carvalho, F. Bellotti, S. de Freitas, S. Louchart, N. Suttie, R. Berta, A. De Gloria, The LM-GM framework for Serious Games Analysis, Int. J. Serious Games. 46 (2015) 391–411. https://pdfs.semanticscholar.org/7df0/20237a6d3995860e7345c77dab28e4d0a001.pdf?_ga=2.238872384.1768826724.1559384891-936227403.1559384891%0Ahttps://seriousgamessociety.org/wp-content/uploads/2016/09/lmgm_framework.pdf%0Ahttp://citeseerx.ist.psu.edu/viewdo.

[65] A. Marczewski, A Revised Gamification Design Framework - Gamified UK - {\#}Gamification Expert, (2017). https://www.gamified.uk/2017/04/06/revised-gamification-design-framework/ (accessed November 24, 2021).

[66] P.W. Stone, Popping the (PICO) question in research and evidence-based practice, Appl. Nurs. Res. 15 (2002) 197–198. https://www.sciencedirect.com/science/article/pii/S0897189702000101 (accessed November 25, 2019).

[67] P. Brereton, B.A. Kitchenham, D. Budgen, M. Turner, M. Khalil, Lessons from applying the systematic literature review process within the software engineering domain, J. Syst. Softw. (2007). doi:10.1016/j.jss.2006.07.009.

[68] C. Wohlin, Guidelines for snowballing in systematic literature studies and a replication in software engineering, in: ACM Int. Conf. Proceeding Ser., 2014. doi:10.1145/2601248.2601268.

[69] A. Drappa, J. Ludewig, Simulation in software engineering training, Proc. - Int. Conf. Softw. Eng. (2000) 199–208. doi:10.1145/337180.337203.

[70] U. Aydan, M. Yilmaz, R. V. O'connor, Towards a serious game to teach ISO/IEC 12207 software lifecycle process: An interactive learning approach, in: Commun. Comput. Inf. Sci., Springer, Cham, 2015: pp. 217–229. doi:10.1007/978-3-319-19860-6_17.

[71] C. Chung, A. Matsuoka, Y. Yang, J. Rubin, M. Chechik, Serious Games for NP-hard Problems: Challenges and Insights, in: Proc. - Int. Conf. Softw. Eng., IEEE Computer Society, 2016: pp. 29–32. doi:10.1145/2896958.2896963.

[72] M.A. Miljanovic, J.S. Bradbury, Robot ON!: A Serious Game for Improving Programming Comprehension, Proc. - Int. Conf. Softw. Eng. 16-May-201 (2016) 33–36. doi:10.1145/2896958.2896962.

[73] M. Xenos, V. Velli, A Serious Game for Introducing Software Engineering Ethics to University Students, Adv. Intell. Syst. Comput. 916 (2020) 579–588. doi:10.1007/978-3-030-11932-4_55.

[74] P. Caserman, S. Göbel, Become a Scrum Master: Immersive Virtual Reality Training to Learn Scrum Framework, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), Springer, Cham, 2020: pp. 34–48. doi:10.1007/978-3-030-61814-8_3.

[75] A. Alexandrova, L. Rapanotti, Requirements analysis gamification in legacy system replacement projects, Requir. Eng. 25 (2020) 131–151. doi:10.1007/s00766-019-00311-2.

[76] A. Przybylek, W. Kowalski, Utilizing online collaborative games to facilitate agile software development, in: Proc. 2018 Fed. Conf. Comput. Sci. Inf. Syst. FedCSIS 2018, 2018: pp. 811–815. doi:10.15439/2018F347.

[77] G. Rodriguez, A. Teyseyre, P. Gonzalez, S. Misra, Gamifying Users' Learning Experience of Scrum, in: Commun. Comput. Inf. Sci., Springer, Cham, 2021: pp.

497–509. doi:10.1007/978-3-030-69143-1_38.

[78] O. Pedreira, F. García, N. Brisaboa, M. Piattini, Gamification in software engineering – A systematic mapping, Inf. Softw. Technol. 57 (2015) 157–168. doi:10.1016/J.INFSOF.2014.08.007.

[79] U. Schafer, Training scrum with gamification: Lessons learned after two teaching periods, in: IEEE Glob. Eng. Educ. Conf. EDUCON, IEEE Computer Society, 2017: pp. 754–761. doi:10.1109/EDUCON.2017.7942932.

[80] F. Larenas, B. Marín, G. Giachetti, Classutopia: A serious game for conceptual modeling design, in: Proc. Int. Conf. Softw. Eng. Knowl. Eng. SEKE, Knowledge Systems Institute Graduate School, 2018: pp. 116–121. doi:10.18293/SEKE2018-145.

[81] H.S. Wankhede, A.W. Kiwelekar, Qualitative assessment of software engineering examination questions with bloom's taxonomy, Indian J. Sci. Technol. 9 (2016). doi:10.17485/ijst/2016/v9i6/85012.

[82] K. Werbach, D. Hunter, The Gamification Toolkit: Dynamics, Mechanics, and Components for the Win, Whart. Digit. Press. (2015) 148. https://books.google.es/books?hl=en&lr=&id=2PU1EAAAQBAJ&oi=fnd&pg=PT10&dq=The+Gamification+Toolkit:+Dynamics,+Mechanics,+and+Components+for+the+Win&ots=MMDzc28J3P&sig=DkaakCu9FNBV7DgUtV10rLVQxfE&redir_esc=y#v=onepage&q=The Gamification Toolkit%3A Dynamics (accessed March 21, 2022).

[83] M. Maarek, S. Louchart, L. McGregor, R. McMenemy, Co-created design of a serious game investigation into developer-centred security, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), Springer, Cham, 2019: pp. 221–231. doi:10.1007/978-3-030-11548-7_21.

[84] J. Bell, S. Sheth, G. Kaiser, Secret ninja testing with HALO software engineering, in: SSE'11 - Proc. 4th Int. Work. Soc. Softw. Eng., 2011: pp. 43–47. doi:10.1145/2024645.2024657.

[85] C.S. Longstreet, K. Cooper, A meta-model for developing simulation games in higher education and professional development training, Proc. CGAMES'2012 USA - 17th Int. Conf. Comput. Games AI, Animat. Mobile, Interact. Multimedia, Educ. Serious Games. (2012) 39–44. doi:10.1109/CGames.2012.6314549.

[86] A. Calderón, M. Ruiz, ProDec: a serious game for software project management training, in: Proc. 8th Int. Conf. Softw. Eng. Adv. Venice, Italy., 2013: pp. 565–570.

[87] A. Marczewski, A Player Type Framework for Gamification Design, Gamified UK. (2015). http://www.gamified.uk/user-types/ (accessed September 4, 2021).

[88] A.C.T. Klock, I. Gasparini, M.S. Pimenta, J. Hamari, Tailored gamification: A review of literature, Int. J. Hum. Comput. Stud. 144 (2020) 102495. doi:10.1016/j.ijhcs.2020.102495.

[89] R. Echeverría, F. Jurado, Using player profiles and learning styles in the design of Educational Games, in: 2015 Int. Symp. Comput. Educ. SIIE 2015, 2016: pp. 141–144. doi:10.1109/SIIE.2015.7451664.

[90] R.A. Bartle, S. de Andrade Silva, Hearts, Clubs, Diamonds, Spades: Players Who Suit Muds Related papers Buy and Share! Social Net work Games and Ludic Shopping, (n.d.).

[91] S. Ludi, J. Collofello, An analysis of the gap between the knowledge and skills learned in academic software engineering course projects and those required in real projects, in: Proc. - Front. Educ. Conf., 2001. doi:10.1109/fie.2001.963881.

[92] I. Garcia, C. Pacheco, F. Méndez, J.A. Calvo-Manzano, The effects of game-based learning in the acquisition of "soft skills" on undergraduate software engineering courses: A systematic literature review, Comput. Appl. Eng. Educ. 28 (2020) 1327–1354. doi:10.1002/cae.22304.

[93] A. Alarcón, N. Martinez, J. Sandoval, Use of learning strategies of SWEBOK© guide proposed knowledge areas, in: Adv. Intell. Syst. Comput., Springer Verlag, 2013: pp. 243–254. doi:10.1007/978-3-642-30867-3_22.

[94] A. De Gloria, F. Bellotti, R. Berta, Serious Games for education and training, Int. J. Serious Games. 1 (2014) 1–15. doi:10.17083/ijsg.v1i1.11.

[95] B.S. Bloom, D.R. Krathwohl, Taxonomy of Educational Objectives: The Classification of Educational Goals, in: Handb. I Cogn. Domain., 1956.

[96] A.M. Toda, A.C.T. Klock, W. Oliveira, P.T. Palomino, L. Rodrigues, L. Shi, I. Bittencourt, I. Gasparini, S. Isotani, A.I. Cristea, Analysing gamification elements in educational environments using an existing Gamification taxonomy, Smart Learn. Environ. 6 (2019). doi:10.1186/s40561-019-0106-1.

[97] G.F. Tondello, R.R. Wehbe, L. Diamond, M. Busch, A. Marczewski, L.E. Nacke, The Gamification User Types Hexad Scale, (2016). doi:10.1145/2967934.2968082.

[98] B. Rammstedt, O.P. John, Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German, J. Res. Pers. 41 (2007) 203–212. doi:10.1016/j.jrp.2006.02.001.

[99] W.S. Ravyse, A. Seugnet Blignaut, V. Leendertz, A. Woolner, Success factors for serious games to enhance learning: a systematic review, Virtual Real. 21 (2017) 31–58. doi:10.1007/s10055-016-0298-4.

[100] W.S. Ravyse, A. Seugnet Blignaut, V. Leendertz, A. Woolner, Success factors for serious games to enhance learning: a systematic review, Virtual Real. 21 (2017)

31–58. doi:10.1007/s10055-016-0298-4.

[101] A. Dimitriadou, N. Djafarova, O. Turetken, M. Verkuyl, A. Ferworn, Challenges in Serious Game Design and Development: Educators' Experiences, Simul. Gaming. 52 (2021) 132–152. doi:10.1177/1046878120944197.

[102] N. Iten, D. Petko, Learning with serious games: Is fun playing the game a predictor of learning success?, Br. J. Educ. Technol. 47 (2016) 151–163. doi:10.1111/bjet.12226.

[103] I. Daoudi, R. Chebil, W.L. Chaari, A novel tool to predict the impact of adopting a serious game on a learning process, in: ICEIS 2018 - Proc. 20th Int. Conf. Enterp. Inf. Syst., SciTePress, 2018: pp. 585–592. doi:10.5220/0006685905850592.

[104] C. Xin, Artificial intelligence application in mobile phone serious game, in: Proc. 1st Int. Work. Educ. Technol. Comput. Sci. ETCS 2009, 2009: pp. 1093–1095. doi:10.1109/ETCS.2009.510.

[105] W. Westera, R. Prada, S. Mascarenhas, P.A. Santos, J. Dias, M. Guimarães, K. Georgiadis, E. Nyamsuren, K. Bahreini, Z. Yumak, C. Christyowidiasmoro, M. Dascalu, G. Gutu-Robu, S. Ruseti, Artificial intelligence moving serious gaming: Presenting reusable game AI components, Educ. Inf. Technol. 25 (2020) 351–380. doi:10.1007/s10639-019-09968-2.

[106] D. Picca, D. Jaccard, G. Eberlé, Natural Language Processing in Serious Games: A state of the art., Int. J. Serious Games. 2 (2015). doi:10.17083/ijsg.v2i3.87.

[107] W.L. Johnson, H.H. Vilhjálmsson, S. Marsella, Serious games for language learning: How much game, how much AI?, in: AIED, 2005: pp. 306–313.

[108] E. Lach, Dynamic difficulty adjustment for serious game using modified evolutionary algorithm, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), Springer, Cham, 2017: pp. 370–379. doi:10.1007/978-3-319-59063-9_33.

[109] G.N. Yannakakis, J. Togelius, A Panorama of Artificial and Computational Intelligence in Games, IEEE Trans. Comput. Intell. AI Games. 7 (2015) 317–335. doi:10.1109/TCIAIG.2014.2339221.

[110] J. Gulati, P. Bhardwaj, B. Suri, Comparative study of personality models in software engineering, in: ACM Int. Conf. Proceeding Ser., 2015: pp. 209–216. doi:10.1145/2791405.2791445.

[111] M. Kharbouch, A. Vizcaino, J.L. Fernandez Aleman, Un marco de diseño de juegos serios basado en reutilización y automatización, in: M.A. Rodríguez Luaces (Ed.), Actas Las XXVIII Jornadas Ing. Del Softw. y Bases Datos (JISBD 2024), Sistedes, 2024. https://hdl.handle.net/11705/JISBD/2024/32.

[112] O. Díaz, M. Genero, J.P. Contell, M. Piattini, Adding relevance to rigor:

Assessing the contributions of SLRs in Software Engineering through Citation Context Analysis, Comput. Sci. Rev. 53 (2024) 100649. doi:10.1016/j.cosrev.2024.100649.

[113] X. Zhou, Y. Jin, H. Zhang, S. Li, X. Huang, A map of threats to validity of systematic literature reviews in software engineering, in: Proc. - Asia-Pacific Softw. Eng. Conf. APSEC, IEEE Computer Society, 2016: pp. 153–160. doi:10.1109/APSEC.2016.031.

[114] C. Wohlin, P. Runeson, P.A. Da Mota Silveira Neto, E. Engström, I. Do Carmo Machado, E.S. De Almeida, On the reliability of mapping studies in software engineering, J. Syst. Softw. 86 (2013) 2594–2610. doi:10.1016/j.jss.2013.04.076.

[115] C. Wohlin, P. Runeson, M. Höst, M.C. Ohlsson, B. Regnell, A. Wesslén, Experimentation in software engineering, Springer-Verlag Berlin Heidelberg, 2012. doi:10.1007/978-3-642-29044-2.

[116] T. Susi, M. Johannesson, Serious Games-An Overview, 2007. www.americasarmy.com; (accessed July 1, 2022).

[117] R. Márquez, A. Vizcaíno, F.O. Garciá, A. Manjavacas, GLOBAL-MANAGER: A serious game for providing training in project manager skills, in: Proc. - 2020 ACM/IEEE 15th Int. Conf. Glob. Softw. Eng. ICGSE 2020, 2020: pp. 127–131. doi:10.1145/3372787.3390443.

[118] V. Basili, C. Caldiera, H.D. Rombach, Goal Question Metric Paradigm, Encycl. Softw. Eng. (Marciniak, J.J., Ed.), John Wiley Sons. 1 (1994) 528–532. http://www.cs.umd.edu/users/basili/publications/technical/T87.pdf (accessed July 3, 2022).

[119] R. Savi, C.G. Von Wangenheim, A.F. Borgatto, A model for the evaluation of educational games for teaching software engineering, in: Proc. - 25th Brazilian Symp. Softw. Eng. SBES 2011, 2011: pp. 194–203. doi:10.1109/SBES.2011.27.

[120] R. Smith, O. Gotel, RE-O-POLY: A game to introduce lightweight requirements engineering good practices, in: Int. Work. Requir. Eng. Train., 2007: pp. 1–5.

[121] M. Alexander, J. Beatty, Effective design and use of requirements engineering training games, in: 2008 Requir. Eng. Educ. Training, REET'08, IEEE Computer Society, 2008: pp. 18–21. doi:10.1109/REET.2008.2.

[122] R.Q. Gonçalves, Development of a game to support the teaching of Requirements Engineering: The Requirements Island, in: IX Simpósio Bras. Jogos e Entretenimento Digit. - SBGames 2010, 2010: pp. 358–361.

[123] E. Knauss, K. Schneider, K. Stapel, A game for taking requirements engineering more seriously, in: 2008 3rd Int. Work. Multimed. Enjoyable Requir. Eng. MERE'08, IEEE Computer Society, 2008: pp. 22–26. doi:10.1109/MERE.2008.1.

[124] J.M. Fernandes, S.M. Sousa, PlayScrum - A card game to learn the scrum agile method, in: 2nd Int. Conf. Games Virtual Worlds Serious Appl. VS-GAMES 2010, 2010: pp. 52–59. doi:10.1109/VS-GAMES.2010.24.

[125] S. Subhash, E.A. Cudney, Gamified learning in higher education: A systematic review of the literature, Comput. Human Behav. 87 (2018) 192–206. doi:10.1016/j.chb.2018.05.028.

[126] A.P. Markopoulos, A. Fragkou, P.D. Kasidiaris, J.P. Davim, Gamification in engineering education and professional training, Int. J. Mech. Eng. Educ. 43 (2015) 118–131. doi:10.1177/0306419015591324.

[127] D. Strmečki, A. Bernik, D. Radošević, Gamification in e-learning: Introducing gamified design elements into e-learning systems, J. Comput. Sci. 11 (2015) 1108–1117. doi:10.3844/jcssp.2015.1108.1117.

[128] F. Llorens-Largo, F.J. Gallego-Duran, C.J. Villagra-Arnedo, P. Compan-Rosique, R. Satorre-Cuerda, R. Molina-Carmona, Gamification of the Learning Process: Lessons Learned, Rev. Iberoam. Tecnol. Del Aprendiz. 11 (2016) 227–234. doi:10.1109/RITA.2016.2619138.

[129] D. Lopez-Fernandez, A. Gordillo, P.P. Alarcon, E. Tovar, Comparing Traditional Teaching and Game-Based Learning Using Teacher-Authored Games on Computer Science Education, IEEE Trans. Educ. 64 (2021) 367–373. doi:10.1109/TE.2021.3057849.

[130] J. Dias, Teaching operations research to undergraduate management students: The role of gamification, Int. J. Manag. Educ. 15 (2017) 98–111. doi:10.1016/j.ijme.2017.01.002.

[131] H. Caton, D. Greenhill, Rewards and penalties: A gamification approach for increasing attendance and engagement in an undergraduate computing module, Int. J. Game-Based Learn. 4 (2014) 1–12. doi:10.4018/ijgbl.2014070101.

[132] P. Fotaris, T. Mastoras, R. Leinfellner, Y. Rosunally, Climbing up the leaderboard: An empirical study of applying gamification techniques to a computer programming class, Electron. J. e-Learning. 14 (2016) 94–110. www.ejel.org (accessed August 27, 2022).

[133] D. Jaccard, K.E. Bonnier, M. Hellström, How might serious games trigger a transformation in project management education ? Lessons learned from 10 Years of experimentations, Proj. Leadersh. Soc. 3 (2022). doi:10.1016/j.plas.2022.100047.

[134] S.A. Andari, O. Cardin, P. Berruet, R. Ardi, Serious Game as Learning Media for Reconfigurable Manufacturing System: A Model Development, in: IFAC-PapersOnLine, 2022: pp. 1675–1680. doi:10.1016/j.ifacol.2022.09.638.

[135] I. García, C. Pacheco, A. León, J.A. Calvo-Manzano, A serious game for teaching the fundamentals of ISO/IEC/IEEE 29148 systems and software engineering – Lifecycle processes – Requirements engineering at undergraduate level, Comput. Stand. Interfaces. 67 (2020). doi:10.1016/j.csi.2019.103377.

[136] M.E. Morales-Trujillo, G.A. García-Mireles, P. Maslova, What can go wrong in a software project? Have fun solving it, Proc. 2018 Fed. Conf. Comput. Sci. Inf. Syst. FedCSIS 2018. 15 (2018) 995–1004. doi:10.15439/2018F300.

[137] W.L. Lee, SCRUM-X: An interactive and experiential learning platform for teaching Scrum, in: 7th Int. Multi-Conference Complexity, Informatics Cybern. IMCIC 2016 7th Int. Conf. Soc. Inf. Technol. ICSIT 2016 - Proc., 2016: pp. 192–197. https://ink.library.smu.edu.sg/sis_research/3378 (accessed November 10, 2021).

[138] J. Fernandes, D. Duarte, C. Ribeiro, C. Farinha, J.M. Pereira, M.M. Da Silva, IThink : A game-based approach towards improving collaboration and participation in requirement elicitation, in: Procedia Comput. Sci., Elsevier, 2012: pp. 66–77. doi:10.1016/j.procs.2012.10.059.

[139] M. Paasivaara, V. Heikkilä, C. Lassenius, T. Toivola, Teaching students scrum using LEGO blocks, in: 36th Int. Conf. Softw. Eng. ICSE Companion 2014 - Proc., Association for Computing Machinery, 2014: pp. 382–391. doi:10.1145/2591062.2591169.

[140] T. Hanghøj, C.E. Brund, Teacher roles and positionings in relation to educational games, in: 4th Eur. Conf. Games Based Learn. 2010, ECGBL 2010, 2010: pp. 116–122. doi:10.2307/jj.608141.9.

[141] M. Romero, M. Usart, Serious games integration in an entrepreneurship Massive Online Open Course (MOOC), in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 2013: pp. 212–225. doi:10.1007/978-3-642-40790-1_21.

[142] M. Klemetti, O. Taimisto, P. Karppinen, The Attitudes of Finnish School Teachers Towards Commercial Educational Games, in: Des. Use Serious Games, 2008: pp. 97–105. doi:10.1007/978-1-4020-9496-5_7.

[143] E. Jääskä, K. Aaltonen, Teachers' experiences of using game-based learning methods in project management higher education, Proj. Leadersh. Soc. 3 (2022). doi:10.1016/j.plas.2022.100041.

[144] S. Björk, J. Holopainen, Patterns in Game Design, in: Game Des. Read. A Rules Play Anthol., Charles River Media, 2005: pp. 410–437.

[145] I. Salman, A.T. Misirli, N. Juristo, Are students representatives of professionals in software engineering experiments?, in: Proc. - Int. Conf. Softw. Eng., IEEE Computer Society, 2015: pp. 666–676. doi:10.1109/ICSE.2015.82.

[146] D. Falessi, N. Juristo, C. Wohlin, B. Turhan, J. Münch, A. Jedlitschka, M. Oivo, Empirical software engineering experts on the use of students and professionals in experiments, Empir. Softw. Eng. 23 (2018) 452–489. doi:10.1007/S10664-017-9523-3/TABLES/3.

[147] D. Checa, I. Miguel-Alonso, A. Bustillo, Immersive virtual-reality computer-assembly serious game to enhance autonomous learning, Virtual Real. 27 (2023) 3301–3318. doi:10.1007/s10055-021-00607-1.

[148] S. Xinogalos, M. Satratzemi, Special Issue on New Challenges in Serious Game Design, Appl. Sci. 13 (2023). doi:10.3390/app13137675.

[149] D. Abbott, O. Chatzifoti, J. Craven, Serious Game Rapid Online Co-design to Facilitate Change Within Education, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 2021: pp. 233–238. doi:10.1007/978-3-030-92182-8_22.

[150] W. Westera, The devil's advocate: identifying persistent problems in serious game design, Int. J. Serious Games. 9 (2022) 115–124. doi:10.17083/ijsg.v9i3.547.

[151] C. Schade, How to Evaluate Serious Games Concepts? A Systematic Prototyping and Testing Approach, in: Proc. Eur. Conf. Games-Based Learn., 2022: pp. 518–525. doi:10.34190/ecgbl.16.1.515.

[152] N. Zagalo, A.P. Oliveira, P. Cardoso, M. Vairinhos, BEATS & UNITS: A STORY-GAME DESIGN FRAMEWORK, Int. J. Film Media Arts. 8 (2023) 52–67. doi:10.24140/ijfma.v8.n1.03.

[153] K. Peffers, T. Tuunanen, M.A. Rothenberger, S. Chatterjee, A design science research methodology for information systems research, J. Manag. Inf. Syst. 24 (2007) 45–77. doi:10.2753/MIS0742-1222240302.

[154] L.W. de Oliveira, S.S. de Oliveira, M.A. Vieira, S. de L.V. dos Santos, S.T. de Carvalho, Design Science Research: Balancing Science and Art in Building a Game Applied to Health, J. Interact. Syst. 13 (2022) 127–138. doi:10.5753/jis.2022.2682.

[155] M. Huseynli, U. Bub, M.C. Ogbuachi, Development of a Method for the Engineering of Digital Innovation Using Design Science Research, Inf. 13 (2022). doi:10.3390/info13120573.

[156] M.W. Easterday, E.M. Gerber, D.G. Rees Lewis, THE PREMISES OF DESIGN RESEARCH, in: Learn. Sci. Conversat. Theor. Methodol. Bound. Spaces, 2022: pp. 76–92. doi:10.4324/9781003089728-11.

[157] I. Bider, E. Perjons, Design science in action: developing a modeling technique for eliciting requirements on business process management (BPM) tools, Softw. Syst. Model. 14 (2015) 1159–1188. doi:10.1007/s10270-014-0412-6.

[158] M.B. Pankowska, Design science research for non-formal education, in: Handb. Res. Enhancing Innov. High. Educ. Institutions, 2020: pp. 587–604. doi:10.4018/978-1-7998-2708-5.ch025.

[159] J.O. De Sordi, Design Science Research Method, in: Des. Sci. Res. Methodol., Springer International Publishing, 2021: pp. 59–78. doi:10.1007/978-3-030-82156-2_5.

[160] F.G.M. Silva, Practical methodology for the design of educational serious games, Inf. 11 (2020). doi:10.3390/info11010014.

[161] L.K.P. Suryapranata, F.L. Gaol, B. Soewito, H.L.H.S. Warnars, G.P. Kusuma, Quality measurement for serious games, in: Proc. - 2017 Int. Conf. Appl. Comput. Commun. Technol. ComCom 2017, 2017: pp. 1–4. doi:10.1109/COMCOM.2017.8167098.

[162] H.M. Wantu, N. Djafri, N. Lamatenggo, M.K. Umar, Learning Models: A Literature Review, Int. J. Soc. Sci. Hum. Res. 06 (2023). doi:10.47191/ijsshr/v6-i5-31.

[163] V.S. Chouhan, S. Srivastava, Understanding Competencies and Competency Modeling — A Literature Survey, IOSR J. Bus. Manag. 16 (2014) 14–22. doi:10.9790/487x-16111422.

[164] M.N. Bin Nordin, S.S. Maidin, M. Rajoo, M. Mahmod, W.N.F.A. Jani, M.P. Yusoh, M.M. Magiman, N.F. Mosbiran, International Frameworks For 21st Century Competences: Comparative Education, Res Mil. 12 (2022) 7332–7344.

[165] K.F. Geisinger, Empirical Considerations on Intelligence Testing and Models of Intelligence: Updates for Educational Measurement Professionals, Appl. Meas. Educ. 32 (2019) 193–197. doi:10.1080/08957347.2019.1619564.

[166] D.R. Krathwohl, A revision of bloom's taxonomy: An overview, Theory Pract. 41 (2002) 212–218. doi:10.1207/s15430421tip4104_2.

[167] G.F.T. and L.E.N. Jeanine Krath, Maximilian Altmeyer, Gamified UK User Type Quiz - HEXAD 12, (2023). https://gamified.uk/UserTypeTest2023/user-type-test.php?lid=.

[168] A. Barczak, H. Woźniak, Comparative Study on Game Engines, Stud. Inform. (2020) 5–24. doi:10.34739/si.2019.23.01.

[169] K. Fleming, J. Abad, L. Booth, L. Schueller, A. Baills, A. Scolobig, B. Petrovic, G. Zuccaro, M.F. Leone, The use of serious games in engaging stakeholders for disaster risk reduction, management and climate change adaption information elicitation, Int. J. Disaster Risk Reduct. 49 (2020). doi:10.1016/j.ijdrr.2020.101669.

[170] N. Zagalo, A.P. Oliveira, P. Cardoso, Beats and Units Framework: A Story-Game Integration Framework for the Ideation Stage of Narrative Design of Serious

Games, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 2021: pp. 172–176. doi:10.1007/978-3-030-92300-6_15.

[171] W. Thomas, A framework for the selection and design of serious games, in: Proc. 14th Int. Conf. Game Based Learn. ECGBL 2020, 2020: pp. 613–619. doi:10.34190/GBL.20.034.

[172] F. Archuby, C. Sanz, C. Manresa-Yee, DIJS: Methodology for the Design and Development of Digital Educational Serious Games, IEEE Trans. Games. 15 (2023) 273–284. doi:10.1109/TG.2022.3217737.

[173] M. Carrión-Toro, M. Santorum, P. Acosta-Vargas, J. Aguilar, M. Pérez, iPlus a user-centered methodology for serious games design, Appl. Sci. 10 (2020) 1–33. doi:10.3390/app10249007.

[174] G.C. Natucci, M.A.F. Borges, The Experience, Dynamics and Artifacts Framework: Towards a Holistic Model for Designing Serious and Entertainment Games, in: IEEE Conf. Comput. Intell. Games, CIG, 2021. doi:10.1109/CoG52621.2021.9619144.

[175] J. Wiemeyer, Towards a generic framework for serious games, in: Adv. Intell. Syst. Comput., 2020: pp. 193–200. doi:10.1007/978-3-030-35048-2_23.

[176] F. de Rosa, A. De Gloria, Design methodology of analytical games for knowledge acquisition, Int. J. Serious Games. 8 (2021) 3–23. doi:10.17083/IJSG.V8I4.456.

[177] D. Abbott, O. Chatzifoti, J. Ferguson, S. Louchart, S. Stals, Serious "Slow" Game Jam - A Game Jam Model for Serious Game Design, in: ACM Int. Conf. Proceeding Ser., 2023: pp. 28–36. doi:10.1145/3610602.3610604.

[178] H.K.H. Abd El-Sattar, A New Learning Theory-based Framework for Combining Flow State with Game Elements to Promote Engagement and Learning in Serious Games, Inf. Sci. Lett. 12 (2023) 2663–2677. doi:10.18576/isl/120660.

[179] V. Garneli, K. Patiniotis, K. Chorianopoulos, Designing multiplayer serious games with science content, Multimodal Technol. Interact. 5 (2021). doi:10.3390/mti5030008.

[180] Y. Ma, F. Vallet, F. Cluzel, B. Yannou, Methodological framework for making the transition from traditional innovation teaching towards serious games, in: Proc. ASME Des. Eng. Tech. Conf., 2020. doi:10.1115/DETC2020-22754.

[181] F. Khorram, M. Taromirad, R. Ramsin, SeGa4Biz: Model-Driven Framework for Developing Serious Games for Business Processes, in: Int. Conf. Model. Eng. Softw. Dev., 2021: pp. 139–146. doi:10.5220/0010198801390146.

[182] B. Lemoine, P. Laforcade, S. George, An Analysis Framework for Designing

Declarative Knowledge Training Games Using Roguelite Genre, in: Int. Conf. Comput. Support. Educ. CSEDU - Proc., 2023: pp. 276–287. doi:10.5220/0011840200003470.

[183] N. Flores, A.C.R. Paiva, N. Cruz, Teaching software engineering topics through pedagogical game design patterns: An empirical study, Inf. 11 (2020). doi:10.3390/info11030153.

[184] G.C. Natucci, M.A.F. Borges, Bridging emotional design and serious games: Towards affective learning design patterns, in: Proc. Eur. Conf. Games-Based Learn., 2021: pp. 863–871. doi:10.34190/GBL.21.052.

[185] N.N. Blackburn, R.E. Cardona-Rivera, OGrES Welcome! Toward a systematic theory for serious game design, in: CHI Play 2021 - Ext. Abstr. 2021 Annu. Symp. Comput. Interact. Play, 2021: pp. 242–248. doi:10.1145/3450337.3483460.

[186] R. Srbecky, M. Krapf, B. Wallenborn, M. Then, M. Hemmje, Realization of a framework for game-based learning units using a unity environment, in: Proc. Eur. Conf. Games-Based Learn., 2021: pp. 682–691. doi:10.34190/GBL.21.057.

[187] C. Fabricatore, D. Gyaurov, X. Lopez, Rethinking Serious Games Design in the Age of COVID-19: Setting the Focus on Wicked Problems, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 2020: pp. 243–259. doi:10.1007/978-3-030-61814-8_19.

[188] A. Theodoropoulos, V. Poulopoulos, G. Lepouras, Towards a Framework for Adaptive Gameplay in Serious Games that Teach Programming: Association Between Computational Thinking and Cognitive Style, in: Adv. Intell. Syst. Comput., 2021: pp. 530–541. doi:10.1007/978-3-030-68198-2_49.

[189] T. Marsh, A. Thomas, E.T. Khoo, Between Game Mechanics and Immersive Storytelling: Design Using an Extended Activity Theory Framework, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 2021: pp. 113–128. doi:10.1007/978-3-030-88272-3_9.

[190] I. Kniestedt, I. Lefter, S. Lukosch, F.M. Brazier, Re-framing engagement for applied games: A conceptual framework, Entertain. Comput. 41 (2022). doi:10.1016/j.entcom.2021.100475.

[191] B.K. Walther, L.J. Larsen, Reflections on ludification: Approaching a conceptual framework – and discussing inherent challenges, Int. J. Serious Games. 8 (2021) 115–127. doi:10.17083/IJSG.V8I3.436.

[192] J. Othlinghaus-Wulhorst, H.U. Hoppe, A Technical and Conceptual Framework for Serious Role-Playing Games in the Area of Social Skill Training, Front. Comput. Sci. 2 (2020). doi:10.3389/fcomp.2020.00028.

[193] N. Kosoris, B. Liu, B. Fain, A Framework for XReality Serious Games, in: Conf.

Proc. - IEEE SOUTHEASTCON, 2022: pp. 793–798. doi:10.1109/SoutheastCon48659.2022.9764129.

[194] R. Hunicke, M. Leblanc, R. Zubek, MDA: A formal approach to game design and game research, in: AAAI Work. - Tech. Rep., 2004: pp. 1–5.

[195] B.M. Winn, The Design, Play, and Experience Framework, in: Handb. Res. Eff. Electron. Gaming Educ., IGI Global, 2011: pp. 1010–1024. doi:10.4018/978-1-59904-808-6.ch058.

[196] F.D. Davis, A technology acceptance model for empirically testing new end-user information systems: Theory and results, Management. Ph.D. (1985) 291. doi:oclc/56932490.

[197] T.S. Andre, C.M. Fidopiastis, T.R. Ripley, A.L. Oskorus, R.E. Meyer, R.A. Snyder, Augmented cognition methods for evaluating serious game based insider cyber threat detection training, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), Springer, Berlin, Heidelberg, 2011: pp. 395–403. doi:10.1007/978-3-642-21852-1_46.

[198] U. Aydan, M. Yilmaz, P.M. Clarke, R. V. O'Connor, Teaching ISO/IEC 12207 software lifecycle processes: A serious game approach, Comput. Stand. Interfaces. 54 (2017) 129–138. doi:10.1016/j.csi.2016.11.014.

[199] A. Calderón, M. Ruiz, R. V. O'Connor, Coverage of ISO/IEC 29110 project management process of basic profile by a serious game, in: Commun. Comput. Inf. Sci., Springer, Cham, 2017: pp. 111–122. doi:10.1007/978-3-319-64218-5_9.

[200] A. Calderón, M. Ruiz, R. V. O'Connor, Coverage of the ISO 21500 standard in the context of software project management by a simulation-based serious game, in: Commun. Comput. Inf. Sci., Springer, Cham, 2017: pp. 399–412. doi:10.1007/978-3-319-67383-7_29.

[201] A. Calderón, M. Ruiz, R. V. O'Connor, A serious game to support the ISO 21500 standard education in the context of software project management, Comput. Stand. Interfaces. 60 (2018) 80–92. doi:10.1016/j.csi.2018.04.012.

[202] T.E. Gasiba, U. Lechner, M. Pinto-Albuquerque, Cybersecurity challenges in industry: Measuring the challenge solve time to inform future challenges, Inf. 11 (2020) 1–31. doi:10.3390/info11110533.

[203] I. García, C. Pacheco, A. León, J.A. Calvo-Manzano, A serious game for teaching the fundamentals of ISO/IEC/IEEE 29148 systems and software engineering – Lifecycle processes – Requirements engineering at undergraduate level, Comput. Stand. Interfaces. 67 (2020) 103377. doi:10.1016/j.csi.2019.103377.

[204] T. Espinha Gasiba, U. Lechner, M. Pinto-Albuquerque, Sifu - a cybersecurity awareness platform with challenge assessment and intelligent coach,

Cybersecurity. 3 (2020) 1–23. doi:10.1186/s42400-020-00064-4.

[205] D. Gelperin, C. Enterprises, Improve requirements understanding by playing cooperative games, Proc. INCOSE Int. …. (2011). http://www.sqetraining.com/sites/default/files/articles/3544311.pdf (accessed November 10, 2021).

[206] V.T. Heikkilä, M. Paasivaara, C. Lassenius, Teaching university students Kanban with a collaborative board game, in: Proc. - Int. Conf. Softw. Eng., IEEE Computer Society, 2016: pp. 471–480. doi:10.1145/2889160.2889201.

[207] D.E.H. Lovgren, J. Li, T.D. Oyetoyan, A data-driven security game to facilitate information security education, in: Proc. - 2019 IEEE/ACM 41st Int. Conf. Softw. Eng. Companion, ICSE-Companion 2019, Institute of Electrical and Electronics Engineers Inc., 2019: pp. 256–257. doi:10.1109/ICSE-Companion.2019.00102.

[208] B. Kutun, W. Schmidt, BPMN wheel: Board game for business process modelling, in: Proc. Eur. Conf. Games-Based Learn., 2019: pp. 1008–1012. doi:10.34190/GBL.19.012.

[209] Z. Li, L. O'Brien, S. Flint, R. Sankaranarayana, Object-oriented Sokoban solver: A serious game project for OOAD and AI education, Proc. - Front. Educ. Conf. FIE. 2015-Febru (2015). doi:10.1109/FIE.2014.7044115.

[210] T. Margaria, A.-L. Lamprecht, Modeling of Games and Game Strategies, Encycl. Educ. Inf. Technol. (2020) 1–12. doi:10.1007/978-3-319-60013-0_209-1.

[211] M.A. Miljanovic, J.S. Bradbury, RoboBUG: A serious game for learning debugging techniques, ICER 2017 - Proc. 2017 ACM Conf. Int. Comput. Educ. Res. (2017) 93–100. doi:10.1145/3105726.3106173.

[212] T. Orehovački, S. Babić, Inspecting quality of games designed for learning programming, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), Springer, Cham, 2015: pp. 620–631. doi:10.1007/978-3-319-20609-7_58.

[213] T. Turchi, D. Fogli, A. Malizia, Fostering computational thinking through collaborative game-based learning, Multimed. Tools Appl. 78 (2019) 13649–13673. doi:10.1007/s11042-019-7229-9.

[214] V. Schiller, ENC#YPTED: An Educational Game for Programming in the Unity Engine, in: Conf. Hum. Factors Comput. Syst. - Proc., Association for Computing Machinery, 2021. doi:10.1145/3411763.3451852.

[215] A.D. De Souza, R.D. Seabra, J.M. Ribeiro, L.E.D.S. Rodrigues, SCRUMI: A Board serious virtual game for teaching the SCRUM framework, in: Proc. - 2017 IEEE/ACM 39th Int. Conf. Softw. Eng. Companion, ICSE-C 2017, Institute of Electrical and Electronics Engineers Inc., 2017: pp. 319–321. doi:10.1109/ICSE-

C.2017.124.

[216] N. Naik, P. Jenkins, Relax, it's a game: Utilising gamification in learning agile scrum software development, in: IEEE Conf. Comput. Intell. Games, CIG, IEEE Computer Society, 2019. doi:10.1109/CIG.2019.8848104.

[217] J. Noll, A. Butterfield, K. Farrell, T. Mason, M. McGuire, R. McKinley, GSD Sim: A Global Software Development Game, in: Proc. - Int. Comput. Softw. Appl. Conf., IEEE Computer Society, 2014: pp. 15–20. doi:10.1109/ICGSEW.2014.12.

[218] A. Jain, B. Boehm, SimVBSE: Developing a game for value-based software engineering, in: Softw. Eng. Educ. Conf. Proc., 2006: pp. 103–111. doi:10.1109/CSEET.2006.31.

[219] K. Shaw, J. Dermoudy, Engendering an Empathy for Software Engineering, in: Conf. Res. Pract. Inf. Technol. Ser., 2005: pp. 135–144.

[220] C.G. Von Wangenheim, R. Savi, A.F. Borgatto, DELIVER! - An educational game for teaching Earned Value Management in computing courses, Inf. Softw. Technol. 54 (2012) 286–298. doi:10.1016/j.infsof.2011.10.005.

[221] A. Droubay, D. Bouchard, Pointer Attack: Lessons Learned in Computer Concept Gamification, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), Springer, Cham, 2020: pp. 179–188. doi:10.1007/978-3-030-63464-3_17.

[222] L.S. Ferro, F. Sapio, Another week at the office (awato) – an interactive serious game for threat modeling human factors, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), Springer, Cham, 2020: pp. 123–142. doi:10.1007/978-3-030-50309-3_9.

[223] L. Goeke, A. Quintanar, K. Beckers, S. Pape, PROTECT – An Easy Configurable Serious Game to Train Employees Against Social Engineering Attacks, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 2020: pp. 156–171. doi:10.1007/978-3-030-42051-2_11.

[224] C. Jordan, M. Knapp, D. Mitchell, M. Claypool, K. Fisler, CounterMeasures: A game for teaching computer security, in: Annu. Work. Netw. Syst. Support Games, 2011. doi:10.1109/NetGames.2011.6080983.

[225] A. Baker, E. Oh Navarro, A. van der Hoek, Teaching Software Engineering using Simulation Games, in: Int. Conf. Simul. Educ. (ICSiE '04), 2003: pp. 9–14. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.124.2661&rep=rep1&type=pdf (accessed November 10, 2021).

[226] R. Luh, M. Temper, S. Tjoa, S. Schrittwieser, H. Janicke, PenQuest: a gamified attacker/defender meta model for cyber security assessment and education, J. Comput. Virol. Hacking Tech. 16 (2020) 19–61. doi:10.1007/s11416-019-00342-x.

[227] K. Beckers, S. Pape, A Serious Game for Eliciting Social Engineering Security Requirements, in: Proc. - 2016 IEEE 24th Int. Requir. Eng. Conf. RE 2016, Institute of Electrical and Electronics Engineers Inc., 2016: pp. 16–25. doi:10.1109/RE.2016.39.

[228] W. Eugene, T. Barnes, J. Wilson, Math fluency through game design, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), Springer, Berlin, Heidelberg, 2013: pp. 189–198. doi:10.1007/978-3-642-39241-2_22.

[229] T. Hainey, G. Baxter, A. Stanton, A serious game to teach rudimentary programming: Investigating content integration, in: Proc. Eur. Conf. Games-Based Learn., Dechema e.V., 2019: pp. 298–307. doi:10.34190/GBL.19.145.

[230] O. Leifler, L. Lindblom, M. Svensson, M. Gramfalt, A. Jonsson, Teaching sustainability, ethics and scientific writing: An integrated approach, in: Proc. - Front. Educ. Conf. FIE, Institute of Electrical and Electronics Engineers Inc., 2020. doi:10.1109/FIE44824.2020.9274220.

[231] C.S. Longstreet, K.M.L. Cooper, Developing a meta-model for serious games in higher education, Proc. 12th IEEE Int. Conf. Adv. Learn. Technol. ICALT 2012. (2012) 684–685. doi:10.1109/ICALT.2012.43.

[232] A. Przybyłek, M. Zakrzewski, Adopting collaborative games into agile requirements engineering, ENASE 2018 - Proc. 13th Int. Conf. Eval. Nov. Approaches to Softw. Eng. 2018-March (2018) 54–64. doi:10.5220/0006681900540064.

[233] M.L. Sanchez-Gordón, R. V. O'Connor, R. Colomo-Palacios, E. Herranz, Bridging the gap between SPI and SMES in educational settings: A learning tool supporting ISO/IEC 29110, in: Commun. Comput. Inf. Sci., Springer, Cham, 2016: pp. 3–14. doi:10.1007/978-3-319-44817-6_1.

[234] A. Stiegler, G. Zimmermann, Gamification in the development of accessible software, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), Springer, Cham, 2014: pp. 171–180. doi:10.1007/978-3-319-07437-5_17.

[235] Ç. Üsfekes, E. Tuzun, M. Yilmaz, Y. MacIt, P. Clarke, Auction-based serious game for bug tracking, IET Softw. 13 (2019) 386–392. doi:10.1049/iet-sen.2018.5144.

[236] D. Valencia, A. Vizcaíno, J.P. Soto, M. Piattini, A serious game to improve students' skills in global software development, in: CSEDU 2016 - Proc. 8th Int. Conf. Comput. Support. Educ., 2016: pp. 470–475. doi:10.5220/0005895904700475.

[237] M. Yilmaz, Virtual Reality-Based Daily Scrum Meetings, in: Encycl. Comput.

Graph. Games, 2018: pp. 1–6. doi:10.1007/978-3-319-08234-9_160-1.

[238] I.S. Elgrably, S.R.B. Oliveira, Gamification and evaluation of the use the agile tests in software quality subjects: The application of case studies, in: ENASE 2018 - Proc. 13th Int. Conf. Eval. Nov. Approaches to Softw. Eng., 2018: pp. 416–423. doi:10.5220/0006800304160423.

[239] T. Gasiba, U. Lechner, M. Pinto-Albuquerque, A. Porwal, Cybersecurity Awareness Platform with Virtual Coach and Automated Challenge Assessment, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), Springer, Cham, 2020: pp. 67–83. doi:10.1007/978-3-030-64330-0_5.

[240] A. Kabil, T. Duval, N. Cuppens, Alert characterization by non-expert users in a cybersecurity virtual environment: A usability study, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), Springer, Cham, 2020: pp. 82–101. doi:10.1007/978-3-030-58465-8_6.

[241] K. Magylaitė, L. Čeponienė, M. Jurgelaitis, T. Danikauskas, A Case Study of Applying Gamification in Teaching Project Management, in: Commun. Comput. Inf. Sci., Springer, Cham, 2020: pp. 321–333. doi:10.1007/978-3-030-59506-7_26.

[242] M.O. Ahmad, K. Liukkunen, Software factory project for enhancement of student experiential learning, in: 16th Int. Conf. Cogn. Explor. Learn. Digit. Age, CELDA 2019, International Association for the Development of the Information Society. e-mail: secretariat@iadis.org; Web site: http://www.iadisportal.org, 2019: pp. 297–306. doi:10.33965/celda2019_201911l037.

[243] M.L. Sanchez-Gordón, R. V. O'Connor, R. Colomo-Palacios, E. Herranz, Bridging the gap between SPI and SMES in educational settings: A learning tool supporting ISO/IEC 29110, in: Commun. Comput. Inf. Sci., Multimedia Tools and Applications, 2016: pp. 3–14. doi:10.1007/978-3-319-20609-7.

[244] M. Maarek, S. Louchart, L. McGregor, R. McMenemy, Co-created design of a serious game investigation into developer-centred security, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 11385 LNCS (2019) 221–231. doi:10.1007/978-3-030-11548-7_21.

[245] D. Valencia, A. Vizcaíno, J.P. Soto, M. Piattini, A serious game to improve students' skills in global software development, CSEDU 2016 - Proc. 8th Int. Conf. Comput. Support. Educ. 1 (2016) 470–475. doi:10.5220/0005895904700475.

[246] T. Turchi, D. Fogli, A. Malizia, Fostering computational thinking through collaborative game-based learning, Multimed. Tools Appl. 78 (2019) 13649–13673. doi:10.1007/s11042-019-7229-9.

[247] A. Przybyłek, M. Zakrzewski, Adopting collaborative games into agile requirements engineering, in: ENASE 2018 - Proc. 13th Int. Conf. Eval. Nov.

Approaches to Softw. Eng., SciTePress, 2018: pp. 54–64. doi:10.5220/0006681900540064.

[248] M.L. Sanchez-Gordón, R. V O'Connor, R. Colomo-Palacios, E. Herranz, Bridging the gap between SPI and SMES in educational settings: A learning tool supporting ISO/IEC 29110, in: Commun. Comput. Inf. Sci., Springer, Cham, 2016: pp. 3–14. doi:10.1007/978-3-319-44817-6_1.

[249] A. Stiegler, G. Zimmermann, Gamification in the development of accessible software, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 8513 LNCS (2014) 171–180. doi:10.1007/978-3-319-07437-5_17.

9

APPENDICES

9.1 Appendix A

Data Items	Values	RQ
General study ID	The identifier of the article	-
DOI Number	The Digital Object Identifier of the article	
Article title	The name of the article	-
Authors names	The set of authors names	-
Source	The database in which the article was published	-
Publication channel	The name of the publication channel and whether the latter is a journal, a book series, a conference, a workshop, or other..	-
Year of publication	The calendar year	-
Research types	Whether the study is an evaluation research, a validation research, a solution proposal, a philosophical paper, an experience report, or an opinion paper	-
Research method	The size of the samples, the number of groups used in the study, and whether the latter were	-

	experimental groups or control group, the statistical test used, and the outcomes of the evaluations.	
Area in SE	Which SWEBOK Area	RQ1
Bloom's levels	The bloom's levels addressed in SGs	RQ2
Game Dynamics	The game dynamics implemented in SGs	RQ2
Game-related approaches	The game mechanics and the gamification elements implemented in SGs	RQ3
Learning mechanics	The learning mechanics implemented in SGs	RQ3
Player profile	The kind of player profiles addressed in SGs for SE, the classification adopted for the player profiles in SGs for SE, the characteristics of the player profiles identified, the adaptations of the gameplay to match the player profiles	RQ4

Table 9: Data Items

Research Process			
Need for the map		Motivate the need and relevance	Yes
		Define objectives and questions	Yes
		Consult target audience to define questions	No
Study ident	Choosing search strategy	Database search	Yes
		Snowballing	Yes
		Manual search	Yes
	Develop search	PICO	Yes
		Consult experts	Yes
		Iteratively improve	Yes

	Search	Keywords from known papers	Yes
		Use standards, encyclopedias	Yes
	Search evaluation	Paper test-set	No
		Expert evaluation	No
		Authors' web pages	No
		Test–retest	No
	Inclusion/exclusion	Identify objective criteria for decision	Yes
		Resolve disagreements among multiple researchers	Yes
		Decision rules	Yes
Extr./Class.	Extraction process	Identify objective criteria for decision	Yes
		Obscuring information that could bias	Yes
		Resolve disagreements among multiple researchers	Yes
	Topic-independent	Research type	Yes
		Research method	Yes
		Study focus	Yes
		Contribution type	No
		Venue type	Yes
	Topic-specific	Emerging scheme	No
		Use of standards, etc.	No
Study validity	Discussion of threats		Yes
	Line diagram		Yes
	Pie diagram		Yes
	Bar plot		Yes
	Bubble plot		Yes
	Venn diagram		No
	Heatmap		Yes

Table 10: Study Assessment Checklist

9.2 *Appendix B*

	SG	Area in SE	Environment	Pricing	Focus
SGs Solutions	XL-CITR [197]	SM		NR	Professional
	Floors [198]	CF		NR	
	Secret Ninja Testing [84]	SEM	Standalone	NR	Professional
	ProDec [86] [199] [200] [201]	SEM	Standalone	NR	Academic
	CSC [202]	SR	Non-digital	NR	Professional
	PlayScrum [124]	SEP	Non-digital	NR	Academic
	iThink [138]	SEP	Non-digital	NR	Academic
	Planning poker [32]	SEP	Standalone	NR	Professional
	Requengin [203]	SEP	Standalone	NR	
	Sifu [204]	SEP	Standalone	NR	
	RU Ouija [205]			NR	
	CENGO [36]	SR	Web	NR	Academic
	Collaborative Kanban Board Game [206]	SR	Non-digital	NR	Professional
	DdSG [207]	CF	Web	NR	Academic
	The Essence of Software Engineering – The Board Game [35]	CF	IoT	NR	Academic
	BPMN Wheel [208]	SEM	Non-digital	NR	Academic
	Classutopia [80]	SEM	Non-digital	Paid	Academic
	Object-Oriented Sokoban Solver [209]	SEPP	Standalone	NR	Academic
	SimSYS GDP [85]	SEPP	Web	NR	

	ConnectIT [210]	MF		NR	
	Robot ON! [72]	SR		NR	
	RoboBUG [211]	SD	Mobile	NR	
	White Crow PM [136]	SEPP	Standalone	NR	Academic
	Light bot & Code Combat [212]	SQ	Standalone	NR	
	Go For It! [38]	ST	Standalone	NR	Academic
	TAPASPlay [213]	SR	Non-digital	NR	Professional
	Ethical_Dilemmas [73]	CF		NR	
	ScrumGame [77]	SEM	Mobile	NR	Professional
	ENC#YPTED [214]	SEM	Standalone	NR	Academic
	SCRUMI [215]	SEM	Web	NR	Academic
	It'sa Game [216]	SEM	Web	NR	Academic
	SCRUMIA [37]	SEM	Non-digital	NR	Academic
	SCRUM-X [137]	SEM	Standalone	NR	Professional
	GSD Sim [217]	SEM	Standalone	NR	Academic
	SimVBSE [218]	SEM		NR	Academic
	SimjavaSP [219]	SD	Web	NR	Academic
	SESAM [39]	SEM	Standalone	NR	Academic
	DELIVER! [220]	SEM	Non-digital	NR	Academic
	Become a Scrum Master[74]	SEM	IoT	NR	Academic
	Pointer Attack [221]	CF	Standalone	NR	Academic

	Another Week at the Office (AWATO) [222]	SEPP	Standalone	NR	Academic
	PROTECT [223]	SEPP	Standalone	NR	Professional
	CounterMeasures [224]	CF		NR	Academic
	SimSE [225]	SEP	Standalone	NR	Academic
	GSD-Aware [33]	SEPP	Standalone	NR	Academic
	PenQuest [226]	SEPP	Non-digital	NR	Academic
SGs Approaches	NR [27]	SEM	Non-digital	NR	Academic
	NR [227]	SR	Standalone	NR	Academic
	NR [71]	SEP	Standalone	NR	Academic
	NR [228]	SEP	Standalone	NR	Professional
	NR [34]		Standalone	NR	Academic
	NR [229]	SEM	Standalone	NR	Academic
	NR [230]	CF		NR	
	NR [231]	SQ		NR	
	NR [83]	SD		NR	
	NR [232]	CF	Standalone	NR	Academic
	NR [233]	SEM	Non-digital	NR	Academic
	NR [234]	CF	Mobile	NR	Academic
	NR [235]	ST	Web	NR	Professional
	NR [236]	SEM		NR	
	NR [237]	SEPP	Web	NR	Professional

	NR [76]	SEM	Web	NR	Professional
	NR [238]	SEM	Non-digital	NR	Academic
	NR [79]	SEM	Standalone	Free	Academic
	NR [139]	SEM		NR	Academic
	NR [75]	SR	Standalone	NR	Professional
	NR [239]	SEPP	Standalone	NR	Professional
	NR [70]	ST	Plugin	NR	Academic
	NR [240]	SEPP	Standalone	NR	Other
	NR [241]	SEM	Web	NR	Academic
	NR [242]	SEM	Non-digital	NR	Other

Table 11: SGs in SE List

SM: Software maintenance, CF: Computing foundations, SEM: Software engineering management, SR: Software requirements, SEP: Software engineering process, SEPP: Software engineering professional practice, MF: Mathematical foundations, SD: Software design, SQ: Software quality, ST: Software testing.

9.3 *Appendix C*

Research Type	Selected Studies	Total
Evaluation research	[227] [202] [203] [243] [237] [77] [76] [75] [239] [223]	10
Experience paper	[242] [199] [200]	3
Opinion paper	[205] [229]	2
Philosophical paper	[27] [210]	2
Solution proposal	[70] [71] [85] [244] [72] [234] [245] [214] [217] [218] [219] [86]	12
Validation research	[197] [198] [84] [201] [228] [124] [138] [32] [204] [34] [36] [206] [207] [35] [208] [80] [230] [209] [231] [211] [136] [212] [232] [38] [246] [73] [215] [216] [238] [79] [37] [137] [139] [33] [224] [225] [69] [220] [74] [221] [222] [240] [241] [226]	44

Table 12: Distribution of selected studies by research type

Environment	Selected Studies	Total
IoT	[246] [74]	2
Mobile	[80] [212] [77]	3
Non Digital	[242] [227] [124] [205] [206] [35] [208] [136] [247] [38] [248] [238] [37] [225] [220] [226]	16
Plugin	[84]	1
Standalone	[70] [198] [199] [200] [201] [32] [203] [204] [207] [209] [72] [211] [245] [73] [237] [214] [79] [137] [33] [217] [39] [86] [75] [221] [222] [239] [223] [240]	28
Web	[202] [138] [36] [83] [235] [215] [216] [76] [219] [241]	10
N/A	[197] [27] [71] [228] [34] [229] [231] [85] [210] [249] [139] [224] [218]	13

Table 13: Distribution of selected studies by Environment

SWEBOK Areas	Selected Studies	Total
Software engineering management	[27] [74] [76] [77][79] [37] [80] [84] [86] [199] [200] [201] [208] [214] [215] [216] [137] [217] [218] [39] [220] [229] [233] [236] [238] [139] [241] [242]	28
Software maintenance	[197]	1

Software testing	[70] [38] [235]	3
Computing foundations	[73] [230] [234] [198] [207] [35] [221] [224] [232]	9
Software requirements	[202] [36] [206] [72] [75] [213] [227]	7
Software design	[83] [211] [219]	3
Software engineering process	[71] [124] [138] [32] [203] [204] [225] [228]	8
Software engineering professional practice	[33] [209] [85] [136] [222] [223] [226] [237] [239] [240]	10
Software quality	[231] [212]	2
Mathematical foundations	[210]	1

Table 14: Distribution of selected studies by SWEBOK area

Publication Channel	Selected Studies	Total
Journal	[198] [201] [202] [32] [203] [204] [213] [37] [220] [33] [226] [27] [235] [75]	14
Conference	[197] [199] [200] [86] [124] [138] [205] [36] [206] [207] [35] [208] [80] [209] [85] [211] [136] [212] [38] [73] [77] [214] [215] [216] [137] [217] [218] [219] [39] [74] [221] [222] [223] E [225] [227] [228] [34] [229] [230] [231] [83] [232] [233] [234] [236] [76] [238] [79] [139] [239] [70] [240] [241] [242]	54
Workshop	[84] [72] [224] [71]	4
Other	[210] [237]	2

Table 15: Distribution of selected studies by publication channels

Success Factors	Selected Studies	Total
Backstory & Production	[197] [70] [84] [202] [36] [35] [208] [83] [72] [211] [136] [38] [213] [236] [214] [216] [238] [37] [137] [33] [225] [217] [218] [219] [39] [86] [75] [74] [221] [222] [223] [240] [241] [226]	34
Realism	[70] [227] [202] [136] [236] [237] [214] [139] [33] [224] [217] [218] [219] [39] [220] [86] [74] [222] [223] [240] [226]	21
AI & Adaptivity	[198] [200] [201] [71] [203] [231] [85] [211] [232] [38] [236] [214] [33] [39] [74] [222] [239]	17
Interaction	[84] [228] [32] [35] [72] [233] [216] [225] [219] [220]	11

Feedback & Debriefing	[83] [72] [211] [136] [232] [38] [237] [77] [214] [215] [216] [238] [37] [137] [139] [33] [218] [219] [39] [220] [221] [239]	21

Table 16: Distribution of selected studies by success factors

Gamification Elements	Selected Studies	Total
Consequences	[225]	1
Voting	[75]	1
Anonymity	[75]	1
Investments	[136] [218]	2
Unlockables	[211] [237]	2
Prizes	[238] [221]	2
Virtual Economy	[136] [33]	2
Innovation Platform	[221] [239]	2
Flow	[38] [233] [214]	3
Fixed Rewards	[83] [38] [235] [74]	4
Theme	[214] [37] [75] [74]	4
Narrative	[83] [238] [37] [75]	4
Development Tools	[232] [38] [214] [224]	4
Badges	[70] [200] [207] [86] [75]	5
Customization	[200] [211] [232] [38] [214] [33]	6
Competition	[200] [136] [213] [77] [215] [238] [220] [75]	8
Learning	[210] [77] [238] [139] [33] [224] [225] [217]	8
Strategy	[83] [210] [73] [77] [238] [37] [225] [217] [220] [223] [226]	11
Quests	[70] [198] [84] [200] [71] [77] [214] [215] [238] [33] [224] [39] [86] [239]	14
Leaderboard	[70] [198] [227] [84] [138] [36] [229] [207] [235] [73] [237] [215] [238] [75]	14
Challenges	[242] [231] [85] [212] [236] [237] [238] [37] [139] [33] [224] [75] [74] [221] [239]	15
Level/Progression	[70] [200] [231] [83] [72] [211] [232] [38] [236] [77] [214] [76] [224] [225] [220] [86] [75] [221] [223]	18
Time Pressure	[200] [72] [136] [232] [38] [235] [236] [237] [77] [215] [216] [37] [139] [39] [220] [86] [74] [241]	19
Time Dependent	[70] [200] [72] [211] [232] [38] [235] [236] [237] [77] [215] [216] [37] [139] [33] [225] [219] [39] [220] [86] [241]	21

Progress/feedback	[227] [200] [27] [229] [35] [83] [210] [211] [136] [38] [213] [236] [73] [214] [215] [216] [137] [139] [225] [39] [220] [86] [74] [222] [223] [241]	26
Points	[242] [70] [198] [227] [84] [200] [124] [138] [36] [229] [207] [35] [231] [83] [136] [232] [38] [233] [213] [235] [236] [237] [238] [33] [224] [39] [75] [222] [223] [241] [226]	31

Table 17: Distribution of selected studies by gamification elements

Game Mechanics	Selected Studies	Total
Design/Editing	[34]	1
Tokens	[236]	1
Pavlovian Interactions	[219]	1
Cooperation	[229] [139]	2
Game turns	[207] [224] [239]	3
Questions & Answers	[214] [37] [75]	3
Competition	[197] [199] [214] [215] [224]	5
Progression	[34] [36] [33] [218] [39]	5
Assessment	[231] [136] [215] [76] [217] [39]	6
Strategy Planning	[36] [216] [137] [218] [39] [220] [240]	7
Collaboration	[199] [34] [234] [235] [215] [137] [139] [220]	8
Time Pressure	[199] [210] [139] [224] [225] [218] [219] [39]	8
Feedback	[214] [215] [76] [37] [137] [139] [217] [218] [219]	9
Tutorial	[85] [210] [72] [212] [232] [73] [79] [37] [39] [74] [222]	11
Resources Management	[197] [234] [235] [216] [137] [139] [224] [225] [217] [218] [39] [220] [240] [241]	14
Rewards/Penalties	[231] [211] [235] [214] [215] [76] [139] [33] [224] [225] [220] [74] [239] [240]	14
Cut Scene/Story	[197] [227] [207] [210] [211] [234] [235] [139] [217] [220] [75] [74] [221] [239] [223] [241]	16
Realism	[197] [198] [71] [211] [235] [73] [137] [33] [225] [217] [218] [219] [39] [220] [75] [221] [239] [223] [241]	19
Role-Play	[242] [71] [35] [210] [72] [232] [215] [79] [37] [139] [224] [225] [217] [218] [219] [75] [221] [223] [240] [241]	20

Table 18: Distribution of selected studies by game mechanics

Learning Mechanics	Selected Studies	Total
Planning	[219]	1
Identify	[240]	1
Cooperation	[213]	1
Question & Answer	[77]	1
Explore	[206]	1
Analyze	[240] [239]	2
Tutorial	[137] [220]	2
Collaboration	[200] [213] [236]	3
Participation	[38] [238] [74]	3
Guidance	[200] [137] [74]	3
Repetition	[214] [75] [221]	3
Assessment	[212] [77] [33] [86]	4
Motivation	[210] [136] [232] [77]	4
Competition	[70] [200] [77] [75]	4
Reflect/Discuss	[206] [213] [86] [75] [74]	5
Feedback	[77] [218] [219] [74] [239]	5
Instruction	[227] [83] [236] [137] [75] [74]	6
Action/Task	[240] [232] [38] [77] [137] [39] [75] [74] [221] [223] [241] [226]	12
Simulation	[240] [242] [227] [228] [203] [204] [85] [136] [236] [214] [137] [139] [33] [217] [218] [219] [39] [220] [86] [223] [226]	21

Table 19: Distribution of selected studies by learning mechanics

Game Dynamics	Selected Studies	Total
Multiple Strategies	[77]	1
Resource Scarcity	[236]	1
Deck/Hand building	[225]	1
Feedback	[77] [85]	2
Collaboration	[76] [200]	2
Conflict	[33] [75]	2
Turn Based	[77] [35]	2
Chance	[220] [86] [226] [80]	4
Negotiation	[76] [79] [37] [86]	4
Limited Actions	[33] [226] [77] [225] [218]	5
Competition	[37] [220] [70] [200] [77]	5
Realism	[236] [139] [33] [220] [86] [226] [240] [39] [217] [70]	10

Teams	[84] [235] [236] [216] [76] [79] [37] [137] [139] [220] [86] [75]	12

Table 20: Distribution of selected studies game dynamics

9.4 *Appendix D*

Effectiveness & Efficiency	EE1- The RAF-SGD minimizes duplication of tasks between stakeholders
	EE2- The RAF-SGD automation mechanism contributes to the effectiveness of the proposed framework.
	EE3- The framework allows the traceability of errors and inconsistencies.
	EE4- The framework helps define the goals of the game design
	EE5- The framework contributes to improving the understanding and delimitation of the design stages of a serious game
	EE6- The RAF-SGD contributes to an efficient and organized workflow in serious game design
	EE7- The RAF-SGD facilitates the transfer of knowledge/artifacts between different serious game designs
	EE8- The framework facilitates the adaptation of the serious game to different approaches (pedagogical, competitions, player, ...)
Quality & Resources	QR1- You perceive that the use of the framework improves the educational quality of the game.
	QR2- You perceive that the application of the framework improves the entertainment quality of the game.
	QR3- You perceive that the framework contributes to the creation of more engaging and effective serious games.
	QR4- You perceive that RAF-SGD contributes to the reduction of the time needed for each phase of the serious game design process.
	QR5- The framework facilitates faster iteration and prototyping in the serious game development cycle.
	QR6- You perceive that the framework facilitates resource optimization, for example, by encouraging knowledge transfer between different serious game projects.
	QR7- The framework's emphasis on promoting reusability contributes to the overall quality of serious games.
Interoperability	IO1- The RAF-SGD successfully identifies and clarifies the responsibilities of stakeholders in different stages of the serious game design process.
	IO2- I see the value in using this process in a multi-disciplinary game design team.
	IO3- I think that the process allows all stakeholders the opportunity to collaborate and provide their input

Technology Acceptance Model (TAM)	**Perceived Usefulness**	PU1- The RAF-SGD framework can be understood quickly without extensive training or guidance.
		PU2- The use of RAF-SGD would improve the learning outcomes of serious games.
		PU3- The RAF-SGD is beneficial in addressing the specific challenges of serious game design compared to other available methodologies.
	Perceived Ease Of Use	PEU1- The RAF-SGD generates and stores temporary artifacts, enabling stakeholders to pause or switch projects without disrupting the design process
		PEU2- The RAF-SGD leads to a straightforward and user-friendly design approach
		PEU3- The RAF-SGD it is simpler and easier to adopt than other serious game design frameworks or methodologies.
	Attitude Toward Using	ATU1- You are inclined to incorporate RAF-SGD into your future serious game development projects.
		ATU2- The RAF-SGD is in line with your personal values and professional goals as a serious game stakeholder.
		ATU3- You perceive that RAF-SGD has the potential to revolutionize the field of serious game design.
	Behavioral Intention To Use	BIU1- You will actively advocate for the adoption of RAF-SGD within your professional network of serious gaming stakeholders.
		BIU2- You would be willing to invest time and resources in learning and implementing RAF-SGD into your organization's serious game development processes.
		BIU3- There are some situations where one might choose not to use RAF-SGD despite recognizing its potential advantages for serious game design.

Table 21: Survey Items and Constructs for Evaluating the RAF-SGD Framework