Smart Computing and Intelligence

Series Editors: Kinshuk · Ronghuai Huang · Chris Dede

Ahmed Tlili
Maiga Chang *Editors*

Data Analytics Approaches in Educational Games and Gamification Systems

Springer

Smart Computing and Intelligence

Series Editors

This book series aims to establish itself as a medium for the publication of new research and development of innovative paradigms, models, architectures, conceptual underpinnings and practical implementations encompassed within smart computing and intelligence.

The scope of the series includes but is not limited to smart city, smart education, health informatics, smart ecology, data and computational analytics, smart society, smart learning, complex systems-chaos, computational thinking, brain computer interaction, natural/computer interaction, humanoid behaviour, and impact of educational psychology on computing.

The cornerstone of this series' editorial policy is its unwavering commitment to report the latest results from all areas of smart computing and intelligence research, development, and practice. Our mission is to serve the global smart computing and intelligence community by providing a most valuable publication service.

More information about this series at http://www.springer.com/series/15620

Ahmed Tlili · Maiga Chang
Editors

Data Analytics Approaches in Educational Games and Gamification Systems

Springer

Editors
Ahmed Tlili
Smart Learning Institute
Beijing Normal University
Beijing, China

Maiga Chang
School of Computing
and Information Systems
Athabasca University
Edmonton, AB, Canada

ISSN 2522-0888 ISSN 2522-0896 (electronic)
Smart Computing and Intelligence
ISBN 978-981-32-9334-2 ISBN 978-981-32-9335-9 (eBook)
https://doi.org/10.1007/978-981-32-9335-9

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Foreword by Alexandra I. Cristea

The Bottom-Up Approach in Education Gamification: *Educational Gamification Analytics*

Online education is 'traditionally' static. However, even since Socrates, it has been always known that learning is much more efficient via interaction. Good classroom teachers always introduce some level of interaction in their teaching. While face-to-face individual tutoring is, for most cases, undoubtedly the ideal form of education, this is not scalable for the current world context. Often, learning has to occur either fully or partially at a distance, due to either lack of time, space, funds or remote geographical or political location. Thus, supporting a better experience for online learning is essential for the modern world.

In the context of interactive learning, game-based learning, serious games and gamified environments have been proposed, as an alternative to the static online education, making it possible to introduce a great variety of interaction between system and learner. These areas are not new—they are here, to a different degree, ever since computers appeared. However, the interest in *gamification*, specifically, has only started around 2011 [2]. It involves extracting game-like elements and introducing them in learning environments (as opposed to introducing some learning content into games, as is done in educational games). The idea behind it is that games are often related to very high motivation, as well as being 'in the flow' [1]—whereas static learning environments struggled, possibly not without reason, with the motivational aspects. However, the exact combination and amount of game elements which are useful and appropriate in a learning context are still an open question.

More recently, with the latest developments in hardware, and the move from CPUs to GPUs, massive data storage, and later processing, became possible in all domains, including learning and online education. As a result, actual usage data from learners, teachers, administrators, staff, etc., can be analyzed to better understand how to design the appropriate interactions between students and systems. Such analysis can be very varied, but is described under the umbrella of *learning analytics*.

These emergent developments have made it possible to exploit *cross-disciplinary* synergies. In particular, moving from the top-down approach of education design, starting with a lesson plan and other pedagogical considerations, and gradually transforming them into a system, can now be further supported and extended via a *bottom-up approach*, such as developed in this book. Learner usage data can inform the learning process, as well as further design of gamified educational experiences, toward what could be called *gamification analytics*.

Thus, this exciting new book brings together a collection of articles on topics related to the timely topic of data analytics in gamification and educational games. This is further structured into three sections, with a number of selected representative papers each.

The first section focusses on generics on this area, starting from a systematic review, analysis of opportunities and applications of gamification in schools. The second section targets academic developments in the area. Excellent ideas are presented, such as in the iMoodle (intelligent Moodle) predicting at-risk students, or the learning analytics dashboard. Academic developments in use in schools are further shown, in the form of word problem-solving and a 3D board game. The third section focusses on learner models for the area, including motivational factors, as well as a design view. The papers are very novel and interesting, with some good ideas for the target audience, both for straightaway usage or implementation, as well as for further research. The book is well structured and readable, although it is a collection of different contributions, and a good amount of empirical and theoretical evidence is provided for the arguments brought forward.

The overall multi-disciplinary area of this book, combining gamification, learning analytics and e-learning, is very important and current, and yet not explored enough. As such, this book comes at the right time, with its collection of contributions from some of the most recent research in the area, and explores thus different facets of this problem.

The book is an absolute must for researchers and practitioners in the area of gamification, learning analytics and e-learning (including the rapidly expanding MOOCs) alike and should support future expansion of this area. This book can act as a reference manual for people studying this area, as it contains a great amount of useful information. I especially am looking forward toward further developments of actual practice, particularly in the form of commercial e-learning systems with gamified interaction support—which are much needed in our knowledge-hungry, constantly learning and upskilling world.

Alexandra I. Cristea
alexandra.i.cristea@durham.ac.uk
Department of Computer Science
Durham University
Durham, UK

References

1. Csikszentmihalyi, M. (1975). Flow: The psychology of optimal experience. Harper & Row.
2. Deterding, S., Sicart, M., Nacke, L., O'Hara, K., Dixon, D. (2011). From game design elements to gamefulness: Defining "Gamification". In *Proceedings of the 2011 Annual Conference Extended Abstracts on Human factors in Computing Systems—CHI EA '11*. p. 2425.

Foreword by Jorge Bacca

Over the last years, the learning analytic (LA) field has been an active area of research. According to Siemens et al. [2, p. 4], LA is "the measurement, collection, analysis and reporting of data about learners and their contexts." This area has benefited from the possibilities that technology brings for collecting a large amount of data from a wide variety of aspects in the teaching and learning processes including aspects at the institutional level (academic analytics—AA) and the advances in machine learning for analyzing large datasets.

Research on LA has focused on a wide variety of contexts such as recommender systems, learning design, MOOCs, mobile learning, student modeling, social network analysis, virtual environments and game-based learning among others. Despite the fact that the game industry has been using and developing Game Analytics (GA) for some years, in the field of game-based learning LA is still an emerging area of research that is now taking advantage of the experience and achievements in GA. The objective of GA and LA are different: While GA focuses on improving the player's engagement, LA focuses on the student's learning outcomes [1]. In that regard, research on LA in game-based learning or serious games analytics is an emerging area that deserves attention from the research community to address questions like: How to apply LA in the development of educational games and gamified activities? Which are the opportunities and challenges of LA and AA in game-based learning? How to manage learner modeling and individual differences in game-based learning with the support of LA?

In line with this context of research, the purpose of this book is to shed some light on the field of LA in game-based learning by providing some answers to the questions mentioned above and contributing to advance this field. This book is therefore divided into five parts with its corresponding chapters described as follows:

Part I—Introduction:

- Chapter 1 by Jina Kang, Jewoong Moon and Morgan Wood summarizes the evolution of research on games and gamification and discusses the potential role of data analytics for identifying individual differences in game-based learning.

Part II—Learning Analytics in Educational Games and Gamification Systems: In this part, the reader will find four chapters dedicated to show the current state and opportunities of LA in game-based learning in different research contexts.

- Chapter 2 by Jewoong Moon and Zhichun Liu presents a systematic literature review of 102 articles to describe current research on Sequential Data Analytics (SDA) in the context of game-based learning.
- Chapter 3 by Dirk Ifenthaler and David Gibson focuses on exploring learning engagement and its relationship with learning performance in the context of challenge-based learning.
- Chapter 4 by Valerie Shute, Seyedahmad Rahimi and Ginny Smith discusses the use of LA in the form of stealth assessment in game-based learning. Moreover, the authors explore the importance of including learning supports and its impact on learning performance when using the Physics Playground game.
- Chapter 5 by Juan Montaño, Cristian Mondragón, Hendrys Tobar-Muñoz and Laura Orozco shows how to use LA for the assessment of computational thinking skills using a web-based tool developed by the authors and called HERA in the context of a gamified activity.

Part III—Academic Analytics and Learning Assessment in Educational Games and Gamification Systems: In this part, the readers will find four chapters that discuss learning assessment in the context of game-based learning with LA.

- Chapter 6 by Mouna Denden, Ahmed Tlili, Fathi Essalmi, Mohamed Jemni, Maiga Chang, Kinshuk and Ronghuai Huang introduces and evaluates the Intelligent gamified Moodle (iMoodle) and its framework that includes an LA mechanism with a dashboard for teachers, a warning system for detecting at-risk students and personalized notifications of learning content.
- Chapter 7 by J. X. Seaton, Maiga Chang and Sabine Graf shows how to adopt LA dashboards in educational games to help players improve their in-game performance, in particular, their metacognitive skills. The authors highlight the advantages of using dashboards in this context.
- Chapter 8 by Abdelhafid Chadli, Erwan Tranvouez and Fatima Bendella aims at developing problem-solving skills by integrating a problem-solving model with a serious game. The authors introduce some metrics at the competence level to evaluate students' skills.
- Chapter 9 by Yu-Jie Zheng, I-Ling Cheng, Sie Wai Chew and Nian-Shing Chen depicts a 3D board game for learning about the human internal organs. The authors present the results of an evaluation study on the effect of the board game on the students' learning experience and learning outcomes.

Part IV—Modeling Learners and Finding Individual Differences by Educational Games and Gamification Systems: In this part, the reader will find three chapters dedicated to the use of LA as support of learner modeling and personal dimensions of the learner such as motivation and learning outcomes in the context of game-based learning and gamified experiences.

- Chapter 10 by Sven Manske, Sören Werneburg and H. Ulrich Hoppe presents an analysis of LA techniques to assess computational thinking competences, and the authors introduce a framework for learning analytics in game-based learning in the context of computational thinking.
- Chapter 11 by Rafael Luis Flores, Robelle Silverio, Rommel Feria and Ada Angeli Cariaga introduces a LA model to identify student motivation in a game-based learning environment.
- Chapter 12 by Ana Carolina Tomé Klock, Isabela Gasparini and Marcelo Soares Pimenta tackles the issue of organizing and clarifying the concepts of gamification to assist the design, development and evaluation of gamified experiences.

Part V—Conclusion:

- Chapter 13 by Ahmed Tlili and Maiga Chang summarizes the objectives of adopting data analytics, the metrics that have been collected and current challenges for the adoption of data analytics in educational games.

Jorge Bacca Ph.D.
Fundación Universitaria Konrad Lorenz
Bogotá, Colombia

References

1. Serrano-Laguna, Á., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., Fernández-Manjón, B. (2017). Applying standards to systematize learning analytics in serious games. *Computer Standards & Interfaces*, *50*(September 2016), 116–123. https://doi.org/10.1016/j.csi.2016.09.014
2. Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Buckingham, S., Ferguson, R. et al. (2011). *Open Learning Analytics: An integrated & modularized platform. Proposal to design, implement and evaluate an open platform to integrate heterogeneous learning analytics techniques. SOLAR - Society for Learning Analytics Research*. Retrieved from https://solaresearch.org/wp-content/uploads/2011/12/OpenLearningAnalytics.pdf

Preface

Educational games, gamification learning systems and learning analytics are gaining an increasing attention from researchers and educators. The educational games and gamification system can get learners engaged and motivated in the learning process; at meanwhile, learning analytics grants a system the capability of understanding the learners' needs, assessing learners' skills and knowledge silently, providing teachers detailed information about their students and warning teachers and administrative personnel to pay attention on the at-risk students. This book covers applications of data analytics approaches and research on human behavior analysis in educational games and gamification systems. In particular, this book discusses the purposes, advantages and limitations of using data analytics approaches in game-based learning environments and applications.

This book talks about the data analytics methods, systems/tools and research for analyzing learners' actions, profiles, records and behaviors stored or happened in educational games and gamified learning systems. As the research progress rapidly, this book can be an up-to-date textbook and reference book for not only post-secondary and academic, but also can be a handbook for educational technology relevant companies and industry.

This book arranges research based on three themes: learning analytics, academic analytics and learning assessment, and learner modeling and individual differences. Each theme covers three to four latest research results related to the data analytics in educational games and gamification systems. The aim is to provide readers with methodologies, evidences and experiments through these researches and help readers get clear picture of how data analytics approaches can help not only students and teachers but everyone in the world.

First, this book starts with Moon and Kang's introduction chapter that helps readers get familiar with the subject areas and leads readers to know the importance of data analytics in educational games and gamification research area.

In the second part, four chapters talk about learning analytics in educational games and gamification systems. Moon and Liu in Chap. 2 explore the use of sequential data analytics in game-based learning and major issues while doing so via a systematic literature review. At the end of the chapter, they propose guidelines for readers to use sequential data analytics properly. Ifenthaler and Gibson then in Chap. 3 bring the concept of challenge-based learning up. They study 8951 students' transaction data and find the learning engagement is positively related to learning performance. Their finding in fact implies the importance of making a learning system like educational game or gamification system capable of catering for the individual learner's needs. Shute, Rahimi and Smith, on the other hand, in Chap. 4 discuss the learning supports and their influences in educational game and present a usability study's of designing and developing stealth assessment in an educational game named Physics Playground. At the end of the chapter, they provide insights of the future of using learning analytics in the games for stealth assessment. In the end of this part in Chap. 5, Montaño, Mondragón, Tobar-Muñoz and Orozco create a gamified platform called HERA. In HERA, students participate in gamified activities that are part of assessment and teachers can know their students via learning traces analysis.

The third part of the book is about the academic analytics and learning assessment in educational games and gamifications. This part also has four chapters. Denden and colleagues in Chap. 6 present an iMoodle that is an intelligent gamified Moodle. iMoodle has a built-in learning analytics plug-in that can provide teachers dashboard for teachers to control the learning process and an early warning system for predicting at-risk students. Their finding shows that iMoodle has a high accuracy rate which is almost 90%. Seaton, Chang and Graf also propose the use of dashboard in an educational game called OMEGA (Online Metacognitive Educational Game with Analytics) in Chap. 7. The dashboard can help players see how their performance and skills change over time and what are their weakness and strengths. With the dashboard, players can see their gameplay performance and habits and find the clues and strategies to improve their in-game performance. As the goal of educational games is to allow players to learn unconsciously while playing and playing educational games more and frequently players should learn more or have their skill better, the dashboard can avoid the players quitting from the gameplay due to stuck in the game and cannot get further progress. In Chap. 8, Chadli, Tranvouez and Bendella are also putting their focus on metacognitive skill, in particularly, problem-solving skill. They not only investigate the improvements of second-grade students' word problem-solving skills with educational game's help, but also propose a competency model to measure student's knowledge levels. At the end of this part, Zheng, Cheng, Chew and Chen in Chap. 9 try to improve game-playing process with additional software and sensors. The game collects students' interaction data and provides instantaneous feedbacks for the students.

The fourth part of this book aims to learner modeling and individual difference finding. This part includes three chapters. Manske, Werneburg and Hoppe first in Chap. 10 propose a framework for designing and evaluating game-based computational thinking environment named ctGameStudio. The proposed framework uses learning analytics to provide the learners' dynamic guidance, scaffolds and feedback properly according to their actual state. Then, Luis Flores, Silverio, Feria and Cariaga in Chap. 11 present a learning analytics model that can measure students' motivation within an educational game, Fraction Hero, based on their in-game data.

The model assesses three motivational factors include goal orientation, effort regulation and self-efficacy. They also find that students have higher in-game motivation than self-perceived motivation toward solving problems. At the end of this part, Chap. 12 organizes and clarifies gamification concepts according to seven properties: personal, functional, psychological, temporal, playful, implementable and evaluative, through a user-centered approach done by Klock, Gasparini and Pimenta.

Finally, the last conclusion chapter is written by Tlili, Chang, Huang and Chang. The chapter summarizes all the presented chapters and also discusses correspondent challenges and future insights while adopting data analytics in educational games and gamification systems.

Beijing, China Ahmed Tlili
ahmed.tlili23@yahoo.com

Edmonton, Canada Maiga Chang
maiga.chang@gmail.com

Acknowledgements

We would like to first thank all of the authors for their valuable contributions to this book by sharing their developed case studies and research outcomes of applying data analytics approaches in educational games and gamification systems. These studies and their reported findings definitely help readers learn the way of adopting data analytics and also give readers stepping stones for further research and development thoughts and insights.

We would also like to thank all of the reviewers who accept to review the submitted chapters and give their constructive comments and suggestions for the authors to further enhance the quality of their book chapters, hence enhancing the overall quality of this book. We really appreciate them for giving their reviews in a timely manner that helps our book to meet the production timeline.

Special thanks also go to the series editors, namely Prof. Kinshuk, Prof. Ronghuai Huang and Prof. Chris Dede, for their comments and guidance to prepare this book, as well as all our colleagues in the Smart Learning Institute of Beijing Normal University, China, and Athabasca University, Canada, for their support to finish this book project.

Dr. Ahmed Tlili
Dr. Maiga Chang

Contents

Part I Introduction

1 Educational Games and Gamification: From Foundations to Applications of Data Analytics 3
Jina Kang, Jewoong Moon and Morgan Diederich

Part II Learning Analytics in Educational Games and Gamification Systems

2 Rich Representations for Analyzing Learning Trajectories: Systematic Review on Sequential Data Analytics in Game-Based Learning Research 27
Jewoong Moon and Zhichun Liu

3 Opportunities for Analytics in Challenge-Based Learning 55
Dirk Ifenthaler and David Gibson

4 Game-Based Learning Analytics in Physics Playground 69
Valerie Shute, Seyedahmad Rahimi and Ginny Smith

5 Learning Analytics on the Gamified Assessment of Computational Thinking 95
Juan Montaño, Cristian Mondragón, Hendrys Tobar-Muñoz and Laura Orozco

Part III Academic Analytics and Learning Assessment in Educational Games and Gamification Systems

6 iMoodle: An Intelligent Gamified Moodle to Predict "at-risk" Students Using Learning Analytics Approaches 113
Mouna Denden, Ahmed Tlili, Fathi Essalmi, Mohamed Jemni, Maiga Chang, Kinshuk and Ronghuai Huang

7 Integrating a Learning Analytics Dashboard in an Online Educational Game 127
J. X. Seaton, Maiga Chang and Sabine Graf

8 Learning Word Problem Solving Process in Primary School Students: An Attempt to Combine Serious Game and Polya's Problem Solving Model 139
Abdelhafid Chadli, Erwan Tranvouez and Fatima Bendella

9 Designing a 3D Board Game on Human Internal Organs for Elementary Students 165
Yu-Jie Zheng, I-Ling Cheng, Sie Wai Chew and Nian-Shing Chen

Part IV Modeling Learners and Finding Individual Differences by Educational Games and Gamification Systems

10 Learner Modeling and Learning Analytics in Computational Thinking Games for Education 187
Sven Manske, Sören Werneburg and H. Ulrich Hoppe

11 Motivational Factors Through Learning Analytics in Digital Game-Based Learning 213
Rafael Luis Flores, Robelle Silverio, Rommel Feria and Ada Angeli Cariaga

12 Designing, Developing and Evaluating Gamification: An Overview and Conceptual Approach 227
Ana Carolina Tomé Klock, Isabela Gasparini and Marcelo Soares Pimenta

Part V Conclusion

13 Data Analytics Approaches in Educational Games and Gamification Systems: Summary, Challenges, and Future Insights 249
Ahmed Tlili and Maiga Chang

Part I
Introduction

Chapter 1
Educational Games and Gamification: From Foundations to Applications of Data Analytics

Jina Kang, Jewoong Moon and Morgan Diederich

Abstract A large number of educational games and gamification systems have been developed over three decades. Research has shown game-based learning (GBL) to be effective in enhancing motivation and improving learner performance. However, we have faced challenges of understanding an individual's learning experience within GBL, since learners bring a unique combination of background, context, and skills with them to the game environments, which yields various responses to the game mechanics. Researchers and practitioners therefore have underscored the need for understanding individual differences within the GBL environments. The growing area of data analytics has created possibilities of identifying individual learners' personalities and their play styles within the system. This chapter first describes how educational games and gamification system have evolved in previous GBL research. We further explore the emergent role of data analytics in advancing current research of educational games and gamification, particularly the recent research efforts of understanding individual differences in GBL.

1 Introduction

Digital games have grown in popularity since the mid-1980s when computers and gaming consoles were first introduced [1]. The number of children and adolescents spending time with games via gaming consoles or other mobile devices has increased (e.g., [2, 3]). In the USA, 38% of students in K-12 reported playing video games on a day in 1999, 52% in 2004, and 60% in 2009 [4]. Instructors and institutions have also

J. Kang (✉) · M. Diederich
Utah State University, Instructional Technology and Learning Sciences, Logan, USA
e-mail: jina.kang@usu.edu

M. Diederich
e-mail: morgan.wood23@usu.edu

J. Moon
Florida State University, Educational Psychology and Learning Systems, Tallahassee, USA
e-mail: jm15g@my.fsu.edu

A. Tlili and M. Chang (eds.), *Data Analytics Approaches in Educational Games and Gamification Systems*, Smart Computing and Intelligence,
https://doi.org/10.1007/978-981-32-9335-9_1

immersed students in learning by using games in the classroom. In response to the prevalent uses of educational games, game-based learning (GBL) has been utilized in a variety of learning environments, which includes educational games to gamification systems, and is not limited to age, culture, or subject matters (e.g., history, language, mathematics, and science). As such, the popularity of educational games has shown the significance of games in the twenty-first century. The increasing ubiquity in the use of games to supplement and further engage learners in various types of learning environments has been considered as a prime tool for learning. GBL can provide a unique learning experience in that learners who interact with the system or peers can gain skills or knowledge including strategic thinking, planning, and contextualizing experiences that may not be easily acquired otherwise [5].

To date, research has confirmed GBL has been used to drive and supplement behavioral change, affective and motivational outcomes, perceptual and cognitive skills, knowledge acquisition, and content understanding [6–10]. The recent interest in the field of GBL is understanding learners' actions traced within the system, which can be used to investigate how they learn in GBL. Using the user-generated data, researchers can give various stakeholders actionable insights to enhance learners' engagement and performance, support better game learning design, and further sustain student retention. Diverse data analytics approaches including learning analytics, serious games analytics, and academic analytics have created such possibilities [11–13]. Researchers and practitioners in the field of GBL have sought to understand differences captured as an individual interacts with the game mechanic elements, which can be adapted to improve the design of learning to fulfill individual learners' needs. In this introductory chapter, we will describe how educational games and gamification designs have evolved in previous GBL research. Besides, this chapter will also review the underlying concepts and implementations of both educational games and gamification designs. We will further propose the role of data analytics in advancing current research of educational games and gamification, particularly the recent research efforts of understanding individual differences in GBL.

2 Educational Games to Gamification

2.1 Overview of Educational Games

Often when thinking of games, we traditionally think of "entertainment" games. The definition of a game, in general, is not universal. Stenros [14] completed a review of game definitions since 1930 and identified over 60 formal definitions. The review highlighted that many game definition components are consistent, such as having rules, players undergoing conflict or making decisions, and mention of a purpose. Thus, the question naturally arises if there is a difference in the primary purpose that distinguishes educational games from other games. Compared to commercial games, in educational games, the basis of design is rooted in balancing learning as well as play [6], instead of solely gameplay—often the focus of an entertainment game mostly resides in what yields a monetary gain. Yet for educational games, while a foundational definition is agreed upon, there are certainly debates on the minutia.

Prior accounts have attempted to define educational games by addressing their characteristics. In 2006, the Summit on Educational Games identified elements that are critical for educational games, specifically that a game should have clear goals, repeatable tasks that build mastery, monitors progress, prompts motivation of a task, and adjusts difficulty through personalization of learning [15]. As another example, scholarship refers to educational games alongside "serious games" which are designed to improve skills and learning performance through training and instruction [16–18]. While serious games and educational games are dissimilar from entertainment games in the fact that they are not created primarily for entertainment, the distinction between serious games and educational games is not as clear. For instance, Djaouti et al. [19] indicated that education-focused games are only one category of serious games, while serious games may include any digital games designed not solely for entertainment [20].

A large number of educational games have been developed in various content areas including computer science, economics, geography, history, language, pathology, physics, biology, astronomy, and ecology. For example, *CRYSTAL ISLAND* is a narrative-centered educational game in which the goal is to support middle school classroom instruction [21]. In the game, a student visits her/his sick father on a remote island in order to save the father and research team members who also suffer from the same sickness. The game is designed for students to learn microbiology, while they are solving this mystery illness in the game. *Alien Rescue* is an online educational game which is designed to immerse middle school students to authentic problem-based learning activities [22]. Students are placed in International Space Station as a young scientist. Learners learn space science, while they collect information to figure out which world in our solar system could be an appropriate home for six alien species whose home planets were destroyed. Aligned with National Science Standards, the game is particularly designed for sixth-grade space science to use in class sessions.

Prior GBL research has shown how the use of educational games contributed to improving learners' performance and knowledge acquisitions (e.g., [7–10]). Researchers explain that educational games should be designed to support learners' opportunities that improve their content and contextual understanding of various cognitive and practical skills. To foster learners' skill acquisition in GBL environments, empirical studies have highlighted that educational games should contain an element of assessment providing both learners and designers with information of how their in-game interactions emerge, relating to their game success, as well as meaningful learning [23]. Although both educational games and gamification aim at improving target learners' outcomes (e.g., enhancing learner motivation and engagement), their designs and underlying assumptions are different. Educational games focus on students' internalizations of their learning experiences through a sequence of game actions, and gamification tends to transplant game elements to non-game contexts, such as incentivizing behavior changes.

2.2 Overview of Gamification

Gamification is another new wave of the field has focused on using game-like attributes in educational contexts. Compared to educational games, gamification refers to the adoptions of generic game components to non-game contexts. Since gamification originated from business and marketing fields, previous scholarly works emphasized their alternative role in applying game components to industries [24, 25]. Whereas educational games aim at enhancing learners' intrinsic motivation within game worlds, gamification describes how integral game elements (e.g., digital badges and competition) outside of game environments promote learners' behavior changes in non-game domains. Gamification is not a game itself but a purposeful approach that utilizes gamified experiences that enhance learners' engagement. Therefore, gamification requires a strategic design that focuses on facilitating learners' engagement through related game mechanics (p. 14) [26]. Research on gamification has highlighted the ways to manipulate environmental conditions that allow learners to perceive game-like circumstances—leading to learners' participatory acts toward surrounding learning tasks [27].

To date, gamification studies have underscored to explicate key game elements that are likely to be integrated into existing educational settings [28, 29]. Scholars believe that those game elements can foster students' learning engagement throughout their gameful experiences [30]. Landers [31] delineated a collection of gamification attributes, such as action language, assessment, conflict, control, environment, game fiction, immersion, and rules/goals. Further, Seaborn and Fels [26] also listed the following game elements as required design components for gamification, namely: point, badge, leaderboard, progression, status, level, reward, and role. In addition to analyzing generic game elements, they expanded their analysis as to how a series of game rewards and gamified activities are contextualized and coordinated [26]. Zichermann and Cunningham's [32] gamification design showed various ways to transform existing e-learning settings to gamified contexts. They identified a list of game mechanics and subsequent case studies regarding how gamification can be adopted and implemented in various training settings. The next section discusses theoretical foundations relevant to learner motivation and engagement in the design of educational games and gamification.

2.3 Theoretical Foundations

Many scholars have asserted that the use of educational games and gamification have several advantages for learning. Games are accessible, reasonably priced, and effective substitutions for traditional classroom activities (e.g., [33–36]). Others discussed limitations to effective learning using games, claiming that games do not support in-depth learning and that both learners and instructors are skeptical of the value of games in the learning environments (e.g., [37–39]). Gamification emphasizes enhancing learners' curiosity and external motivation. However, migrating game elements into existing education settings do not guarantee learners' engaging attitude when it does not belong to any strategies that facilitate the learners' goal accomplishments. Nevertheless, the growing popularity over the past two decades of games and the widespread adoptions of gamification designs reflect the growing interest in utilizing games and gamification for learning.

Research, therefore, has underscored the importance of stimulating learners' engagement and motivation when designing educational games (e.g., [6, 40]) and gamification systems (e.g., [41, 42]). Several studies raised their questions of which psychological attributes can provoke learners' engaging acts [24, 29, 43] and explored underlying theoretical frameworks from fundamental motivation theories (e.g., [44–48]) and motivation design models (e.g., ARCS (Attention, Relevance, Confidence, Satisfaction) model in [49, 50]). Although a traditional lens of behaviorism and cognitivism explained students' learning actions and information processing, a lack of understanding has existed regarding their enhanced motivation and active attitude in learning in GBL. Especially, a collection of motivation theories better explains how students' motivation can be managed and facilitated by understanding the underlying mechanics of human nature. Motivation theories together with their design models have sought to examine what ways can support students' self-regulated and mindful actions by considering their internal dynamics in motivation.

Both intrinsic and extrinsic motivations are considered to have an impact on determining learner behavior and learning outcomes [44, 51]. Okan [52] pointed out that intrinsically motivated students were willing to learn a subject and to use what they learned more frequently afterward. In contrast, extrinsic motivation involves an external reward or threat. Some researchers argued that extrinsic motivators could distract learners from learning more about curricular content outside of the classroom [46, 53]. Prior research stated the role of game design that can promote learners' motivation both intrinsically and extrinsically. Empirical studies showed that learners using a game tend to learn more and to become more intrinsically motivated during a problem-solving process than the traditional classroom learning environment (e.g., [54, 55]). Those studies also highlighted that proper extrinsic rewards should be considered in order to motivate learning when designing games by including some form of diegetic extrinsic reward while also balancing extrinsic types of motivation with intrinsic motivation.

In addition to the notion of extrinsic and intrinsic motivation, a few theories have also contributed to understanding why gamification enhances learners' motivation level. First, prime motivation theories highlighted the significant uses of extrinsic rewards to promote learners' motivation. For example, the expectancy-value theory of motivation presents how a series of expectancy-value constructs enhance learners' belief systems. The theory explains that students' efficacy and outcome expectations toward resultant actions facilitate their belief systems. If students can achieve a sense of successful learning experiences, learners can increase confidence in mastery learning. The enhanced expectation, therefore, better promotes engagement in future tasks. Specifically, the theory listed a series of antecedents that can strengthen students' belief systems, such as direct experience, vicarious experience, and verbal persuasion. Also, the theory depicts how belief systems promote their task persistence and motivation [43, 55]. In this theory, a learner's ability and expectancy belief are both critical components to leverage the expectancy value. Because the combination of the learner's ability and belief level determines their expectancy-value relation, it is essential to identify learners' expectancy of success, ability belief, and subjective values are perceived [43, 55]. The chain of the three constructs above explains how gamification design controls the weighted value of extrinsic rewards.

As such, design and implementation of games and gamification are of great importance. We wish to discuss different attributes of well-designed games and gamification systems, terminology, and how it impacts learner motivation and engagement in the next section.

2.4 Design and Implementation

Educational Games Several studies have proposed in what ways educational games promote students' motivation and learning. Dede et al. [56] asserted that educational games encouraged students to perform better in academic settings. Barab et al. [57] suggested that enthusiasm and motivation should be inherent in educational games in order to support students' active learning. Many attempts have been made to define desirable game design attributes for educational games, including 36 learning principles of video games that can affect how people learn [58]. Wilson et al. [59] also provided 16 key gaming attributes necessary for learning including (a) adaptation, (b) assessment, (c) challenge, (d) conflict, (d) control, (e) fantasy, (f) interaction, (g) language/communication, (h) location, (i) mystery, (j) pieces/players, (k) progress/surprise, (l) representation, (m) rules/goals, (n) safety, and (o) sensory stimuli, whereas there were only few game elements discussed in the early literature such as Malone and Lepper's four elements [60]: "challenge, curiosity, control, and fantasy" (pp. 228–229).

Other game elements, such as incentive systems, aesthetics, and narrative designs also contribute to a successful game. These elements synergistically work together to ensure that the overall gameplay is engaging and provide an effective platform in which knowledge acquisition can occur. This must work seamlessly with the narrative design, which encompasses the story in which in-game conflicts will arise and the problems to be solved. These problems and conflicts are built around the learning objectives [6]. Each problem resolved is then typically provided some form of reinforcement. If difficulties occur, then reinforcement and game support is provided. If problems continue to occur, some games have adaptable skill levels, in which the level of difficulty is adjusted to maintain a state of "flow." Flow is the state in which the feeling of enjoyment is obtained when an individual's knowledge levels and given challenges is well-balanced to accomplish a task that is intrinsically motivating [61, 62]. Tarng and Tsai [63] suggested that various situations, themes, or narratives in the game environment could be considered influential factors—contributing to learners' attitudes. When a game integrates content, narrative, and gameplay, it could have an impact on the relationship between learning and engagement [21]. Specifically, Rowe et al. [21] demonstrated that narrative elements play an essential role in the relationship between learning and engagement. The study showed in the game, *CRYSTAL ISLAND*, the narrative motivated learners to solve their tasks, which were not only simple but also sufficient in not distracting learners from the learning goals. They finally highlighted that a well-designed story and elements are necessary in order to lead learners to concentrate on games and tasks.

As an increasing capability in telemetry and computer data processing continues, researchers have been interested in new ways of assessment and feedback in games. Plass et al. [6] proposed a generalizable theory of a successful game. A simple loop is created from a challenge, the game giving a response to the learner's actions, then providing feedback looping to another challenge. As learners receive feedback within the game context to promote better gameplay and tailor learning behaviors and objectives, game designers also need to receive feedback to better inform their learners and improve game design. Timely feedback is a fundamental component that learners need to attain during the process of a game in order to improve learners' motivation and engagement [64–66]. Nadolski and Hummel [67] offered a proof of concept of a retrospective cognitive feedback (RCF) that is characterized by more simple and effective feedback to players in real time based on difficulties in information technology (IT) administration game—administered to vocational students who wish to go into that field. The game is designed for students to learn skills to clarify clients' needs and continue to meet the clients' expectations throughout the five phases of developing their IT system. During each task, students can ask questions to game characters and receive feedback. A total of 110 students were randomly assigned to each of two conditions: RCF group and non-RCF group. In the finding of this study, the RCF group showed higher learning performance compared to the other group. However, the result from pre-/post-motivation questionnaires did not show any significance difference between both groups. They identified that this is a promising first phase and identified ways to better improve the game while also maintaining and further implementing dynamic feedback for educational games.

Ifenthaler et al. [23] highlighted the importance of tracing changes during the learning process and providing learners with requisite feedback while playing a game. The interaction information provided by such clickstreams help identify inadequate behaviors and further improve the game design. In addition, assessment can be utilized to provide critical information on how learning objectives are being met and received. Both quantitative (e.g., pre-/post-test scores, log data) and qualitative data (e.g., interview, observation) can be used to explore personalities, player types, learning strategies which can then in turn be used to inform adaptive feedback in-game via different techniques, including machine learning algorithms. Inappropriate challenge and task design may negatively impact how the learning aspects are received.

Gamification Systems Beyond prior implementations of GBL, research has underscored the extensive and interdisciplinary role of game-element adoptions in various fields. Hence, research on gamification has highlighted the identification of how multiple game mechanics systematically promoted learners' engagement and motivation in non-game contexts.

In the field of e-learning, assorted attempts existed in adopting gamification design. Prior research benchmarked gamification designs that aim to foster learner motivation through reward exchanging systems. Several reviews of the massive open online course (MOOC) portrayed how gamification has been implemented into the course design. They specifically proposed how extra credits and digital badges helped learners to draw their attention [68, 69]. It was found that acquiring digital badges facilitated learners' engaging acts through certifying their accomplishments. Also, some studies investigated the progress bar that aimed to guide students to identify their learning phases [70]. The progress bar is designed to notify learners' progression and encourage learners to monitor their performance reported based on the information of gaps between a learning goal and their current status. For example, Ibáñez, Di-Serio et al.'s [71] case study portrayed how gamification design promoted students' engagement when teaching C-programming. They designed their gamified e-learning tutorial *Q-Learning-G* that enables students to achieve various gameful experiences. The findings showed that certain rewards in the platform significantly enhanced students' high involvement in their project implementations. Notably, the students tended to be engaged in collecting in-system credits and badges when attending a series of learning activities.

Furthermore, gamification designs also exist in encompassing various formats of organizational training under workplace settings [72, 73]. Gamification research has suggested motivational design principles specifically for improving the productivity of human resources in the fields of business and marketing. Oprescu et al. [74] reported a collection of workplace gamification design principles. This study mapped the principles with expected learning outcomes. For example, using persuasive elements aims at provoking learners' initiatives derived from their enhanced satisfaction. Specifically, this study showed a case of the Google incentive system that allows users to contribute to their surrounding social contexts inherently. The facilitative element led to employees' behavior changes and continuous learning experiences throughout social dynamics. Further, Rauch [75] introduced gamification design cases to enterprise-related practices. This study reported how the corporation Oracle adopted a gamified online forum in promoting employees' engagement in their production.

Seminal scholarly works also confirmed that gamification designs promote learners' behavior modification and cognitive awareness in health-related contexts [76, 77]. Emerging wearable technologies enable users to access real-time data—explaining individuals' lifestyle and routine behaviors. Recent reviews of health education [25, 78] also support that gamification design fosters user behavior intentions by facilitating their proactive attitude when managing their physical acts. Specifically, gamification aims at promoting users' awareness when changing their behavior routines tailored to their healthcare needs and patterns. Hamari and Koivisto [79] investigated users' perceptions of gamified exercise service—relating to the understandings of their health. The findings confirmed that social factors play a critical role in maintaining users' sustainable motivation and, therefore, gamification design should consider users' social relations when adopting online gamification environments.

Overall, variant design and implementation cases in both the educational game and gamification research demonstrate that it is vital to identify how game elements, their implementations, and learner contexts emerge. Hence, research has demanded variant forms of analytic frameworks that guide systematic understandings of GBL contexts. Further knowledge will be gained through several data analytics examples to better understand learners and GBL environments (i.e., educational games and gamification), which will be explored in next.

3 Data Analytics Approaches in Educational Games and Gamification Systems

A large number of empirical studies primarily depend on self-reported data from surveys, questionnaires, and pre-/post-test data to examine the benefits of GBL (e.g., [80]). Recent researchers have pointed out existing limitations—including external validity issues of these studies, in which a game environment is mainly considered as a "black box" (p. 17) [11]. Data are thereby collected only before or after learners interact with the game environment. As digital game-based technologies have grown, researchers have paid more attention to the area of data analytics, which has created possibilities of capturing users' behaviors in a game beyond the traditional performance assessment [13, 81]. Compared to data obtained by human-provided data, user-generated data that are captured automatically through their gameplay are less subjective and erroneous [11, 82, 83]. Therefore, the information on such as how many tasks and how fast the tasks are completed can be collected without interrupting users in the GBL environments, in which GBL can be considered as a "white box." As user-generated data becomes a prevalent feature in GBL environments, researchers can interpret learners' repeated actions as behaviors. It is also essential to understand what types of learner actions or behaviors can lead to better learning performance. Ifenthaler [84] argued one challenge of an analytics approach is the limited taxonomy of metrics for different educational games and gamification systems. Recently, there has been a stream of efforts to develop GBL specific metrics—evolved separately from the entertainment game industry—that appropriately measure learner performance depending on the purpose of the game and the systems that capture gameplay traces [85]. The analytics applications are therefore purposed to identify a pattern or trend of gameplay and track users' decision-making processes [13, 81, 85, 86]. Researchers inform educators and game developers of these insights to support better learning design and improve the skills and performance of learners in a GBL environment. Such efforts further produce actionable strategies for addressing academic issues, such as retention or success rates at an institution. The following sections first describe different data analytics approaches in GBL, including its goals and tools: (1) learning analytics that mainly informs learners' behaviors and (2) academic analytics that addresses academic issues. Lastly, the applications of modeling individual learner differences in GBL environments are discussed.

3.1 Learning Analytics for Educational Games and Gamification Systems

In recent years, many researchers have defined learning analytics in various viewpoints. One of the widely known definitions was announced at the 1st International Conference on Learning Analytics and Knowledge [87]: "Learning analytics is the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs." There have been multiple terms emerged to define diverse data analytics approaches in educational games and gamification systems including serious games analytics (e.g., [11]) and game learning analytics (e.g., [88]), academic analytics, and educational data mining [89]. The fundamental concept of all terms is the underlying data-driven process to benefit education [88]. Learning analytics is mainly purposed to provide dynamic pedagogical information to optimize learning and the environments such as the learning management system, intelligent tutoring system, educational games, and gamification systems.

Increasing interests of GBL research and its data-mining techniques have drawn scholars' attention in defining the role of learning analytics in both game and gamification systems. In game industry, game analytics is the applications of data analytics for improving game design and sales in the industry of commercial games [90]. The main purpose is to understand users' gameplay behaviors and detect glitches or errors to improve user experience, which ultimately yields a monetary gain. In comparison, learning analytics in educational games and gamification systems is purposed to identify learners' gameplay processes, classify their knowledge, motivation, or behavior, and assess learning performance (e.g., [91]). An analytics system within a game or gamified system tracks and sends dynamic learner-behavior data, such as the decision-making process back to the learning analytics framework. The observable gameplay information can be used to improve the game design, produce real-time interventions during learners' gameplay, or assess learner performance [23, 81, 85, 86]. Clearly, there is a similar intention to understand learners and improve their learning experience with GBL environments. Both approaches support learners' knowledge acquisition or skill development (e.g., [11, 92]).

3.2 Academic Analytics for Educational Games and Gamification Systems

Academic analytics has been adopted mainly in higher education as an application to address specific issues at an institutional level, such as student retention [93]. Numerous factors might cause dropout decisions in colleges and universities including a lack of motivation, a wrong choice of course or major, a lack of academic skills, and a lack of institutional support services (e.g., [94, 95]). Institutions in higher education have started collecting dynamic student data captured in a learning management system or content management system. They exploited academic analytics to trace learning process and predict students' experience and performances to produce actionable strategies. Such insights provide real-time feedback on the students' learning status, strengths, and weaknesses, and further early remediation actions that contribute to decreasing student retention and, as a consequence, increasing their success rates.

Diverse ways in data analytics approaches have been adopted in preventing students from discontinuing higher education by embedding various gamification elements such as digital badges, extra credits, and leaderboard (see more details in Sect. 2.2). In enhancing student retention, Mah [96] highlighted the needs of digital badges for measuring and operationalizing academic skills to integrate them into algorithms that predict student success. They proposed an analytic model of digital badges by synthesizing three aspects: learning analytics, digital badges, and academic skills, which provide students with personalized feedback including required skills to earn digital badges and visualizations of learning paths and progress. Individual student data including demographic information, prior GPA, and results from freshmen survey such as the Learning and Study Strategies Inventory [97] can be further used to improve the predictive algorithm of the model.

Design-based research has revealed various ways of academic analytics and has been adopted in the fields of higher education. A collection of studies in the ICT-FLAG project demonstrated how academic analytics efficiently supported learning activities in virtual learning environments [98, 99]. The project aimed at building a comprehensive information communication technology (ICT) framework through formative assessment, learning analytics, and gamification for educational stakeholders including teachers and academic program managers. To support academic program managers, the system provided data, such as opinion-mining results by natural language processing (NLP), academic performance, and the dropout ratio of the enrolled students. Androutsopoulou et al. [100] portrayed how academic analytics was used for decision-making in developing the e-participation platform that fosters citizen's involvement in political participation under gamified design. This online platform initiated multiple analytic methods used to better understand qualitative data, such as opinion mining, sentiment analysis, and argumentation mining, and further implemented policy modeling based on synthesized results from data analytics. Citizen data collected through three data-mining phases (i.e., data management, knowledge processing, and collaboration support) enabled policymakers and designers to understand how specific social issues were identified via citizens' argumentation results.

According to Long and Siemens [89], the role of academic analytics is more at an institutional, regional, or international level and primarily to support institutional decision-making, while learning analytics focuses on learners and their learning processes. The distinction between learning analytics and academic analytics exists in terms of the role, and however, it has gradually worn down as the researchers have mixed two terms across various target audiences, levels, and objects of analyses. The common interests are to understand what learners do in educational games and gamification systems, investigate the effectiveness of learning environments using gameplay traces, and implement the findings to improve the system design. Under this presumption, one notable trend in the field of GBL research is understanding individual learners' differences via various data analytics. In the next section, we discuss the recent efforts of building an understanding of personalities and player types in GBL.

3.3 Modeling Individual Differences in Educational Games and Gamification Systems

Behavior, personality, aspirations, and actions guide the way each individual interacts with the world [101]. This unique combination of personal context and situations provides a rich and complex set of ideas, beliefs, and ultimately tendencies that drive what one does. In the field of GBL, researchers and practitioners have sought to understand individual differences within the game environments [102]. Learners each bring their combination of background, context, skills, and expectations with them to a game and therefore have various responses to game mechanics [103]. Therefore, GBL should be designed to fulfill individual learners' needs. This requires understanding differences as they interact with the game mechanic elements. Personality has been considered as one indicator of individual differences, which are associated with individuals' playing styles (e.g., [104]), game actions (e.g., [105]), or learning strategies (e.g., [106]). Monterrat et al. [107] highlighted the needs of adaptive gamified systems that can provide individual learners with personalized experiences to improve their engagement. In such studies, the researchers have claimed that learner actions traced within the game environments can reveal and predict these individual differences. Models aim to build understanding in three main areas: personalities, player types, and motivational factors.

Personality The source and impact of personality is a subject of debate among psychologists, biologists, and behaviorists. Disagreements between these scholars often are rooted in defining how unique the person is or how to classify a person effectively. A common and often used personality model within the GBL field is the Five-Factor Model (FFM) that is often assessed via the Big Five (BFI) [108]. This theory places personality on five gradient dimensions (i.e., the Big Five): extroversion, agreeableness, openness, conscientiousness, and neuroticism [109, 110]. Each person falls somewhere between: (1) openness: inventive and curious versus consistent and cautious, (2) conscientiousness: efficient and organized versus careless and easy-going, (3) extraversion: outgoing and energetic versus solitary and reserved, (4) agreeableness: friendly and companionate versus challenging and detached, and (5) neuroticism: sensitive and nervous versus secure and confident. Neither side of each spectrum is preferred over the other but by understanding where individual places within each trait, researchers can identify tendencies, predispositions, and behaviors that are similar or dissimilar to other personalities and personality types.

Player Type When it comes to player types, Bartle [104] and Ferro et al. [111] have similar views. Bartle [104] highlighted four main player types. These types are derived from the interaction with the game world and with other players, which are labeled: killer, achiever, socializer, and explorer. For instance, those who are high aptitudes in acting and prefer exploring the world are labeled achievers, whereas the opposite quadrant where a player is more focused on interacting with and focusing on other players is considered a socializer. The concepts of Ferro et al. [111] echo similar views, and however, their player type labels better define how a player interacts with the underlying gamified system rather than their behavior. Player types are labeled: dominant, objectivists, humanists, inquisitive, and creative. Each player type explains how the player interacts with the game world and other game elements and how they take advantage of the game mechanics.

Modeling Personalities and Player Type Several studies depicted how personality concepts can be adopted to player modeling in GBL research. Tlili et al. [112] analyzed a total of 19 studies in which personality was discussed in the development of a learning or educational game. They stated that personality within an educational game system context is still a relatively new area of study with more studies being completed since 2016. Studies on personality used behavioral observations and self-report surveys. For instance, Denden et al. [108] claimed the limitation of a subjective method, such as self-report data, which is not likely to gather the learners' actual feeling or experiences in the use of the system. This study proposed a learning analytics framework for modeling extraversion and openness via a player's game behavior traces using *Naïve Bayes* classifier algorithm. They identified that a game may not be conducive for all personality types and that development should consider environments that match the personality.

Ghali et al. [113] investigated whether the use of multimodal physiological data and BFI traits better detect students' success or help needed when playing an educational game. The research team created a computer-based educational game based on drawing a *Lewis Diagram* for college students. The chemistry lesson of the game aimed at procuring a correct *Lewis Diagram* at varying levels of difficulty. The goal was to understand if the physical behaviors and personality of the individual could be used to predict success in the game. They utilized 40 participants' gameplay traces in which electroencephalogram, eye tracking, and facial expression were all collected during the gameplay. In addition, the participants' pre- and post-test, personality quiz scores, and BFI test results were collected. To build a model that predicts student success and their levels of help needed, they tested different algorithms, including support vector machine and logistic regression models. Although the inclusion of BFI features appeared not to improve the model accuracy in this study, this attempt is worthwhile to consider multimodal data and traits variables to design an evidence-driven GBL system.

Understanding player types can be implemented in-game by creating an adaptable game path as Monterrat et al. [114] explored. Working under the theory that people have different reactions and expectations to game elements and mechanics based on player types, they conducted a quasi-experimental study. Using their gamified online learning environment, *Project Voltaire*, a total of 59 French middle schoolers joined this study. The gamified system implemented the following elements: (1) bright stars, indicating players' game mastery, (2) a leaderboard, and (3) a mnemonic sharing feature which enables the student to write a method to remember the rule and share the technique used by another student. The system used the gameplay traces to build players' profiles which were then utilized for adapting an individual player's interface by selecting gaming features shown next. A collection of questionnaires (i.e., BrainHex typology; [115]) were given to students to identify player types, task complexity, and students' enjoyment. The finding of the study highlighted including gamification features does not always yield positive learning outcomes, rather considering players' preferences and profiles are essential to reduce the complexity level of certain features. While the adaptation process was not found to improve engagement, it laid a critical foundation for future research as little other work has been done in the area.

In the review of personality modeling studies, Denden et al. [108] noted only three out of eleven studies contained a learning feature in their gaming systems. Empirical studies using games in a non-educational context also have shown the efforts of modeling player personalities or player types based on different gameplay traces, which further provided similar insights that players' personality affects their gameplay style or preferences on game genres in educational contexts (e.g., [105, 116, 117]). Understanding players' personality and player types unlocks a promising area of study for games and gamification systems. As researchers seek to more fully identify and embrace the relationship between player personalities and their play types and game design and mechanics through data analytics, more adaptive GBL environments could be developed that amplify learning outcomes. As the field is still growing, the majority of the relevant works are based on a topology specific to massively multiplayer online role-playing game (MMORPG). Developing a player model typology that can be applied for different types of GBL systems still remains a challenge [114].

4 Discussion

A recent review from the *Horizon Report: 2019 Higher Education Edition* [118] included games and gamification as a topic required to be scaled or considered failed, whereas the reports from 2012 to 2014 viewed digital games and gamification as a promising tool for learning. They noted its little impact on an academic institution due to multiple reasons, including a lack of campus budget and limited institutional support. Also, the 2019 *Horizon Report* labeled the concept of adaptive learning as another "fail or scale" topic. Although emerging studies underscored the potentials of adaptive learning, people believe that its technologies are still at the early stage. In light of these trends, the field of GBL has also increasingly considered adaptivity and personalization. This direction aims at designing rich learning experiences for student success. It is possible that the movement around learning analytics in games and gamification systems takes a step further in developing adaptive GBL systems in cost-effective ways.

With regard to identifying personalized design in GBL environments, current research has proposed various player models that seek to explore individual differences based on various factors. In several studies, researchers attempted to explore how personality, observable behaviors, and in-game actions can be modeled in relation to proposed game mechanics. Recent studies have been tailored to giving the right response to the learner's actions and provide real-time feedback using numerous machine learning techniques. In addition to these rising data-mining approaches, various theories have also contributed to understanding the relationship between personalities and player type within GBL. This effort is designed to optimize game support and surroundings based on players' behavior dynamics.

Although increasing adoptions of learning analytics and data mining drove a new movement in GBL research, there are also concerns as well. Limited design and analytics frameworks (e.g., gameplay topology) for different gaming systems remain a significant challenge. As mentioned in Sect. 3.3, more studies have done in recent years, indicating the potentials of this growing field. Since the mid-1980s where various educational games were started developing, many studies have shown the effectiveness of educational games and gamification systems. Yet, relentless implementations of educational games and gamifications without contextualization also raised a question regarding their effectiveness. We hope a new step toward evidence-centered design [119] in GBL research gets academics' attention back to gaming and gamifications in education.

5 Conclusion

This introductory chapter recapitulated how educational games and gamification research evolved, mainly focusing on its theoretical elements and conceptual frameworks. Further, this chapter included how previous GBL design and implementation issues were reviewed in both educational games and gamification contexts. Since identifying students' psychological attributes via in-game systems has been increasingly necessitated, various studies adopted diverse approaches of learning analytics in technology-enhanced learning environments. Lastly, this chapter informed potential roles of data analytics by introducing several studies to simulate a personalized GBL environment in consideration of learners' personality and their play types.

References

1. Young, M. F., Slota, S., Cutter, A. B., Jalette, G., Mullin, G., Lai, B., et al. (2012). Our princess is in another castle: A review of trends in serious gaming for education. *Review of Educational Research, 82,* 61–89.
2. Chassiakos, Y. L. R., Radesky, J., Christakis, D., Moreno, M. A., & Cross, C. (2016). Children and adolescents and digital media. *Pediatrics, 138,* 5.
3. Pew Research Center. (2018). *5 facts about Americans and video games*.
4. Rideout, V. J., Foerh, U. G., & Roberts, D. F. (2010). *Generation M2: Media in the lives of 8- to 18-year-olds*.
5. Watkins, R., Leigh, D., Foshay, R., & Kaufman, R. (1998). Kirkpatrick plus: Evaluation and continuous improvement with a community focus. *Educational Technology Research and Development, 46,* 90–96.
6. Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist, 50*(4), 258–283.
7. Admiraal, W., Huizenga, J., Akkerman, S., & Dam, G. T. (2011). The concept of flow in collaborative game-based learning. *Computers in Human Behavior, 27*(3), 1185–1194.
8. Charles, M. T., Bustard, D., & Black, M. (2009). Game inspired tool support for e-learning processes. *Electronic Journal of e-Learning, 7*(2), 100–110.
9. Charles, M. T., Bustard, D., & Black, M. (2011). Experiences of promoting student engagement through game-oriented learning framework. In *Serious game and edutainment applications*. New York: Springer.
10. Kanthan, R., & Senger, J. L. (2011). The impact of specially designed digital games-based learning in undergraduate pathology and medical education. *Archives of Pathology and Laboratory Medicine, 135*(1), 135–142.
11. Loh, C. S., Sheng, Y., & Ifenthaler, D. (2015). Serious games analytics: Theoretical framework. In C. S. Loh, Y. Sheng, & D. Ifenthaler (Eds.), *Serious games analytics: Methodologies for performance measurement, assessment, and improvement* (pp. 3–29). Cham: Springer International Publishing.
12. Schmidt, R. A., & Lee, T. D. (2011). *Motor control and learning: A behavioral emphasis* (5th ed.). Champaign, IL, US: Human Kinetics.
13. Wallner, G., & Kriglstein, S. (2013). Visualization-based analysis of gameplay data—A review of literature. *Entertainment Computing, 4,* 143–155.
14. Stenros, J. (2017). The game definition game: A review. *Games and Culture., 12*(6), 499–520.
15. Federation of American Scientists. (2006). *Summit on educational games: Harnessing the power of video games for learning*, Washington, DC.
16. Abt, C. C. (1970). *Serious games*. United Press of America.
17. Sawyer, B. (2009). Foreword: From virtual U to serious game to something bigger. In U. Ritterfeld, M. Cody, & P. Vorderer (Eds.), *Serious games: Mechanisms and effects* (pp. xi–xvi). Routledge.
18. Zyda, M. (2005). From visual simulation to virtual reality to games. *IEEE Computer, 38*(9), 25–32.

19. Djaouti, D., Alvarez, J., Jessel, J.-P., & Rampnoux, O. (2011). Origins of serious games. In M. Ma, A. Oikonomou, & L. C. Jain (Eds.), *Serious games and edutainment applications* (pp. 25–43). New York: Springer.
20. Michael, D., & Chen, S. (2006). *Serious games: Games that educate, train, and inform*. Thomson Course Technology.
21. Rowe, J. P., Shores, L. R., Mott, B. W., & Lester, J. C. (2011). Integrating learning, problem solving, and engagement in narrative-centered learning environments. *International Journal of Artificial Intelligence in Education, 21*(1–2), 115–133.
22. Liu, M., Horton, L., Olmanson, J., & Toprac, P. (2011). A study of learning and motivation in a new media enriched environment for middle school science. *Educational Technology Research and Development, 59,* 249–265.
23. Ifenthaler, D., Eseryel, D., & Ge, X. (2012). *Assessment in game-based learning: Foundations, innovations, and perspectives*. New York: Springer.
24. Huotari, K., & Hamari, J. (2012). Defining gamification: A service marketing perspective. In *Proceeding of the 16th International Academic Mindtrek Conference*.
25. Johnson, D., Deterding, S., Kuhn, K.-A., Staneva, A., Stoyanov, S., & Hides, L. (2016). Gamification for health and wellbeing: A systematic review of the literature. *Internet Interventions, 6,* 89–106.
26. Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action. *International Journal of Human Computer Studies, 73*(C), 14–31.
27. Glover, I. (2013). Play as you learn: Gamification as a technique for motivating learners. In *Edmedia+ innovate learning*.
28. Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work?—A literature review of empirical studies on gamification. In *HICSS*.
29. Hamari, J., Shernoff, D., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior, 54,* 170–179.
30. Raczkowski, F. (2014). *Making points the point: Towards a history of ideas of gamification*.
31. Landers, R. N. (2014). Developing a theory of gamified learning: Linking serious games and gamification of learning. *Simulation & Gaming, 45*(6), 752–768.
32. Zichermann, G., & Cunningham, C. (2011). *Gamification by design: Implementing game mechanics in web and mobile apps*. O'Reilly. ISBN 978-1-449-39767-8.
33. Belanich, J., Sibley, D., & Orvis, K. L. (2004). *Instructional characteristics and motivational features of a PC-based game* (ARI Research Report 1822). U.S. Army Research Institute for the Behavioral and Social Sciences.
34. Driskell, J. E., & Dwyer, D. J. (1984). Microcomputer videogame based training. *Educational Technology, 24*(2), 11–17.
35. Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research and Development, 44*(2), 43–58.
36. Smith, P., Sciarini, L., & Nicholson, D. (2007). The utilization of low cost gaming hardware in conventional simulation. In *Proceedings of the Interservice/Industry Training, Simulation, & Education Conference* (pp. 965–972). Orlando, FL: National Defense Industrial Association.
37. Egenfeldt-Nielsen, S. (2004). Practical barriers in using educational computer games. *On the Horizon, 12*(1), 18–21.
38. Egenfeldt-Nielsen, S. (2006). Overview of research on the educational use of video games. *Digital Kompetanse, 3*(1), 184–213.
39. Prensky, M. (2003). *Digital game-based learning*. McGraw-Hill.
40. Ke, F., Xie, K., & Xie, Y. (2016). Game-based learning engagement: A theory- and data-driven exploration. *British Journal of Educational Technology, 47*(6), 1183–1201.
41. De Sousa Borges, S., Durelli, V. H., Reis, H. M., & Isotani, S. (2014). A systematic mapping on gamification applied to education. In *Proceedings of the 29th Annual ACM Symposium on Applied Computing*.

42. Looyestyn, J., Kernot, J., Boshoff, K., Ryan, J., Edney, S., & Maher, C. (2017). Does gamification increase engagement with online programs? A systematic review. *PLOS One, 12,* 3.
43. Landers, R. N., Bauer, K. N., Callan, R. C., & Armstrong, M. B. (2015). Psychological theory and the gamification of learning. In *Gamification in education and business* (pp. 165–186). New York: Springer.
44. Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Springer.
45. Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior, 26,* 331–362.
46. Maehr, M. L. (1976). Continuing motivation: An analysis of a seldom considered educational outcome. *Review of Educational Research, 46,* 443–462.
47. Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science, 5,* 333–369.
48. Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology, 85*(4), 571–581.
49. Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *Journal of Instructional Development, 10,* 3.
50. Keller, J. M. (2009). *Motivational design for learning and performance: The ARCS model approach*. Springer Science & Business Media.
51. Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming, 33,* 441–467.
52. Okan, Z. (2003). Edutainment: Is learning at risk? *British Journal of Educational Technology, 34,* 255–264.
53. Greeno, J., Collins, A., & Resnick, L. (1996). Cognition and learning. In D. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 15–46). Macmillan.
54. Liu, M., Toprac, P., & Yuen, T. (2009). What factors make a multimedia learning environment engaging: A case study. In R. Zheng (Ed.), *Cognitive effects of multimedia learning* (pp. 173–192). Hershey, PA: Idea Group Inc.
55. Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81.
56. Dede, C., Ketelhut, D., & Nelson, B. (2004). *Design-based research on gender, class, race, and ethnicity in a multi-user virtual environment*. Annual Meeting of the American Educational Research Association.
57. Barab, S., Thomas, M., Dodge, T., Carteaux, R., & Tuzun, H. (2005). Making learning fun: Quest Atlantis, a game without guns. *Educational Technology Research and Design., 53,* 86–107.
58. Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. Palgrave/Macmillan.
59. Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C. S., Estock, J. L., et al. (2009). Relationships between game attributes and learning outcomes: Review and research proposals. *Simulation & Gaming, 40,* 217–266.
60. Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning and instruction* (Vol. 3, pp. 223–253). Hillsdale.
61. Choi, D., & Kim, J. (2004). Why people continue to play online games: In search of critical design factors to increase customer loyalty to online contents. *CyberPsychology & Behavior, 7*(1), 11–24.
62. Csikszentmihalyi, I. S. (1992). *Optimal experience: Psychological studies of flow in consciousness*. Cambridge: Cambridge University Press.
63. Tarng, W., & Tsai, W. (2010). The design and analysis of learning effects for a game-based learning system. *Engineering and Technology, 61,* 336–345.

64. Charles, D., Charles, T., McNeill, M., Bustard, D., & Black, M. (2011). Game-based feedback for educational multi-user virtual environments. *British Journal of Educational Technology, 42,* 638–654.
65. Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education, 1,* 3–31.
66. Thurlings, M., Vermeulen, M., Bastiaens, T., & Stijnen, S. (2013). Understanding feedback: A learning theory perspective. *Educational Research Review, 9,* 1–15.
67. Nadolski, R. J., & Hummel, H. G. (2017). Retrospective cognitive feedback for progress monitoring in serious games. *British Journal of Educational Technology, 48*(6), 1368–1379.
68. Collazos, C. A., González, C. S., & García, R. (2014). Computer supported collaborative MOOCs: CSCM. In *Proceedings of the 2014 Workshop on Interaction Design in Educational Environments*.
69. Dale, S. (2014). Gamification: Making work fun, or making fun of work? *Business Information Review, 31*(2), 82–90.
70. Tan, C. T. (2013). Towards a MOOC game. In *Proceedings of the 9th Australasian Conference on Interactive Entertainment: Matters of Life and Death*.
71. Ibáñez, M.-B., Di-Serio, A., & Delgado-Kloos, C. (2014). Gamification for engaging computer science students in learning activities: A case study. *IEEE Transactions on Learning Technologies, 7*(3), 291–301.
72. Kapp, K. M. (2012). *The gamification of learning and instruction*. San Francisco: Wiley.
73. Richter, G., Raban, D. R., & Rafaeli, S. (2015). *Studying gamification: The effect of rewards and incentives on motivation gamification in education and business* (pp. 21–46). New York: Springer.
74. Oprescu, F., Jones, C., & Katsikitis, M. (2014). I play at work—Ten principles for transforming work processes through gamification. *Frontiers in Psychology, 5,* 14.
75. Rauch, M. (2013). Best practices for using enterprise gamification to engage employees and customers. In *The International Conference on Human-Computer Interaction*.
76. King, D., Greaves, F., Exeter, C., & Darzi, A. (2013). *'Gamification': Influencing health behaviours with games*. London: SAGE Publications Sage UK.
77. Pereira, P., Duarte, E., Rebelo, F., & Noriega, P. (2014). A review of gamification for health-related contexts. In *The International Conference of Design, User Experience, and Usability*.
78. Edwards, E. A., Lumsden, J., Rivas, C., Steed, L., Edwards, L., Thiyagarajan, A., et al. (2016). Gamification for health promotion: Systematic review of behaviour change techniques in smartphone apps. *British Medical Journal Open, 6,* 10.
79. Hamari, J., & Koivisto, J. (2015). Why do people use gamification services? *International Journal of Information Management, 35,* 419–431.
80. Calderón, A., & Ruiz, M. (2015). A systematic literature review on serious games evaluation: An application to software project management. *Computers & Education, 87,* 396–422.
81. Loh, C. S. (2012). Information trails: In-process assessment of game-based learning. In D. Ifenthaler, D. Eseryel, & X. Ge (Eds.), *Assessment in game-based learning: Foundations, innovations, and perspectives* (pp. 123–144). New York: Springer.
82. Fan, X., Miller, B. C., Park, K.-E., Winward, B. W., Christensen, M., Grotevant, H. D., et al. (2006). An exploratory study about inaccuracy and invalidity in adolescent self-report surveys. *Field Methods, 18,* 223–244.
83. Quellmalz, E., Timms, M., & Schneider, S. (2009). Assessment of student learning in science simulations and games. In *Proceedings of the Workshop on Learning Science: Computer Games, Simulations, and Education*. National Academy of Sciences.
84. Ifenthaler, D. (2015). Learning analytics. In J. M. Spector (Ed.), *The SAGE encyclopedia of educational technology*. SAGE Publications.
85. Reese, D. D., Tabachnick, B. G., & Kosko, R. E. (2015). Video game learning dynamics: Actionable measures of multidimensional learning trajectories. *British Journal of Educational Technology, 46,* 98–122.
86. Linek, S. B., Öttl, G., & Albert, D. (2010). Non-invasive data tracking in educational games: Combination of logfiles and natural language processing. In L. G. Chova & D. M. Belenguer (Eds.), *Proceeding of International Technology, Education and Development Conference*.

87. Long, P. (2011). *LAK'11: Proceedings of the 1st International Conference on Learning Analytics and Knowledge.* ACM.
88. Freire, M., Serrano-Laguna, Á., Iglesias, B. M, Martínez-Ortiz, I., Moreno-Ger, P., & Fernández-Manjón, B. (2016). Game learning analytics: Learning analytics for serious games. In *Learning, design, and technology: An international compendium of theory, research, practice, and policy* (pp. 1–29).
89. Long, P., & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review, 46*(5), 30.
90. Seif El-Nasr, M., Drachen, A., & Canossa, A. (2013). *Game analytics: Maximizing the value of player data*. New York: Springer.
91. Hämäläinen, W., & Vinni, M. (2010). Classifiers for educational technology. In C. Romero, S. Ventura, M. Pechenizkiy, & R. S. J. d. Baker (Eds.), *Handbook of educational data mining* (pp. 54–74). CRC Press.
92. Bellotti, F., Kapralos, B., Lee, K., Moreno-Ger, P., & Berta, R. (2013). Assessment in and of serious games: An overview. *Advances in Human-Computer Interaction, 2013*, 11.
93. OECD. (2013). *Education at a glance 2013: OECD indicators*.
94. Goldfinch, J., & Hughes, M. (2007). Skills, learning styles and success of first-year undergraduates. *Active Learning in Higher Education, 8,* 259–273.
95. Yorke, M., & Longden, B. (2007). *The first-year experience in higher education in the UK*. Report on Phase 1 of a project funded by the Higher Education Academy.
96. Mah, D.-K. (2016). Learning analytics and digital badges: Potential impact on student retention in higher education. *Technology, Knowledge and Learning, 21,* 285–305.
97. Weinstein, C. E., & Palmer, D. (1990). *LASSI-HS user's manual*. H&H.
98. Guitart, I., & Conesa, J. (2016). Adoption of business strategies to provide analytical systems for teachers in the context of universities. *International Journal of Emerging Technologies in Learning (iJET), 11*(07), 34–40.
99. Guitart, I., Conesa, J., & Casas, J. (2016). *A preliminary study about the analytic maturity of educational organizations*. Paper presented at the 2016 International Conference on Intelligent Networking and Collaborative Systems (INCoS).
100. Androutsopoulou, A., Karacapilidis, N., Loukis, E., & Charalabidis, Y. (2018). Combining technocrats' expertise with public opinion through an innovative e-participation platform. In *IEEE Transactions on Emerging Topics in Computing*.
101. Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt Press.
102. Khenissi, M. A., Essalmi, F., & Jemni, M. (2015). Kinshuk: Learner modeling using educational games: A review of the literature. *Smart Learning Environments, 2,* 6.
103. Yee, N. (2006). Motivations for play in online games. *Cyberpsychology, Behavior, and Social Networking, 9,* 772–775.
104. Bartle, R. A. (2004). *Designing virtual worlds*. USA: New Riders Publishing.
105. Tekofsky, S., Spronck, P., Plaat, A., Van Den Herik, J., & Broersen, J. (2013). Play style: Showing your age. In *2013 IEEE Conference on Computational Intelligence in Games (CIG)* (pp. 1–8).
106. Pavalache-Ilie, M., & Cocorada, S. (2014). Interactions of learner's personality in online learning environment. *Procedia—Social and Behavioral Sciences, 128,* 117–122.
107. Monterrat, B., Desmarais, M., Lavoué, E., & George, S. (2015). A player model for adaptive gamification in learning environments. In *Proceedings of International Conference on Artificial Intelligence in Education (AIED)* (pp. 297–306).
108. Denden, M., Tlili, A., Essalmi, F., & Jemni, M. (2018). Implicit modeling of learners' personalities in a game-based learning environment using their gaming behaviors. *Smart Learning Environments, 5,* 29.
109. McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52,* 81–90.
110. Goldberg, L. R. (1990). An alternative "description of personality": The big-five factor structure. *Journal of Personality and Social Psychology, 59,* 1216–1229.

111. Ferro, L. S., Walz, S. P., & Greuter, S. (2013). Towards personalised, gamified systems: An investigation into game design, personality and player typologies. In *Proceedings of the 9th Australasian Conference on Interactive Entertainment: Matters of Life and Death* (Article No. 7).
112. Tlili, A., Essalmi, F., Jemni, M., Kinshuk, & Chen, N.-S. (2016). Role of personality in computer based learning. *Computers in Human Behavior, 64*, 805–813.
113. Ghali, R., Ouellet, S., & Frasson, C. (2016). LewiSpace: An exploratory study with a machine learning model in an educational game. *Journal of Education and Training Studies, 4*(1), 192–201.
114. Monterrat, B., Lavoué, E., & George, S. (2017). Adaptation of gaming features for motivating learners. *Simulation & Gaming, 48*(5), 625–656.
115. Nacke L. E., Bateman, C., & Mandryk. R. L. (2011). BrainHex: Preliminary results from a neurobiological gamer typology survey. In *ICEC* (pp. 288–293).
116. Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., et al. (2018). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*(1), 84–96.
117. Bunian, S., Canossa, A., Colvin, R., & El-Nasr, M. S. (2018). Modeling individual differences in game behavior using HMM. In *Proceedings of the Thirteenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE-17)* (pp. 158–164).
118. Alexander, B., Ashford-Rowe, K., Barajas-Murph, N., Dobbin, G., Knott, J., McCormack, M., et al. (2019). *Horizon Report: 2019 Higher Education Edition*. EDUCAUSE.
119. Gibson, D., & de Freitas, S. (2016). Exploratory analysis in learning analytics. *Technology, Knowledge and Learning., 21*(1), 5–19.

Part II
Learning Analytics in Educational Games and Gamification Systems

Chapter 2
Rich Representations for Analyzing Learning Trajectories: Systematic Review on Sequential Data Analytics in Game-Based Learning Research

Jewoong Moon and Zhichun Liu

Abstract This chapter focuses on sequential data analytics (SDA), which is one of the prominent behavior analysis frameworks in game-based learning (GBL) research. Although researchers have used a variety of SDA approaches in GBL, they have provided limited information that demonstrates the way they have employed those SDA approaches in different learning contexts. This study used a systematic literature review to demonstrate findings that synthesize SDA's empirical uses in various GBL contexts. In this chapter, we recapitulate the characteristics of several SDA techniques that salient GBL studies have used first. Then, we address the underlying theoretical foundations that explain the proper uses of SDA in GBL research. Lastly, the chapter concludes with brief guidelines that illustrate the way to use SDA, as well as reveal major issues in implementing SDA.

1 Introduction

In game-based learning (GBL) research, a question exists regarding how to capture a wide spectrum of students' learning trajectories during their gameplay [1]. Compared to the emerging learning analytics (LA) and educational data mining (EDM) fields, GBL research highlights primarily the interpretation of students' behavioral data while engaged in gameplay. Researchers require iterative design actions to use evidence-centered design (ECD) in GBL studies [2, 3]. During the phases of ECD, understanding students' learning trajectories is the key to establish and corroborate game design rationales that are associated strongly with their learning outcomes. Further, tracing students' learning trajectories also can help researchers examine the students' performance unobtrusively [4, 5].

Several researchers in GBL have examined prominent factors as precursors of students' learning performance by tracking their behavioral changes

J. Moon (✉) · Z. Liu
Educational Psychology and Learning Systems, Florida State University, Tallahassee, USA
e-mail: jewoong.moon@gmail.com

Z. Liu
e-mail: liulukas91@gmail.com

A. Tlili and M. Chang (eds.), *Data Analytics Approaches in Educational Games and Gamification Systems*, Smart Computing and Intelligence,
https://doi.org/10.1007/978-981-32-9335-9_2

during a game [6, 7]. Students' behavioral changes usually indicate their sequential patterns, which refer to a series of gameplay actions intended to accomplish tasks in a game [8, 9]. Identifying students' gameplay actions is also believed to indicate their mindful learning processes, including decision-making, problem-solving, and affective status during gameplay [8, 10].

To envision students' learning trajectories clearly through their behavior patterns, GBL research has validated and adopted sequential data analytics (SDA) increasingly [8, 11, 12]. Principally, SDA seeks to identify the meaningful associations between a series of game actions and learning outcomes. While prior evaluation frameworks in GBL relied largely on estimating performance differences among groups of learners, SDA pays more attention to capturing hidden causal associations between salient game actions and each student's learning performance, respectively. Thus, SDA is a powerful tool for researchers who attempt to discover which students' game actions are likely to promote their learning outcomes [6, 13, 14].

Although many studies in GBL primarily demonstrated the effects of either digital games or gamified learning applications on students' learning performance [15, 16], few researchers yet have aggregated and synthesized the findings of the way previous GBL studies implemented SDA in different circumstances. Moreover, prior work in GBL has not differentiated the types of SDA depending upon each technique's features and associated GBL design cases.

In response to the aforementioned issues, this chapter explores the underlying issues and procedures used when implementing SDA in GBL research. First, to facilitate the readers' understanding, the chapter explains how SDA has been introduced and adopted in different learning contexts. Further, the chapter describes the ways to conduct SDA and offers examples of the multiple analysis techniques used to portray how learners behave in GBL environments. To collect and analyze the data, this study carried out a systematic literature review that depicted varied SDA's characteristics extensively. During the discussion, the chapter addresses a few key issues in implementing SDA in GBL research. There are two research questions relevant to the scope of this chapter: (1) *How has SDA been used in GBL research? and* (2) *Which key analytics in SDA have been used in GBL research?*

2 Method

2.1 Procedure

A systematic search of multiple online bibliographic databases (i.e., ERIC, IEEE Explore, ScienceDirect, ACM Library, and ISI Web of Science) was conducted on SDA in the GBL environments. In addition to academic bibliographic databases, Google Scholar was also used to provide a wide coverage of relevant studies. This study also examined the reference lists of seminal articles and traced productive authors' work to expand the initial inclusion.

Synonyms of the keywords were used because both GBL and SDA have many related but different expressions. Search terms included combinations of "game-based learning," "educational game," "serious game," "game analytics," "sequential mining," "sequential analysis," "sequence analysis," "sequential data mining," and "sequential pattern mining." If the database provided a thesaurus (e.g., ERIC), an additional search was also made.

The initial search returned 932 results, and the researchers conducted the first round of screening at the title and abstract level. The initial search result was read by two researchers independently to see if both educational game and SDA appear. If only SDA appears, the article is selected only if it is a methodological or commentary publication that informs the application of SDA in GBL. As a result, 129 articles were screened in their entirety. Finally, based on the inclusion and exclusion criteria below, 102 articles were maintained for coding. The flowchart (Fig. 1) shows the search procedure.

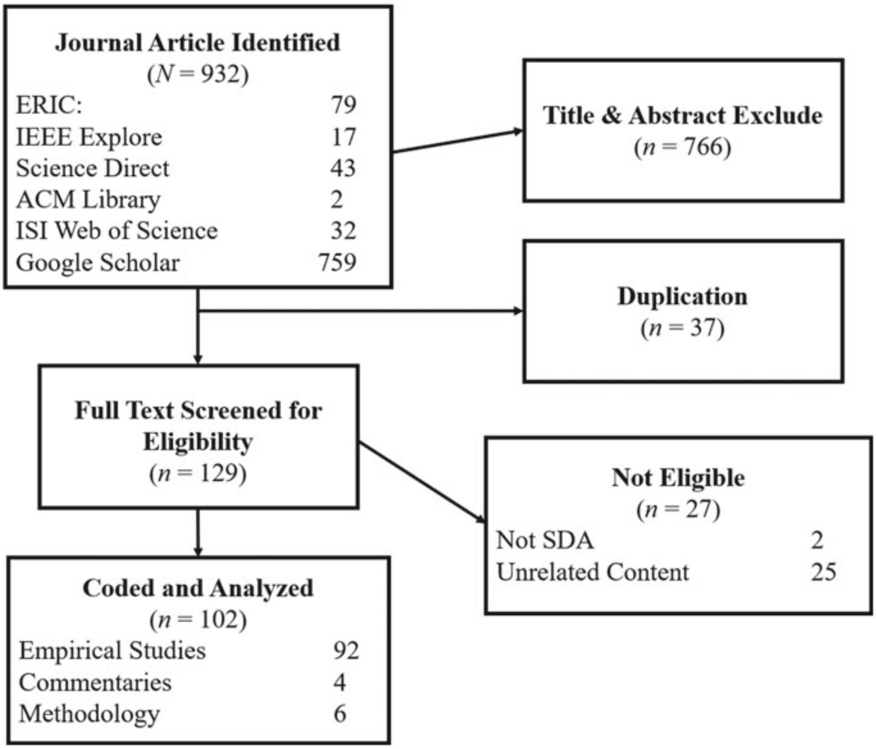


Fig. 1 Study identification flow diagram

2.2 Inclusion and Exclusion Criteria

The inclusion and exclusion criteria were as follows:

(1) Environment relevance: Studies were required to be conducted in a digital GBL environment in which the instruction had to be delivered via game-like interactions. Studies on sports games environments were excluded.
(2) Method relevance: Studies were required to use sequential analytical approach(es) to examine participants' in situ data (e.g., game behavior, affective states, biometric data). Studies that used only other analytical approaches (e.g., cluster analysis) were excluded.
(3) Content relevance: Studies were required to use sequential analytic approach(es) to draw meaningful conclusions. Studies that only focus on adaptivity and usability without clear presentation of SDA were excluded.
(4) Language and quality: Studies were required to be empirical studies written in English and published in peer-reviewed journals, as chapters, or in refereed conference proceedings.
(5) Because this study is a review of a method used widely, several non-empirical articles (e.g., commentaries on game analytics, methodological articles, cases/simulations from other related disciplines) were included but reviewed separately.

2.3 Coding Procedure

This study particularly focused on keeping high reliability of the literature reviews to include high-quality articles that would provide key themes regarding SDAs. The researchers established the initial coding framework that classified seminal articles based on the aforementioned criteria. In particular, this study adopted a qualitative coding framework proposed by constant comparative method [17]. This study paid special attention to maintaining the consistency of the content analysis results by using two coders, who aimed at inter-reliability checking. Two individual coders independently coded the articles. Through in-depth discussions, the coders attained 100% agreement through iterative refinements of the coding results.

3 Findings

3.1 How Has SDA Been Used in GBL Research?

3.1.1 Trends of SDA in the Literatures

The figure shows the trend of SDA publication in the literature (Fig. 2). It shows a growing trend of the empirical studies over the years: Before 2000, very few empirical studies used SDA; between 2000 and 2010, one or two empirical studies were published each year; after 2010, SDA has been more frequently used by researchers. This trend agrees with the growing trend of game analytics GBL in general [18]. However, compared to the massive body of GBL literature, SDA application is still under-studied. Although SDA has been used frequently in many other fields (e.g., economics, behavioral psychology, linguistics), and GBL has been studied long before this decade, many early educational games generally do not capture enough students' behavioral data that can be analyzed with SDA approach. As a result, SDA is less used. Thanks to the advancement of technology of data capturing and storage, GBL researchers today can study the learning experience at a finer granularity with in situ data. Therefore, it is likely that SDA will be used more frequently in GBL in the future.

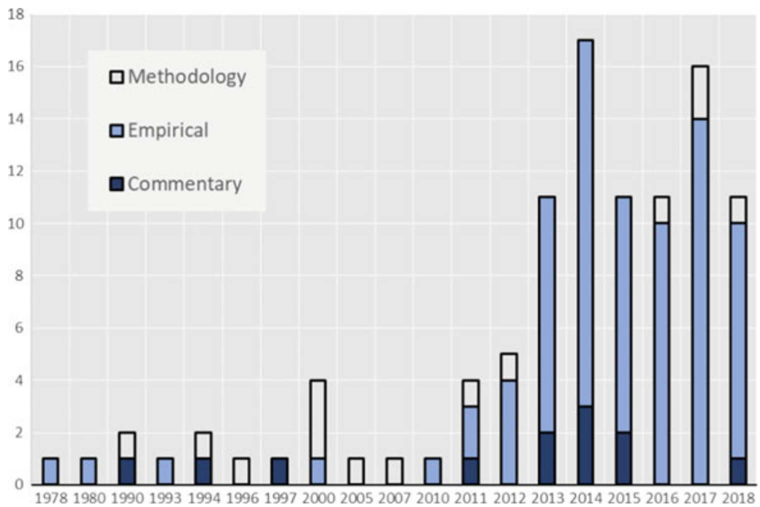


Fig. 2 Tentative trend of SDA publication in the literature

3.1.2 The Advantages of Using SDA

Identifying and understanding users' behavioral patterns in computer-mediated environments have been a major interest of most informatic studies. Early researchers attempted to extract a series of information paths that users experience frequently during their explorations of given information systems [19]. These researchers have used SDA to identify interesting associations among observable variables in data mining. While some early SDA approaches in prior studies were criticized due to their possible commitments of a Type 1 error [20]. SDA has been used widely because it is able to collect micro-level behavioral changes in interactions better [21]. Researchers have stated the possibility of a Type 1 error commitment in implementing SDA because SDA largely relies on the distribution of behavior event data. Basically, SDA flexibly sets its required sample sizes in data analysis. In other words, it indicates that SDA generates its result by repeated behavior events of the same sample. The higher repetition of data observation and analysis a researcher has in the same sample, the higher probabilities of a Type 1 error the data analysis has. Specifically, due to uncorrected z scores in normal distributions, an inflated Type 1 error is a danger, which leads to wrong estimation of a significance in identifying key behavior transition paths. To avoid this issue, a group of researchers suggested the required number of events to analyze the statistical significances of each behavior transition path, respectively.

Despite the fact that there is a likelihood of Type 1 errors in the statistical significance testing of SDA, micro-level SDA investigation still largely contributes to monitoring in-depth and rich representations of behavioral changes during users' interactions. Further, fine-grained SDA also provides baseline data that gauges students' future behaviors and relevant learning support designs [22]. In the wide-ranging spectrum of the SDA field, there are two notable SDA approaches most researchers have conducted: (1) sequential analysis in human observations [20] and (2) sequential pattern mining (SPM) [23]. Sequential analysis in human observations originated from a rudimentary computational model in behavioral analyses. The sequential analysis is based on the statistic that focuses on inferring prominent human decisions or actions as unobservable external stimuli. In using sequential analysis, researchers attempt to condense the scope of their outcome measures, such as human behavior events or actions. Then, the researchers extract the most frequent actions associated with certain stimuli. If the frequency of behavior events and actions from the analysis increases significantly, it suggests a possible association between the outcome measures and specific stimuli. Thus, sequential analysis focuses on estimating the probability of a single variable that is likely to influence the outcome in all series of behavior sequences. Since Bakeman and Brownlee [24] introduced sequential analysis, particularly to collect observable and explicit interactions among people's behaviors, this technique has been fundamental in encouraging researchers' versatile adoption of sequential analysis in different learning environments.

Compared to sequential analysis, which seeks to identify unobserved human behaviors, SPM attempts primarily to capture notable behavior sequences. SPM is designed to collect a variety of series of explicit actions in behavior combinations. In data mining research [15, 19], SPM refers to the approach in which a set of multiple behavioral events in a computer system is identified. The sequences in this analysis can be computer logs archived automatically [6] or behavioral codes [25] that human observers label. Specifically, the logs archived are computer-based trigger events saved in databases, while the behaviors human observers code are generated manually using in-depth transcriptions in qualitative research. By comparison to sequential analysis, SPM aggregates the total number of data-sequence combinations. When implementing SPM, most researchers have used different types of computer algorithms that extract the frequent use of multiple sets of sequence combinations. Kang et al. [6] used SPM to infer the most frequent set of action sequences during students' play in the science game *Alien Rescue*. The researchers adopted a *C-SPADE* algorithm [19] that emphasizes identifying temporal associations as well as chains of gameplay sequences. In addition, Taub et al. [4] employed the *SPAM* algorithm to portray students' gameplay sequences that represent their scientific reasoning skills in GBL. The algorithm in this study emphasized demonstrating all sequences of students' gameplay within the boundary the researchers set.

3.1.3 Applying SDA in GBL Research

GBL research has employed SDA to examine meaningful game behaviors or sequences in response to the nature of SDA techniques mentioned above. In underlying GBL research, emerging topics, such as stealth assessment [1, 26] and serious game analytics [22], have been key notions that explain the importance of in-depth and quantitative analytics. Researchers have used various methods to estimate students' meaningful behaviors following improvements in GBL studies. Under a few key evaluation frameworks in GBL, researchers emphasize using implicit evaluation approaches that prevent students themselves from being aware that they are being assessed during their gameplay. In accordance with the nature of implicit assessment in GBL studies, SDA has been particularly useful to explore the in situ learning contexts that may influence students' gameplay patterns. Compared to other analytics, SDA is able to depict better the way a game evokes certain learning actions. Those learning actions can be an indicator that helps us understand the way students attain meaningful learning experiences and what experiences they may undergo in a game. The following characteristics of GBL demonstrate why SDA is especially helpful in analyzing such data.

SDA and Narrative Design in GBL

In GBL, learners interact with educational games in various ways to develop concepts, learn skills, understand rules, apply knowledge, and solve problems [27]. If we treat all of the interactions as events, a sequence of events can be mapped to represent the learning experience. From a qualitative inquiry perspective, it is important to describe learners' lived experiences in GBL environments [28]. Therefore, rather than treating an educational game as a "black box," researchers should strive to understand the way learners' gameplay leads to learning outcomes. Although a pure narrative research design is uncommon in GBL research, it provides a good way to study gameplayers' learning experiences (i.e., gaming experience). SDA is a useful tool with which to narrate "the story" of gameplay. Both sequential analysis and SPM can help researchers describe the experience (e.g., the frequency of actions, the transition probability between actions) and generate insights (e.g., patterns and notable sequences that emerge).

Although using SDA alone is not narrative research, adopting this approach in GBL is consistent with narrative design. As mentioned before, the primary purpose of the narrative design is to tell the story of people's lives [29]. SDA allows GBL researchers to represent and analyze gamers' learning experiences quantitatively. By describing the sequences and discovering their attributes, researchers can understand the nature of the experience and the way it prepares students. In addition, the narrative design in GBL focuses on representing an individual's experience in chronological order [30]. One of the SDA's most important characteristics is that the sequence of "events" is ordered in a time series. This helps researchers describe a learner's trajectory and understand thereby the way a particular trajectory may lead to a certain learning outcome.

SDA in Inquiry-Based and Discovery Learning

Inquiry-based and discovery learning are common approaches in designing GBL experiences. Educational games often use engaging storytelling to establish the context and propose meaningful problems to the learners [31]. For example, the game *Crystal Island* is designed as a narrative-centered learning environment. To promote knowledge of microbiology, students are asked to play the role of a medical researcher to solve multiple puzzles in an epidemic illness on a remote island [4, 32]. The game's narrative nature provides students with a self-regulated learning experience with which they can acquire the target knowledge through interactions with different modules of the game (e.g., investigative actions, inventory collection, learning resources, NPC dialogue, and game logs).

Because inquiry and discovery learning emphasize the learning interaction significantly, it makes sense to monitor students' actions and their learning trajectories [33]. Fortunately, if the actions occur in the digital GBL environment, computerized systems can capture and record a history log with very high fidelity. If not captured by the computerized system, researchers also can use qualitative observations to capture the actions. Once the sequence of gaming actions is obtained, researchers can use it to accomplish multiple goals by using SDA (e.g., capturing in situ learning contexts, predicting future behaviors, providing personalized suggestions).

3.1.4 SDA Objectives

Capturing In Situ Learning Contexts

In accordance with the key nature of GBL research, SDA is effective in elucidating in situ learning contexts in which students' interactions occur during gameplay [34]. Generally, GBL highlights the examinations of students' adaptive processes when they attend to the game rules and contextual limitations given [35]. Prior studies using SDA have demonstrated clearly the ways in which GBL research seeks to monitor students' behavioral changes during gameplay. Taub et al. [36] employed multi-channel data mining with SDA to identify learners' cognitive- and metacognitive-self-regulatory learning processes. The study implemented the game *Crystal Island*, which is designed to promote students' scientific reasoning skills via exploratory learning. The study sampled 50 students' eye-tracking responses associated with their game sequence logs. The study findings stressed the importance of game sequence mining that collects all combinations of game behaviors that are associated strongly with meaningful learning. Another study by Taub et al. [4] implemented SDA to depict all of the processes in the way students exploited their self-regulated learning strategies by testing students' gameplay patterns in *Crystal Island*. This study sought to determine efficient game behaviors that reached the goal of a single game task and then examined the way in situ game contexts influenced their efficient game behaviors. In addition, Kinnebrew et al. [37] used SDA to determine the affordances of game events that are most likely to be associated with students' learning contexts. They used the game *SURGE Next*, which addresses major physics concepts related to Newton's law. The students in this game were supposed to identify different types of forces that influenced game results. This study aimed at identifying how gameplay data provides researchers with clues to the potential baseline performance that categorizes learners' differences in gameplay. By capturing in situ data, researchers are believed to understand students' contextual adaptation acts, indicating students' engaged behaviors.

Collecting Baseline Data for Future Prediction

SDA has been used not only to capture in situ learning contexts during gameplay, but also to collect baseline data to predict students' future gaming actions. In GBL, prediction is a persistent research goal that gauges students' future learning behaviors in a game. In particular, when designing an educational game, identifying students' typical interactions during gameplay is vital when adopting the design of an adaptive learning system. Generally, an adaptive learning system underscores the responsiveness of a system that adjusts either the level or types of formative feedback.

With respect to an intelligent system's adaptability [38], several researchers have proposed that identifying students' routinized behaviors in a learning environment is necessary to offer sufficient background about the way to provide proper scaffolding in a timely manner. Sun and Giles [9] emphasized sequence prediction as a key category that explains the way human high-order reasoning takes place. To build sequence prediction in an intelligent system, gauging users' prior sequential patterns is indispensable.

Among many GBL studies, several researchers designed different types of prediction models associated with students' baseline gameplay data. Kinnebrew et al. [37] implemented SDA to build their prediction model, which clusters students' gameplay patterns according to their game performance. To corroborate their initial game interaction design in the game *SURGE Next*, they triangulated the findings of students' prior knowledge, learning outcomes, and gameplay behaviors. By implementing SDA, the study collected 65 differential patterns during iterative mining processes. Although the scope of the analysis was largely the demonstration of different game behavior patterns based on the students' prior knowledge, it is noticeable that the study was specifically designed to identify basic game behavior patterns, which is essential to design adaptive game level changes and learning support. In addition, making predictions based on the baseline data enabled the researcher to make the best use of the understanding of the in situ learning trajectory. Inferences also can be made for further analysis (e.g., clustering and regression). Chen [2] displayed how SDA can be employed to establish the design framework of competition-driven educational games. The key design question in this study was how to design game interactions that consider both characteristics of peer competition and task-based learning. This study illustrated the way students used the mini-game *Pet-Master*, which focuses on students' animal-raising skills. To perform the skills required in the game, they needed to use basic math computations and Chinese idioms during gameplay. The study finding reported that the students tended to employ their competition-driven behaviors in the early stages of their gameplay. This study showed a notable behavior cycle in that students were likely to switch their gameplay stages from social dimensions to an economic system. The result of the study indicates that identifying the game behavior cycle was useful to understand the way students are likely to act adaptively in each step of all the game interactions (i.e., peer competition → strengthening the power of their surrogate → finding an equipping system → attending to an economic system).

Providing Personalized Learning Experiences

SDA has been useful to address the way personalized GBL should be designed for GBL research. After predicting students' game actions in GBL, researchers design adaptive learning support to elicit their gameplay to promote meaningful learning. Associated with this issue, the notion of personalized learning has been a key design idea in that a learning system is believed to provide adaptive learning support based on students' prior learning paths and their changes in affective state [39]. The learning system is encouraged to propose different types of scaffolding and external visual stimuli based on both occurrence frequency and types of students' learning actions. Researchers have assessed temporal associations between particular student actions and the timing of using personalized learning support in GBL environments. Through the system framework, GBL can provide contextual feedback that may help students perform their game tasks effectively based on their improvement level. This personalized learning framework reveals the way GBL researchers consider the gradual increase in task complexity based on students' game actions in the personalized system. Students' game actions are vital clues to address the way GBL supports learners' meaningful learning process adaptively. Relevant to this issue, SDA can acquire the entire sequential occurrence of various gameplay actions and identify the relation between gameplay patterns and learning outcomes.

With an understanding of in situ learning trajectories and salient prediction based on baseline data, researchers can use SDA to provide learners with personalized learning experiences in a game. Personalization based on sequential data considers both the context and history of learning and emphasizes the person's experience. Hwang et al. [40] explored the interrelation between students' English listening performance and behavior patterns in a problem-based learning game. In this study, SDA was used to show students' gameplay patterns that are associated with their problem-solving solutions in learning English. The study focused on designing a personalized learning support that considers students' gameplay paths necessary to their problem-solving. With the help of SDA, the researchers extracted the notable combinations of students' explicit game behaviors and proposed a game design framework that considers students' problem-solving approaches, which are represented by their gameplay paths. Other case studies [41–43] also have exploited SDA results to build adaptive learning support systems. For example, Andres et al. [41] evaluated students' behavioral sequences in applying physics concepts in the game *Physics Playground*. The study collected the students' behavior sequences in computed logs that were associated with their problem-solving in physics. They emphasized depicting which set of behavioral sequences indicates students' affective states of their gameplay. The study findings contributed to determining the way a game system detects potential affective variables that may influence their learning automatically.

3.2 Which Key Analytics in SDA Have Been Used in GBL Research?

3.2.1 Data Source and Behavior Coding

Behavior Coding Scheme

Although SDA is one of the quantitative techniques, it is rooted in qualitative inquiry, as noted previously (i.e., narrative research). Therefore, a major data source is observational data based on coding schemes. A behavior coding scheme is a human observation guideline that illustrates which explicit actions should be measured by human observation in accordance with the goals of the study's research questions. The field of SDA has allowed GBL researchers to establish behavior coding schemes that demonstrate the entire list of observable variables. Researchers have used behavior coding schemes due to several reasons. First, a behavior coding scheme helps researchers reliably specify each game behavior state that multiple behavior analysis coders can capture. Detailed descriptions of the coding scheme focus on explaining explicit features of a certain behavior. For example, Hou [25] used a refined behavior coding scheme that features iterative behavior coding steps. They implemented three successive steps of behavior analyses. After they archived all student players' in-game actions for behavior coders' references, two experts in GBL research were joined to build an exploratory coding scheme including students' meaningful in-game behaviors, such as all possible game motions, events, and interactions. Through the axis coding from a qualitative research framework, they distilled 10 major behavior categories. At the last phase of the behavior coding, trained behavior coders labeled the behavior logs of students' behaviors in the archived data. To ensure the coding scheme's reliability, the study checked the *Kappa* coefficient to indicate the inter-reliability of behavior observations among multiple coders. Chang et al. [44] also showed their systematic design of a behavior coding scheme to identify study participants' peer interaction behaviors when using a game. To capture students' social interactions, they recorded the students' behaviors when playing a game. They sampled a total of 3600 interaction actions from the students of 21 groups. This study also reported a *Kappa* coefficient to ensure their high inter-reliability of behavior coders. Those approaches to report the inter-reliability of coders demonstrate how behavior analyses according to their coding schemes were systematically implemented. Second, a behavior coding scheme is also key in quantifying qualitative data [45]. Using a coding scheme gives an opportunity to transform observational to measurable data, such as state and static events. The data analysis hosted by a behavior coding scheme enables researchers to investigate whether an intervention increased the tendency of certain target behavior by indicating numerical changes of the behavior in the coding scheme. Bakeman and Quera [46] demonstrated their sequential data interchange standard (SDIS), including three data types (untimed event, timed event, and interval). Specifically, in their classification, *untimed events* are a kind of static behavior type that displays the frequency of each action in a time frame. Differently, *interval* type refers to the time duration, indicating how long certain behavior lasts. Prior GBL studies have used two kinds of behavior data to examine which in-game action sequences appeared. In addition, researchers aimed at estimating how long learners maintain the state of a specific action when playing a game over time. Conclusively, those two types of behavioral data allow researchers to conduct various association analyses either to test statistical significance of the associations or to illustrate how a set of gameplay patterns appears.

Data Types

Researchers have used several types of behavioral data when they have employed SDA in GBL studies [46–48]. SDA emphasizes the temporal association between two independent behavior states and attempts to identify hidden relations among multiple behavior variables. Further, the technique synthesizes the occurrences of behaviors and simulates students' general learning trajectories during GBL. In prior GBL studies, researchers have adopted behavior variables, including students' in-game action, explicit body actions during gameplay, and groups' game actions as behavior variables. Those types of behavior variables could refer to students' cognitive, affective, and/or metacognitive states and indicate the occurrence of meaningful learning.

The measurement of behavior variables has varied in GBL studies. First, some researchers have employed human observations to evaluate students' behaviors based on a certain behavior coding scheme the researchers developed conceptually in advance. Human observations collected by multiple behavior coders can make it easier to reveal hidden patterns in learning sequences during students' gameplay. It is also likely to generate qualitative themes underlying students' reactions to a game. Prior studies have proven that using behavior coding schemes designed systematically yields reliable measures. Ocumpaugh et al. [47] proposed a systematic behavior coding manual, *BROMP* (Baker Rodrigo Ocumpaugh Monitoring Protocol), that is designed to measure students' explicit behaviors in a classroom setting and has been exploited in various educational data mining studies. For example, several studies have used *BROMP* to capture students' work context, actions, utterances, and facial expressions accompanied by gestures. To replace the necessity of stealth learning indicators, *BROMP* has been introduced to adopt multi-data sensor analysis, which does not rely only on students' learning achievement.

Other GBL research tends to use computer-log analyses that can extract and rearrange all game sequences automatically [5, 6]. By comparison to using human observations, analyzing computer logs requires GBL researchers to represent either a single log or a certain loop of multiple logs as combinations of learning actions in GBL. While a log itself may not include any meaning, the researcher can identify the associations among computer logs and the relevant conceptual variables they show explicitly. Martínez and Yannakakis [5] demonstrated how they generated and defined their behavior logs in gameplayers' log files. First, they defined three major game log types (performance, navigation, and, physiological events). Through iterative dimension reduction as data refinement, they collected a total of 41 gameplay features from the collection of game logs. Afterward, they implemented sequence mining to extract key gameplay sequence patterns. This study was designed to collect multimodal information from players and furnish the features of each game log types to build a predictor of players' affective status in their gameplay. Another study [49] collected students' performance logs with their timestamps to illustrate how students' discovery learning occurred at two specified learning conditions. Their computer logs were used for indicating how students reacted to prompted questions of their learning environment system. The logs were examined to identify whether students understood their learning task regarding the control of variable strategy (CVS). Kang et al. [6] also used their game logs to compute the frequency of major game sequences. Before implementing sequence mining, the researchers defined several log types, which refer to students' meaningful interactions of tools to support their scientific inquiries (e.g., sharing cognitive load, supporting cognitive process, and supporting out-of-reach activities). This study inclusively arranged students' multiple navigation log data and implemented sequential pattern mining.

3.2.2 Analytics Approaches

Behavior Frequency Analysis

Although behavior frequency analysis technically does not include any SDA features, estimating the frequency of game behaviors helps GBL researchers gauge the extent to which students are likely to perform certain game actions associated with either game interactions or events. This analysis focuses only on demonstrating the ratio of certain game actions in the total of game interaction variations. Those studies have adopted behavior frequency analysis as a preliminary analytic technique that captures salient game features to narrow the scope of further sequential analysis. GBL researchers have employed this analysis to determine the way students' game actions tend to occur. However, this analysis has limited ability to explain hidden associations between the occurrences of game actions and the particular period of game interactions. Some studies by Hou [25, 50] have reported the results of behavior frequency analysis. Hou [25] attempted to explain potential gender differences in game patterns and reported the proportion of each in-game behavior on which the students acted, respectively. In addition to reporting the descriptive statistical findings, this study also performed a simple ANOVA to investigate whether there was a statistically significant difference between genders. Further, Hou [50] also employed a behavior frequency analysis that depicted the distribution of game behaviors students used. This study adopted the analysis to explore whether there is a notable tendency in the game actions to be investigated in detail.

Progressive Sequential Analysis

Since coined the term *progressive sequential analysis*, several GBL studies have adopted progressive sequential analysis that grasps students' gradual changes in game behaviors over time during gameplay. In comparison to behavioral frequency analysis, this approach highlights temporal associations in each game behavior students perform. Although the analysis itself does not address the statistical significance of associations among students' game behaviors, it is helpful in portraying sequential connections among the game behaviors. Specifically, a progressive sequential analysis encourages GBL researchers to scrutinize the way students evolve their game sequences associated with learning goals. Hou [50] conducted a progressive sequential analysis to identify students' behavioral transactions that occur by learning phases in a problem-based learning game in English literacy. This study divided the game into three learning phases and then investigated the way students change their behaviors when they encounter each phase during gameplay. This approach has been employed using cluster analysis to examine different transaction patterns according to learning anxiety level. Progressive sequential analysis also has been implemented in qualitative analyses to infer which learning contexts are likely to influence students' behavior changes over time. Li and Liu [51] employed a progressive sequential analysis with in-depth content analysis and explored various transaction types of collaborative problem-solving skills in students' online discussions.

Transitional Probability Matrix

While a behavior frequency analysis reports only the frequency with which certain behaviors occur, a transitional probability matrix allows researchers to identify the extent to which they can estimate whether particular action states may trigger certain actions. Under the *hidden Markov* chain theorem, a stochastic statistical table represents this matrix. As supervised learning in data mining [52], the *hidden Markov* chain underlies the inter-dependency of behavior states. Specifically, the theorem presumes that behavior states in the model influence each other. In the theorem, key behavior patterns in observations are invisible, but each state in outcome behavior states is shown that indicates the likelihood of major behavior states. The *hidden Markov* chain consists of two types of probabilities: *transition* and *emission*. Once GBL researchers highlight the probability distribution of sequential patterns, the *transition probabilities* among various behavior states explain the way transitions among game behaviors take place with a certain probability. On the other hand, the *emission probabilities* show how likely it is that one game behavior promotes conclusively the outcome behaviors on which the researchers may focus.

GBL studies have made several attempts to use transitional probability matrices to explain hidden relations among game behaviors. Chen [2] explored primary school students' gameplay patterns related to their learning actions in the context of a competition-based game. This study used a transitional probability matrix to extract salient game sequences most students were likely to perform. The matrix can filter in salient game actions that exceed a *Z*-score distribution and imply that the behaviors are statistically significant. The study findings showed that two such combinations of behavior states (Learning/Pet-feeding → Item-Shopping → Competing) appeared to be meaningful during gameplay. Snow et al. [53] employed the analysis of transitional probability matrices to infer students' choice patterns in accordance with their reading abilities over time. By using transitional probability matrices, the researchers examined whether low-ability students' regulatory behaviors progressed. To extract meaningful game sequences in all variations of behaviors, this study conducted a residual analysis of each behavior state. The results demonstrated that low-ability students' regulatory behaviors tended to use generative practice game in comparison to those of high-ability students.

Lag Sequential Analysis

Many researchers have used lag sequential analysis (LSA) in behavioral psychology studies. This analysis focuses primarily on identifying a particular chain of behavior sequences statistically. Gottman et al. [20] indicated that this analysis investigates associations in sequenced series of dichotomous behavior states. Researchers usually carry out a chi-squared test to confirm a statistical difference that indicates a particular association among two different behavior states in various combinations of behavior sequences.

In favor of the features aforementioned, GBL researchers have implemented LSA particularly to explore the way certain game interactions are likely to promote the occurrences of certain outcome variables. GBL supposes largely that students' explicit actions intended to solve problems in tasks in their gameplay are associated with meaningful learning. GBL researchers believe that students' learning actions and relevant affective states can be labeled by behavior coding and have attempted to elucidate highly probable connections between students' game experiences and certain learning states, such as engagement and motivation. For example, Hou [25] employed a LSA to demonstrate the sequences in a total of 100 student players' gameplay patterns. By examining the adjusted residuals of each behavior transaction in a Z-distribution, the study found 21 statistically significant game sequences that occurred during students' gameplay. Sun et al. [10] sampled a total of 2362 behavioral codes and then implemented a LSA to extract salient game sequences the students demonstrated and cluster them according to multiple group differences (e.g., flow, anxiety, and boredom).

In comparison to behavior frequency analysis and sequential pattern mining (SPM), LSA has been exploited largely to determine whether a particular behavior association is statistically meaningful. While behavior frequency analysis and SPM are designed generally to portray frequent occurrences of behaviors and their combinations, LSA concentrates on identifying a particular chain that is statistically significant. The statistical findings in the analysis usually are deemed a major causal factor in the outcome variables during the analysis.

Sequential Pattern Mining

Sequential pattern mining (SPM) is among the algorithmic processes that archive a salient set of behavior associations. Since the field of learning analytics emerged, SPM has been adopted in a wide array of informatics studies. Codocedo et al. [54] defined SPM as data analysis that identifies notable patterns in symbols, sets, or events. Lin et al. [55] stated that SPM functions as a decision-maker that discovers new patterns from various perspectives. A series of the analysis procedures in SPM emphasizes pattern identification in time series data. While *LSA* seeks primarily to determine statistically significant associations among behavior states, SPM's entire goal is to describe frequent occurrences of actions. Thus, SPM decomposes all of the variations in action states labeled by systematic behavior coding and then archives all cases of behavior combinations that take place.

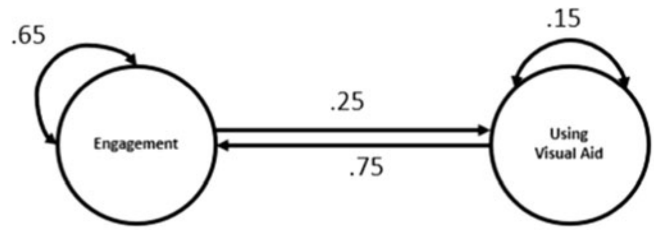
To employ SPM, researchers must use several major data algorithms, such as *generalized sequential pattern* (GSP) [23], *sequential pattern discovery using equivalence classes* (SPADE) [19], and *frequent pattern-projected SPM* (FreeSpan) [56]. First, GSP is a prominent algorithm that computes the number of occurrences of the unit for the analysis very simply. The unit of the algorithm's analysis may refer to a unique behavior state on which a researcher focuses. By using the a priori-based rule [57], this algorithm can generate easily multiple candidate sequence combinations that occur frequently in time series data. The SPADE algorithm also is designed to collect frequent sequences.

This approach has been highlighted specifically because it arranges ID-based sequences in the table vertically. This mining technique draws a table that includes the name of a certain event and its frequency. *FreeSpan* projects a small set of sequence databases and allows the database to increase by adding subsequent fragments of the data. This algorithm has been used because it can process sequential data faster than the a priori-based GSP can, which concentrates primarily on reducing the number of data transaction paths.

GBL researchers' academic interest in SPM has increased steadily. This interest focuses specifically on identifying students' paths in decision-making and capturing behavior patterns that may refer to their game interactions in approaches to problem-solving. To indicate the students' improvement with associated sequences during the game, some studies have attempted to cluster groups by students' learning outcomes. For example, Kang et al. [6] employed a serious game, Alien Rescue, for elementary school students. They adopted SPM with the SPADE algorithm to identify the most frequent game sequences that the students performed. Further, they grouped students by their learning performance. Based on the two groups in the study, the study visualized path diagrams that indicated different sequential patterns a group of students demonstrated. Kinnebrew et al. [37] adopted differential sequence mining (DSM), which implements group clustering to reduce the noise in data preprocessing in SPM. This approach is similar to the adoption of cluster analysis with SPM. However, DSM includes cluster analysis as one of the steps required during data mining. The study divided the participants into two groups and then illustrated their sequential patterns based on their prior learning achievement.

3.2.3 Interpreting and Visualizing Results

SDA is used in GBL research to portray students' learning sequences in different ways. Relevant to sequential analysis in human behaviors, researchers have attempted to demonstrate multiple path diagrams that represent the direction and probability of a single transaction between two independent behavior states [10, 25, 58]. In GBL research, this path diagram depicts the way students change their behavior state to achieve the game's goal. Figure 3 is an example path diagram drawn from sequential analysis of human observations. The arrow denotes a single transaction, indicating



	Engagement	Using Visual Aid
Engagement	.65	.25
Using Visual Aid	.75	.15

Fig. 3 A path diagram based on a transition probability matrix

that one behavior occurs with a certain probability depending on another behavior. As Fig. 3 shows, engagement follows a student using visual aids with a probability of 0.75, while engagement follows the previous engagement status with a probability of 0.65. On the other hand, using visual aids follows engagement with a 0.25 probability, although using visual aids drives another using visual aid behavior state with a 0.15 probability. The path diagram shows by this result, and the matrix table demonstrates the way the findings drawn from sequential analysis can be visualized and interpreted in empirical studies.

On the other hand, SPM adheres to archiving students' major behavior patterns. In students' gameplay, SPM lists either students' frequent behavior-log combinations or salient action patterns that appear to indicate students' attempts to solve game tasks. Although SPM has limited ability to capture hidden associations among multiple behaviors in the pattern the algorithm computed statistically, SPM still is able to map which game stage challenges students and suggest whether embedding scaffolds are needed. SPM usually implements a decision tree diagram that provides an overview of which adaptation should be provided in each stage of students' gameplay. The *IF-THEN* rule in a decision tree diagram helps researchers emphasize providing additional learning support in certain game events that are likely to challenge students.

Interestingly, students' behavior transactions in GBL studies are not always linear; rather, they may be compound because multiple behavior states are interconnected and occur concurrently in students' gameplay. In particular, when students encounter ill-structured game tasks in their play, they are inclined to explore their surrounding circumstances first and attempt to test latent problem-solving solutions while still examining other problems. The behaviors in which students engage to reach their game goal vary and the behavior associations tend to be complex.

3.2.4 Practical Guidelines for Using SDA in GBL Research

The table summarizes six different SDAs with short description, examples, and existing tools. The purpose of this table is to map some techniques that are used most frequently and their examples. It is not exhaustive and not intended to capture all of these SDAs' technical details (Table 1).

4 Discussion

4.1 Uses of SDA in GBL Research

As presented above, we identified three main objectives of using SDA in GBL research: (1) capturing in situ learning context, (2) collecting baseline data for future prediction, and (3) providing personalized learning experiences. Although they appear to be separate objectives, they build upon each other. Capturing in situ learning context is the foundation of the other two objectives because SDA provides a rich representation of the learning trajectory and further analyses are possible only with the meaningful data. At this level, the in situ learning context is represented from a descriptive perspective [4, 25]. Next level of SDA is to use baseline data (i.e., in situ learning context) to make further predictions and draw inferences. For example, if the pattern of behavior is identified, the next possible step(s) can be predicted [8, 22]. Furthermore, based on the correlation between students' behavior sequence pattern and learning outcomes, we can predict the possible outcome given the observed sequence [34, 37]. Finally, based on predictions and inferences, SDA can help to design adaptive learning experiences and personalized support to optimize the learning trajectory. With SDA, scaffolding in GBL can be done at a finer grain level comparing to overall analyses (e.g., Bayesian Network). Hwang et al. [31] argued that based on the identifying students' problem-solving style, additional support should be designed to facilitate the diverse needs of each type of learners.

SDA is a promising technique that can be applied in GBL design and research. The literature also shows an increasing trend of the empirical articles. However, we noticed that most of the work collected in this current review is only at the first level. Predictions and inferences (i.e., level 2) are conducted post hoc instead of a priori. Therefore, the results from SDA analyses may not necessarily transfer beyond the participants. Level 3 objective is based on both level 1 representation and level 2 prediction. Although theoretical papers are published, and small-scale usability examples are presented, we did not see any full example of using SDA for designing adaptivity in GBL.

Table 1 Practical guidelines of six SDAs in GBL research

SDA	Description	Examples	Tools
Behavior frequency analysis	Investigating a simple distribution of behaviors. ANOVA or chi-squared can be used to examine the difference between groups	Andres et al. [41] Hou [25] Hou [50] Neuman et al. [59]	*Software*: GSEQ; BORIS; Observer XT
Sequential analysis (Lag = 1)	Investigating directional transition probabilities among behaviors. Adjusted residual often is used to determine whether a correlation exists Transition probability distribution also can be investigated through Markov chain or hidden Markov model (HMM)	Hsieh et al. [34] Hou [50]	*Software*: GSEQ; BORIS; Observer XT; SADI R packages for HMM
Lag sequential analysis	A general approach of sequential analysis (lag $\geq$ 1). For example: the sequence is A $\rightarrow$ B $\rightarrow$ C. The lag 1 transition is A $\rightarrow$ B or B $\rightarrow$ C. The lag 2 transition is A $\rightarrow$ C. Behaviors are assumed to be sequenced, but not necessarily at equal time intervals	Biswas et al. [60] Jeong et al. [61] Wallner [21] Yang et al. [16]	*Software*: GSEQ; BORIS; Observer XT; SADI *R packages*: HMM; dempmixS4 Sequential; behavseq
Sequential pattern mining	Discovering a set of sequences measured with respect to particular criteria (e.g., frequency, length). Popular algorithms include GSP, SPAM, SPADE, and C-SPADE	Kang et al. [6] Kinnebrew et al. [37]	*R package*: arulesSequences *Free software*: SPMF
Differential sequence mining	Measuring the similarity or difference in behavior patterns between two sets of sequences	Kinnebrew and Biswas [8] Kinnebrew et al. [37] Sabourin et al. [26] Loh et al. [22]	*R packages*: TraMineR; arulesSequences; cluster

4.2 Implementing SDA in GBL

Implementing SDA in GBL research is not only about feeding data to models. As Baker and Inventado [18] pointed out, most educational data mining (EDM) and learning analytics (LA) researchers use learning science and educational theories to guide their selection of analyses techniques and aim to feed back to the theory with the results. SDA should be a systematic research approach guided by theoretical frameworks or specific focus. The first step is to determine the objective and scope of the research which have great implications on what data should be collected and how the results should be interpreted.

With determining the objective and scope of the research, the following things should be considered: (1) what is the data source (e.g., game action, keystroke, utterance, facial expression, biophysical information, interaction among peers)? (2) What data is going to be collected (e.g., selected behavior based on theoretical framework)? (3) How the data is going to be collected (e.g., human observation, automated log file)? These questions should be answered thoroughly before applying SDA.

After collecting the data, cleaning data is usually a major task of SDA. Although studies generally do not report this process, according to general EDM practice, data cleaning is essential to prepare the data for analyses [62]. Similar to establishing a coding scheme in qualitative inquiries, SDA data cleaning can also be an iterative process. Guided by theories, cleaning the data involves (1) formatting the data, (2) omitting irrelevant information, (3) computing variables, and (4) dealing with missing data.

With cleaned data, the researchers can then choose different SDA techniques based on the proposed research questions. The question can range from a simple descriptive question about what behaviors happened to an exploratory question about what the patterns emerged.

4.3 Limitation of Using SDA in GBL

Although SDA in GBL seems to be a promising analytical and mining approach to understand the in situ learning data, the application of the technique might be limited by the following two challenges. First, SDA requires a large volume of data and sometimes high computational power. Although the quantity of analyzable data has increased over the years [18], not all researchers are well-equipped with the ability to access fine-grain data required by SDA easily. Even if the data can be captured, cleaning analyzing the data might consume a lot of computational power. Second, SDA is often performed as post hoc analysis. Therefore, it is challenging to ensure the validity of the results without cross-validating with the participants. In addition, the participants may not even recall some certain behaviors because the data is captured at a fine granularity. Another issue with post hoc analysis is if the scope of the study is biased, data collection will be biased which in turn leads to an unvalidated biased result. Whereas the first challenge is relatively easy to solve because it is almost completely at the hardware level, the second one can be tricky because it relies on the carefully planning and scoping beforehand, information triangulation, and awareness of bias. The section below will highlight two important things to be mindful about when using SDA in GBL research.

4.4 Key Issues in Implementing SDA

4.4.1 Examining Implicit Behaviors with SDA

Commonly, researchers model the learning sequence with the logged game interactions. As mentioned above, computerized systems usually archive the data automatically. However, in some cases, modeling the observed sequence alone is insufficient, because many other variables (e.g., metacognition and affective states) also may affect the game interaction observed. Without introducing these variables, the sequence or patterns observed may not have a clear meaning. In addition, some behaviors or relations (e.g., off-task behavior and dialogue relation) are not manifested explicitly in the game interaction observed. Thus, examining the implicit behaviors/relations can help researchers map the learning trajectory better.

As a result, it is important to examine the variables that also affect the game interaction observed. Because the game interaction is examined with SDA, it also makes sense to examine these variables from a time series perspective. For example, Biswas et al. [60] measured students self-regulated learning skills in gaming interactions with the *hidden Markov model* (HMM). In addition to the activities observed in the game environment (i.e., Betty's Brain, a learn-by-teaching ITS), they constructed three hidden states of problem-solving processes (i.e., information gathering, map building, and monitoring). Based on the HMM probabilistic transition between hidden and observed events, students who learned with teachable agents demonstrated a better metacognitive behavior pattern than did those who learned by themselves. Martínez and Yannakakis [5] proposed a method of multimodal sequential pattern mining. Physiological signals data (i.e., blood volume pulse and skin conductance) were recorded in addition to the game logs. By examining the data from multiple sources, the sequential pattern obtained predicted the user's effect better compared to single modal data. Although combining data from multiple sources or introducing additional variables to the sequential analytics may be complicated and time-consuming, this approach provides more information compared to SDA only on interactions observed and it is easier to frame the discussion based on theoretical foundations.

Another approach is to examine the implicit interactions in game interactions. For example, it may be important to examine off-task behaviors, not only in the GBL experience, but also in general educational research. First, off-task behavior sometimes indicates inattention [63]. Further, students may collaborate with each other sometimes when they are off task. These are all important pieces of information for researchers, because they either can provide an explanation for failure or indication of treatment integrity as a reliability threat. Similarly, when students are not playing, we cannot assume that time freezes. It makes sense to code the off-task behaviors, or away-from-keyboard (AFK) behaviors as well, which may be as simple as a pause. Unfortunately, the authors did not find any GBL study that analyzed these types of behavior.

4.4.2 Post hoc Analysis in SDA and Establishing Causality

Although SDA provides a rich representation of the learning trajectory, it is important to note that it does not provide sufficient evidence to conclude the causation between specific sequence patterns and learning outcomes. Typically, researchers not only describe the learning sequence and discover patterns, but conduct post hoc analyses as well [14, 50, 52]. For example, researchers may categorize learners into multiple groups based on their learning achievement (e.g., high versus low performance groups). Subsequently, they may try to establish a sequential model for each category and compare the differences among groups (e.g., frequency of a behavior and/or pattern, the probability of transition between states). However, there is a potential logical fallacy (i.e., cum hoc ergo propter hoc) when drawing further causal conclusions.

Like observational studies, SDA cannot provide rigid causal relations between variables. Normally, a causal relation is established with the results of randomized controlled trials. If a causal relation exists, one must identify clearly: (1) the cause, (2) relation between cause and effect, and (3) that there is no alternative explanation of the effect [64]. If one attempts to draw any causal conclusion based only on the relation between sequential data and the outcome, there is no alternative explanation for the outcome. Even if the sequence or pattern occurs before the outcome and it seems to occur step by step, the sequence may not necessarily lead to the outcome. Thus, because there is no alternative explanation, a causal relation cannot be established. Similarly, if the "potential cause" is the descriptive data in the sequence model (e.g., frequency of behavior or transition probability between states), the effect of the entire trajectory, other instances, and students' psychological states is all neglected. Therefore, both researchers and readers of SDA should be very cautious about drawing such causal conclusions.

A conservative but safer approach to report post hoc analyses in SDA is to remind the readers of the potential of a logical fallacy. If the goal is to provide implications of causal inferences, to differentiate groups, the researchers should examine not only the sequence per se, but all a priori information available. Similar approaches can be found in studies that have adopted a retrospective cohort design, which will not be discussed in detail here [65]. Based on a closer look at the data, the conclusion should shed light on the possible causes of learning outcomes. Subsequently, randomized controlled trials should be used to examine the proposed causes and the learning outcomes.

5 Conclusion

SDA's primary purpose in many GBL studies has been to identify hidden behavior associations that lead to students' meaningful learning through certain gameplay interactions. Through a systematic literature review, this chapter explored current research using SDA in the context of GBL. Generally, GBL requires students to perform given game tasks and change their actions adaptively based on the surrounding game contexts they encounter. Using SDA not only reveals students' learning sequences, but also provides background channel data that reinforce an adaptive learning system. This chapter also addressed key GBL design features that explain why SDA is effective. Researchers have attempted to measure largely to what extent students are engaged with a game's narratives. By comparison to learning engagement in students' gameplay, researchers have noted that conducting SDA is effective in gauging the effect of the quality of a game narrative's design on students' engagement. In addition, SDA is associated with learning design principles, such as discovery and inquiry learning that elicit students' self-regulated explorations and help them achieve their learning goal.

Although SDA has been limited in confirming the causalities among students' game actions associated with their learning trajectory, there is a clear indication that SDA is able to collect a variety of information datasets that may refer to students' game behaviors related to the occurrence of meaningful learning. SDA research has been employed with a variety of analytic approaches, such as behavior frequency analysis, progressive sequential analysis, transitional probability matrix, lag sequential analysis, and sequential pattern mining. While some SDAs emphasize demonstrating sequential patterns and frequent occurrences of actions primarily, others tend to reveal statistically significant associations between two independent game behavior states. To highlight the salient association among behaviors in SDA, depicting multiple transactions of students' game behaviors in GBL also has been considered as a way to visualize information. This chapter demonstrated an example path diagram to explain the way sequential paths can be interpreted.

References

1. Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. *Computer Games and Instruction, 55*(2), 503–524.
2. Chen, Z.-H. (2014). Exploring students' behaviors in a competition-driven educational game. *Computers in Human Behavior, 35,* 68–74.
3. Ke, F., & Shute, V. (2015). Design of game-based stealth assessment and learning support. In *Serious games analytics* (pp. 301–318). New York: Springer.
4. Taub, M., Azevedo, R., Bradbury, A. E., Millar, G. C., & Lester, J. (2018). Using sequence mining to reveal the efficiency in scientific reasoning during STEM learning with a game-based learning environment. *Learning and Instruction, 54,* 93–103.
5. Martínez, H. P., & Yannakakis, G. N. (2011). Mining multimodal sequential patterns: A case study on affect detection. In *13th International Conference on Multimodal Interfaces* (pp. 3–10).

6. Kang, J., Liu, M., & Qu, W. (2017). Using gameplay data to examine learning behavior patterns in a serious game. *Computers in Human Behavior, 72,* 757–770.
7. Loh, C. S., Sheng, Y., & Ifenthaler, D. (2015). Serious games analytics: Theoretical framework. In *Serious games analytics* (pp. 3–29). New York: Springer.
8. Kinnebrew, J. S., & Biswas, G. (2012). Identifying learning behaviors by contextualizing differential sequence mining with action features and performance evolution. In *International Educational Data Mining Society*.
9. Sun, R., & Giles, C. L. (2001). Sequence learning: From recognition and prediction to sequential decision making. *IEEE Intelligent Systems, 16*(4), 67–70.
10. Sun, J. C.-Y., Kuo, C.-Y., Hou, H.-T., & Lin, Y.-Y. (2017). Exploring learners' sequential behavioral patterns, flow experience, and learning performance in an anti-phishing educational game. *Journal of Educational Technology & Society, 20*(1).
11. Biswas, G., Kinnebrew, J. S., & Segedy, J. R. (2014). Using a cognitive/metacognitive task model to analyze students learning behaviors. In *International Conference on Augmented Cognition* (pp. 190–201).
12. Liao, C. C., Chen, Z.-H., Cheng, H. N., & Chang, T.-W. (2012). Unfolding learning behaviors: A sequential analysis approach in a game-based learning environment. *Research & Practice in Technology Enhanced Learning, 7*(1).
13. Hsieh, Y.-H., Lin, Y.-C., & Hou, H.-T. (2016). Exploring the role of flow experience, learning performance and potential behavior clusters in elementary students' game-based learning. *Interactive Learning Environments, 24*(1), 178–193.
14. Hung, Y. H., Chang, R. I., & Lin, C. F. (2016). Hybrid learning style identification and developing adaptive problem-solving learning activities. *Computers in Human Behavior, 55,* 552–561.
15. Drachen, A., Thurau, C., Togelius, J., Yannakakis, G. N., & Bauckhage, C. (2013). Game data mining. In *Game analytics* (pp. 205–253). New York: Springer.
16. Yang, T.-C., Chen, S. Y., & Hwang, G.-J. (2015). The influences of a two-tier test strategy on student learning: A lag sequential analysis approach. *Computers & Education, 82,* 366–377.
17. Kolb, S. M. (2012). Grounded theory and the constant comparative method: Valid research strategies for educators. *Journal of Emerging Trends in Educational Research and Policy Studies, 3*(1), 83.
18. Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In *Learning analytics* (pp. 61–75). New York: Springer.
19. Zaki, M. J. (2000). Sequence mining in categorical domains: Incorporating constraints. In *9th International Conference on Information and Knowledge Management* (pp. 422–429).
20. Gottman, J., Gottman, J. M., & Roy, A. K. (1990). *Sequential analysis: A guide for behavioral researchers*. Cambridge University Press.
21. Wallner, G., & Kriglstein, S. (2015). Comparative visualization of player behavior for serious game analytics. In *Serious games analytics* (pp. 159–179). New York: Springer.
22. Loh, C. S., Li, I.-H., & Sheng, Y. (2016). Comparison of similarity measures to differentiate players' actions and decision-making profiles in serious games analytics. *Computers in Human Behavior, 64,* 562–574.
23. Srikant, R., & Agrawal, R. (1996). Mining sequential patterns: Generalizations and performance improvements. In *International Conference on Extending Database Technology* (pp. 1–17). New York: Springer.
24. Bakeman, R., & Brownlee, J. R. (1980). The strategic use of parallel play: A sequential analysis. In *Child development* (pp. 873–878).
25. Hou, H.-T. (2012). Exploring the behavioral patterns of learners in an educational massively multiple online role-playing game (MMORPG). *Computers & Education, 58*(4), 1225–1233.
26. Sabourin, J. L., Shores, L. R., Mott, B. W., & Lester, J. C. (2013). Understanding and predicting student self-regulated learning strategies in game-based learning environments. *International Journal of Artificial Intelligence in Education, 23*(1–4), 94–114.
27. Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research, 86*(1), 79–122.

28. Bogdan, R., & Biklen, S. K. (1997). *Qualitative research for education*. MA: Allyn & Bacon Boston.
29. Connelly, F. M., & Clandinin, D. J. (1990). Stories of experience and narrative inquiry. *Educational Researcher, 19*(5), 2–14.
30. Cortazzi, M. (2014). *Narrative analysis*. Routledge.
31. Hwang, G.-J., Chiu, L.-Y., & Chen, C.-H. (2015). A contextual game-based learning approach to improving students' inquiry-based learning performance in social studies courses. *Computers & Education, 81,* 13–25.
32. Lester, J. C., Ha, E. Y., Lee, S. Y., Mott, B. W., Rowe, J. P., & Sabourin, J. L. (2013). Serious games get smart: Intelligent game-based learning environments. *AI Magazine, 34*(4), 31–45.
33. Woo, Y., & Reeves, T. C. (2007). Meaningful interaction in web-based learning: A social constructivist interpretation. *The Internet and Higher Education, 10*(1), 15–25.
34. Hsieh, Y.-H., Lin, Y.-C., & Hou, H.-T. (2015). Exploring elementary-school students' engagement patterns in a game-based learning environment. *Journal of Educational Technology & Society, 18*(2), 336.
35. Vandercruysse, S., & Elen, J. (2017). Towards a game-based learning instructional design model focusing on integration. In *Instructional techniques to facilitate learning and motivation of serious games* (pp. 17–35). New York: Springer.
36. Taub, M., Mudrick, N. V., Azevedo, R., Millar, G. C., Rowe, J., & Lester, J. (2017). Using multi-channel data with multi-level modeling to assess in-game performance during gameplay with Crystal Island. *Computers in Human Behavior, 76,* 641–655.
37. Kinnebrew, J. S., Killingsworth, S. S., Clark, D. B., Biswas, G., Sengupta, P., Minstrell, J., et al. (2017). Contextual markup and mining in digital games for science learning: Connecting player behaviors to learning goals. *IEEE Transactions on Learning Technologies, 10*(1), 93–103.
38. Sabourin, J., Mott, B., & Lester, J. (2013). Discovering behavior patterns of self-regulated learners in an inquiry-based learning environment. In: Lane H.C., Yacef K., Mostow J., Pavlik P. (Eds), *Artificial Intelligence in Education AIED. Lecture Notes in Computer Science* (Vol. 7926). Springer, Berlin, Heidelberg.
39. Lin, C. F., Yeh, Y.-C., Hung, Y. H., & Chang, R. I. (2013). Data mining for providing a personalized learning path in creativity: An application of decision trees. *Computers & Education, 68,* 199–210.
40. Hwang, G.-J., Hsu, T.-C., Lai, C.-L., & Hsueh, C.-J. (2017). Interaction of problem-based gaming and learning anxiety in language students' English listening performance and progressive behavioral patterns. *Computers & Education, 106,* 26–42.
41. Andres, J. M. L., Rodrigo, M. M. T., Baker, R. S., Paquette, L., Shute, V. J., & Ventura, M. (2015). Analyzing student action sequences and affect while playing physics playground. In *International Conference on Educational Data Mining*.
42. Serrano-Laguna, Á., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., & Fernández-Manjón, B. (2017). Applying standards to systematize learning analytics in serious games. *Computer Standards & Interfaces, 50,* 116–123.
43. Shih, W.-C. (2017). Mining learners' behavioral sequential patterns in a Blockly visual programming educational game. In *International Conference on Industrial Engineering, Management Science and Application (ICIMSA)* (pp. 1–2).
44. Chang, C. J., Chang, M. H., Liu, C. C., Chiu, B. C., Fan Chiang, S. H., Wen, C. T., et al. (2017). An analysis of collaborative problem-solving activities mediated by individual-based and collaborative computer simulations. *Journal of Computer Assisted Learning, 33*(6), 649–662.
45. Heyman, R. E., Lorber, M. F., Eddy, J. M., & West, T. V. (2014). *Behavioral observation and coding*.
46. Bakeman, R., & Quera, V. (2011). *Sequential analysis and observational methods for the behavioral sciences*. Cambridge University Press.
47. Ocumpaugh, J., Baker, R. S., & Rodrigo, M. M. T. (2015). *Baker Rodrigo Ocumpaugh monitoring protocol (BROMP) 2.0 technical and training manual*. New York, NY and Manila, Philippines: Teachers College, Columbia University and Ateneo Laboratory for the Learning Sciences.

48. Ocumpaugh, J., Baker, R. S. d., Gaudino, S., Labrum, M. J., & Dezendorf, T. (2013). *Field observations of engagement in reasoning mind* (pp. 624–627).
49. Gobert, J. D., Sao Pedro, M. A., Baker, R. S., Toto, E., & Montalvo, O. (2012). Leveraging educational data mining for real-time performance assessment of scientific inquiry skills within microworlds. *Journal of Educational Data Mining, 4*(1), 104–143.
50. Hou, H.-T. (2015). Integrating cluster and sequential analysis to explore learners' flow and behavioral patterns in a simulation game with situated-learning context for science courses: A video-based process exploration. *Computers in Human Behavior, 48,* 424–435.
51. Li, C.-H., & Liu, Z.-Y. (2017). Collaborative problem-solving behavior of 15-year-old Taiwanese students in science education. *Eurasia Journal of Mathematics, Science and Technology Education, 13*(10), 6677–6695.
52. Wen, C.-T., Chang, C.-J., Chang, M.-H., Chiang, S.-H. F., Liu, C.-C., Hwang, F.-K., et al. (2018). The learning analytics of model-based learning facilitated by a problem-solving simulation game. *Instructional Science, 46*(6), 847–867.
53. Snow, E. L., Jackson, G. T., & McNamara, D. S. (2014). Emergent behaviors in computer-based learning environments: Computational signals of catching up. *Computers in Human Behavior, 41,* 62–70.
54. Codocedo, V., Bosc, G., Kaytoue, M., Boulicaut, J.-F., & Napoli, A. (2017). A proposition for sequence mining using pattern structures. In: Bertet K., Borchmann D., Cellier P., Ferré S. (Eds) *Formal concept analysis. ICFCA 2017. Lecture Notes in Computer Science* (Vol. 10308, pp. 106–121).
55. Lin, T.-J., Duh, H. B.-L., Li, N., Wang, H.-Y., & Tsai, C.-C. (2013). An investigation of learners' collaborative knowledge construction performances and behavior patterns in an augmented reality simulation system. *Computers & Education, 68,* 314–321.
56. Han, J., Pei, J., Mortazavi-Asl, B., Chen, Q., Dayal, U., & Hsu, M.-C. (2000). FreeSpan: Frequent pattern-projected sequential pattern mining. In *ACM SIGKDD* (pp. 355–359).
57. Agarwal, R., & Srikant, R. (1994). Fast algorithms for mining association rules. In *20th International Conference on Very Large Data Bases (VLDB)* (pp. 487–499).
58. Tsai, M.-J., Huang, L.-J., Hou, H.-T., Hsu, C.-Y., & Chiou, G.-L. (2016). Visual behavior, flow and achievement in game-based learning. *Computers & Education, 98,* 115–129.
59. Neuman, Y., Leibowitz, L., & Schwarz, B. (2000). Patterns of verbal mediation during problem solving: A sequential analysis of self-explanation. *The Journal of Experimental Education*, *68*(3), 197–213.
60. Biswas, G., Jeong, H., Kinnebrew, J. S., Sulcer, B., & Roscoe, R. (2010). Measuring self-regulated learning skills through social interactions in a teachable agent environment. *Research and Practice in Technology Enhanced Learning, 5*(02), 123–152.
61. Jeong, A. C. (2010). Assessing change in learners' causal understanding using sequential analysis and causal maps. *Innovative assessment for the 21st century* (pp. 187–205). Boston, MA: Springer.
62. Vialardi, C., Bravo, J., Shafti, L., & Ortigosa, A. (2009). Recommendation in higher education using data mining techniques. In *International Working Group on Educational Data Mining*.
63. Rowe, J. P., McQuiggan, S. W., Robison, J. L., & Lester, J. C. (2009). Off-task behavior in narrative-centered learning environments. In *AIED* (pp. 99–106).
64. Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
65. Mann, C. (2003). Observational research methods. Research design II: Cohort, cross sectional, and case-control studies. *Emergency Medicine Journal, 20*(1), 54–60.

Chapter 3
Opportunities for Analytics in Challenge-Based Learning

Dirk Ifenthaler and David Gibson

Abstract This study is part of a research programme investigating the dynamics and impacts of learning engagement in a challenge-based digital learning environment. Learning engagement is a multidimensional concept which includes an individual's ability to behaviourally, cognitively, emotionally, and motivationally engage in an on-going learning process. Challenge-based learning gives significant freedom to the learner to decide what and when to engage and interact with digital learning materials. In light of previous empirical findings, we expect that learning engagement is positively related to learning performance in a challenge-based online learning environment. This study was based on data from the Challenge platform, including transaction data from 8951 students. Findings indicate that learning engagement in challenge-based digital learning environments is, as expected, positively related to learning performance. Implications point toward the need for personalised and adaptive learning environments to be developed in order to cater for the individual needs of learners in challenge-based online learning environments.

1 Introduction

Challenge-based learning is a pedagogical concept that incorporates aspects of collaborative problem-based learning and contextual teaching and learning while focusing on current real-world problems. Problems vary in terms of their structure. Jonassen [1] classifies problems on a continuum from well-structured to ill-structured. Well-structured problems have a well-defined initial state, a known goal state or solution, and a constrained set of known procedures for solving a class of problems. In contrast, the solutions to ill-structured problems are neither predictable

D. Ifenthaler (✉) · D. Gibson
Curtin University, Bentley, WA, Australia
e-mail: dirk.ifenthaler@curtin.edu.au; dirk@ifenthaler.info

D. Gibson
e-mail: david.c.gibson@curtin.edu.au

D. Ifenthaler
University of Mannheim, Mannheim, BW, Germany

A. Tlili and M. Chang (eds.), *Data Analytics Approaches in Educational Games and Gamification Systems*, Smart Computing and Intelligence,
https://doi.org/10.1007/978-981-32-9335-9_3

nor convergent because they often possess aspects that are unknown. Additionally, they possess multiple solutions or solution strategies or often no solutions at all [2]. Jonassen [3] reiterates that the structure of a problem often overlaps with complexity: Ill-structured problems tend to be more complex, especially those emerging from everyday practice, whereas most well-structured problems tend to be less complex. The complexity of a problem is determined by the number of functions or variables it involves; the degree of connectivity among these variables; the type of functional relationships between these properties; and the stability of the properties of the problem over time [4]. Simple problems are composed of few variables, while ill-structured problems may include many variables that may interact in unpredictable ways. When the conditions of a problem change, a person must continuously adapt his or her understanding of the problem while searching for new solutions, because the old solutions may no longer be viable. Static problems are those in which the factors are stable over time while ill-structured problems tend to be more dynamic [5]. Hence, in order to successfully solve complex and ill-structured problems, the person involved in problem-solving must be able to view and simulate the dynamic problem system in its entirety imagining the events that would take place if a particular action were to be performed [6]. It has been argued convincingly that games can serve as situated problem-solving environments, in which players are immersed in a culture and way of thinking [7, 8].

In this article, we describe the foundations of challenge-based learning and provide an overview of the Curtin Challenge digital learning (Challenge) platform. We then present an assessment and analytics framework linked with Challenge. A case study then demonstrates the analytics capabilities focussing on learning engagement before we conclude with implications and future work.

2 Challenge-Based Learning

The term challenge-based learning arose in the U.S. in the early 2000s with the support of innovative technology groups such as Apple Education, the New Media Consortium, The Society for Information Technology and Teacher Education, and the U.S. Department of Education Office of Educational Technology. Challenge-based learning builds on the practice of problem-based learning, but with an exclusive focus on real-world problems being creatively addressed by diverse collaborative teams. In addition, several key distinctions add relevancy and urgency for students, especially when combined with game-inspired methods such as badges, levels, points, transparent goals and clear progress-related feedback in self-paced learning [9–12].

The pedagogical approach of challenge-based learning adds game-based elements, which creates increased self-empowerment for individuals in teams by making explicit the learning process and higher order goals (not the solutions), providing assessable progress indicators of group process evolution and product quality based on the PL-C-PS framework (rather than focusing on product delivery timelines and expert-only scored quality feedback as in traditional assignments), and utilising exogenous rewards, awards and recognition that go beyond the current context [13].

For example, a team selected as one of the best in the world this year for a solution in water quality, might receive award certificates and recommendation letters that enhance their resumes, increase their opportunities for advanced studies and give the team members bragging rights for their successful collaborative efforts. Game-based additions to challenge-based learning might also include engaging, fun, light-heartedness and wit embedded into self-guided learning experiences [14]; so a challenge-based approach can include these aspects of game-based learning even though the purposes of the engagement are serious for both the learners and the real-world recipients of the team-based solutions and efforts.

Online global learning challenges engage students' curiosity and desire to learn by making central the solving of open-ended problems as a member of a self-organising and self-directing international team [15]. In particular, when delivered as a mobile learning experience using an application platform developed at Curtin University in Western Australia, such challenges can integrate twenty-first century tools, require collaboration, and assist students in managing their time and work schedules, while effectively scaling to large numbers of students.

Research on challenge-based learning is beginning to show impacts such as increased engagement, increased time working on tasks, creative application of technology, and increased satisfaction with learning [16].

3 Challenge

The Challenge platform (http://challenge.curtin.edu.au) is specifically designed to engage learners in solving real-world problems in a social learning environment, with unobtrusive data collection enabling seamless demonstration and assessment of learning outcomes. The platform is being developed to support both individual and team-based learning in primarily open-ended ill-structured problem-solving and project-based learning contexts. Challenge can also support self-guided learning, automated feedback, branching storylines, self-organising teams, and distributed processes of mentoring, learning support and assessment.

A challenge is regarded as a collection of learning artefacts and corresponding learning tasks linked to specific learning outcomes or competences to be demonstrated. Figure 1 shows four of several challenges that have been utilised by over 25,000 students.

From a design perspective, Career, Leadership and English Challenges have been planned for higher education students whereas Global Discovery focusses on a more general audience. Career Challenge includes 14 modules including Who am I?; How do I get to know my industry?; Decision-making strategies; Resumes; Cover letters; Selection criteria; Interviews; Drive your career; Workplace rights and responsibilities; etc. Average completion time is about 1 h per module. The design features of each module contain 'activities' including one to three different learner

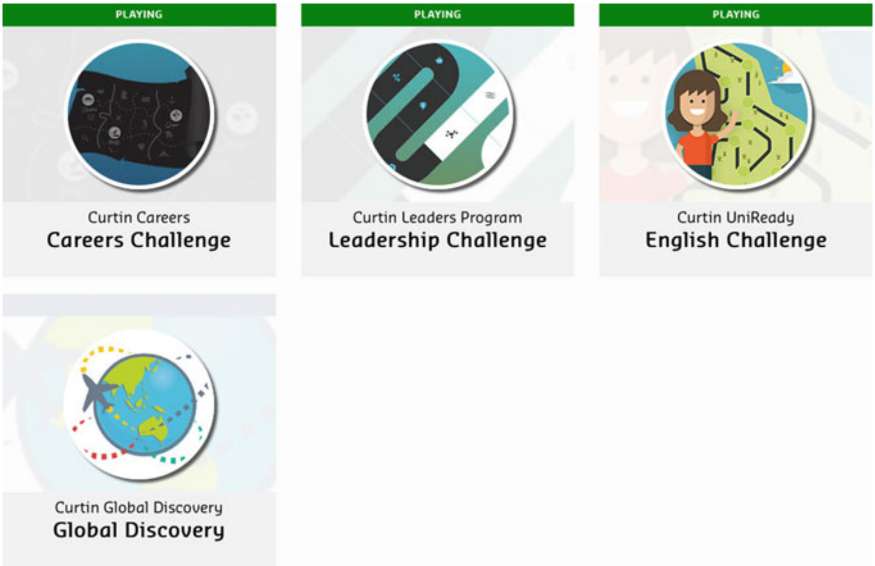


Fig. 1 Curtin challenge platform provides a hub of possible learning opportunities

interactions or 'tasks.' For example, the module *Who am I* in the Career Challenge is a collection of five activities containing learning interactions, such as choosing from among options, writing a short response to a prompt, spinning a wheel to create random prompts, creating, organising, and listing ideas, or matching items. Figure 2 shows an example activity focussing on selection criteria. Learners interact by dragging specific selection criteria to different categories of selection criteria. Immediate feedback is provided through green lines as correct relation or red line as incorrect relation.

Authoring content for the Challenge platform requires collaboration among discipline experts, digital instructional designers, and technologists. The authoring team needs skills in systems thinking, mental models, game-based learning and digital delivery technologies in addition to the pedagogical and content knowledge of instruction in a field of knowledge. Curtin University meets this challenge by forming flexible teams of people from learning and teaching as well as the faculties and larger community to undertake authoring and implementing digital learning on the platform.

The Challenge platform is now of sufficient maturity to extend its reach beyond current students. It is envisaged that new collaborations will be established with other educational institutions that will enable instructors and researchers to share the platform and learning pathways, with learners anywhere in the world; enable new challenge pathways to be developed by educators anywhere for use by learners everywhere; and drive high quality research to inform the future of learning.

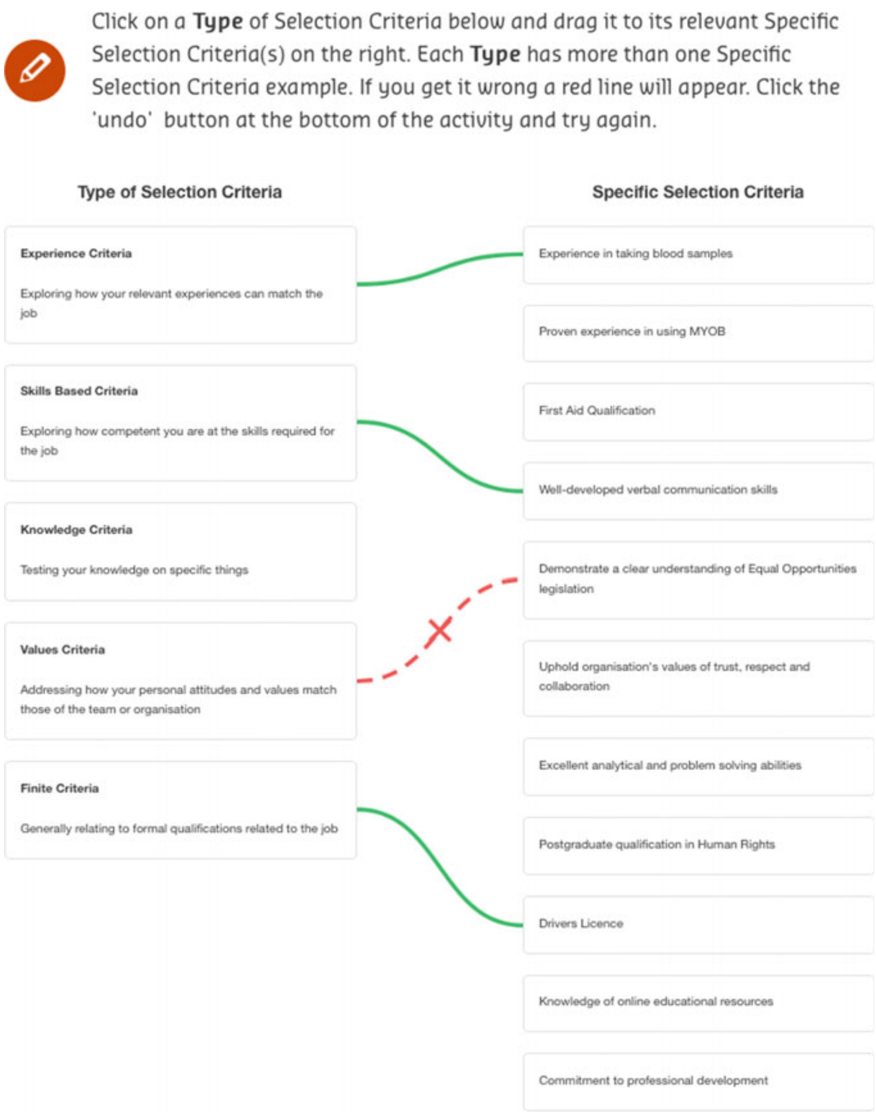


Fig. 2 Task example in the Career Challenge

4 Analytics in Challenge

Research on learning analytics has drawn a lot of attention over the past five years [17]. Learning analytics use static and dynamic information about learners and learning environments—assessing, eliciting, and analysing it—for real-time modelling, prediction, and support of learning processes as well as learning environments [18]. Only recently, serious games analytics has been introduced which focuses on improving game-play and game design as well as optimising learning processes and outcomes [19]. Serious games analytics converts learner-generated information into actionable insights for real-time processing [20]. Metrics for serious games analytics are similar to those of learning analytics and ideally include the learners' individual characteristics (e.g., socio-demographic information, interests, prior knowledge, skills, and competencies) and learner-generated game data (e.g., time spent, obstacles managed, goals or tasks completed, navigation patterns, social interaction, etc.) [20–22].

The application of serious games analytics opens up opportunities for the assessment of engagement within game-based learning environments. The availability of real-time information about the learners' actions and behaviours stemming from key decision points or game-specific events provide insights into the extent of the learners' engagement during game-play. The analysis of single action or behaviour and the investigation of more complex series of actions and behaviours can elicit patterns of engagement, and therefore provide key insights into learning processes [13].

The data traces captured by the challenge platform are highly detailed, with many events per learning activity, which when combined with new input devices and approaches brings the potential for measuring indicators of physical, emotional and cognitive states of the learner. The data innovation of the platform is the ability to capture event-based records of the higher frequency and higher dimensional aspects of learning engagement, which is in turn useful for analysis of the effectiveness and impact on the physical, emotional and cognitive layers of learning caused or influenced by the engagements. This forms a high-resolution analytics base on which people can conduct research into digital learning and teaching as well as into how to achieve better outcomes in scalable digital learning experiences [23].

The process of turning session log files and process stream data into indicators has been recently summarised in Griffin and Care [24] which also notes several precursor research projects with results related to digital learning. Further, a process of exploratory data analysis is required based on post hoc analysis of real people using an appropriately designed digital space to learn. The growing field of learning analytics focused on learning and learners (as opposed to teaching, institutional progress, curriculum and other outcomes) is exploring and expanding the knowledge base concerning the challenges and solutions of the layered and complex analyses required nowadays for a better understanding of the impact of digitally enhanced learning spaces on how people learn—we refer to this as *analytics for learning*.

For the case study described next, a basic educational data mining approach has been utilised [25]. Raw data of the relevant Challenge and cohort were selected and pre-processed including cleaning and matching with external data sources (e.g., student background information). Next, data were transformed focussing on time-based events linked to specific learning activities and related performance. Simple natural language algorithms were applied to open-text responses (including word count, use of language). Standard regression analyses were applied to answer the research hypotheses.

5 Case Study on Learning Engagement

This case study sought to investigate the dynamics of learning engagement in a challenge-based digital learning environment using a data analytics approach. The context of the present study is set in the *Curtin Challenge*. A learner interacts with Challenge content by pointing, clicking, sliding items, vocalising, taking pictures and drawing as well as watching, listening, reading and writing as in typical digital learning environments.

Learning engagement is generally regarded as the time and effort an individual invests on a specific learning activity [26]. Further, learning engagement is a multidimensional concept and understood as the individual's ability to behaviourally, cognitively, emotionally, and motivationally interact with learning artefacts in an on-going learning process [27]. A generally accepted assumption is that the more students engage with a subject matter or phenomenon in question, the more they tend to learn [28]. This assumption is consistent with the theory of self-regulated learning [29] and concepts of engagement [30]. Accordingly, learning engagement is positively linked to desirable learning outcomes or learning performance [31]. Several studies focussing on learning engagement support the assumption that higher engagement of a learner corresponds with higher learning outcomes [32]. However, most of these studies have been conducted in face-to-face learning environments. Accordingly, a confirmation of these findings in digital learning environments is still lacking.

In light of previous empirical findings on learning engagement [33–37], we expect that learning engagement is positively related to learning performance in a challenge-based digital learning environment. Attributes of learning engagement in such a learning environment are conceptualised through several actions: (a) launching a specific activity (task), (b) spending active time on the task, (c) entering a written response, and (d) finishing a task. The learning performance measured in this study is computed by the number of correct answers in a subset of tasks designed with embedded feedback to the student. The hypotheses of this study focus on the attributes of learning engagement and its relation to learning performance specifically in the Career Challenge. We assume that launching specific activities (tasks) is related to the learning performance in challenge-based digital learning environments (Hypothesis 1). Further, we assume that spending active time on tasks is related to learning performance (Hypotheses 2). Also, we expect that the length of written responses is related to the learning performance (Hypothesis 3). The final assumption focusses on the relationship between finishing tasks and learning performance (Hypothesis 4).

5.1 Case Method

The data set of the Career Challenge consists of 52,675,225 rows of raw data containing information of $N_C = 8951$ students (3571 male; 5380 female) with an average age of $M = 25.72$ years (SD $= 6.64$). In a period of 24 months (January 2016–January 2018), students spent a total of 10,239 h interacting with the Career Challenge. The students in the sample stem from various backgrounds and study programmes as well as.

Raw data from the Career Challenge were cleaned and transformed into a transaction data set in which each row represents an event of one user. The dependent variable *learning_performance* (LP) was computed as the number of correct answers in an activity. The variables reflecting attributes of learning engagement were computed as follows: *launching_task* (LT) as the number of activities started by a student; *time_on_task* (TT) as the duration in seconds spent in an activity; *written_response* (WR) as the number of words submitted by a student; *finishing_task* (FT) as the number of activities finished by a student.

5.2 Case Findings

In order to test the above presented four hypotheses, regression analyses were computed to determine whether attributes of learning engagement (i.e., launching task, time on task, written response, finishing task) were significant predictors of learning performance in challenge-based digital learning environments.

Table 1 shows zero-order correlations of attributes of learning engagement and learning performance for the Career Challenge. All correlations were significant at $p < 0.001$. High positive correlations were found between launching task (LT; $M = 6.73$; SD $= 8.95$) and learning outcome (LP; $M = 8.38$; SD $= 13.19$), time on task (TT; $M = 4118.09$; SD $= 6623.88$), as well as written response (WR; $M = 166.92$; SD $= 284.62$). Moderate positive correlations were found for written response and learning outcome as well as time on task. Low positive correlations were found for the remaining variable combinations.

The linear regression analysis for the Career Challenge is presented in Table 2, yielding a ΔR^2 of 0.713 ($F(4, 8950) = 5568.79$, $p < 0.001$). Clearly, the number of activities started by a student (LT; $\beta = 0.80$, $p < 0.001$) positively predicted the learning performance. In addition, the number of activities finished by a student (FT; $\beta = 0.04$, $p < 0.001$) and the number of words submitted by a student (WR; $\beta = 0.13$,

Table 1 Zero-order correlations, means and standard deviations of attributes of learning engagement and learning performance for the Career Challenge

	Zero-order *r*				
	LT	TT	WR	FT	LP
LT	–				
TT	0.771***	–			
WR	0.724***	0.685***	–		
FT	0.355***	0.290***	0.331***	–	
LP	0.839***	0.628***	0.660***	0.340***	–
M	6.73	4118.09	166.92	1.24	8.38
SD	8.95	6623.88	284.62	4.40	13.19

***$p < 0.001$; LP = learning outcome; LT = launching task; TT = time on task; WR = written response; FT = finishing task; $N_C = 8951$

Table 2 Regression analyses predicting learning performance by attributes of learning engagement for the Career Challenge

	R^2	ΔR^2	*B*	SE *B*	β
LP	0.713	0.713			
LT			1.177	0.015	0.80***
TT			0.001	0.001	−0.09***
FT			0.115	0.018	0.04***
WR			0.006	0.001	0.13***

***$p < 0.001$; LP = learning performance; LT = launching task; TT = time on task; FT = finishing task; WR = written response; $N_C = 8951$

$p < 0.001$) positively predicted the learning performance. In contrast, the duration students spent on a task (TT; $\beta = -0.09$, $p < 0.001$) was inversely related to learning performance.

In sum, the four hypotheses are accepted for the Career Challenge, confirming significant relationships between attributes of learning engagement and learning performance.

5.3 *Case Discussion*

The analytic results showed that learning engagement in challenge-based digital learning environments is significantly related to learning performance. These findings support previous studies conducted in face-to-face situations [34, 38, 39]. Significant attributes predicting the learning performance of the student appeared to be the number of activities started and the number of activities finished by a student. This is a reflection of active engagement with the learning environment [33]. At the same time, better learners seem to spend less time on a specific task in the Career Challenge. This may be interpreted as a reflection of existing prior knowledge or a progression towards an advanced learner [40]. Another significant indicator predicting learning performance in the Career Challenge was the number of words submitted in open-text activities. On a surface level, these findings are also related to studies conducted in writing research and clearly reflect the impact of the variation in learning engagement [36, 41].

Limitations of this case study include the restricted access of student data, for example, course load, past academic performance, or personal characteristics, for linking additional data to the reported engagement and performance measures. Combining such additional data in the future will provide a more detailed insight into the multidimensional concepts to be investigated. Second, the Career Challenge does not presently include an overall performance measure which has been validated against an outside criterion. Accordingly, a revision of the learning and assessment design should include additional or revised measures which follow accepted criteria or competence indicators. However, without the externally validated benchmarks, there is sufficient available data which can be used to improve the existing learning design through algorithms focussing on design features and navigation sequences of learners [42–44]. Third, as we included the analysis of open-text answers in our analysis model, this approach is limited by the overall potential of the simple approaches used in natural language processing (NLP). Further development of a future analysis will include a focus on deeper levels of syntactic complexity, lexical sophistication, and quality of writing as well as a deep semantic analysis compared to expert solutions [45, 46].

6 Conclusion

The Challenge platform is being developed to support both individual and team-based learning in primarily open-ended ill-structured problem-solving and project-based learning contexts [47]. The platform can also support self-guided learning, automated feedback, branching story lines, self-organising teams, and distributed processes of mentoring, learning support and assessment [48, 49].

The data traces captured by the Challenge platform are highly detailed, with many events per learning activity. The data and analytics innovation of the Challenge platform is the ability to capture event-based records of higher frequency with the potential to analyse higher dimensional aspects of learning engagement, which we believe may be in turn useful for analysis of the embedded learning design's effectiveness and impact on the physical, emotional and cognitive layers of learning caused or influenced by digital engagements. The data from the challenge-based learning platform forms a high-resolution analytics base on which researchers can conduct studies into learning analytics design [44, 50]. In addition, research on how to achieve better outcomes in scalable digital learning experiences is expected to grow [23, 49].

There are multiple opportunities arising from analytics of digitally delivered challenge-based learning. Analyses of the learning performance transcript, even when automated and multileveled, is a mixture of *conditional and inferential interpretation* that can utilise several frames of reference while adding layers of interpreted evidence, insights concerning the complexity and additional dimensionality to our understanding of the performance and our ability to re-present the performance in the light of our understandings [48]. Practitioners, for example, learning designers, may use the detailed data traces to inform changes required in the design of individual activities or the flow of the story line [44]. Tutors may use the analytics data to monitor and adjust interactions with specific modules or tasks in real-time. For educational researchers, the detailed trace data can provide insights into navigation patterns of individual learners and linking them with individual characteristics or learning performance. Data scientists may use the same data to apply advance analytics algorithms using A/B testing or other analytics approaches.

Future research will focus on the analysis of several large extant data sets from the Challenge platform. Currently, the possibility of adaptive algorithms based on learning engagement and learning performance are being investigated. Such algorithms will enable meaningful microanalysis of individual performance as well as personalised and adaptive feedback to the learner whenever it is needed.

Acknowledgements This research is supported by Curtin University's UNESCO Chair of Data Science in Higher Education Learning and Teaching (https://research.curtin.edu.au/unesco/).

References

1. Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology Research and Development, 45,* 65–94.
2. Funke, J. (2012). Complex problem solving. In N. M. Seel (Ed.), *The encyclopedia of the sciences of learning* (Vol. 3, pp. 682–685). New York, NY: Springer.
3. Jonassen, D. H. (2011). *Learning to solve problems. A handbook for designing problem-solving learning environments*. New York: Routledge.
4. Funke, J. (1991). Solving complex problems: Exploration and control of complex problems. In R. J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 185–222). Hillsdale, NJ: Lawrence Erlbaum.
5. Seel, N. M., Ifenthaler, D., & Pirnay-Dummer, P. (2009). Mental models and problem solving: Technological solutions for measurement and assessment of the development of expertise. In P. Blumschein, W. Hung, D. H. Jonassen, & J. Strobel (Eds.), *Model-based approaches to learning: Using systems models and simulations to improve understanding and problem solving in complex domains* (pp. 17–40). Rotterdam: Sense Publishers.
6. Eseryel, D., Ifenthaler, D., & Ge, X. (2013). Validation study of a method for assessing complex ill-structured problem solving by using causal representations. *Educational Technology Research and Development, 61,* 443–463.

7. Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave-Macmillan.
8. Eseryel, D., Ge, X., Ifenthaler, D., & Law, V. (2011). Dynamic modeling as cognitive regulation scaffold for complex problem solving skill acquisition in an educational massively multiplayer online game environment. *Journal of Educational Computing Research, 45,* 265–287.
9. Ifenthaler, D., Bellin-Mularski, N., & Mah, D.-K. (Eds.). (2016). *Foundations of digital badges and micro-credentials*. New York, NY: Springer.
10. Gibson, D. C., Ostashewski, N., Flintoff, K., Grant, S., & Knight, E. (2013). Digital badges in education. *Education and Information Technologies, 20,* 403–410.
11. Ifenthaler, D. (2011). Intelligent model-based feedback. Helping students to monitor their individual learning progress. In S. Graf, F. Lin, Kinshuk, & R. McGreal (Eds.), *Intelligent and adaptive systems: Technology enhanced support for learners and teachers* (pp. 88–100). Hershey, PA: IGI Global.
12. Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education, 38,* 698–712.
13. Ge, X., & Ifenthaler, D. (2017). Designing engaging educational games and assessing engagement in game-based learning. In R. Zheng & M. K. Gardner (Eds.), *Handbook of research on serious games for educational applications* (pp. 255–272). Hershey, PA: IGI Global.
14. Prensky, M. (2001). *Digital game-based learning*. New York, NY: McGraw-Hill.
15. Harris, D., & Nolte, P. (2007). *Global challenge award: External evaluation year 1 2006–2007*. Montpelier, VT: Vermont Institutes Evaluation Center.
16. Roselli, R., & Brophy, S. (2006). Effectiveness of challenge-based instruction in biomechanics. *Journal of Engineering Education, 95,* 311–324.
17. Ifenthaler, D. (2017). Are higher education institutions prepared for learning analytics? *TechTrends, 61,* 366–371.
18. Ifenthaler, D., & Widanapathirana, C. (2014). Development and validation of a learning analytics framework: Two case studies using support vector machines. *Technology, Knowledge and Learning, 19,* 221–240.
19. Seif El-Nasr, M., Drachen, A., & Canossa, A. (Eds.). (2013). *Game analytics. Maximizing the value of player data*. London: Springer.
20. Loh, C. S., Sheng, Y., & Ifenthaler, D. (2015). Serious games analytics: Theoretical framework. In C. S. Loh, Y. Sheng, & D. Ifenthaler (Eds.), *Serious games analytics: Methodologies for performance measurement, assessment, and improvement* (pp. 3–29). New York, NY: Springer.
21. Berland, M., Baker, R. S., & Bilkstein, P. (2014). Educational data mining and learning analytics: Applications to constructionist research. *Technology, Knowledge and Learning, 19,* 205–220.
22. Gibson, D. C., & Clarke-Midura, J. (2015). Some psychometric and design implications of game-based learning analytics. In P. Isaias, J. M. Spector, D. Ifenthaler, & D. G. Sampson (Eds.), *E-Learning systems, environments and approaches: Theory and implementation* (pp. 247–261). New York, NY: Springer.
23. Gibson, D. C., & Jackl, P. (2015). Theoretical considerations for game-based e-learning analytics. In T. Reiners & L. Wood (Eds.), *Gamification in education and business* (pp. 403–416). New York, NY: Springer.
24. Griffin, P., & Care, E. (Eds.). (2015). *Assessment and teaching of 21st Century skills: Methods and approach*. Dordrecht: Springer.
25. Baradwaj, B. K., & Pal, S. (2011). Mining educational data to analyze students' performance. *International Journal of Advanced Computer Science and Applications, 2,* 63–69.
26. Kuh, G. D. (2009). What student affairs professionals need to know about student engagement. *Journal of College Student Development, 50,* 683–706.
27. Wolters, C. A., & Taylor, D. J. (2012). A self-regulated learning perspective on student engagement. In S. Christenson, A. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 635–651). Boston, MA: Springer.
28. Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: testing the linkages. *Research in Higher Education, 47,* 1–32.

29. Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41,* 64–70.
30. Fredricks, J. A., & McColskey, W. (2012). The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. In S. I. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 763–781). New York, NY: Springer.
31. Klein, S. P., Kuh, G. D., Chun, M., Hamilton, L., & Shavelson, R. (2005). An approach to measuring cognitive outcomes across higher education institutions. *Research in Higher Education, 46,* 251–276.
32. Carini, R. M. (2012). Engagement in learning. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 1153–1156). Boston, MA: Springer.
33. Kirschner, F., Kester, L., & Corbalan, G. (2011). Cognitive load theory and multimedia learning, task characteristics and learning engagement: The current state of the art. *Computers in Human Behavior, 27,* 1–4.
34. Chen, I.-S. (2017). Computer self-efficacy, learning performance, and the mediating role of learning engagement. *Computers in Human Behavior, 72,* 362–370.
35. Miller, B. W. (2015). Using reading times and eye-movements to measure cognitive engagement. *Educational Psychologist, 50,* 31–42.
36. Miller, B. W., Anderson, R. C., Morris, J., Lin, T. J., Jadallah, M., & Sun, J. (2014). The effects of reading to prepare for argumentative discussion on cognitive engagement and conceptual growth. *Learning and Instruction*, *33*, 67–80.
37. Flowerday, T., & Shell, D. F. (2015). Disentangling the effects of interest and choice on learning, engagement, and attitude. *Learning and Individual Differences, 40,* 134–140.
38. Lin, W., Wang, L., Bamberger, P. A., Zhang, Q., Wang, H., Guo, W., et al. (2016). Leading future orientations for current effectiveness: The role of engagement and supervisor coaching in linking future work self salience to job performance. *Journal of Vocational Behavior, 92,* 145–156.
39. Pourbarkhordari, A., Zhou, E. H. I., & Pourkarimi, J. (2016). How individual-focused transformational leadership enhances its influence on job performance through employee work engagement. *International Journal of Business and Management, 11,* 249–261.
40. Ifenthaler, D., & Seel, N. M. (2005). The measurement of change: Learning-dependent progression of mental models. *Technology, Instruction, Cognition and Learning, 2,* 317–336.
41. Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48,* 163–189.
42. Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist, 57,* 1439–1459.
43. Agrawal, R., Golshan, B., & Papalexakis, E. (2016). Toward data-driven design of educational courses: A feasibility study. *Journal of Educational Data Mining, 8,* 1–21.
44. Ifenthaler, D., Gibson, D. C., & Dobozy, E. (2018). Informing learning design through analytics: Applying network graph analysis. *Australasian Journal of Educational Technology, 34,* 117–132.
45. Crossley, S. A. (2013). Advancing research in second language writing through computational tools and machine learning techniques. *Language Teaching, 46,* 256–271.
46. Ifenthaler, D. (2014). AKOVIA: Automated knowledge visualization and assessment. *Technology, Knowledge and Learning, 19,* 241–248.
47. Eseryel, D., Law, V., Ifenthaler, D., Ge, X., & Miller, R. B. (2014). An investigation of the interrelationships between motivation, engagement, and complex problem solving in game-based learning. *Journal of Educational Technology & Society, 17,* 42–53.
48. Gibson, D. C., & Ifenthaler, D. (2018). Analysing performance in authentic digital scenarios. In T.-W. Chang, R. Huang, & Kinshuk (Eds.), *Authentic learning through advances in technologies* (pp. 17–27). New York, NY: Springer.

49. Gibson, D. C. (2018). Unobtrusive observation of team learning attributes in digital learning. *Frontiers in Psychology, 9,* 1–5.
50. Ifenthaler, D. (2017). Learning analytics design. In L. Lin & J. M. Spector (Eds.), *The sciences of learning and instructional design. Constructive articulation between communities* (pp. 202–211). New York, NY: Routledge.

Chapter 4
Game-Based Learning Analytics in Physics Playground

Valerie Shute, Seyedahmad Rahimi and Ginny Smith

Abstract Well-designed digital games hold promise as effective learning environments. However, designing games that support both learning and engagement without disrupting flow [1] is quite tricky. In addition to including various game design features (e.g., interactive problem solving, adaptive challenges, and player control of gameplay) to engage players, the game needs ongoing assessment and support of players' knowledge and skills. In this chapter, we (a) generally discuss various types of learning supports and their influence on learning in educational games, (b) describe stealth assessment in the context of the design and development of particular supports within a game called Physics Playground [2], (c) present the results from recent usability studies examining the effects of our new supports on learning, and (d) provide insights into the future of game-based learning analytics in the form of stealth assessment that will be used for adaptation.

1 Introduction

Play is often talked about as if it were a relief from serious learning. But for children, play is serious learning. —Fred Rogers

As noted in the quote above, Mr. Rogers, along with many others before him, recognized the crucial link between play and learning. If true, then why are our schools more like factories than playgrounds? Before explaining this reality, first imagine the following: Public schools that apply progressive methods—such as individualizing instruction, motivating students relative to their interests, and developing collaborative group projects—to achieve the goal of producing knowledgeable and skilled lifelong learners. The teachers are happy, they work hard, and are valued

V. Shute (✉) · S. Rahimi · G. Smith
Florida State University, Tallahassee, USA
e-mail: vshute@fsu.edu

S. Rahimi
e-mail: sr13y@fsu.edu

G. Smith
e-mail: glsmith2@fsu.edu

A. Tlili and M. Chang (eds.), *Data Analytics Approaches in Educational Games and Gamification Systems*, Smart Computing and Intelligence,
https://doi.org/10.1007/978-981-32-9335-9_4

by the community. In addition, they hold leadership roles in the school and work individually and collectively to figure out the best ways to reach and teach their students. These same teachers create new textbooks and conduct research to see whether their methods worked. School days are structured to allow teachers time to meet and discuss their findings with colleagues.

Is this an ideal vision of schools of the future? Yes and no. According to Ravitch [3], the image above describes several model public schools in the USA in the 1920s and 1930s, inspired by John Dewey's vision of education (e.g., the Lincoln School at Teachers College in New York, and the Winnetka, Illinois, public schools). These schools were engaging places for children to learn and were attractive places for teachers to teach; they avoided the monotonous routines of traditional schools [4].

So, what happened to these exciting experiments of educational reform, and more importantly, what lessons can we learn from them? First, according to Kliebard [5], they failed because the techniques and founding ideas were misapplied by so-called experts who believed that mass education could be accomplished cheaply, employing low-paid and poorly trained teachers who would either follow their manuals or stand aside while students pursued their interests. Second, they failed because the reforms rejected traditional subject-matter curricula and substituted vocational training for the 90% of the student population who, at the time, were not expected to seek or hold professional careers (see [6], "The Elimination of Waste in Education"). Finally, this period also saw mass IQ testing (e.g., [7]) gaining a firm foothold in education, with systematic use of Terman's National Intelligence Test in senior and junior high schools. The testing was aimed specifically at efficiently assigning students into high, middle, or low educational tracks according to their supposedly innate mental abilities.

In general, there was a fundamental shift to practical education going on in the country during the early 1900s, countering "wasted time" in schools and abandoning the classics as useless and inefficient for the masses. Bobbitt, along with some other early educational researchers and administrators such as Ellwood and Ayers [5], inserted into the national educational discourse the metaphor of the school as a "factory." This metaphor has persisted to this day; yet if schools were actual factories, they would have been shut down years ago.

How can we counter this entrenched school-as-factory metaphor? One idea that has garnered a lot of interest lately is to use well-designed digital games as learning environments. Over the past couple of decades, research in game-based learning demonstrates educational games are generally effective learning tools (e.g., [8–10]). When people play well-designed games, they often lose track of time—i.e., experience the state of flow [1]. Teachers try to engage students with learning materials, but the engagement is usually not comparable to that experienced with good video games [10, 11]. Digital game-based learning can be defined as digital activities with goals, interaction, challenges, and feedback that are designed to integrate learning with gameplay.

There is no archetype for game-based learning. That is, games vary by content (e.g., level of narrative, subject matter), design (e.g., 2D, 3D, amount, and quality of graphics), genre (e.g., first-person shooter games, puzzles), and player configuration (e.g., single player, multiplayer, competitive, and cooperative). The complicated part is designing games that support learning and engagement without disrupting flow [1]. For example, Habgood, Ainsworth, and Benford [11] suggest that when the learner is still figuring things out in the game (e.g., learning the basic game mechanics) providing learning content at that point is not a good idea.

Research on game-based learning also recommends the use of learning supports or scaffolds to aid in student knowledge and skill acquisition and transfer, specifically using a mixture of supports in the game, delivered via various modalities [12]. Players may need different types of learning support at different points during gameplay (e.g., more scaffolding at the beginning of the game) or they may prefer a different type of support (e.g., one might not want to see a solution, but instead just receive a hint). However, the research on learning supports and scaffolding used in learning environments in general is conflicted. Some researchers (e.g., [13]) note that learning environments that allow for full autonomy (i.e., student control), without explicit supports, can be more engaging and effective environments than those without such freedom. Clark, Tanner-Smith, and Killingsworth [8] concluded from their meta-analysis that extra instruction (after gameplay, in the form of learning support) did not produce any significant learning differences between game and non-game conditions where compared. However, Wouters and van Oostendorp [14] conducted a meta-analysis on the topic and, overall, found a positive, moderate effect of learning supports ($d = 0.34$, $z = 7.26$, and $p < 0.001$), suggesting the use of learning supports in games can, in fact, improve learning.

The challenge in the design of game-based learning is not just on how to integrate learning through various design features and supports, but also on how to accurately assess the player's knowledge and skills, in real time, and at an actionable grain size. The use of analytics, specifically, stealth assessment [15] built through evidence-centered design [16] is one possible solution. Evidence-centered design (ECD) is a framework to build valid assessments and generate estimates of student performance. It consists of conceptual and computational models working together. The three major models include the competency model, the evidence model, and the task model. The competency model is comprised of everything you want to measure during the assessment. The task model identifies the features of selected learning tasks needed to provide observable evidence about the targeted unobservable competencies. This is realized through the evidence model, which serves as the bridge between the competency model and the task model.

Stealth assessment is a specialized implementation of ECD, where assessment is embedded so deeply into the learning environment it is invisible to the learners [17]. Stealth assessment for game-based learning begins with a student immersed in gameplay, producing a myriad of performance data, all captured in the log file. Next, the automated stealth assessment machinery measures the observables from the logfile data. It then outputs the results of the analysis to the student model (i.e., an individualized competency model based on each student's data) which then provides estimates about the current state of the competencies for each individual student. These estimates are used to provide personalized feedback and other types of learning support to the player who continues to play the game and produce more performance data. Thus, the stealth assessment provides real-time estimates as the cycle continues (for more, see [17]).

The use of analytics in the form of stealth assessment has many benefits. In a well-designed video game, with embedded stealth assessment, students are fully engaged in the experience. Student performance during this level of engagement enables more accurate extraction of students' knowledge and skills. Test anxiety can cause students to perform below their actual ability on tests. Because it is designed to be unobtrusive, stealth assessment frees students from the anxiety of traditional tests and thus improves the reliability and validity of the assessment (e.g., [18, 19]). Another benefit is that the stealth assessment can provide information about students' competencies at a fine grain size. When compared with conventional assessments like multiple-choice formats that yield a single summative score at the end, stealth assessment delivers more valid, reliable, and cumulative information about a student's knowledge and/or skill development. Its automation means teachers do not need to spend time on tedious tasks such as calculating scores and deriving grades. Finally, stealth assessment models, once developed and validated, can be recycled in other learning or gaming environments through the adjustment of the evidence and task models to the particular game indicators (e.g., [20]).

While stealth assessment can provide accurate, detailed information about student performance, it can also provide adaptive support. For example, different types of learning support can be employed and tailored, per student. That is, the what, how, and when of learning supports can be fit to the current needs of individuals. Effectively integrating the assessment and associated supports relies on an iterative design and testing process, with an eye toward adaptivity—where the supports are available or delivered at the right time, and in the right form to maximally enhance learning. Figure 1 illustrates the flow of events in the game, based on information captured in the log file, automatically scored, and accumulated via the stealth assessment's models.

In our design-based research project, we aim to develop and test a methodology for crafting valid and engaging game-based assessments and dynamically linking those assessments to in-game learning supports (i.e., an adaptive algorithm and ongoing feedback; see link 4 in Fig. 1). This methodology will contribute to the design of next-generation learning games that successfully blur the distinction between assessment and learning and harness the power of gameplay data analytics.

In this chapter, we (a) review the literature on various learning supports and their influence on learning and performance in educational games, (b) describe our own experiences with stealth assessment and the design and development of different learning supports within a game called Physics Playground [2], (c) present the results from recent usability studies examining the effects of our new supports on learning, and (d) provide insights into the future of game-based learning analytics in the form of stealth assessment that can be used for adaptation.

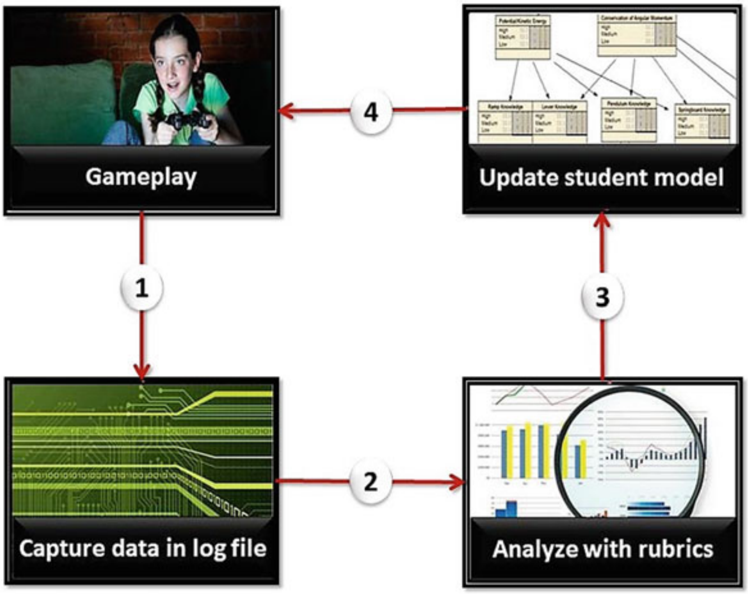


Fig. 1 Stealth assessment cycle

2 Review of the Effects of Learning Supports in Games

Many kinds of learning supports have been used and tested in educational games and other kinds of learning environments. Overall, the results are mixed.

2.1 Types of Supports

In their meta-analysis, Wouters and van Oostendorp [14] identified 24 different types of learning supports and grouped them into ten categories. Here, we limit the focus to three of their categories of support: modeling, advice, and modality. The category of *modeling* includes supports that provide an explication or illustration of how to solve a problem or perform a task in the game. The two most common supports within the modeling category are: (1) scaffolding [21] and (2) worked examples (or expert solutions) [22]. The main purpose of scaffolding is to focus attention via the simplification of the problem at hand [23]. This can be accomplished by providing constraints to the problem that increase the odds of a learner's effective action as they focus attention on specific features of the task in an otherwise complex stimulus field. The main purpose of worked examples is to clearly demonstrate a solution to the task, via human or computer. One possible criticism of this category of support is that learners can replicate a shown solution without having to think about the concepts used to solve the problem.

The category of *advice* (e.g., [24]) refers to support that is intended to guide the learner in the right direction without revealing the solution (as occurs with worked examples). All types of advice (contextualized, adaptive or not) that are game-generated can be grouped under this category. Many popular role-playing games provide advice or hints through characters that players encounter in the game world. These characters can give textual hints during dialogs with the player. Other games allow players to buy hints with earned game rewards like coins or points. Generally, including hints/advice in games is intended to provide support for struggling players, but do they help learning? That likely depends on the type of hint provided (e.g., abstract vs. concrete), and how it is presented (e.g., [25]).

Finally, *modality* [26, 27] [12], like the name indicates, comprises learning supports provided via different modalities (e.g., auditory, visual, textual). Each type can positively or negatively affect learning. For example, Moreno and Mayer [27] found learners remembered more educational content, showed more transfer, and rated more favorably virtual reality environments that used speech rather than on-screen text to deliver learning materials. Providing materials via different channels, or multimodality, is an important component of successful educational games [12]. Ritterfeld and colleagues found that multimodality positively affects knowledge gains in both short-term (i.e., immediate posttest) and long-term (i.e., delayed posttest) evaluations.

2.2 Timing of Supports

The two main questions about learning supports concern what to present (described above), and when to make it available. Csikszentmihalyi [1] claimed that learners learn best when they are fully engaged in some process—i.e., in the state of flow. Inducing flow involves the provision of clear and unambiguous goals, challenging yet achievable levels of difficulty, and immediate feedback (e.g., [28]). Based on flow theory, a task that is too difficult can be frustrating and/or confusing while a task that is too easy may be boring, thus the optimal state (of flow) resides between the two. Similarly, Vygotsky's zone of proximal development (ZPD) suggests that learning is at its best when the learning materials are just at the outer edges of students' existing level of understanding and ability [29]. Considering these two aspects of deep learning—facilitating the state of flow and providing materials compatible with learners' ZPDs—adaptive learning environments such as games can be used to facilitate both by adapting to learners' current competency state(s).

In this section, we define adaptivity—related to the timing of supports—as the ability of a device to alter its behavior according to changes in the environment. In the context of instructional environments, adaptivity can help to provide personalized instruction for different learners and facilitate the state of flow throughout the learning process. An adaptive learning environment should monitor various (and often evolving) characteristics of learners then balance challenges and ability levels to improve learning (for more details on adaptivity in learning environments, see [30]).

One way to include adaptivity in educational games is to use micro-adaptation [31, 32]. This approach entails monitoring and interpreting the learner's particular behaviors, as with stealth assessment. Micro-adaptivity then may provide the learner with appropriate learning supports and/or adjust various aspects of the game (e.g., level difficulty) based on the student model estimates without disrupting the state of flow [31]. Adaptive games can adapt challenges to the current estimated levels of player's knowledge and skills [1], [29] and provide formative feedback [33] and other types of support in unobtrusive ways [34].

In summary, previous findings suggest that the content of the supports, as well as the timing of their availability/delivery, should be carefully designed according to the game features to achieve specific instructional purposes. Cognitive supports are needed in the game to bolster deep conceptual learning. In Physics Playground, this means helping students move from a qualitative, informal understanding of physics to a deeper, more conceptual, and formal understanding. In support of this approach, Hatano asserts that conceptual knowledge gives "meaning to each step of the skill and provides criteria for selection among alternative possibilities for each step within the procedures" ([35], p. 15). Without a pairing between concepts and procedures, students develop only routine expertise, which is the ability to solve a narrowly defined set of predictable and often artificial (school-based) problems. Routine expertise is not very helpful outside of the school setting because it cannot be adjusted for and/or applied to real-life or unexpected situations (see [35, 36]).

We are interested in supporting adaptive expertise, which requires a student to develop conceptual understanding which, in turn, allows that student to invent new solutions to problems and even new procedures for solving problems. However, providing such support in games is more complicated than in other types of interactive learning environments. Cognitive support in games must reinforce emerging concepts and principles to deepen learning and engender transfer to other contexts, but without disrupting engagement while learners are immersed in gameplay.

We now present a case study illustrating how we have been incorporating and testing various supports in our game called Physics Playground.

3 *Physics Playground*—Evolution of Learning Supports

In this section, we elaborate on the process we have gone through to design, develop, test, and revise our learning game, *Physics Playground* (PP). From its inception, *PP* has gone through various changes which led to the development of different versions of the game. For simplicity, we refer to the first version of *PP* as *PPv1*, and to the current version of *PP* (with new task types, learning supports, an incentive system, open student model, and other features) as *PPv2*. Finally, if what we are referring to is general, we simply use the term *PP*.

3.1 The Original Physics Playground—PPv1

PP is a two-dimensional physics game designed to enhance physics understanding [2]. The goal in *PP* is simple—hit a red balloon using a green ball. *PPv1* includes only one type of game level: *sketching*. Using a mouse or stylus, players draw objects on the screen, create simple machines (i.e., ramp, lever, pendulum, or springboard), and target the red balloon with the green ball (see Fig. 2).

As shown in Fig. 2, the solution for the level called *Chocolate Factory* is a ramp affixed to the top part of the level using a pin and including an adequate slope which can guide the ball to the balloon.

In *PPv1*, we used stealth assessment technology [15] to measure player's conceptual understanding of physics related to: (1) Newton's laws of force and motion, (2) potential and kinetic energy, and (3) conservation of angular momentum [37]. Also, *PPv1* was used to measure non-cognitive competencies such as persistence [38] and creativity. Across multiple studies, we consistently found that (1) *PP* can foster motivation and improve learning and (2) the embedded stealth assessment measures are reliable and valid—significantly correlated with external measures (see [38]). Our primary goal, however, has always been improving physics understanding in a fun way—without disrupting flow. To that end, we took a step further to design and develop a new version of *PP* with a broader scope and adaptive learning supports.

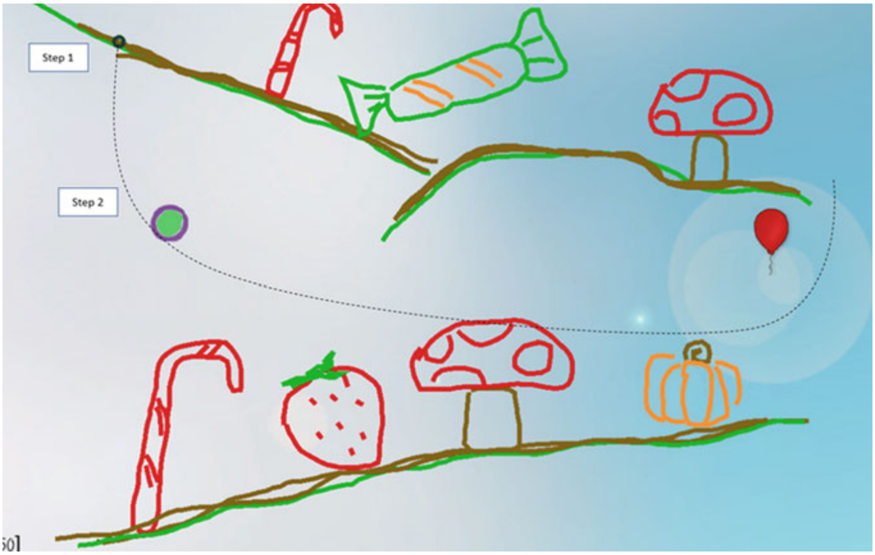


Fig. 2 *Chocolate Factory* level in *PPv1*

3.2 *The Current Physics Playground—PPv2*

The first step we took to develop *PPv2* was to redefine our previously rather sparse physics competency model. The new physics competency model (see Fig. 3) was guided by the Next Generation Science Standards (NGSS) and designed through an iterative process with the help of two physics experts.

New Task Type. The expanded competency model required the addition of new game tasks to the task model to elicit the new evidence. We needed to accurately measure students' proficiency levels per concept with the stealth assessment, so we designed a new task type, *manipulation levels.* In manipulation levels, drawing is disabled, and new features are used to move the ball to the balloon. The new features include (1) sliders related to mass, gravity, and air resistance, (2) the ability to make the ball bounce by clicking the bounciness checkbox, and (3) new sources of exerting external force (e.g., puffer, and static and dynamic blowers) to solve a level. For example, Fig. 4 shows a manipulation level called *Plum Blossom*. In a manipulation level, students get an initial situation with a predefined value for each slider. Then, students can manipulate the variables (i.e., sliders) to solve the level. When the *Plum Blossom* level is played initially, the ball falls, due to gravity, and it is not possible to elevate the ball and hit the balloon. To solve *Plum Blossom*, the player must change the gravity value to zero and use the blue puffer on the left side of the ball to exert a little force. With no gravity, the ball moves slowly to the right and hits the balloon. We designed and developed 55 new manipulation levels targeting various physics concepts in our physics understanding competency model.

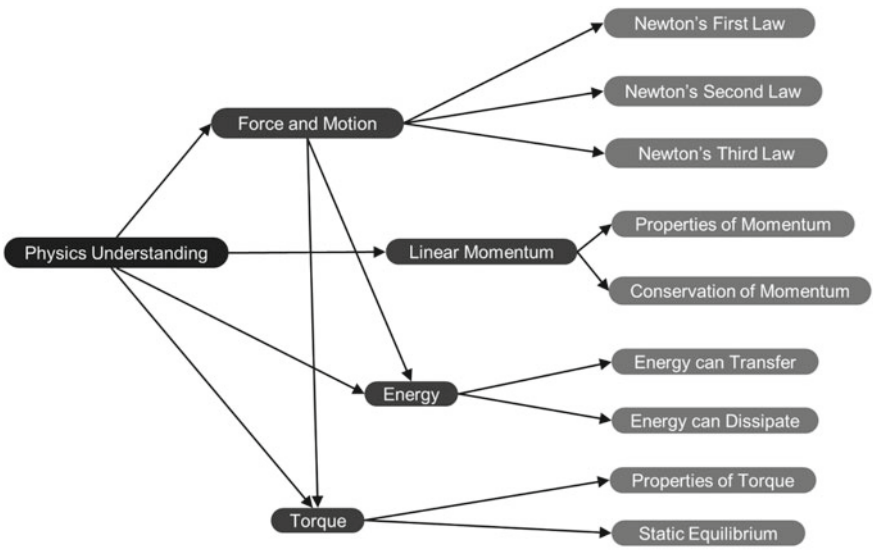


Fig. 3 Competency model for physics understanding in PPv2

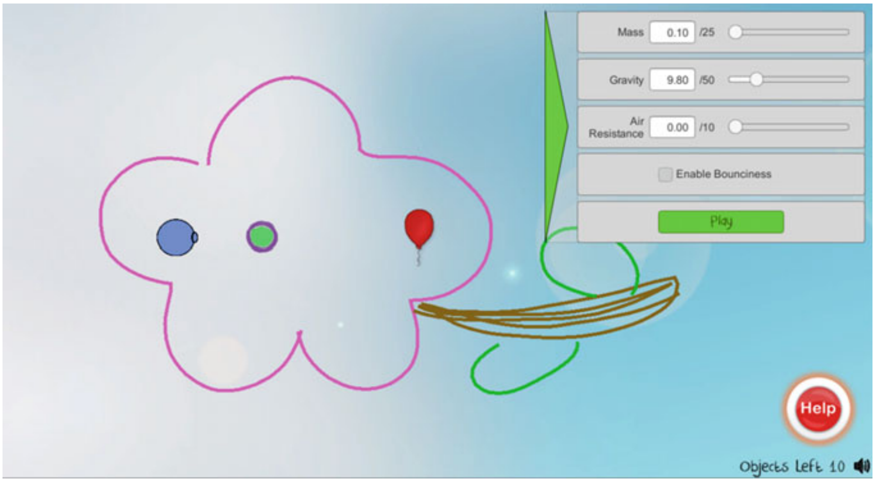


Fig. 4 *Plum Blossom* level in *PPv2*

We tested the new task type in our first usability study. Based on our observations and interviews, students enjoyed playing both sketching and manipulation levels. For the sketching levels, students enjoyed drawing on the screen and inventing creative solutions. However, sketching levels were reported as more difficult than manipulation levels by students. For the manipulation levels, students liked the direct maneuvering of the physics variables and the ability to see immediate results of the change in variables. They also liked that they were not limited by their ability to accurately draw and could focus more on controlling the movement of the ball.

Along with new task types, we also developed other features for the game, such as new game tutorials, the help interface and support content, and an incentive system.

Game Tutorials. Originally, the game tutorials were interactive videos, placed in two separate playgrounds—sketching tutorials and manipulation tutorials. The tutorials introduced essential game tools relevant to our two task types. Students watched how to do something and then had an opportunity to try it. Usability testing revealed that the tutorials were not particularly effective. They were too long, and students could not accurately recall the information later when playing the game. Based on these results and several rounds of revision, the tutorials are now interactive levels with on-screen instructions. Sketching tutorials illustrate how to draw simple machines. For example, in Fig. 5, you can see the lever tutorial, with on-screen, step-by-step instructions. If students follow the instructions, they can easily solve the level, get a silver coin ($10), and move to the next tutorial. Manipulation tutorials show how to use the puffer/blower (that can exert a one-time and small force or a constant force), sliders (i.e., for mass, gravity, and air resistance), and the bounciness function.

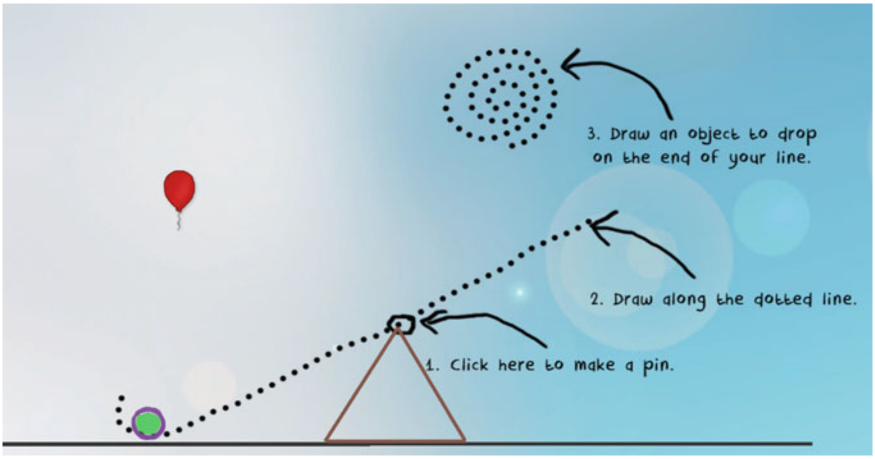


Fig. 5 Lever tutorial level in *PPv2*

Learning Support. When designing the learning supports for *PP,* we had two major components to develop: (1) the location in the game and user interface for the supports and (2) the content and type of supports to offer.

Support Location and Interface. In the first version of *PPv2*, the learning supports were accessed via the "Support Kit" button located on the left side of the screen. Clicking on the button opened the support menu (Fig. 6). However, in the first usability study, students generally did *not* voluntarily click the button to open the

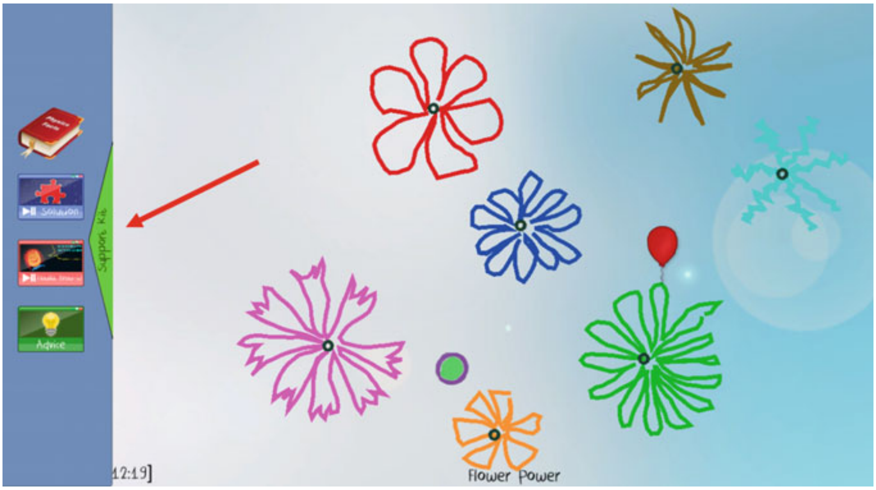


Fig. 6 Old support menu in *Flower Power* level in *PPv2*

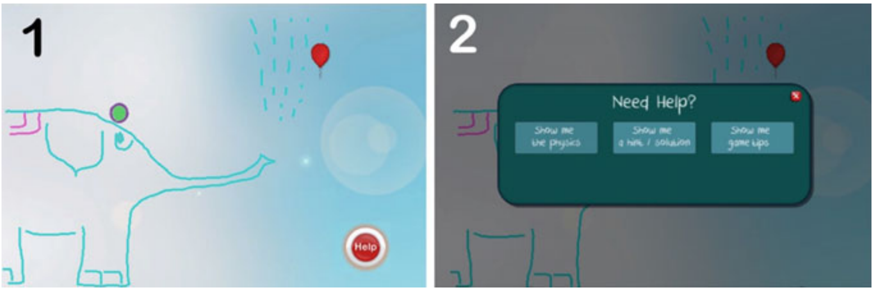


Fig. 7 New "Help" button (left) and help menu (right) in *PPv2*

help menu. Consequently, we decided to revise the color, name, and position of the button to make it clear and visually appealing. Thus, we designed a "Help" button.

The current support interface of the game begins in a level with a player clicking the help button, located in the lower-right corner of the screen (Fig. 7). This triggers a pop-up window showing three options: "Show me the physics", "Show me a hint/solution," and "Show me game tips."

The first two options provide two different paths: learning support or gameplay support. "Show me the physics" comprises the modality-related, content-rich learning supports where students can learn about physics phenomena via multiple representations. "Show me a hint/solution" focuses on game action-oriented, problem solution modeling. Finally, *Show me Game Tips* is where students find game rules and tutorial reminders. Below are descriptions of each of these support options, including their development process.

Support Content. In parallel with designing and developing the support interface, we developed numerous learning supports for *PPv2*: (1) worked examples, (2) animations, (3) interactive definitions, (4) formulas, (5) Hewitt videos, (6) glossary, and (7) hints. In line with Wouters and van Oostendorp's categorization [14], our *worked examples* serve the function of modeling; our *hints* focus on advice; and our *animations*, *formulas*, *Hewitt videos*, and *glossary* promote conceptual understanding via dynamic modalities (i.e., each physics concept in the game can be presented across multimodal representations of the targeted physics knowledge). We designed, developed, tested, and revised these learning supports across three usability studies. Each usability study focused on a different set of supports.

Show me the Physics. Clicking *Show me the Physics* leads the student to the physics support page showing the following options: "Animation", "Definition," "Formula," Hewitt video," and "Glossary" (note that the formula option is not present if the concept does not have an associated formula or equation, see Fig. 8).

Animations. The animations integrate gameplay and support for learning. The team reviewed all the game levels, both sketching and manipulation, focusing on how the level was solved and the competencies with which it was linked. A separate animation has been or will be developed for each intersection of solution agent (i.e., simple machine) and competency. The new support videos utilize the game levels

Fig. 8 "*Show me the Physics*" menu in *PPv2*

to illustrate the physics concepts through failed and successful solution attempts. Narration and on-screen text with video pauses provide an overlay of the physics involved. The new physics animations, with narration, connect the physics concepts with how they are applied in the game to solve a level.

Interactive Definitions. Originally, this was an online document entitled, Physics Facts, which when clicked, led to a non-interactive physics term list, showing definitions and short examples. The results of the first usability test showed students did not like or use this support. They reported it as an intensive reading, that lacked visuals and/or interactions, and was not at all like the other game components. Based on these results, we transformed the boring, static support into an interactive, drag-and-drop quiz. Players now, interactively, construct definitions of terms, like a Cloze task [39]. Clicking definition opens a window showing an incomplete definition with five blanks, five options, and a relevant animation of the term/concept. Students drag each of the five phrases to the correct blanks within the definition. If the dragged phrase is not correct, it snaps back to its original place. When the blanks are correctly filled, a congratulation message pops up and displays the complete definition of the term.

- *Formulas*. In collaboration with the physics experts, we created annotated mathematical formulas for the physics terms. Clicking on the formula option reveals the formula, along with a short explanation of each component/variable.
- *Hewitt videos.* Hewitt videos allowed students to watch a short (1–2 min) physics video developed by Paul Hewitt explaining the primary concept related to the level. The physics experts helped select the most relevant videos for the game competencies. With Paul Hewitt's permission, the team edited the length of the videos to make them illustrate a targeted competency.
- *Glossary*. The glossary provides brief explanations of 28 physics terms. The terms have been selected, edited, and revised by the physics experts.

Show me a Hint or Solution. Clicking on this option takes the student to either a worked example or a hint—both of which are linked to the specific level being played.

- *Worked examples.* Worked examples are videos of expert solutions of game levels. All worked examples are less than a minute long with the majority being less than 30 s. We created at least one worked example for each game level and solution agent (130 + levels—both task types). From our first and second usability studies, we found that students liked worked examples and selected this support more frequently than any of the other types. However, this support enabled students to solve levels without thinking or problem solving first. Consequently, our new incentive system (discussed in detail later) charges for viewing this support.
- *Hints*. In the first version of PPv2, this support was called *Advice.* When this support was selected, it triggered a short, general hint for solving a level (e.g., "Remember that a larger force will cause an object to accelerate faster"). Results of the first usability test showed this support was not effective. Students commented that the advice was too vague and thus unhelpful. So, we replaced the original advice with level-specific physics solution hints (e.g., "Try drawing a ramp").

Show me Game Tips. If students are playing the game for an extended period of time, they will likely forget some of the game mechanics and ways to draw different simple machines (e.g., ramp or lever). Consequently, we developed a support related to gameplay—show me game tips. When students select this support, a window opens with tabs that each contains game play reminders (Fig. 9).

- *"Controls" and "Simple Machines."* These only appear when the player is in a sketching level. When a student clicks on the "Controls" tab, a scrollable page pops up showing game mechanics (i.e., nudge, draw an object, and delete an object for a sketching level, etc.). When a student clicks on the "Simple Machines" tab, images

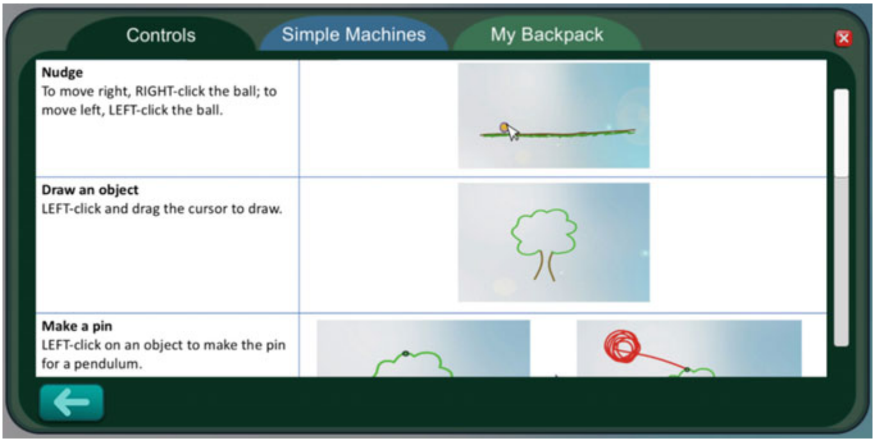


Fig. 9 "*Show me Game Tips*" menu in *PPv2*

of the four simple machine tutorials (i.e., lever, pendulum, ramp, and springboard) appear. Each image is clickable and can be enlarged. By viewing the larger images, learners can quickly see how to create the agents without going through the full tutorials again.
- *Tools*. This option only appears when the player is in a manipulation level. Here players view rules for the sliders and a short explanation about other tools available (i.e., puffers and blowers).
- *My Backpack*. In both sketching and manipulation levels, "Show me Game Tips" includes "My Backpack." A screenshot from "My Backpack" will be shown with textboxes pointing at different parts of "My Backpack" explaining the various functions.

Incentive System. To encourage student performance and use of learning supports, we added an incentive system in *PPv2*. Most of the incentive system is contained within *My Backpack* (accessed via the top left corner of the level selection area in *PPv2*). When clicked, *My Backpack* provides information about progress in the game, as well as a space to customize game play (Fig. 10). That is, two progress bars—one for sketching levels and one for manipulation levels—show how many levels the student has solved and how many remain. A money bag displays their current balance with a drop-down function that shows the amount of gold and silver coins they have collected so far. The "Physics" tab shows the estimated competency level for each targeted physics concept (based on real-time stealth assessment), and the Store tab provides options to change the background music, background image, or ball type. This customization is an additional component of the incentive system and must be purchased by students with the money they make in the game.

Each level in the game has a "par" that is based on the degree of difficulty of the level. Each level was scored on two difficulty indices, game mechanics and physics concepts. A composite score was used to create the par. For sketching levels, the par is based on the minimum number of objects used in a solution. For manipulation levels, the par is based on attempts (i.e., each time a slider adjustment is made and the "Play" button clicked). If the player's solution is at or under par, a gold coin (worth $20) is given, and otherwise, a silver coin (worth $10) is awarded to the player. In Fig. 11, you can see that the player has collected eight gold coins and two silver coins, and the amount of money is $180.

Fig. 10 My Backpack views—physics estimates (left) and store (right)

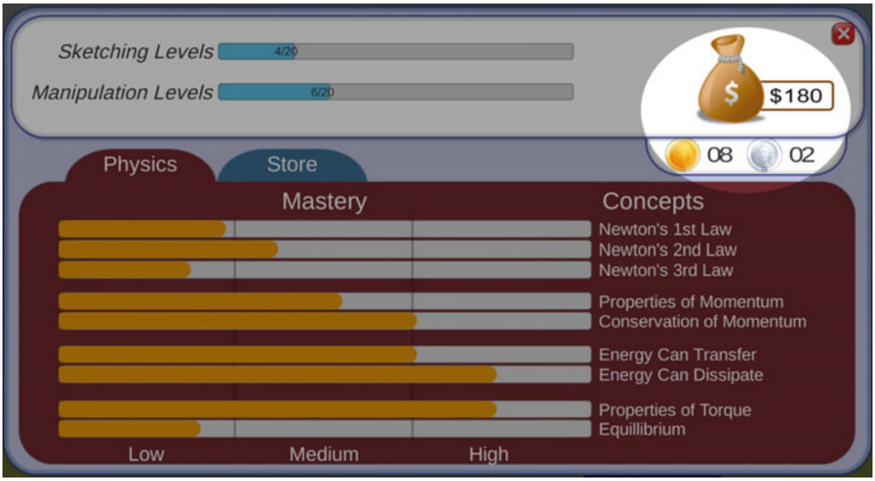


Fig. 11 Money bag and coins collected in *PPv2*

4 Testing the New Supports and Test Items—a Usability Study

The purpose of our most recent usability study was to investigate the effectiveness of the new animations when combined with gameplay, and pilot test a set of new near-transfer test items we developed as an external measure of physics understanding. For these purposes, we selected two minimally overlapping concepts in our competency model: energy can transfer (EcT) and properties of torque (PoT).

4.1 New Learning Supports—Physics

The new learning supports we included in *PPv2* for this study consist of seven new physics animations explaining the EcT and PoT concepts. The production of these supports was an outcome of our previous usability studies.

4.2 Measures

Physics Understanding Test. We created two physics test forms (Form $A = 14$ items; Form $B = 14$ items) each of which included 10 near-transfer test items (new for this study), and 4 far-transfer test items (used in prior studies). Each item included in the test targeted either EcT or PoT (see Figs. 12 and 13 for examples of a near- and far-transfer item).

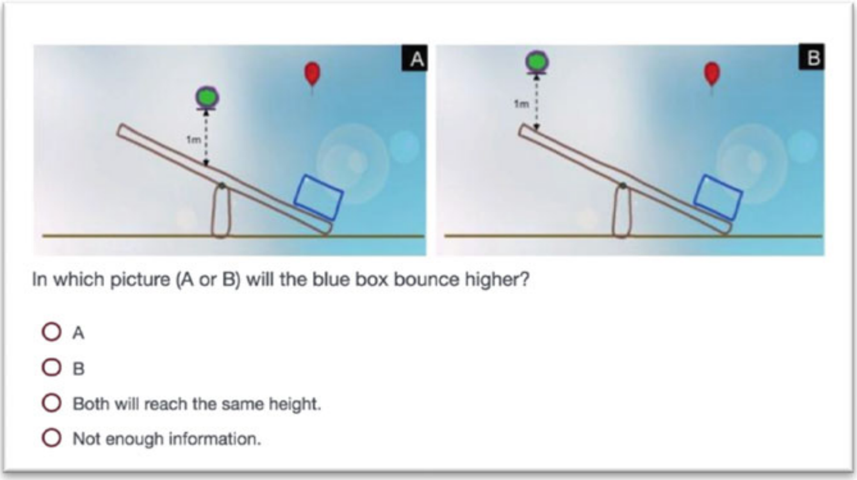


Fig. 12 An example of our PoT near-transfer test items. The answer is B

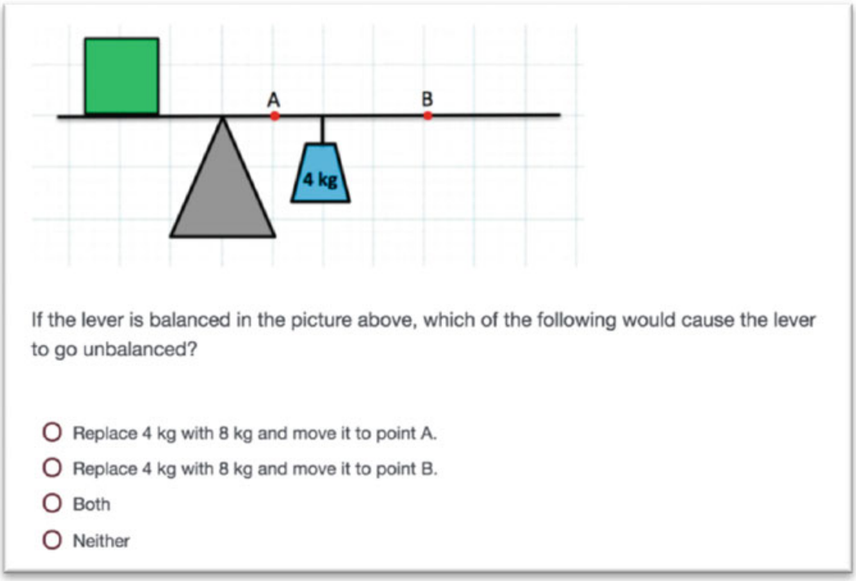


Fig. 13 An example of our PoT far-transfer test items. The answer is B

Game and Learning Supports Satisfaction Survey. To evaluate students' satisfaction of the game and our new learning supports, we used a 16-item Likert-scale questionnaire, developed in house, with two parts: (1) game satisfaction and (2) learning supports' satisfaction.

4.3 Method

Participants. Our convenience sample included 14 students (6 seventh graders, 8 eighth graders; 6 females, and 8 males) from the School of Arts and Sciences (SAS) in Florida. They were compensated with a $10 gift card upon completion of the study. All students played the same version of the game.

PP **Levels Selected**. We selected 30 sketching levels (a mixture of levels with PoT or EcT as their primary physics concept) with variable difficulty levels. We also included the new set of sketching tutorial levels. In total, students had 35 levels to complete.

Procedure. Students first completed an online demographic questionnaire followed by the pretest (Form *A*). Next, students played the game, individually, for 75 min. Student gameplay was monitored by six researchers. The researchers allowed students to access the learning supports (worked examples, physics animations, and game tools) freely during the first 20 min. For the following 55 min, students were only allowed to access the "physics supports" (i.e., our new animations), and the researchers prompted the students to access them every 8 min or after completing three game levels. At the end of the 75 min of gameplay, students completed the posttest (Form *B*) and the game and learning supports satisfaction questionnaire.

4.4 Results

Despite the limitations of this usability study (i.e., small sample size, short gameplay time, and lack of control group), we obtained some useful findings that can help us improve the game for future, larger studies. We first examined the near-transfer items and identified a few problematic items. Then, we examined the mean differences between the various subsets of the pretest and posttest. Finally, we looked at the game and learning supports satisfaction questionnaire to see how the students felt about the game in general and the learning supports in particular.

Item Analysis. Cronbach's α for the EcT near-transfer items (both pre- and posttest items) was 0.61, and the α calculated for the PoT near-transfer items (pre- and posttest items) was 0.38. We found three items with zero mean variability (either all the students got those items wrong or right) and three items showing near-zero mean variability (only 1 or 2 students got those items right). These items have been revised for future use. It is expected that when we pilot test these revised items and have a larger sample size, we will obtain a higher reliability for these items.

Physics Understanding. To assess students' physics understanding, we analyzed the pretest and posttest relative to their sections as follows: (1) near-transfer EcT tests scores, (2) near-transfer PoT test scores, (3) overall near-transfer test scores (with both EcT and PoT items combined), (4) overall far-transfer test scores, and (5) overall pretest and posttest scores with all the items included (near and far-transfer). Then we conducted several paired-sample t-tests to examine the differences between the means coming from these subsets, and several correlational analyses to examine the relationships between these subsets in the pretest and posttest. Table 1 summarizes our findings.

As shown in Table 1, students scored significantly higher on the posttest compared to the pretest ($M_{pre} = 0.57$, $M_{post} = 0.63$, $t(13) = -2.20$, $p < 0.05$, Cohen's $d = 0.60$). In addition, the near-transfer pretest significantly correlated with the near-transfer posttest ($r = 0.53$, $p < 0.05$).

Game and Learning Supports Satisfaction. To get a sense about students' overall satisfaction from the game and the learning supports, we analyzed responses to the questionnaire which followed the posttest. We divided the results into two parts: game satisfaction (Likert-scale, 1 = strongly disagree to 5 = strongly agree; see Table 2), and learning supports satisfaction (Likert-scale, 1 = strongly disagree to 5 = strongly agree; see Table 3).

As shown in Table 2, students really liked the game on average ($M = 4.24$, $SD = 0.62$). This finding is consistent with our previous findings in other research studies (e.g., [40]). Also, students agreed that the game helped them learn some physics ($M = 3.93$, $SD = 1.07$).

Table 3 shows that students found the learning supports satisfying and useful ($M = 3.99$, $SD = 0.51$) and reported the new animations helped them learn physics ($M = 3.79$, $SD = 1.19$). Moreover, males and females equally enjoyed the game and the supports.

Having a small sample size and one-group pretest–posttest design can only provide preliminary insights. The overall results from this usability study suggest we are

Table 1 Descriptive statistics, paired-sample *t*-tests, and correlations for physics measures ($n = 14$)

	Pretest		Posttest		Paired-sample *t*-test (pre and post)		Correlation (pre and post)	
Measures	*M*	*SD*	*M*	*SD*	*t (13)*	sig.	*r*	sig.
EcT	0.44	0.25	0.54	0.16	−1.71	0.11	0.51	0.06
PoT	0.76	0.16	0.76	0.22	0.00	1.00	0.20	0.49
Near transfer	0.60	0.12	0.65	0.18	−1.61	0.13	0.53	0.04*
Far transfer	0.48	0.15	0.57	0.18	−1.44	0.17	0.05	0.87
All items	0.57	0.07	0.63	0.09	−2.20	0.04*	0.22	0.44

Note The means are standardized averages
*Significant at the $p < .05$. EcT = near-transfer EcT items. PoT = near-transfer PoT items

Table 2 Likert-scale game satisfaction questionnaire ($n = 14$)

Items	*M*	*SD*
I enjoyed the game very much	4.57	0.85
I thought the game was boring (RC)	4.71	0.83
The game did not hold my attention (RC)	4.29	1.20
I thought I performed well in the game	4.00	0.56
I was pretty skilled at playing the game	3.71	0.83
I put a lot of effort into solving levels	4.43	0.76
The game helped me learn some physics	3.93	1.07
Physics is fun and interesting	4.36	1.15
I'd like to play this game again	4.21	1.19
I'd recommend this game to my friends	4.14	1.29
Game satisfaction scale	4.24	0.62

Note RC = reverse coded

Table 3 LS satisfaction questionnaire ($n = 14$)

Items	*M*	*SD*
The "level solutions" helped me solve the levels	4.14	0.86
The "physics supports" helped me learn physics	3.79	1.19
The supports were generally annoying (RC)	4.14	1.23
The supports were pretty easy to use	4.21	0.70
The supports did not help me at all (RC)	4.00	1.18
I'd rather solve levels without supports (RC)	3.64	1.50
LS satisfaction scale	3.99	0.51

Note RC = reverse coded

on the right path. However, we have revised our near-transfer items (based on item analysis results) and will conduct more pilot testing on those items before using them in larger studies. Also, we will collect more qualitative data on our new learning supports with further rounds of revisions as needed. The reflection on students' learning experiences prepares us for the next phase of the project—implementing an adaptive algorithm into the game. Next, we discuss the remaining steps needed to include adaptation using game-based learning analytics in *PPv2*.

5 Testing Game-Based Learning Analytics in *Physics Playground*

Shute, Ke, and Wang [17] listed ten steps—derived from multiples studies conducted relative to stealth assessment—to include accurate measurement and adaptation in PP:

1. Develop the full competency model (CM) of the targeted knowledge, skills, or other attributes based on full literature and expert reviews
2. Select or develop the game in which the stealth assessment will be embedded
3. Identify a full list of relevant gameplay actions/indicators/observables that serve as evidence to inform CM and its facets
4. Design and develop new tasks in the game, if necessary
5. Create a Q-matrix to link actions/indicators to relevant facets of target competencies to ensure adequate coverage (i.e., enough tasks per facet in the CM)
6. Establish the scoring rules to score indicators using classification into discrete categories (e.g., solved/unsolved, very good/good/ok/poor relative to quality of the actions). This becomes the "scoring rules" part of the evidence model (EM)
7. Establish statistical relationships between each indicator and associated levels of CM variables (EM)
8. Pilot test Bayesian networks (BNs) and modify parameters
9. Validate the stealth assessment with external measures
10. Include adaptation of levels and/or support delivery in the game.

At the time of writing this chapter, we have completed steps 1 through 8 with the new version of PP. That is, we have revised/elaborated the competency model of physics understanding, (b) created task types and associated levels that provide the evidence we need to assess students' physics understanding via stealth assessment, (c) developed and tested a variety of learning supports to help students enhance their physics knowledge during gameplay, and (d) set up an incentive system that can boost students' motivation to use the learning supports in the game. In the coming months, to complete the 10-step guideline mentioned above, we will add and test online adaptation [41] in *PP* for the selection of levels and learning supports delivery.

Level Selection. During gameplay, students provide a plethora of data (stored in a log file). The data are analyzed by the evidence identification (EI) process—in real time. The results of this analysis (e.g., scores and tallies) are then passed to the evidence accumulation (EA) process, which statistically updates the claims about relevant competencies in the student model—e.g., the student is at a medium level regarding understanding the concept of Newton's first law of motion. Using the stealth assessment results in *PP*, and based on an adaptive algorithm (see [19]), the system will pick the next level for the student. The best next level for a student is one with a fifty-fifty chance of success based on the student's prior performance in the game. In other words, the next level presented to the student will likely be in his/her ZPD [29].

Learning Supports Delivery. Currently, and in line with the game design notion of learner autonomy in game play, we allow players to access the help voluntarily. We will be testing the incentive system in an upcoming study, to see if it works as intended (i.e., fosters use of physics supports and reduces abuse of worked examples). However, we have also developed a quit-prediction model that uses gameplay data in the log file as the basis to make inferences about when a player is seriously struggling and about to quit the game [42]. The model is based on high-level intuitive features that are generalizable across levels, so it can now be used in future work to automatically trigger cognitive and affective supports to motivate students to pursue a game level until completion. To move toward game-directed learning support adaptivity, we plan to include some simple rules that accompany the quit-prediction model to determine when to deploy supports and which supports to choose.

6 Conclusion

Designing learning games, capable of assessing and improving student learning, has serious challenges. For one, integrating just-in-time learning supports that do not disrupt the fun of the game is a hurdle we are actively trying to surmount. In this chapter, we discussed the importance of including learning supports and their influence on learning and performance in educational games, described our own experiences with stealth assessment and the design and development of different learning supports in *PP*, presented the results from a recent usability study examining the effects of our new supports on learning (with promising results on our new learning supports and game satisfaction), and provided insights into the next steps of game-based learning analytics via stealth assessment. Finally, we will continue to design, develop, and test adaptivity of game levels students play in *PP* and the learning supports they receive.

The central research study in our design and evaluation of learning support components, including adaptive sequencing, is expected to yield principles that designers of other educational games can use. Again, we aim to come up with a methodology for developing game-based assessments and dynamically linking those assessments to in-game learning supports. As we formalize the design process and share it, other researchers and designers are able to utilize the methodology.

Through the use of game-based learning and stealth assessment, learning analytics can be used to both measure *and* support student learning in an engaging way. Harnessing the power of data generated by students in game play activities enables more accurate assessments of student understanding and misconceptions than one-off summative evaluations (e.g., final score). Better estimations of student struggles and achievements can lead to better individualized instruction and more motivated students, paving the way for new educational paradigms that replace the school-as-factory metaphor.

Acknowledgements We wish to express our gratitude to the funding by the US National Science Foundation (NSF #037988) and the US Department of Education (IES #039019) for generously supporting this research.

References

1. Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper and Row.
2. Shute, V. J., & Ventura, M. (2013). *Stealth assessment: Measuring and supporting learning in video games*. Cambridge, Massachusetts: The MIT Press.
3. Ravitch, D. (2000). *Left back: A century of failed school reforms*. New York: Simon & Schuster.
4. Shute, V.J. (2007). Tensions, trends, tools, and technologies: Time for an educational sea change. In: C. A. Dwyer. (Eds.), The future of assessment: Shaping teaching and learning. (pp. 139–187). Lawrence Erlbaum Associates, Taylor & Francis Group, New York, NY.
5. Kliebard, H. (1987). *The struggle for the American curriculum, 1893–1958*. New York: Routledge and Kegan Paul.
6. Bobbitt, J. F. (1912). The elimination of waste in education. *The Elementary School Teacher., 12,* 259–271.
7. Lemann, N. (1999). The IQ meritocracy: Our test-obsessed society has Binet and Terman to thank—or to blame [Electronic version]. *Times, 153,* 115–116.
8. Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A Systematic review and meta-analysis. *Review of Educational Research., 86,* 79–122.
9. Castell, S. D., Jenson, J. (2003). Serious play: Curriculum for a post-talk era. *Journal of the Canadian Association for Curriculum Studies, 1*.
10. Prensky, M. (2001). *Digital game-based learning*. New York: McGraw-Hill.
11. Habgood, M. P. J., Ainsworth, S. E., & Benford, S. (2005). Endogenous fantasy and learning in digital games. *Simulation & Gaming., 36,* 483–498.
12. Ritterfeld, U., Shen, C., Wang, H., Nocera, L., & Wong, W. L. (2009). Multimodality and interactivity: Connecting properties of serious games with educational outcomes. *CyberPsychology & Behavior, 12,* 691–697.
13. Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education, 84,* 740–756.
14. Wouters, P., & van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers & Education, 60,* 412–425.
15. Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 503–524). Charlotte, NC: Information Age Publishers.
16. Mislevy, R. J., Steinberg, L. S., Almond, R. G.(2003). Focus article: On the structure of educational assessments. *Measurement: Interdisciplinary Research & Perspective, 1*, 3–62.
17. Shute, V., Ke, F., & Wang, L. (2017). Assessment and adaptation in games. In P. Wouters & H. van Oostendorp (Eds.), *Instructional techniques to facilitate learning and motivation of serious games* (pp. 59–78). Cham: Springer.
18. DiCerbo, K. E., & Behrens, J. T. (2012). Implications of the digital ocean on current and future assessment. In R. Lissitz & H. J. Jiao (Eds.), *Computers and their impact on state assessment: Recent history and predictions for the future* (pp. 273–306). Charlotte, NC: Information Age.
19. Shute, V. J., Hansen, E. G., & Almond, R. G. (2008). You can't fatten a hog by weighing it—Or can you? Evaluating an assessment for learning system called ACED. *International Journal of Artificial Intelligence in Education, 18,* 289–316.

20. Sun, C., Shute, V. J., Stewart, A., Yonehiro, J., Duran, N., D'Mello, S. Toward a generalized competency model of collaborative problem solving. Computers & Education. (in press).
21. Barzilai, S., & Blau, I. (2014). Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education, 70,* 65–79.
22. Lang, J., O'Neil, H. (2008). *The effect of presenting just-in-time worked examples for problem solving in a computer game*. In Presented at the American Educational Research Association Conference, New York.
23. Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences, 13,* 423–451.
24. Leutner, D. (1993). Guided discovery learning with computer-based simulation games: Effects of adaptive and non-adaptive instructional support. *Learning and Instruction, 3,* 113–132.
25. O'Rourke, E., Ballweber, C., Popovií, Z.(2014). Hint systems may negatively impact performance in educational games. In *Presented at the Proceedings of the first ACM conference on Learning @ scale conference*. Atlanta, GA, USA March 4.
26. Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction, 15,* 313–331.
27. Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology, 94,* 598–610.
28. Cowley, B., Charles, D., & Black, M. (2008). Toward an understanding of flow in video games. *Computers in Entertainment, 6,* 1–27.
29. Vygotsky, L. S. (1978). *Mind in society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.
30. Shute, V. J., & Zapata-Rivera, D. (2012). Adaptive educational systems. In P. Durlach (Ed.), *Adaptive technologies for training and education* (pp. 7–27). New York, NY: Cambridge University Press.
31. Kickmeier-Rust, M. D., & Albert, D. (2010). Micro-adaptivity: Protecting immersion in didactically adaptive digital educational games: Micro-adaptivity in digital educational games. *Journal of Computer Assisted learning, 26,* 95–105.
32. Shute, V. J., Graf, E. A., Hansen, E. G.(2005). Designing adaptive, diagnostic math assessments for individuals with and without visual disabilities. In L. PytlikZillig, R. Bruning, M. Bodvarsson (Eds.), *Technology-based education: Bringing researchers and practitioners together*. pp. 169–202. Information Age Publishing, Greenwich, CT.
33. Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research., 78,* 153–189.
34. Peirce, N., Conlan, O., Wade, V.(2008). Adaptive educational games: Providing non-invasive personalised learning experiences. In: *2008 Second IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning*, pp. 28–35. IEEE, Banff, AB, Canada.
35. Hatano, G. (1982). Cognitive consequences of practice in culture specific procedural skills. *Quarterly Newsletter of the Laboratory of Comparative Human Cognition., 4,* 15–18.
36. Verschaffel, L., Luwel, K., Torbeyns, J., & Van Dooren, W. (2009). Conceptualizing, investigating, and enhancing adaptive expertise in elementary mathematics education. *European Journal of Psychology of Education, 24,* 335.
37. Shute, V., Ventura, M., Kim, Y. J., & Wang, L. (2014). Video games and learning. In W. G. Tierney, Z. Corwin, T. Fullerton, & G. Ragusa (Eds.), *Postsecondary play: The role of games and social media in higher education* (pp. 217–235). Baltimore, MD: John Hopkins University Press.
38. Shute, V. J., D'Mello, S., Baker, R., Cho, K., Bosch, N., Ocumpaugh, J., et al. (2015). Modeling how incoming knowledge, persistence, affective states, and in-game progress influence student learning from an educational game. *Computers & Education, 86,* 224–235.
39. Taylor, W. L. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism Bulletin, 30,* 415–433.
40. Shute, V. J., Ventura, M., & Kim, Y. J. (2013). Assessment and learning of qualitative physics in Newton's playground. *The Journal of Educational Research., 106,* 423–430.

41. Lopes, R., & Bidarra, R. (2011). Adaptivity challenges in games and simulations: A survey. *IEEE Transactions on Computational Intelligence and AI in Games., 3,* 85–99.
42. Karumbaiah, S., Baker, R. S., Shute, V. (2018). *Predicting quitting in students playing a learning game*. In 11th International Conference on Educational Data Mining, pp. 1–10. Buffalo, NY.

Chapter 5
Learning Analytics on the Gamified Assessment of Computational Thinking

Juan Montaño, Cristian Mondragón, Hendrys Tobar-Muñoz and Laura Orozco

Abstract Learning Analytics (LA) can be applied to many aspects of learning. In particular, LA can be applied to ease the process of assessment as it can help teachers to understand the state and evolution of their students in order to intervene on their learning routes. In this chapter, we show the use of LA in the assessment of Computational Thinking (CT), understood as the set of thought processes involved in the use of computational agents (such as computers). This assessment process tends to be of high complexity for teachers, as it requires a high amount of trained human resources per student; and it may be cumbersome for students, due to the uncertainty that might be involved into the assessment objectives and the increase in anxiety levels presented during the assessment process. We created a gamified platform (called Hera) where students can participate in a gamified activity as part of the assessment established by the teacher. After this, the teacher can gain insight into their students by analyzing the resulting Learning Traces. The chapter shows the framework used for developing the assessment strategies used within the platform, an overview of the platform and the results of an experiment conducted with it on a real CT learning classroom using the popular programming tool "Scratch".

1 Introduction

Computational Thinking is the thought processes involved in formulating problems and their solutions so that the solutions are represented in a form that can be effectively

J. Montaño · C. Mondragón · H. Tobar-Muñoz (✉) · L. Orozco
Department of Systems, Universidad Del Cauca, Popayán, Cauca, Colombia
e-mail: fabian@unicauca.edu.co

J. Montaño
e-mail: juanmontano@unicauca.edu.co

C. Mondragón
e-mail: mrcristian@unicauca.edu.co

L. Orozco
e-mail: lmorozco@unicauca.edu.co

A. Tlili and M. Chang (eds.), *Data Analytics Approaches in Educational Games and Gamification Systems*, Smart Computing and Intelligence,
https://doi.org/10.1007/978-981-32-9335-9_5

carried out by an information-processing agent [1] such as an electronic computer. This type of thinking is very relevant for people to learn on the current age of information as it helps people to be more efficient and effective in their daily contexts. Thus, CT should be taught and assessed in formal learning environments [2]. The concept of Computational Thinking (CT) has been a very trending topic in Computer Science. By doing so, a considerable amount of research was done about the development of Computational Thinking in young children, finding ways to encourage the development of CT inside the classroom [3]. However, as the scope of the term keeps broadening, the assessment of CT keeps turning into a more complex matter, and thus creating the need to define a proper assessment framework. Moreover, because of the complexity in the teaching and learning process of CT, some issues as the high amount of highly trained teachers and the required time to make a proper assessment grows exponentially with the number of students involved.

The use of Learning Analytics (LA) has proven to be a key component in the improvement of the learning environment itself [4]. Therefore, its use into the assessment process of the CT development could ease some of the difficulties found in the assessment process. Also, Gamification techniques used in learning environments have shown to increase engagement for learning and assessment [5]. Hence, a gamified platform for the assessment of CT which also helps teachers with LA should be of utility for teachers as it should ease the task of assessment while providing the teachers with valuable information about the state and evolution of their students. This study could prove helpful as the use of gamification, with LA techniques, in order to assist the assessment process of CT is an area with a lack of research, and thus this approach could benefit the development of CT courses in a two-dimension way, addressing both: the students' and teachers' perspectives. This study aims to develop a gamified web platform (called HERA) with the purpose of easing the CT assessment processes in CT courses—with the use of some LA techniques—while also using gamification techniques in order to address some of the difficulties found in the assessment process itself such as a lack of motivation from the students [6].

In this paper, we describe and discuss the use of assessment analytics in order to improve the assessment processes of a CT course. In our observations, we discover that there was an impact on the learning processes while using Learning Analytics with a gamified platform, proving to be a useful tool in enhancing a formal learning environment.

This chapter is structured as follows: Related works are described in Sect. 2. Our methodology for the design of our online platform and the implementation of the Learning Analytics is described in Sect. 3. An overview of the gamified platform is mentioned in Sect. 4. The Analysis of the Learning Traces is reported on Sect. 5. And finally, in Sects. 6 and 7 we present the discussion and conclude the chapter.

2 Related Works

In order to make an insightful analysis about the use of learning traces, and how they can be used to improve the assessment process of Computational Thinking, it is necessary to look at previous research done in the matter. Also, in this chapter, a brief review of previous gamification approaches aimed for assessment is shown.

2.1 The Assessment of Computational Thinking

The concept of Computational Thinking can be seen as a very broad term, one of its most cited definitions was provided by Wing [1] as: "*the thought processes involved in formulating problems and their solutions so that the solutions are represented in a form that can be effectively carried out by an information-processing agent*".

The assessment of Computational Thinking, especially in children, has been a very active research field. One of the most influential works is provided by Resnik and Brennan [7]. In which they describe how the use of interactive programming tools (such as Scratch[1]) could help kids in the process of learning key concepts of Computational Thinking, while learning computational practices and developing and sharing their own computational perspectives.

The complexity involved in the development of Computational Thinking, and the push for Computer Science (CS) related courses into formal environments, reveals the need for a standardization of the curriculums and assessment processes [5, 8]. The research done by Werner [9] shows the need for a curriculum definition as a big challenge for CT-related courses to gain widespread use. Additionally, research suggests that the use of visual programming languages (such as Scratch) could greatly ease the early adoption of CT concepts and practices from the infancy developing stages [10–14]. Also, many works into standardizing the CT assessment processes suggest that approaching CT assessment into three separate dimensions (concepts, practices, and perspectives) would be beneficial, as each one addresses in a unique way the mechanisms involved into CT [2].

Finally, this suggests that the research on the field is only in an early definition stage, and that multiple "empirical" approaches are prevalent on the field [9, 15, 16]. This might suggest that a LA approach to the assessment process, based on a heavily used and recognized CT assessment framework, could add some insight into the field.

[1]Scratch is a visual programming language that enables the creation and sharing of multimedia content [13].

2.2 Gamification

The concept of Gamification and its applications in a learning environment has been also a very widespread subject of study. Meanwhile, some research has been done studying the use of particular Gamification techniques (like badges, leader-boards, etc.) [5, 17], there has been some research related to how the use of a more involved gamified environment, through the use of storytelling and role-playing, could benefit the students' engagement levels [5]. The work proposed by Nicholson [18] suggests that the implementation of some key components into the gamified environment, with a user-center design philosophy in mind, could greatly improve the engagement levels of its students, through the generation of user meaningful content.

Also, the process to implement a gamified platform it's a very important step into the success of the platform itself. Therefore, the work of Morschheuser [19] which seeks to study and synthesize the best practices for the implementation process, which includes a deep analysis of the target user base, a deep involvement of the targeted users into the design and iteration process, and a continuous review of the platform, resulting in a method for gamification design.

The studies proposed by Liu and Chu [20] allow observing a great impact between the use of gamification—if it is highly related to educational contents—, motivation and levels of interest, which has a high correlation with a better academic performance in general. Because of this, it is worthy to mention the work proposed by Nicholson [18]. This framework is presented as a set of basic principles that allow the creation of a gamified environment, facilitating the creation of value and meaning for users regarding the educational contents which are included in a specific environment. These principles are given as follows:

- **Theory of organismic integration**: It is responsible for exploring how, and in what quantity, different types of external motivations can be integrated with the underlying activity (activity to be gamified) and internalized in the user's consciousness.
- **Situational relevance**: Involves, in some way, the user in the process of selection of goals in order to facilitate correlation between those goals and goals that the user has previously internalized.
- **Situational motivational affordability**: Suggests that a user will have higher levels of intrinsic motivation if there is a relationship between the subject of study and the context of the student.
- **Universal design for learning**: Defines guidelines for the creation and design of content. The former under the premise that students should be responsible for demonstrating their competence in learning processes.
- **Content generated by the player**: Allows the content developed by the player to extend the life of a game and allow designers to see how creative users can be with the toolkits provided.

2.3 Learning Analytics

Other approaches focus on the uses of machine learning to analyze code patterns from student-submitted work and predict their future performance and finally their final exam grade; these works suggest that recommendation systems, based on feedback loops, could improve the students learning processes [21–23].

Finally, it is noteworthy that the implementation of LA techniques in the assessment process could not only benefit the students but also the teachers, heavily reducing time and human resources needed for properly manage a CT course [24]. The use of LA could result in a helpful tool in order to analyze and manage the assessment processes involved in our study.

A summary of the previously discussed papers is shown in Table 1

3 Method

We developed a digital gamified platform to assist the assessment process of Computational Thinking in young children. In order to do that, we carried out two main processes: the design of the gamification system and gathering the learning traces.

Table 1 Table of related works

Paper	Issues	Techniques used	Outcomes
[2]	CT assessment	Multiple assessment systems focused on cognitive and not cognitive aspects of CT	Mixed, mostly positive
[9]	CT assessment	Design and implementation of a performance assessment tool	Mixed, but mostly positive
[11, 24]	CT assessment with LA	Automatic code valuation through software	Positive
[8]	CT assessment	Use of digital ink for CT cognitive assessment	Mixed
[21, 22]	LA through code data mining	Data mining techniques to detect positive and negative code patterns	Mixed, but mostly positive
[23]	LA code logging pre-processing	Low-level processing of students' code log data to find learning patterns	Mixed, but mostly positive

3.1 Implementing Meaningful Gamification

In order to design a successful gamification platform, it is necessary to develop an environment in which a meaningful interest, between the user and the platform, could grow. Based on the method provided by Morschheuser [19] the development of a gamification platform was conducted following the next seven phases:

- **Preparation**: To start the process of designing the gamified assessment platform, we observed the learning environment of young children (with ages varying between 12 and 17). The students were taking a Computational Thinking course. All children had, at least, a basic knowledge of Scratch and programming in general. In our observation period, we noticed a lack of engagement in some of the students towards the course's content, and also there were some critical issues in the assessment process, because the ratio of students versus teachers were so high, and thus making the process tedious and inefficient. Those were the critical issues to address with our platform.
- **Analysis**: In this part of the process, we needed to define the context where the gamification should be developed. And thus, we defined a user target base consisting of young children between 12 and 17 years old, identified the lack of engagement between the user target base and the course content as their main need, and thus proposed the course's engagement levels as our success metric for the gamified platform.
- **Ideation**: We gathered and refined several ideas in brainstorming sessions. Finally, we settled down on the creation of several gamified strategies to simplify the assessment results to the students, while gathering the learning traces needed. A gamified strategy is a method with the purpose of gamifying an assessment process. The strategies must focus on one or two of the CT concepts or practices, and their goal is to show students their successes and mistakes in a fun and meaningful way.
- **Design**: Once the strategies were created, we conducted a rapid prototyping and iteration process, in which we validated the strategies.
- **Implementation**: Then, we developed a Web App (Called "Hera") which helps teachers on the automation of the assessment process, while also gathering the desired traces in a database (for further analysis), and simplifying the assessment results for the students, through the gamified content.
- **Evaluation**: Once the app was developed and tested in class, we assessed the benefits of the app's usage in the classroom through the criteria mentioned earlier. We tested the anxiety levels of our students, with the use of the works proposed by Liebert and Morris [25].

3.2 Gathering the Learning Traces

To collect the Learning Traces (LT) data for our study, we develop our process based on the learning analytics cycle proposed by Clow [26] (Fig. 1). In order to create a

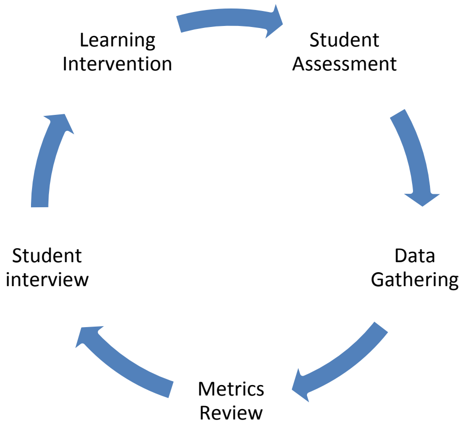


Fig. 1 Our learning analytics process

cyclical process, we first gathered some basic profile info from our students, coupled with some questions related to the familiarity the students had with their computer usage. With that initial info, we created a starting assessment for the students in order to assess the level of previous CT of the students.

Our cycle starts with the student assessment, in which the students receive an assessment previously design and solved by the teacher; when finished, the assessment results are gathered and processed in a data base, the captured data contains code-related info such as dead scripts, duplicate scripts, total blocks, total scripts; and even more specific data such as the criteria described in the CT framework suggested by Resnick [7]. The platform gathers data on each criterion in the following way:

- Abstraction and pattern recognition: when all the needed programming blocks are used and when the system detects user-created blocks and the use of clones.
- Data collection: usage of input blocks, variables, and sensors.
- Parallelism and synchronization: usage of two threads starting with a green flag block or more than one type of event.
- Flow control: usage of simple blocks (if, for, while), use of complex blocks and use of nested blocks.
- Information analysis: usage of the basic logical operators, complex logical operators, and nested logical operators
- Decomposition of problems: usage of two threads per sprite.
- Algorithmic thinking: usage of sequences.

After this, an analysis on the data gathered was conducted by the teacher, followed by a student interview in which assessment feedback is given, and finally the teacher creates individual interventions on the students learning route. These parts of our cycle are described in more depth on latter sections in this chapter.

4 Gamified Platform

Hera[2] is a web app developed to ease the assessment process in Computational Thinking courses through the processing and analysis of Scratch code. In the app, a teacher can create a course and include several challenges (which are the equivalent to the course's assessments) that must be completed by the course's students. Every challenge must have a Scratch scenario for the students to complete, paired with the challenge's solutions provided by the teacher—which would be used as a reference point for comparison with the students' submissions. And also, every challenge must define which of the CT concepts and practices should be assessed.

Then, the student must log in into the app, check the course, and complete the challenges. The app's interfaces are shown in Figs. 2 and 3. Before the students start their challenges, the app shows them a mission to complete with the submission of the code. Once the student has finished a Scratch scenario of a course's challenge, they must submit the Scratch's project id into the app. Then the app would gather all the necessary traces from the code and store into a database. The app makes an analysis of the submitted code and shows a graphic representation, based on the assessment criteria defined for the challenge, which represents the automated part of the assessment process.

The graphic representation of the student's assessment results is our main mechanism for gamifiyng the platform, because it let us show the students a 'game-like'

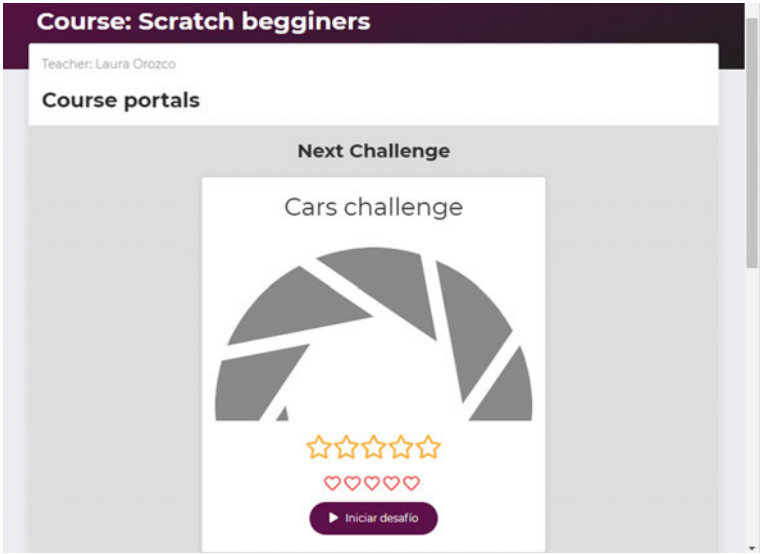


Fig. 2 Course overview main page

[2]http://heratest.azurewebsites.net.

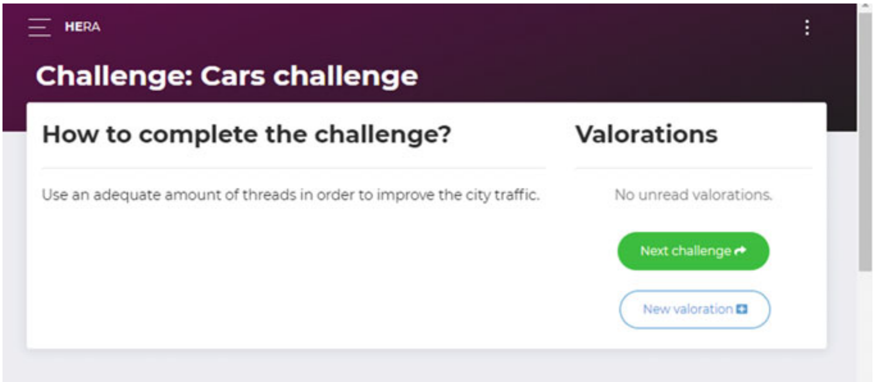


Fig. 3 Challenge overview main page

representation of the benefits and consequences of their use of the CT concepts and practices, this mechanism is shown in Fig. 4.

For example, a graphic representation for the assessment of the correct use of threads and parallel programming would be represented in a series of trucks which must deliver some packages from left to right (Fig. 5). Then, only the students which uses a correct number of threads considering the challenge's available resources (in this case the number of path lanes) would optimize the number of trucks on the road while preventing traffic. This is similar to how programmers use threads to optimize the use of CPU cores available while preventing throttling. After all is set and done, the students must discuss their conclusions with their peers and the teacher, which must enable them to reinforce the concept and let the students develop their own CT perspectives by thinking of real-world scenarios in which they could use those concepts and practices.

Later on, the teacher can make a manual review of the students' submitted code and post their approval. By doing this, the app gives a student some prizes such as badges or achievements based on the manual assessment results. The student can also view and compare the progress made by their pairs. Also, the teacher can access

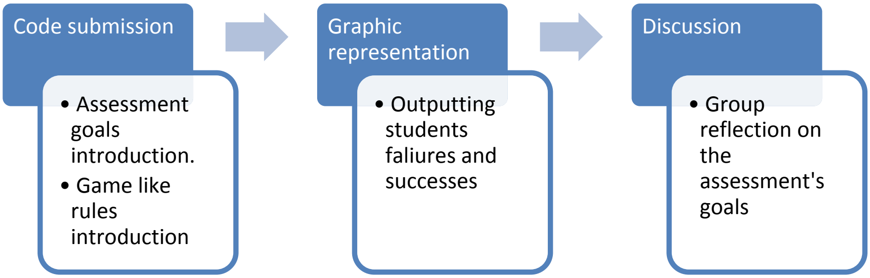


Fig. 4 Gamified assessment process

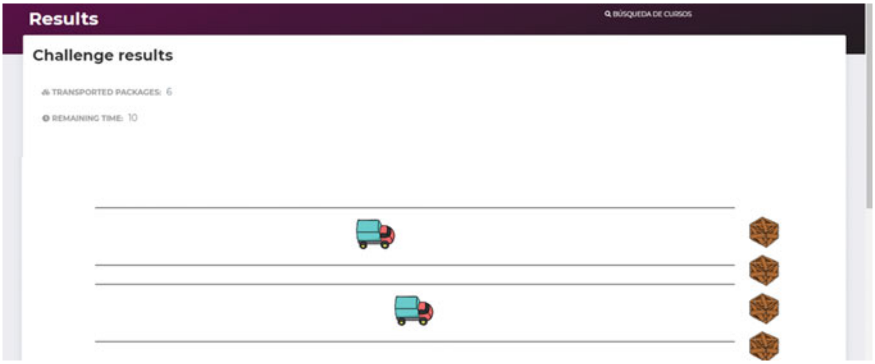


Fig. 5 Example of an assessment graphic representation

to an overview of all of the course's traces and input the suggested challenges to be completed by the students, and thus, changing the learning environment to better fit the student's progress into the course.

5 Learning Traces Analysis

Over the course of our experiment, we used HERA to assist the assessment process of three CT courses. The assessment processes were carried out on a southwestern Colombian high school. The first course had a total of 12 students, with ages between 12 and 17. The second had a total of 22 students, with ages between 12 and 17, and the third one had a total of 24 students, with ages between 12 and 17. It is noteworthy that all of the students had a basic knowledge of programming with Scratch and thus, every course had the same challenge curriculum. The criteria used in the challenge assessment were our Learning Traces. Those were grouped into seven main categories described as follows:

- **Abstraction and pattern recognition**: Which focuses on not having unused code, the use of functions in the code, and the use of clones of blocks of code (a specific functionality of the Scratch environment).
- **Flow control**: Assessment of the correct use of every control instruction (such as if and for statements), and also the adequate use in nesting those statements.
- **Input control**: Assessment of the adequate use of statements designed to capture user input into the code, the naming of variables, and the use of non-user-defined variables.
- **Data analysis**: Assessment of the treatment and transformation of the data through the use of data transformation blocks or statements, and also their adequate nesting if necessary.
- **Parallelism and threading**: Assessment of the adequate use of threading and multi-tasking enabling blocks.

- **Problem-solving**: Assessment of the student's ability to decompose a problem into multiple smaller ones in order to address them more easily.
- **Algorithmic thinking**: Assessment of the student's ability to develop sequences of tasks that would be translated into blocks of code, in order to solve a problem.

Every time a student submits their code in order to finish a challenge, the app analyzes the submitted code and gathers the number of occurrences found of each of the criteria described above and store the info into the database.

The analysis of the gathered learning traces consisted of the automatic generation of the following statistics:

- The average, median, and mode of all the CT criteria occurrences of a student per challenge.
- The average, median, and mode of all the CT criteria occurrences of the course per challenge.
- The average of every CT criterion occurrence among the students.
- The average of all the CT criteria occurrences of the course.

Once the course session is done, the app gathers all the analyzed data, including the student's profile info, and makes a comparison between the student performance against the teacher's solution for the challenge, and also their peers solutions; letting the course's teacher review the obtained metrics of all of the submitted code and plan an intervention on the learning route. Then, the teacher decides on adding a new set of challenges into the course, for each student, according to their learning needs.

For our experiment, a subset of the metrics obtained in one of the courses is portrayed in Fig. 6. A detailed review of those metrics allows teachers to observe a

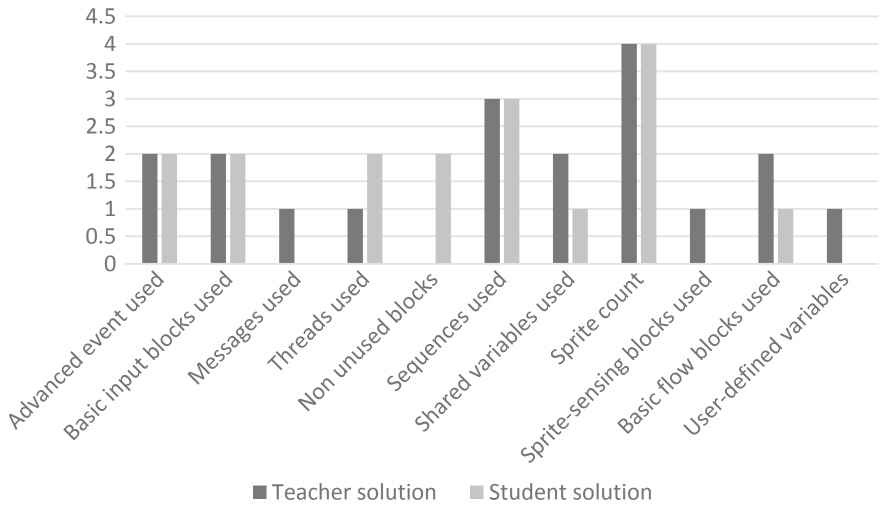


Fig. 6 Project comparisons between student and teacher solutions

course's major trends, thus, making a great reference in order to be compared with any student's individual results.

An overview of all of the analyzed traces are as follows:

- Advanced event use.
- Nested logical operator uses.
- Data input blocks use.
- Basic logical operators use.
- Clone use.
- Events use.
- List use.
- Complex logical operators use.
- Correct message use.
- Multiple threads use.
- Non-unused blocks.
- Sequences use.
- Shared variable use.
- Sprite sensing use.
- Use two green flag blocks.
- Medium complexity blocks use.
- Nested flux control blocks use.
- Simple complexity blocks use.
- User defined blocks.
- Variable declaration and usage.
- Non-user-defined variables usage.

As an example of our intervention process, the teacher observed two main trends in a student's code submissions, which were: having unused blocks and adequate use of flow control blocks. Then the teacher proceeded to interview the student about the thought process involved during his previous assessments, which helped the teacher found out that the student did not understand the consequences of leaving unused code in his developments. Once the teacher had a clear opinion on the difficulties to be addressed, an adequate intervention, by setting the student's next challenge, which focused on a bigger code project so chunks of unused code would make significantly harder to debug for errors and introduce new changes. Finally, the teacher approaches the student with valuable feedback in order for him to finish the challenge successfully, explaining to him the advantages to have a clean and well-structured code. This process is done with every student in the course, once per session.

6 Discussion

Observing the courses of our experiment, we noticed that the use of an online gamified platform had an impact on the students' behavior. Based on student feedback, we conclude that there was an eagerness by the students to use the platform, although it could have been by the novelty of the platform itself, and thus a study with a prolonged time span is suggested for further research.

Additionally, we observed competition between the students while comparing their assessment results into the app, which cause them to improve their results in order to "beat" their peers. Therefore, a competitive multiplayer aspect could be integrated directly into the app, letting the students easily compare their assessment results and promote healthy competition in the classroom.

Teachers also had a significant improvement while using the application. Based on their feedback, they note that a significant amount of time is reduced in the process of preparing their class content and assessments. Also, the app helps the teachers know what the students are doing at any time during their class sessions.

It is noteworthy that, for this experiment, there was a small amount of data to be analyzed. Thus, the interventions made into the learning environments, with the generated metrics, were made manually by the teacher. However, the data being bigger a deeper and automatic intervention should be made in order to ease the work of the teacher. The early detection of learning difficulties, or the use of bad coding practices, was one of the benefits of the learning traces analysis provided in the app.

As mentioned previously [9] the lack of a defined CT curriculum is one of the main challenges in order to CT courses to gain widespread use. However, assessing CT related courses by reviewing student-submitted code has proven to be a great tool, because the insight provided about the students' thought processes involved. Additionally, the works provided by Grover [2] suggest that a standard question quiz is not helpful in order to assess CT. Additionally, it is very important to pair the automatic assessment process with external interventions [11]—usually provided by the teachers.

Because of the way the app let teachers focus on students with learning difficulties, they can quickly engage with them and address the issues directly, reinforcing the concepts that were misunderstood or applied incorrectly. Therefore, the use of an online platform should not be used as a replacement for the teacher, but as a tool to improve the learning environments.

7 Conclusions

In this chapter, we have discussed the use of learning analytics and gamification on the assessment process of Computational Thinking (CT). Over the course of our experiment we implemented a web app, called HERA, made to ease the assessment process by automating the code review, gathering learning traces based on the student implementations of the CT concepts and practices described earlier, and processing those traces in order to make metrics that allow teachers to make insightful interventions into their learning environments.

The usage of the app consisted of the definition and submission of CT course assessments, called challenges. Those would be completed by the students and then submitted into the app. The learning traces gathered was done on every student's challenge submission, where the app would analyze the submitted code in order to find the number of occurrences of the CT components and practices. Once the info was gathered, the app would display the metrics needed in order to help the teacher make an adequate intervention, by selecting the student's next challenge, into their learning environment.

The app allowed the Teacher to observe the students' evolution in an easy way while allowing students to be assessed in a fun way by the means of gamification. As the app gathers relevant data, it helps teachers, in a semi-automated way, to be insightful into their students' performance thus, allowing meaningful interventions on the learning route of each student.

In conclusion, not only we found relevant the use of learning analytics into the assessment process [3], but also the use of an automated platform which could benefit the learning processes into formal learning environments.

However, there were some limitations in the implementation and use of our experiment, mainly in the amount of analyzed data. The latter suggests that a similar study with larger amounts of data could result in a system that could automatize the intervention process.

Overall, the use of Learning Analytics had an impact, in both students and teachers. Based on the results it seems like a viable tool for the use in formal environments. However, it is noteworthy to mention that the platform should be used as a tool and not as a replacement because it enhances the teacher's ability to reach their student and it is not intended to substitute a teacher's capabilities.

Future work includes the assessment of the application Hera itself, to measure subjectively or objectively, how students get motivated and what implications the app has on the efficiency for teachers. Also, the app is intended to be used as a relief on the nervousness often experimented by the assessed students, hence, observations on these matters are still pending.

In any case, the application is currently undergoing development and it can be accessible through the web. We request interested readers to contact the authors in order to get the URL and access to the platform (which is in the Spanish language in its current version).

References

1. Wing, J. (2008). Computational thinking and thinking about computing. In *IPDPS Miami 2008—22nd IEEE International Symposium on Parallel and Distributed Processing. Symposium, Program CD-ROM,* July, 2008 (pp. 3717–3725).
2. Grover, S., & Pea, R. (2013). Computational thinking in K − 12 : A review of the state of the field, *42*(1), 38–43.
3. Selby, C., Dorling, M., & Woollard, J. (2015). Evidence of assessing computational thinking. In *IFIP TC3 Working Conference A New Culture of Learning: Computing and Next Generations* (pp. 232–242).
4. Ellis, C. (2013). Broadening the scope and increasing the usefulness of learning analytics: The case for assessment analytics. *British Journal of Educational Technology, 44*(4), 662–664.

5. Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work?—A literature review of empirical studies on gamification. In *Proceedings of the Annual Hawaii International Conference on System Sciences*.
6. Ochoa, S., Guerrero, L., Pino, J., Collazos, C., & Fuller, D. (2003). Improving learning by collaborative testing. *Student-Centered Learning Journal*.
7. Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. In *Annual American Education Research Association Meeting,* Vancouver, BC, Canada (pp. 1–25).
8. Ambrósio, A. P., Xavier, C., & Georges, F. (2014). Digital ink for cognitive assessment of computational thinking.
9. Werner, L., Denner, J., & Campe, S. (2012). The fairy performance assessment : Measuring computational thinking in middle school. In *Proceedings of the 43rd ACM Technical Symposium on Computer Science Education—SIGCSE '12* (pp. 215–220).
10. Weese, J. L. (2016). Mixed methods for the assessment and incorporation of computational thinking in K-12 and higher education (pp. 279–280).
11. Moreno-león, J., Rey, U., Carlos, J., & Román-gonzález, M. (2016). *Comparing computational thinking development assessment scores with software complexity metrics* (pp. 1040–1045).
12. Seiter, L., & Foreman, B. (2013). *Modeling the learning progressions of computational thinking of primary grade students* (pp. 59–66).
13. Maloney, J., Resnick, M., Rusk, N., Silverman, B., & Eastmond, E. (2010). The scratch programming language and environment. *ACM Transactions on Computing Education, 10*(4), 1–15.
14. Zhong, B., Wang, Q., Chen, J., & Li, Y. (2016). An exploration of integrated assessment for computational thinking (122).
15. Grover, S., & Pea, R. (2013). Computational thinking in K–12 a review of the state of the field. *Educational Research*, *42*(1), 38–43.
16. Grover, S. (2017). Assessing algorithmic and computational thinking in K-12 : Lessons from a middle school classroom, 269–288.
17. Mora, A., Riera, D., Gonzalez, C., & Arnedo-Moreno, J. (2015). A literature review of gamification design frameworks. In *VS-Games 2015—7th International Conference on Games and Virtual Worlds Serious Application*.
18. Nicholson, S. (2012). A user-centered theoretical framework for meaningful gamification. *Games + Learning Social* 1–7.
19. Morschheuser, B., Werder, K., Hamari, J., & Abe, J. (2017). How to gamify? A method for designing gamification. In *Proceedings of the 50th Annual Hawaii International Conference on System Sciences (HICSS),* Hawaii, USA, January 4–7, 2017 (pp. 1–10).
20. Liu, T. Y., & Chu, Y. L. (2010). Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. *Computers & Education, 55*(2), 630–643.
21. Berland, M., Baker, R. S., & Blikstein, P. (2014). Educational data mining and learning analytics: Applications to constructionist research, 205–220.
22. Blikstein, P., Worsley, M., Piech, C., Sahami, M., Cooper, S., & Koller, D. (2014). Journal of the learning programming pluralism: Using learning analytics to detect patterns in the learning of computer programming, 37–41.
23. Mcdowell, C. (2013). A first step in learning analytics: Pre-processing low-level alice logging data of middle school students, *5*(2), 11–37.
24. Moreno-león, J., & Carlos, R. J. (2017). On the automatic assessment of computational thinking skills: A comparison with human experts, 2788–2795.
25. Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports, 20*(3), 975–978.
26. Clow, D. (2012). The learning analytics cycle. In *Proceedings 2nd International Conference Learning Analytics and Knowledge—LAK '12*, p. 134.

Part III
Academic Analytics and Learning Assessment in Educational Games and Gamification Systems

Chapter 6
iMoodle: An Intelligent Gamified Moodle to Predict "at-risk" Students Using Learning Analytics Approaches

Mouna Denden, Ahmed Tlili, Fathi Essalmi, Mohamed Jemni, Maiga Chang, Kinshuk and Ronghuai Huang

Abstract Online learning is gaining increasing attention by researchers and educators since it makes students learn without being limited in time or space like traditional classrooms. Particularly, several researchers have also focused on gamifying the provided online courses to motivate and engage students. However, this type of learning still faces several challenges, including the difficulties for teachers to control the learning process and keep track of their students' learning progress. Therefore, this study presents an ongoing project which is a gamified intelligent Moodle (iMoodle) that uses learning analytics to provide dashboard for teachers to control the learning process. It also aims to increase the students' success rate with an early warning system for predicting at-risk students, as well as providing real-time interventions of supportive learning content as notifications. The beta version of iMoodle was tested for technical reliability in a public Tunisian university for three months and few bugs were reported by the teacher and had been fixed. The post-fact technique was also used to evaluate the accuracy of predicting at-risk students. The obtained result highlighted that iMoodle has a high accuracy rate which is almost 90%.

M. Denden (✉) · F. Essalmi · M. Jemni
Research Laboratory of Technologies of Information and Communication & Electrical Engineering (LaTICE), Tunis Higher School of Engineering (ENSIT), University of TUNIS, Tunis, Tunisia
e-mail: mouna.denden91@gmail.com

F. Essalmi
e-mail: fathi.essalmi@isg.rnu.tn

M. Jemni
e-mail: mohamed.jemni@fst.rnu.tn

A. Tlili · R. Huang
Smart Learning Institute of Beijing Normal University, Beijing, China
e-mail: ahmed.tlili23@yahoo.com

M. Chang
School of Computing and Information Systems, Athabasca University, Athabasca, Canada
e-mail: maiga.chang@gmail.com

Kinshuk
University of North Texas, 3940 N. Elm Street, G 150, Denton, TX 76207, USA
e-mail: kinshuk@ieee.org

A. Tlili and M. Chang (eds.), *Data Analytics Approaches in Educational Games and Gamification Systems*, Smart Computing and Intelligence,
https://doi.org/10.1007/978-981-32-9335-9_6

1 Introduction

Distance educational systems have gained increasing use within institutions in the twenty-first century since they offer e-learning options to students and improve the quality of traditional courses in classrooms. These e-learning systems, such as Modular Object-Oriented Dynamic Learning Environment (Moodle), provide students different types of activities, such as preparation of assignments and quizzes, and engagement in discussions using chats and forums. Moodle is one of the most well-known free and open-source e-learning platforms which allows the development of interactive and simple online courses and experiences [1].

However, the distributed nature of distance learning has raised new challenges. For instance, unlike classrooms, it becomes much harder for teachers in distance learning to supervise, control and adjusts the learning process [2]. In massive open online courses, where thousands of students are learning, it is very difficult for a teacher to consider individual capabilities and preferences. In addition, the assessment of course outcomes in Learning Management Systems (LMSs) is a challenging and demanding task for both accreditation and faculty [1]. Anohina [3] stated that it is necessary to provide an intelligent system with adaptive abilities so it could effectively take the teacher role. Researchers suggested using Learning Analytics (LA) for representing important information about students online [2]. In this context, Siemens [4] defined LA as "the use of intelligent data, learner-produced data, and analysis models to discover information and social connections, and to predict and advise on learning". Learning analytics is recently a hot topic among researchers and educators where various groups, societies, and journals are encouraging the research in LA field and the practice in higher education [1].

LA is often integrated into online learning environments, including Moodle, through the use of plugins. However, plugins usually require a considerable effort, most often involving programming, to adapt or deploy them [2]. This can limit their use by teachers. In addition, to the best of our knowledge, no plugin is reported online which provides real-time interventions to students for a better learning process. Additionally, several studies highlighted the effectiveness of applying gamification in online learning environments to motivate and engage students [5, 6]. Gamification refers to the use of game design elements, such as badges and points, in non-gaming contexts [7].

Therefore, this paper presents an intelligent gamified Moodle (iMoodle), based on a newly developed online LA system named Supervise Me in Moodle (SMiM), which: (1) provides dashboards for teachers to easily help them supervise their students online; (2) predicts at-risk students who might fail to pass their final exams. Specifically, the use of some game design elements might help in predicting students' with lower performance and who can be at-risk of failing to pass their final exams; and, (3) provides real-time interventions, as notifications, by providing supportive learning content for students while learning.

The rest of the paper is structured as follows: Sect. 2 conducts a literature review about gamification and learning analytics. Section 3 presents the implemented framework of the gamified iMoodle with the use of SMiM system. Section 4 explains the experimental procedure for evaluating iMoodle and discusses the obtained results. Finally, Sect. 5 makes a conclusion with a summary of the findings, limitations and potential research directions.

2 Related Work

2.1 Gamification

Various approaches were proposed in the literature to motivate students and increase their learning outcomes. One of these approaches is gamification which refers to the use of the motivational power of digital games via the application of game design elements, such as badges and leaderboard, in non-gaming context to engage and motivate users [7]. According to Kapp [8], gamification is defined as "using game-based mechanics, aesthetics and game thinking to engage people, motivate action, promote learning, and solve problems". Many researchers discussed the effectiveness of gamification in educational contexts [5, 9, 10]. For instance, Kim, Song, Lockee and Burton [5] stated that gamification is an effective instructional approach that is able to increase students' motivation and engagement, enhance their learning performance and promote collaboration skills. Brewer et al. [11] also found that the application of gamification in a learning environment has helped in increasing the percentage of task completion from 73 to 97%.

Several game design elements were reported in the literature that can be integrated into educational contexts, but the most commonly used ones are Points, Badges and Leaderboards (PBL) [12]. In this context, Garcia et al. [13] investigated the efficiency of gamification by implementing PBL into programming course. They found that students' performance in programming tests increased by using a gamified environment compared to a non-gamified environment. Similarly, an experiment study by Hew et al. [14] at an Asian university reported that the integration of points, badges and leaderboard have a positive impact on students' motivation and engagement to involve more in difficult tasks. Barata et al. [15] also included game design elements like points, levels, leaderboard, challenges and badges to gamify a Master's level college course and found that gamification can be an effective tool to enhance students' attendance and participation,

Additionally, the implemented game design elements, such as points and progress bar, can also give an overview of students' progress and performance in a given course. Therefore, several researchers suggested the use of these elements to motivate students and also to provide teachers with feedback about their students' performance. This can further help them predict at-risk students [6, 16]. For example, the number of the collected badges from the submitted activities and students' rank on the leaderboard, which is based on their collected number of points from their interactions with the learning environment, are indicators of students' performance in the course, hence they can be used to help the system predict the students with low performance (at-risk of failing or dropping a class).

2.2 Learning Analytics in Moodle

Learning analytics has emerged as a very promising area with techniques to effectively use the data generated by students while learning to improve the learning process. Van Barneveld et al. [17] defined LA as "the use of analytic techniques to help target instructional, curricular, and support resources to support the achievement of specific learning goals". Powell and MacNeill [18] identified five potential purposes of LA as follows: (1) provide students feedback about their learning progress compared to their colleagues; (2) predict at-risk students; (3) help teachers plan interventions when needed; (4) enhance the designed courses; and, (5) support decision making when it comes to administrative tasks.

Moodle offers several learning analytics tools to assess students' performance and to help in evaluating different skills and competencies. For example, GISMO [19] is a visualization tool for Moodle which is used by teachers to analyze the learning process of all students. It is incorporated within Moodle as an additional block. It generates graphical representations to evaluate students' behaviors, based on their log data. MOCLog [19] analyzes online students' interactions and provides summative statistical reports for both students and teachers to enable them to better understand the educational process. Analytics and Recommendations [20] uses visualization techniques, namely colors and graphs, to provide information regarding students' involvement in each activity of online course as well as recommendations to students so that they can improve their attainment. LAe-R [21] is a plugin which is based on the concept of assessment rubrics technique. LAe-R has various grading levels and criteria that are associated with students' data identified from the analysis of their online interactions and learning behaviors. At-risk student reporting tool [22] provides information for teachers, based on a decision tree model, about students who might be at risk of failing a course.

All the above presented LA tools in Moodle focus mostly on offering various criteria which help teachers in assessing design aspects of the effectiveness of their provided online courses for improving their quality and for identifying opportunities for interventions and improvements. However, despite the fact that predicting at-risk students early in the semester can increase academic success [23], only one tool focuses on doing so (i.e., At-risk student reporting tool). In particular, this tool simply reports the at-risk students to teachers without providing them a medium for interventions to help these students. In addition, most of the above-presented tools are in the form of plugins which usually require a considerable effort, most often involving programming, to adapt or deploy them [2]. To overcome these difficulties, a new iMoodle is developed where its framework is described in the next section. iMoodle differs from Moodle by having a built-in LA system, namely SMiM, which easily helps teachers control the online learning process without going through the complicated process of installing different plugins to achieve different objectives (since every plugin has its own objective). iMoodle also differs from Moodle by providing students real-time interventions and support as notifications as well as predicting at-risk students.

3 Framework of the Intelligent Gamified Moodle (iMoodle)

Figure 1 presents the framework of the implemented gamified iMoodle [24]. iMoodle aims to predict at-risk students as well as model students' personalities to provide them personalized interventions. Specifically, the student's personality, as an individual difference, was considered in this research due to its importance and influence on the learning process and behaviors of students [25]. Therefore, modeling the students' personalities, for instance, whether they are extrovert or introvert, can enhance their learning outcomes and specifically provide more appropriate interventions for them if they are at-risk [26]. However, this paper mainly focuses on predicting at-risk students, and personality modeling is beyond its scope. As shown in Fig. 1, during the learning process, the students' traces are collected in an online database and automatically analyzed in order to extract knowledge and provide real-time interventions.

A learning analytic system SMiM is developed and integrated into iMoodle in the Moodle block form where teachers can easily access it and keep track of their

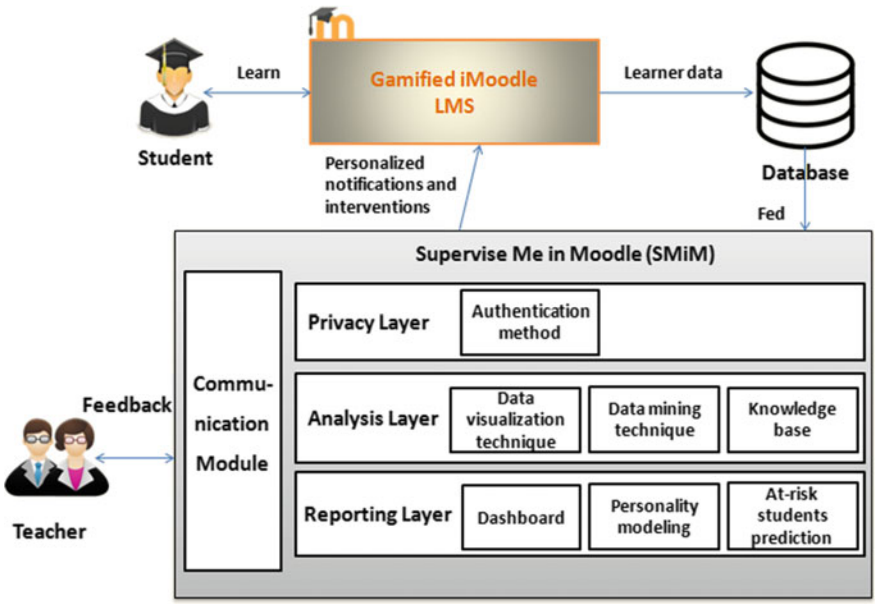


Fig. 1 The developed iMoodle Framework

students in each enrolled course. SMiM has three layers, namely: (1) privacy layer keeps students' traces safe; (2) analysis layer uses both data mining and visualization techniques to extract useful information for teachers; and, (3) reporting layer predicts at-risk students, implicitly model personality based on the students log data, and provides reports and real-time interventions while learning. Each of these layers as well as the gamified iMoodle are explained in the next subsequent sections.

3.1 Gamified iMoodle

To enhance students' learning motivation and engagement, gamification was applied in our iMoodle. Specifically, to have an effective application of gamification, the self-determination theory was applied while designing our gamifed iMoodle. This theory is one of the motivational theories which is widely and successfully applied in gamified learning environments [13]. It is based on the fulfillment of students' different psychological needs [27, 28], namely: (1) need for competence refers to the motivation to overcome challenges and achieved success. This can be satisfied using game design elements which provide feedback about students' success to trigger the feeling of competence and challenge; (2) need for autonomy refers to self-direction and freedom of choices. This can be satisfied using game design elements which allow students to be in charge and make their own decisions; and, (3) need for social relatedness refers to the feeling of connectedness and being a part of a group. This can be satisfied using game design elements which can trigger the feeling of relatedness within students. Table 1 presents the selected and implemented game design elements in our iMoodle, their descriptions, and how they are related to the three psychological needs.

Table 1 Implemented Game design elements in the gamified iMoodle

Psychological needs	Game design elements and description	Matching psychological needs to game elements
Competence	*Points*: numerical presentation of student's performance	They give an immediate feedback about students' progress and performance in the course
	Leaderboard: a board that shows students' rank based on their collected points	
	Progress bar: shows student's progress in a course	
	Badges: virtual rewards	
Autonomy	*Badges*: virtual rewards	It provides a freedom of choice for students to display or hide their awarded badges on their profiles
Social relatedness	*Chat*: instantaneous online discussion	It provides social support

3.2 SMiM

The three main layers of the SMiM learning analytics system are detailed below.
Privacy Layer. This layer aims to keep the online students' privacy safe with the login and password authentication method. In this context, to access the reports and information provided by SMiM, the teacher should have his/her session already active on iMoodle (i.e., the teacher has already entered his/her credentials to access iMoodle and chosen his/her courses). If not, the teacher will be redirected to the authentication interface. This keeps the information regarding students safe where only authorized teachers can have access to it. In particular, the student's password is encrypted and stored within the online database. In addition, the Secure Sockets Layer (SSL) protocol is used to ensure a secured communication of students' data within iMoodle. Furthermore, since the collected data and the obtained analytics results, recommendations and interventions should have a pre-defined time for how long they are going to be stored and used [29], the collected traces and generated reports are stored for a pre-defined period (one academic year) before they are automatically deleted.
Analysis Layer. This layer aims to analyze the students' collected data in order to extract useful information for teachers, predict at-risk students and generate real-time interventions for them. Specifically, SMiM uses both data visualization and data mining techniques to analyze these traces. Data visualization is the use of computer-supported, interactive, visual representations of abstract data to amplify cognition. This can be achieved, for example, using tables, charts and histograms. In this context, SMiM uses data visualization to provide statistical reports for teachers to control the learning process and keep track of their students. Data mining, on the other hand, is the process of applying a computer-based methodology for discovering knowledge from data. In this context, SMiM uses association rules mining based on Apriori algorithm, to predict early in the semester at-risk students within iMoodle who would likely fail their final exams of a particular course, hence increase academic success by providing early support.

Association rule mining discovers relationships among attributes in databases, producing if-then statements concerning attribute-values. An $X \Rightarrow Y$ association rule expresses a close correlation between items (attribute-value) in a database with values of support and confidence as survey by Shankar and Purosothmana [30]. In particular, Apriori Algorithm is used to find these association rules. It has two important variables: Minimum Support Threshold which is a support of an association pattern is the percentage of task-relevant data transaction for which the pattern is true (see equation a) and Minimum Confidence Threshold which is defined as the measure of certainty associated with each pattern (see equation b) [31].

$$\text{(a)}\quad \text{Support}\ (X \Rightarrow Y) = \frac{\text{Number of tuples containing both } X \text{ and } Y}{\text{Total number of tuples}}$$

$$\text{(b)}\quad \text{Confidence}\ (X \Rightarrow Y) = \frac{\text{Number of tuples containing both } X \text{and } Y}{\text{Number of tuples containing } X}$$

The Apriori algorithm developed within SMiM was first applied on previous learning dataset (knowledge base) from a public university in Tunisia which contains the final exam grades of students in a course and their learning behaviors within a classic Moodle. This was to extract the predictive association rules to detect at-risk students in iMoodle. In particular, based on a literature review, two types of factors are found that can help in predicting at-risk students namely, demographic and performance/behavior [32–34].

Demographic factors describe the students' background and profile to identify the probability of students to successfully complete a course. However, since iMoodle aims to be used in both online and blended learning, demographic data would not work particularly well in this case because students can be from anywhere in the world. Performance/behavior factors, on the other hand, consider students' actions in a course, such as what they viewed or submitted, as well as their performance on activities/assignments based on the assigned grades from the teacher.

Based on student performance/behavior, we selected five factors to help in at-risk students' identification, namely: (1) Number of acquired badges which highlights the number of conducted learning activities, since every time a student finishes a learning activity, he/she gets a badge. This factor has been often used, for instance, by Billings [34], Xenos et al. [35] and Macfadyen and Dawson [36]; (2) Activities grades which refer to the value assigned by teachers to assignments and quizzes requested and delivered by students. In particular, if a student did not deliver an activity before its deadline, he/she receives a grade of zero. Also, if a teacher has not given the grade yet, this activity is not considered. In particular, the learning activities can be various assignments or quizzes that should be answered. This factor has been often used for designing early at-risk students' warning systems, for example, by Macfadyen and Dawson [36] and Arnold and Pistilli [37]; (3) Student's rank on the leaderboard which is based on the acquired number of points from his/her interaction with iMoodle (i.e., doing activities, participating in chat and forums, access to resources, etc.). For instance, if a student does not complete all the required activities and have low interaction with iMoodle, his/her score will be very low, hence he/she will be ranked at the bottom. Specifically, this factor presents an engagement trigger and an indicator of predicting at-risk students as highlighted by Liu et al. [38]; (4) Course progress which can be seen in the progress bar. It refers to the number of activities realized from the total of activities requested in a course. This factor has been recommended by Khalil and Ebner [16] to help in predicting at-risk students who have not completed the requested activities; and, (5) Forum and chat interactions which refer to students' participation in online discussions, such as the number of posts read, posts created and replies. This factor has been often used by Liu et al. [38] and Khalil and Ebner [16].

Reporting Layer. After the analysis process is done (within the analysis layer), the reporting layer provides the generated reports and the automatic real-time interventions as follows:

Dashboard: SMiM provides dashboards within iMoodle for teachers to aid them control the learning process online and keep track of their students. This dashboard highlights the number of completion rate of each learning activity and quiz in each course, form, and chat interactions, the number of badges earned by each student, the progress of each student in the course and his/her rank on the leaderboard based on their collected number of points. For instance, as shown in Fig. 2, SMiM shows teachers the completion rate of each learning activity in the "Méthodologie de Conception Orientée Objet" (MCOO) course. This can help them keep track of their students' progress online, hence not move to the next learning activity until they ensure that all their students have done the first one. Also, when the teacher clicks on each assignment, iMoodle shows the percent of students who got over and under the average grade. In particular, if students are at-risk, iMoodle provides real-time interventions, as notifications, by suggesting additional learning content support for them to further enhance their knowledge. The details regarding these provided supportive notifications are automatically stored in the database for future uses. Not only that, an interface is also shown for teachers where they can directly communicate with those students to help them pass the learning activities which they did not correctly finish.

At-risk students prediction: Through the use of predictive modeling techniques, it is possible to forecast students' success in a course and identify those that are at-risk. Therefore, iMoodle, based on SMiM system, uses a predictive model (discussed in the analysis layer) as an early warning system to predict at-risk students in a course and inform the teacher. Teachers can then communicate with the at-risk students and provide them the required support for improving their performance in the course. Figure 3 presents examples of strong association rules obtained after running the Apriori algorithm. It is seen that the confidence of the association rules is very high (100%). In particular, the "forum and chat interactions" factor was excluded because over 75% of students did not use the forum and chat facilities. Finally, Fig. 4 presents the detected at-risk students based on the obtained association rules.

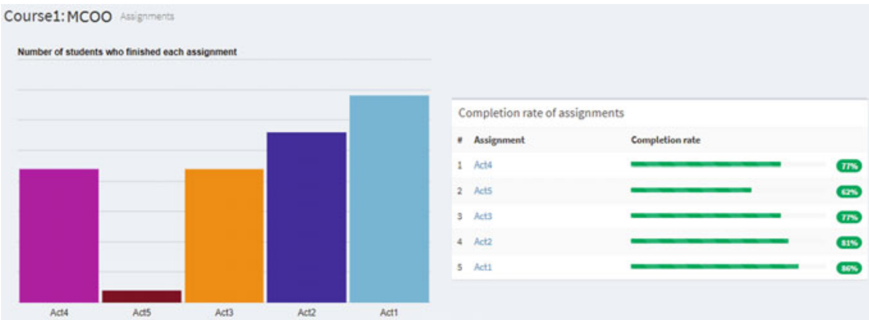


Fig. 2 Completion rate dashboard of learning activities within a given course

Association Rule	Confidence
assignments.low,quiz.low =>failure	100 %
Badges.low,quiz.low =>failure	100 %
Badges.low,assignments.low =>failure	100 %
assignments.low,quiz.low,rank.low,progress.low =>failure	100 %
Badges.low,rank.low,progress.low =>failure	100 %

Fig. 3 Examples of the obtained strong association rules

Modeling at risk students

Show 10 entries Search:

Fist name	Last name	Email	Phone
ab[redacted]	ba[redacted]	ba[redacted]@gmail.com	[redacted]
ach[redacted]	ha[redacted]	ha[redacted]@gmail.com	[redacted]
am[redacted]	ay[redacted]	ay[redacted]@gmail.com	[redacted]
ach[redacted]	ha[redacted]	ha[redacted]@gmail.com	[redacted]

Fig. 4 Identified at-risk students in a given course

4 Evaluation

An experiment was conducted to evaluate the technical reliability of the beta version of iMoodle. This experiment also evaluates the accuracy rate of iMoodle using SMiM in predicting at-risk students.

4.1 Experimental Design

The beta version of the iMoodle based on the built-in SMiM system was technically evaluated to test and enhance it if there were any bugs. In this context, the developed iMoodle was used for three months, in a public Tunisian university. The teacher was then requested to give a report highlighting the technical issues that were faced when using iMoodle. The feedback given by the teacher was then used to further work on the beta version and make it stable for future uses.

The post-fact technique was also used to mainly evaluate the accuracy of iMoodle in predicting at-risk students. This technique uses data from past events to understand a phenomenon. In this case, the data from a finished course on a classic Moodle was analyzed using the predictive model within iMoodle. The obtained at-risk students were then verified based on their exam grades to evaluate the accuracy rate.

4.2 Results

While the teacher reported that the developed iMoodle based on SMiM system helped her easily control the learning process and communicate with her students, several technical issues were found. For instance, the teacher reported that the automatic notification for students to provide additional supportive learning contents did not work for some learning activities. She also reported that some options within iMoodle (e.g., activate/deactivate notifications) should be disabled from the students' learning sessions in order to not affect the learning process. These technical issues were fixed in our iMoodle stable version.

Table 2, on the other hand, presents the obtained results of the accuracy rate of predicting at-risk students within iMoodle. In particular, the number of correct results shows the number of students who are correctly identified within iMoodle in comparison with their final exams grades. The intervention layer within iMoodle, in this particular experiment, has no impact since the experiment is conducted using previous dataset and not from a current learning process. The efficiency of iMoodle in reducing the number of at-risk students is beyond the scope of this paper.

As shown in Table 2, the accuracy rate of iMoodle in predicting at-risk students is almost 90%, which can be considered as sufficiently high. This means that our system is efficient in the prediction process. Particularly, only seven students were not correctly identified (i.e., they were at-risk but iMoodle identified them as not, and vice versa).

The obtained accuracy rate result was compared with other similar works, including the developed plugin for detecting at-risk students. For instance, Kotsiantis et al. [39] found that the accuracy rate of their system range between 63% and 83%. The prediction system of Da Silva et al. [22] had an accuracy of 85%. Liu et al. [38] and Khalil and Ebner [16], however, did not mention the accuracy rate of their systems in predicting at-risk students. To conclude, the developed gamifed iMoodle

Table 2 Accuracy rate of predicting at-risk students within iMoodle

Course	Number of students	Number of correct results	Number of wrong results	Accuracy
MCOO	61	54	7	88.52%

based on SMiM system has a better accuracy rate than the previous systems (which have mentioned their accuracy rates). Particularly, it can be deduced that the used factors, namely number of acquired badges, activities grades (in both assignments and quizzes), student's rank on the leaderboard and course progress provide efficient combination for the at-risk identification.

It should be noted that it is very difficult to correctly identify all students since some students might alter their behaviors and put more effort to study outside of iMoodle (which cannot be detected) or fail the exam due to unforeseen events, such as becoming ill at the time of the exam.

5 Conclusion

This paper presented a new gamified and intelligent version of Moodle (iMoodle) which aims to help teachers control the learning process online and keep track of their students. iMoodle provides, based on a built-in LA system called SMiM, a dashboard for teachers to help them understand the learning process and make decisions. It also provides an early warning system by detecting at-risk students, based on various factors extracted from the literature, using association rules mining. Finally, iMoodle provides automatic personalized supportive learning content as notifications for students based on their behaviors online. The beta version of iMoodle was tested for three months during the first semester and several technical issues were identified and fixed. Furthermore, the predictive model was evaluated and the obtained results highlighted that iMoodle has a high accuracy rate in predicting at-risk students.

Despite the promising results, there were some limitations of the experiment which should be acknowledged and further investigated. For instance, the effectiveness of the iMoodle in learning was not evaluated. Also, the detection process of at-risk students was from only one course which has limited number of students (only 61 students). Future research work could focus on: (1) using the iMoodle and compare its impact on learning outcomes and technology acceptance with a classic Moodle; (2) investigating the efficiency of iMoodle using the intervention layer in reducing the number of at-risk students and increasing academic success, in comparison with a classic Moodle; and, (3) further develop iMoodle to provide as well personalized interventions based on students' personalities.

References

1. Yassine, S., Kadry, S., & Sicilia, M. A. (2016). *A framework for learning analytics in moodle for assessing course outcomes*. In Global Engineering Education Conference (pp. 261–266).
2. Vozniuk, A., Govaerts, S., & Gillet, D. (2013). *Towards portable learning analytics dashboards*. In 13th International Conference on Advanced Learning Technologies (pp. 412–416).
3. Anohina, A. (2007). Advances in intelligent tutoring systems: Problem-solving modes and model of hints. *Journal of Computers Communications & Control, 2*(1), 48–55.

4. Siemens, G. (2010). What are learning analytics?. Retrieved August 12, 2016, from http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/.
5. Kim, S., Song, K., Lockee, B., & Burton, J. (2018). What is gamification in learning and education?. In *Gamification in learning and education* (pp. 25–38). Springer, Cham.
6. Gañán, D., Caballé, S., Clarisó, R., & Conesa, J. (2016). Analysis and design of an eLearning platform featuring learning analytics and gamification. In *2016 10th International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS)*, IEEE (pp. 87–94).
7. Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). Gamification: Using game-design elements in non-gaming contexts. In *Proceedings of the CHI 2011*. Vancouver, BC, Canada.
8. Kapp, K. M. (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education*. Wiley.
9. Andrade, F. R. H., Mizoguchi, R., & Isotani, S. (2016). The Bright and Dark Sides of Gamification. In *Proceedings of the International Conference on Intelligent Tutoring Systems*, *9684*, 1–11.
10. Villagrasa, S., Fonseca, D., Redondo, E., & Duran, J. (2018). *Teaching case of gamification and visual technologies for education*. Gamification in Education: Breakthroughs in Research and Practice: Breakthroughs in Research and Practice, p. 205.
11. Brewer, R., Anthony, L., Brown, Q., Irwin, G., Nias, J., & Tate, B. (2013). Using gamification to motivate children to complete empirical studies in lab environments. In *12th International Conference on Interaction Design and Children* (pp. 388–391). New York.
12. Dichev, C., & Dicheva, D. (2017). Gamifying education: what is known, what is believed and what remains uncertain: A critical review. *International Journal of Educational Technology in Higher Education, 14*(1), 9.
13. Garcia, J., Copiaco, J. R., Nufable, J. P., Amoranto, F., & Azcarraga, J. (2015). Code it! A gamified learning environment for iterative programming. In *Doctoral Student Consortium (DSC)-Proceedings of the 23rd International Conference on Computers in Education (ICCE)* (pp. 373– 378).
14. Hew, K. F., Huang, B., Chu, K. W. S., & Chiu, D. K. (2016). Engaging Asian students through game mechanics: Findings from two experiment studies. *Computers & Education, 92,* 221–236.
15. Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2013). Engaging engineering students with gamification. In *5th international conference on Games and virtual worlds for serious applications (VSGAMES), IEEE* (pp. 1–8).
16. Khalil, M., & Ebner, M. (2016). Learning analytics in MOOCs: Can data improve students retention and learning? In *EdMedia + Innovate Learning. Association for the Advancement of Computing in Education (AACE)* (pp. 581–588).
17. Van Barneveld, A., Arnold, K. E., & Campbell, J. P. (2012). Analytics in higher education: Establishing a common language. *EDUCAUSE learning initiative, 1,* 1–11.
18. Powell, S., & MacNeill, S. (2012). Institutional readiness for analytics. *JISC CETIS Analytics Series, 1*(8).
19. Dietz-Uhler, B., & Hurn, J. E. (2013). Using learning analytics to predict (and improve) student success: A faculty perspective. *Journal of Interactive Online Learning, 12*(1), 17–26.
20. Sampayo, C. F. (2013). *Analytics and Recommendations*. In Moodle Docs. Retrieved from https://moodle.org/plugins/view.php?plugin=block_analytics_recommendations.
21. Petropoulou, O., Kasimatis, K., Dimopoulos, I., & Retalis, S. (2014). LAe-R: A new learning analytics tool in Moodle for assessing students' performance. *Bulletin of the IEEE Technical Committee on Learning Technology, 16*(1), 1–13.
22. Da Silva, J. M. C., Hobbs, D., & Graf, S. (2014). Integrating an at-risk student model into learning management systems. In *Nuevas Ideas en Informática Educativa TISE* (pp. 120–124).
23. Marbouti, F., Diefes-Dux, H. A., & Madhavan, K. (2016). Models for early prediction of at-risk students in a course using standards-based grading. *Computers & Education, 103,* 1–15.
24. Tlili, A., Essalmi, F., & Jemni, M., Chang, M., & Kinshuk. (2018). iMoodle: An Intelligent Moodle Based on Learning Analytics. In *Intelligent Tutoring System* (pp. 476–479).

25. Tlili, A., Essalmi, F., Jemni, M., Kinshuk, & Chen, N. S. (2016). Role of personality in computer based learning. *Computers in Human Behavior, 64,* 805–813.
26. Santos, O. C. (2016). Emotions and personality in adaptive e-learning systems: an affective computing perspective. In *Emotions and personality in personalized services* (pp. 263–285). Springer, Cham.
27. Lombriser, P., Dalpiaz, F., Lucassen, G., & Brinkkemper, S. (2016). *Gamified requirements engineering: Model and experimentation*. Springer.
28. Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Springer.
29. Tlili, A., Essalmi, F., Jemni, M., Kinshuk., & Chen, N. S. (2018). A complete validated learning analytics framework: Designing issues from data preparation perspective. *International Journal of Information and Communication Technology Education (IJICTE), 14*(2), 1–16.
30. Shankar. S. & Purosothmana, T. (2009). Utility sentient frequent itemset mining and association rule mining. A literature survey and comparative study. *International Journal of Soft Computing Applications, 4*, 81–95.
31. Chan, C. C. H., Ming-Hsiu, L., & Yun-chiang, K. (2007). Association rules mining for knowledge management: A case study of library services. In *Proceedings of the 9th WSEAS international conference on Mathematical methods and computational techniques in electrical engineering* (pp. 1–6).
32. Wolff, A., Zdrahal, Z., Nikolov, A., & Pantucek, M. (2013). Improving retention: Predicting at-risk students by analysing clicking behaviour in a virtual learning environment. In *Proceedings of the third international conference on learning analytics and knowledge* (pp. 145–149).
33. Levy, Y. (2007). Comparing dropouts and persistence in e-Learning courses. *Computers & Education, 48*(2), 185–204.
34. Billings, D. M. (1987). Factors related to progress towards completion of correspondence courses in a baccalaureate nursing programme. *Journal of Advanced Nursing, 12*(6), 743–750.
35. Xenos, M., Pierrakeas, C., & Pintelas, P. (2002). A survey on student dropout rates and dropout causes concerning the students in the course of informatics of the Hellenic Open University. *Computers & Education, 39*(4), 361–377.
36. Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education, 54*(2), 588–599.
37. Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 267–270).
38. Liu, D. Y. T., Froissard, J. C., Richards, D., & Atif, A. (2015). An enhanced learning analytics plugin for Moodle: Student engagement and personalised intervention.
39. Kotsiantis, S. B., Pierrakeas, C. J., & Pintelas, P. E. (2003). Preventing student dropout in distance learning using machine learning techniques. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems* pp. 267–274.

Chapter 7
Integrating a Learning Analytics Dashboard in an Online Educational Game

J. X. Seaton, Maiga Chang and Sabine Graf

Abstract The goal of educational games is to allow players to learn unconsciously while playing. The more a player plays an educational game, the more their learning and their skills can increase. Just like in other games, players in educational games may encounter situations where they feel like they cannot make further progress like passing a level or completing a quest. If players are stuck in an educational game, then they may choose to quit playing the game, which also means that they quit learning. Especially if players quit early, the effect of the educational game will be limited and not last for too long. Therefore, providing players with information on how to improve their performance, such as when and how to play the game, which parts or skill improvement is needed to overcome a challenge and go further in the game, can help to encourage them to play the game more often and continuously. This chapter discusses how the research team integrated a learning analytics dashboard into an educational game so that the players can see their game play performance and habits, and find clues and strategies to improve their in-game performance. The proposed dashboard provides players with a variety of information that will allow them to see how their performance and skills change over time, what their weakness and strengths are and much more. This chapter talks about the design of learning analytics dashboards for educational games and explains the use of the proposed dashboard to help players improve their in-game performance through use cases with 3-month simulated gameplay data.

1 Introduction

Educational games have the potential to make learning more engaging because, unlike traditional media, games are interactive. Educational games do not just present

J. X. Seaton · M. Chang (✉) · S. Graf
Athabasca University, Athabasca, Canada
e-mail: maigac@athabascau.ca

S. Graf
e-mail: sabineg@athabascau.ca

A. Tlili and M. Chang (eds.), *Data Analytics Approaches in Educational Games and Gamification Systems*, Smart Computing and Intelligence,
https://doi.org/10.1007/978-981-32-9335-9_7

players with information, but problems for them to solve [1]. Part of what makes a game fun is that the in-game problems are challenging [1]. By framing a learning objective around such a challenge, it can be integrated into a game in a fun way. In this way, an educational game allows the player to learn by playing. Learning is implicit from the feedback they receive about the actions they have taken or choices they have made in the game [2].

However, the mere incorporation of an educational game into a learning environment is not guaranteed to increase student motivation to learn [3]. The initial novelty of an educational game can increase motivation, but that interest will fade over time as the game becomes familiar to the students [4]. Therefore, it is important for educational games to include motivational techniques that encourage continuous play [3]. An effective motivational technique in education is to highlight a learner's accomplishments [5], thus learning analytics dashboards which visualize a player's improvement could be motivational.

It is difficult for players in any game, including educational games, to connect the feedback they see in a game to how they can further improve their in-game performance. When they are leveling up, usually they are seeing more challenging and difficult quests or problems. It is common to see that they are learning through the loop of failing and re-attempting. Feedback about a failed attempt can help them play the game better [6]. Incorporating learning analytics into educational games can demonstrate to players how their gameplay is connected to their improvement in the game, provide them with an opportunity to analyze their gameplay and the effects of their playing habits on their in-game performance, and find strategies to improve their gameplay. In turn, by improving their gameplay, players automatically and implicitly improve their learning in the educational game.

The aim of this chapter is to demonstrate how learning analytics can be incorporated into an educational game. The educational game featured in this chapter focuses on improving players' metacognitive skills by playing short subgames against other players. The learning analytics dashboard presented in the chapter uses two types of charts: line graphs and scatter plots. Line graphs are used to show a player how their metacognitive skills have improved over time, and while they have played. The scatter plot visualizes how the player's performance in subgames is affected by the time of day or how long a player plays in a single sitting. The dashboard has been evaluated in a proof of concept evaluation with three months of simulated gameplay data, which demonstrates the benefits of the information presented in the dashboard.

This chapter first reviews related works on how learning analytics have been incorporated into educational games. Then it presents a general overview of the educational game where the learning analytics dashboard was implemented, followed by a description of the dashboard. Next, a proof of concept evaluation of the proposed dashboard is presented. Finally, the conclusion section summarizes the findings and discusses future work.

2 Related Works

Applying learning analytics to educational games is an emerging field. Much of the motivation to incorporate learning analytics into educational games has been to understand how to use educational games for assessment [7]. Games can appear as black boxes that do not give instructors much information about the player's learning process [8]. Thus, although teachers are open to including educational games in the classroom, they are hesitant to use games to assess learning [9]. Learning analytics is seen as a way to open the black box by providing aggregate data about where players are struggling [8], and the common mistakes made by learners [7].

Educational games track and log a variety of information about players that can be used for learning analytics to work [10]. For example, an event log with a timestamp, information about when players login or reach a goal can determine how long players are playing or how long it took them to reach their goals. Additional gameplay data specific to the educational game include the player's scores, position, or decisions made in the game can all provide meaningful information about how the player progressed through the game [11].

However, while games could technically log a lot of data, it can be challenging to add learning analytics into them because typical game design often discards any variables not necessary for gameplay to optimize the performance of the game [12]. Therefore, integrating learning analytics into an existing game often means that the data collected is limited and may not provide the desired information.

Loh et al. [13] encountered this issue when incorporating learning analytics into an existing commercial game *Neverwinter Nights*. *Neverwinter Nights* was modified to be an educational game, where the path a student took through the game could tell the instructor something about the process the student used to complete his or her works. The modified educational game intercepted and logged gameplay data to create learning analytics reports for instructors. The reports contained both individual and aggregate information about the students' paths to help the instructor assess individual student's learning progress and identify common issues that students faced. One issue with integrating learning analytics into *Neverwinter Nights* was that some of the gameplay data was not descriptive enough. For example, the fact a player got a new item could be seen from the log, but how the player got the item is not known because it was not relevant to gameplay.

Educational games that are designed for the inclusion of learning analytics from the ground up can create variables and data specifically suited for learning analytics. Activities directly related to learning objectives can then be logged for later analysis. For example, the game CMX is an educational Massively Multiplayer Online Role Playing Game (MMORPG) that teaches computer programming [9]. With respect to learning analytics, the game creates reports for instructors about how players are progressing within the game. Instructors can see how many learning activities students have completed, how many errors they made, how many times a player has logged in, how long a player has played, and how many times a player has interacted with another player. To aid in assessment, instructors can also create a report that assigns students a percent score by comparing the students' data to a sample of ideal game behavior.

While many educational games use learning analytics to provide teachers with additional information, some games provide such additional information to learners. For example, the educational game eAdventure [14], which is an educational game plugin designed for edX courses, provides students with reports that assess their learning. Due to the very high number of students in edX's massively open online courses, it is very difficult to provide students with individual assessments from teachers. Therefore, eAdventure uses the existing learning analytics tools offered in edX courses to provide learners with some additional information on how they are doing in the game, including how much time they are spending playing the game; the time it took them to finished the game (or a subsection of the game), and their score in the game.

The application program interface xAPI can also aid in designing educational games that support learning analytics for assessment [8]. By using xAPI, game designers can determine what data is relevant to learning, log the data during gameplay, formatted according to xAPI specifications, and then generate different learning analytic reports. The reports can feature a variety of information and can be configured to display reports relevant to students, teachers, or administrators. For example, the game Countrix, which is an educational game about geography, utilizes xAPI to log information about student errors to create a report for the players in real-time about their error rate [11].

The learning analytics dashboard presented in this chapter differs from those discussed in that the focus is not on assessing a player's learning progress. The purpose of the dashboard is to motivate playing by helping players understand how they can perform better in the game. As players improve in an educational game, they are implicitly learning and improving in the areas targeted by the game. Thus, supporting higher in-game performance translates into supporting learning. In such environments, players are playing the game for enjoyment and not necessarily for learning. Therefore, the players might appreciate information about how to improve their in-game performance more than an assessment report of their learning progress.

The learning analytics dashboard introduced in this chapter also has been designed for the inclusion of learning analytics rather than adding existing learning analytics tools or features afterwards. As such, it benefits from using a wide variety of data to provide players with game-specific information on their performance and play habits as well as allows them to create their own custom visualizations to analyze their performance and play habits.

3 Overview of Game

The educational game designed by the research team is aiming to improve players' metacognitive skills. Metacognition is the understanding of a person's own cognition and thought processes [15]. The game targets four skills that are essential to metacognition: (1) problem solving, (2) associative reasoning, (3) organization and planning, and (4) monitoring/checking work for accuracy.

The game has ten subgames and each of them targets the improvement of a metacognitive skill. Players are playing matches against other players. In each match, players are playing three subgames and both players are scored by how they performed individually and against their opponent. For each subgame played, a performance score is calculated that shows how well the player played that subgame. The player is also compared to their opponent by adding up the performance score for each of the three subgames they played in a match. The player with the highest sum of performance scores is the declared winner of the match. The winner receives points and the loser loses points, which allow players to be ranked against other players based on game performance. There is no limit to how many matches a player can play in a play session. A play session is defined from when a player logs into the game to when they log out or are inactive for ten minutes.

A metacognitive skill score in a subgame is calculated based on the performance score as a percentage value compared to the highest possible performance a player can get. The score represents the metacognitive skill level reached in the subgame. The player's overall score for a particular metacognitive skill comes from the highest scores in all subgames associated with the same metacognitive skill.

Besides the points players get for winning a match, several other motivational features have been included in the game to encourage players to continue playing. Players can unlock 48 badges that are linked to game activities, such as logging in for consecutive days in a row, winning matches, and using the learning analytics dashboard. Players can also earn currency every time they play a match, which they can then use to upgrade a robot avatar that represents them in the game. The game also features a leaderboard that can rank players against each other. As mentioned, players can be ranked by points, but additionally, player can be ranked on the leaderboard by other metrics, such as their metacognitive skill score, or how much currency they have.

4 Learning Analytics Dashboard

The purpose of the learning analytics dashboard is to show players how they can improve their performance in subgames, and consequently, improve their metacognitive skills. A variety of information is tracked about players for the learning analytics dashboard. This information includes, when players start and end a play session; when players start and end a match, when players start and end a subgame; which subgames players played; which metacognitive skill is associated with the subgames played; how the players performed in subgames; and the players metacognitive skill score after a subgame played. This information can be used to determine how often players login, how long they play, how many matches they play in a play session, what time in a day they usually play, how they performed in a subgame, and how their metacognitive skill score changes after playing a subgame.

There are two charts that have been adopted by the proposed learning analytics dashboard: (1) line graphs, which visualize metacognitive skill scores; and (2) scatter plots, which visualize the performance scores. The dashboard also offers players (1) a "Brain" tab, which visualizes metacognitive skills; and, (2) a "Game" tab, which visualizes performance in subgames (see Fig. 1).

Through the "Brain" tab, a player can select which metacognitive skills (i.e., a single skill or a group of skills) he or she wants to see and in the "Game" tab, the player can select which subgames (i.e., a single subgame or a group of subgames) performance should be displayed so he or she can check it out. In addition, in the "Brain" tab, each metacognitive skill can be exploded to show the performance scores of the subgames related to the respective metacognitive skill. The purpose of the exploded view is to demonstrate how subgame performances impact the player's metacognitive skill score. Moreover, players can filter based on a time frame using a sliding time frame bar so they can focus on the visualized data within a particular time frame.

The line graphs (in Fig. 1) visualize a player's scores over time to show how the player has improved. The player can check their improvement over days, play sessions, or matches played. Seeing the growth over days can give players a general overview of how they have improved over time; seeing the growth over play sessions or matches can give them more details about how they improved when they have multiple play sessions in a day or multiple matches in a play session.

While the "Brain" tab allows players to see the improvement for each metacognitive skill, at subgame level (i.e., either in the "Game" tab or when a metacognitive

Fig. 1 Line graph of metacognitive skills improvement with problem-solving exploded

skill is exploded in the “Brain” tab), the actual metacognitive skill scores in the particular subgame is shown, providing more details about how well a player did in those particular subgames. For example, Fig. 1 shows such a chart with the problem-solving skill exploded and the other three skills displayed, but not exploded. The metacognitive skills that are not exploded show one line each visualizing how the player’s skill has changed over three months. Whereas, problem-solving instead has two lines, one line for each subgame that contributes to the calculation of the player’s problem-solving score. The subgame lines are red to show that they are associated with problem-solving, but have different line dash patterns so that they can be distinguished from each other.

The scatter plot visualizations focus on showing how performance is affected by playing habits. There are two views: performance by time in a day and performance by matches played in a session. The first view (as shown in Fig. 2) displays how the player performed per metacognitive skill or subgame at different times of the day. The x-axis is the time of day a subgame was played and the y-axis is the performance player got playing the subgame. The purpose of this visualization is to help a player identify if they perform better at different time of a day.

For example, Fig. 2 shows a visualization of a player that has played problem-solving subgames between 8:00 am and 10:00 pm over three months. Points that are close horizontally, represent subgames that were played around same time of day. When grouping subgames by metacognitive skill in the scatter plot, points for the same metacognitive skill are drawn in the same color, but different shapes are used for different subgames. Because both subgames 1 and 2 are associated with

Fig. 2 A scatter plot depicting a player performing better in the evening

problem-solving skill, the points have two shapes: circle for subgame 1 and triangle for subgame 2. Both points are red to show that both are associated with problem-solving.

The second view of the scatter plots (see Fig. 3) shows how performance changes over multiple matches played in a play session. The x-axis shows the time and day the session took place. The y-axis shows when the subgame was played within the session (in minutes), with 0 on the y-axis representing the start of the session. Each point represents the player playing a subgame during a play session. Points that are line up vertically represent a play session. The color of the point indicates the performance of the player had in that subgame—a darker color indicating higher performance and a lighter color indicating lower performance. The purpose of the visualization is to show if a player's performance changes by playing multiple subgames in one session.

In Fig. 3, for example, we can see that in the first session on January 18th, three subgames were played. The three points are light because the player had their lowest performances in those games within the time frame that was selected on the bottom of the screen with the time frame bar. Conversely, the last three subgames in the last session on January 23rd have a dark green colored points indicating that the player had their highest performance in those subgames within the selected time frame.

Fig. 3 Scatter plot of subgame performance by matches played in session

5 Proof of Concept Evaluation

The purpose of the proof of concept evaluation is to verify whether the learning analytics dashboard can give players meaningful information about how they can improve their in-game performance. Three months of simulated gameplay data were created. The evaluation uses four use cases to evaluate the resulting visualizations and to explain how they benefit players. The four use cases include: (1) a player not performing well in one of the four metacognitive skills, (2) a player that plays sometimes very often and sometimes rarely, (3) a player that performs better at a certain time of a day, and (4) a player whose performance increases after re-familiarizing themselves with the game.

Metacognitive skills from one area do not necessarily translate to the others [16]. Therefore, first use case deals with visualizing a player that is lower in one of the four metacognitive skill areas. Figure 1 depicts a line graph that displays a player's metacognitive skills over time. The depicted player has lower scores in games that target problem-solving. The problem-solving skill line is exploded to show that subgame 1 and 2 contribute to the skill score. The player can use this information to determine that he or she needs to develop strategies to improve his or her performance in subgame 1 and 2. Showing the player that both subgames target the same metacognitive skill will also indicate that strategies that work in one game could apply to the other.

Skill development is dependent on many factors, but an important element is regular practice [17]. The second use case demonstrates how the connection between regular practice and high performance can be visualized and noticed by players. Figure 4 depicts a player that plays subgame 8 only a few times in November, then plays it frequently in the month of December, and then infrequently again in January. Although his or her skill improves across the three months, there is a greater improvement in the month where he or she plays often and less improvement in the months where the player plays only a few times. This shows him or her that if he or she wants to improve his or her in-game performance faster, he or she should play the game often rather than erratically.

Performance on some types of cognitive tasks, such as those associated with metacognitive skills, can be varied by time of day [18]. Figure 5 shows the scatter plot which visualizes the player's performance in subgame 5 by the time of day it was played. When he or she is looking at this chart, he or she can see that his or her performance is relatively low in the morning and during the day, and increases towards the evening. Therefore, this chart can help a player to identify which times are better for him or her to play the game. For example, if this player identified that he or she needs to improve his or her Planning and Organization skill and subgame 5 is associated to the skill (according to the dashboard shown in Fig. 1). Figure 5 shows the player that he or she may improve their score by increasing the number of subgames played in the evening.

Performance in cognitive tasks is influenced by familiarity with the task [19]. Players may perform poorly at cognitive task in the beginning or after a longer

Fig. 4 A line graph depicting a player that played less in November and January but more in

Fig. 5 A scatter plot depicting a player that performs better in the evening

break, not because they are unskilled, but because they are unsure about what they need to do. In the context of the game designed by the research team, this could mean that a player might need to warm-up by playing multiple matches and subgames in one play session.

The last use case deals with a visualization of a player who performs better after they have played a few matches to re-familiarize themselves with the subgames. In Fig. 3 the chart depicts a scatter plot of performances in subgame 3 based on how many games were played in a play session. The player's performance increases consistently within a play session, which can be seen by how the points become darker in a session. Between frequent play sessions, the darkness of the points in the beginning of a new session is similar in darkness to the points at the end of the previous session, which indicates that the player performance remains stable between short breaks in play. However, after a longer lapse in play, such as the gap between January 19th and 22nd, the points become much lighter indicating a dropping performance.

Seeing the performance of subgames played in the same session can help players identify if they need to play more games after an absence to re-familiarize themselves with the subgames as well as how their performance changes in sessions with multiple subgames. For example, a player could use this dashboard in tandem with the one featured in Fig. 4 to identify if a plateau in performance could be overcome by playing more subgames in one play session.

6 Conclusion

This chapter presented how to adopt learning analytics into an educational game. The proposed learning analytics dashboard provides a way for players to see and analyze their game play habits and allows them to understand how those habits may affect their in-game performance. With the dashboard, players can be made aware of how to improve their in-game performance. The designed dashboard was evaluated through a proof of concept evaluation with a 3-month gameplay simulated dataset by considering four use cases. The evaluation showed that the dashboard can provide players with meaningful feedback about how their play habits affect their in-game performance and with a useful tool to build strategies to improve their in-game performance.

Future work will focus on players' perceived usability and acceptance toward the learning analytics dashboard. The research team also plans to investigate how players use the dashboard while playing, and if their play habits change after using the dashboard. Players' acceptance of the dashboard will be explored by analyzing players' usage rates and by administering questionnaires to collect their self-reported satisfaction towards the dashboard. The collected data will be analyzed with statistical approaches.

References

1. Gee, J. P. (2013). *Good video games + good learning* (2nd ed.). New York: Peter Lang Publishing.
2. Song, M., & Zhang, S. (2008). EFM: A model for educational game design. In *Technologies for e-learning and digital entertainment* (pp. 509–517).
3. Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Der Spek, E. D. (2013, February). A meta-analysis of the cognitive and motivational effects of serious games. *The Journal of Educational Psychology*, 1–17 (Advance Online Publication).
4. Wang, A. I. (2015, March). The wear out effect of a game-based student response system. *Computers and Edcuation, 82*, 217–227.
5. Keller, J. M. (1987). Strategies for stimulating the motivation to learn. *Performance Improvement, 26*(8), 1–7.
6. Hauge, J. B., Manjón, B. F., Berta, R., Padrón-Nápoles, C., Giucci, G., Westera, W., et al. (2014, July). Implications of learning analytics for serious game design. In *Proceedings of IEEE 14th International Conference on Advanced Learning Technologies (ICALT)* (pp. 230–232). IEEE.
7. Loh, C. S. (2013, January). Improving the impact and return of investment of game-based learning. *The International Journal of Virtual and Personal Learning Environments,4*(1), 1–15.
8. Alonso-Fernandez, C., Calvo, A., Freire, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2017, April). Systematizing game learning analytics for serious games. In *Proceedings of Global Engineering Education Conference (EDUCON)* (pp. 1106–1113). IEEE.
9. Malliarakis, C., Satratzemi, M., & Xinogalos, S. (2014, July). Integrating learning analytics in an educational MMORPG for computer programming. In *Proceedings of IEEE 14th International Conference on Advanced Learning Technologies (ICALT)* (pp. 233–237). IEEE.
10. Serrano-Laguna, A, Torrente, J., Moreno-Ger, P., & Manjón, F. B. (2012, December). Tracing a little for big improvements: Application of learning analytics and videogames for student assessment. *Procedia Computer Science, 15*, 203–209.
11. Serrano-Laguna, Á., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., & Fernández-Manjón, B. (2017, February). Applying standards to systematize learning analytics in serious games. *Computer Standards and Interfaces, 50*, 116–123.
12. Loh, C. S. (2012). Information trails: In-process assessment of game-based learning. In *Assessment in Game-Based Learning: Foundations, Innovations, and Perspectives* (Chap. 8, pp. 123–144). New York: Springer Science + Business Media.
13. Loh, C. S., Anantachai, A., Byun, J., & Lenox, J. (2007, July). Assessing what players learn in serious games: In situ data collected, information trails, and quantitative analysis. In *Proceedings of 10th International Conference for Computer Games: AI, Animation, Mobile, Education & Serious Games (CGAMES 2007)* (pp. 25–28).
14. Freire, M., del Blanco, Á., & Fernández-Manjón, B. (2014, April). Serious games as edX MOOC activities. In *Proceedings of Global Engineering Education Conference (EDUCON)* (pp. 867–871). IEEE.
15. Flavell, J. H. (1979, October). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*(10), 906–911.
16. Schraw, G. (1998, March). Promoting general metacognitive awareness. *Instructional Science, 26*, 113–125.
17. Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In *The Cambridge handbook of expertise and expert performance* (pp. 685–705).
18. Goldstein, D., Hahn, C. S., Hasher, L., Wiprzycka, U. J., & Zelazo, P. D. (2018, January). Time of day, intellectual performance, and behavioral problems in morning versus evening type adolescents: Is there a Synchrony effect? *Personality and Individual Differences, 42*(3), 431–440.
19. Collie, A., Maruff, P., Darby, D. G., & McStephen, M. (2003, March). The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test-retest intervals. *The Journal of the International Neuropsychological Society, 9*, 419–428.

Chapter 8
Learning Word Problem Solving Process in Primary School Students: An Attempt to Combine Serious Game and Polya's Problem Solving Model

Abdelhafid Chadli, Erwan Tranvouez and Fatima Bendella

Abstract Mathematics learning has become one of the most researched fields in education. Particularly, word or story problem solving skills have been gaining an enormous amount of attention from researchers and practitioners. Within this context, several studies have been done in order to analyze the impact that serious games have on learning processes and, in particular, on the development of word problem solving skills. However, little is known regarding how games may influence student acquisition of the process skills of problem solving. In a first attempt, this theoretical paper deals with word problem solving skill enhancement in second-grade school children by means of a practical educational serious game that addresses general and specific abilities involved in problem solving, focusing on how different parts of a solution effort relate to each other. The serious game is based on Polya's problem solving model. The emphasis of using the specific model was on dividing the problem solving procedure into stages and the concentration on the essential details of a problem solving process and the relationships between the various parts of the solution.

1 Introduction

Many researchers pointed that mathematics proficiency is important in personal independent thinking ability, competitive selection, and student reasoning abilities [1]. In traditional mathematics method teaching, teachers usually provide instruction of mathematics concepts by using abstract examples and words [2]. Therefore, students

A. Chadli
Computer Science Department, Ibn Khaldoun University Tiaret, Tiaret, Algeria
e-mail: abdelhafith.chadli@univ-tiaret.dz

E. Tranvouez
University of Aix-Marseille LIS-UMR CNRS, 7020 Marseille, France
e-mail: erwan.tranvouez@lis-lab.fr

F. Bendella (✉)
Computer Science Department, University USTO MB of Oran, Oran, Algeria
e-mail: fatima.bendella@univ-usto.dz

A. Tlili and M. Chang (eds.), *Data Analytics Approaches in Educational Games and Gamification Systems*, Smart Computing and Intelligence,
https://doi.org/10.1007/978-981-32-9335-9_8

encounter many difficulties in acquiring what is taught, and consequently, this causes them to memorize most mathematical concepts without understanding [3]. Furthermore, the literature suggests that the traditional face-to-face teaching approach makes it difficult for teachers to afford individualized instruction for each student [4, 5].

Some students think that mathematics is a difficult and tedious lesson, so they are less interested in learning, resulting in turns in low mathematical problem solving ability. Particularly, in the process of problem solving (steps taken by student in order to solve mathematical problem), children often meet difficulties; in particular, mathematical word problems make students feel that they are too abstract and vague. Despite the growing attention directed at problem solving skills (the student's abilities related to problem solving. See Fig. 9), teachers often have trouble in teaching students how to approach problems and how to make use of proper mathematical tools [6]. The difficulties stem in part from the fact that the teaching methods are inadequate and limited, and the difficulty is greater for elementary school teachers who are not subject matter (mathematics) teachers. Thus, many students struggle with arithmetical problem solving with negative effects on their motivational attitude to this learning area. However, besides that, many studies showed that some types of problems are inherently difficult for students. For example, Stern and Lehrndorfer [7] reported that word problems depicting the comparison of quantities have been shown to be difficult for elementary school children in several studies.

There are two conditions for successful mathematical word problem solving; students not only have to demonstrate considerable comprehension ability in terms of language, but also need relevant knowledge on how different parts of a solution effort relate to each other. This shows that in solving word problems in mathematics, it is necessary, first, to understand the narrative and then to concentrate on the essential details of the problem solving process and the relationships between the various parts of the solution in order to solve problems successfully. Students with average and lower learning achievement may not be so because they are unfamiliar with computations, but for they have problems designing process skills of problem solving and relating different parts of the problem solution with each other, resulting in difficulties in solving problems. In order to improve children's mathematics abilities, many scholars and educators seek to use technology to serve as the medium to elevate learning interest of students.

Regarding mathematical education, many concepts are difficult to understand for students of primary education. This could be related to the high degree of complexity and the level of abstraction of such concepts as well as the students' limited experience [8]. As an alternative, the use of technology can make significant help to the teaching of mathematics [8, 9]. Many researchers pointed out that the students' difficulty of understanding scientific concepts can be confronted using serious games. For example, Klopfer and Yoon [10] demonstrated in their study that one of the high-order cognitive abilities that serious games can contribute to is developing problem solving skills.

Over the last years, there has been greater research on the influence of computer games on mathematics instruction. People who play computer games are often engaged in problem solving and task performance [11]. A considerable number of researchers have shown interest in using game characteristics (e.g., challenges, points, rewards, surprising events, and competition) to enhance learning [12–14]. Up to now, a considerable amount of the literature published over the last two decades shows that playing video games can improve a wide variety of perceptual and cognitive processes [15].

Recently, a number of studies have been conducted regarding the effectiveness of game-based learning in mathematics skill enhancement [16–19]. However, all researches dealt with understanding of concepts of numbers and numbering or understanding the meanings of word problems in mathematics in terms of language. Most importantly, game-based learning studies often fail to use theoretical foundations [20]. Given the lack of consistent empirical evidence with respect to the incorporation of learning theories into the design of game-based learning, this paper aims to present a serious game that can help students in overcoming problem solving difficulties due to poor comprehension. Thus, students can understand semantics in word problems in mathematics with a major focus on the problem solving process based on Polya's strategy [21] (a four-step action plan designed by Polya to solve problems).

According to the experts in Algerian primary education, the expected second-grade math learning objectives that are associated with addition and subtraction are summarized in the following points:

- Model addition and subtraction with place value.
- Recall addition and subtraction facts.
- Use different methods to develop fluency in adding and subtracting multi-digit numbers.
- Add and subtract whole numbers to 1000.
- Solve multi-digit addition and subtraction problems.
- Use mental math strategies to add and subtract.

Furthermore, in the field of cognitive psychology, one common approach divides the types of thinking into problem solving and reasoning. This idea motivated us to align the serious game with one-step additive word problem solving because it is essential for the development of student's thinking as stated by Hepworth et al. [22].

The research question addressed in this theoretical paper was how mathematical word-based problem solving process could be improved by using a serious game that incorporates Polya's problem solving strategy?

We organize this paper as follows. First, we describe Polya's model and its using scope in game-based learning systems as well as the evolvement of serious games in word problem achievement. Second, we present the system design and describe the game scenarios of each problem solving stage. In the third part, we describe the analytics of student in-game assessment. We finally conclude by assessing the adopted strategy and identify future work directions.

2 Review of the Literature

Problems during mathematical problem solving are often caused by students' inabilities to be mindful of their thinking processes, especially with word problems [23]. Diverse researches have pointed out that a majority of students are not skilled to find out the required information in a problem and then transform it into mathematical concepts, plans, representations, and appropriate procedures [24–26]. Moreover, students have trouble combining concepts into mathematical information in a significant way and transferring the conceptual aspects into actual solutions to the problems [27]. In addition, they have less ability to control their thinking solving process to solve problems, particularly in relating different parts of word problem solution. This deficit gives the impression to students that mathematical word problem solving skills are difficult to master and they do not have the competence to solve them [28]. Polya's problem solving strategy may be an essential tool to overcome this deficit.

Due to these issues, solving word problems is difficult, particularly taking into account the student's inability to control their thought processes and apply concepts and procedures to the problems. To remedy this situation, students need to elaborate and develop the skills and practices that will give meaning to the problems and let them interpret the problems cautiously. Furthermore, to master the problem solving process skills, students must acquire knowledge of specific problem solving strategies [25, 29].

2.1 Polya's Problem Solving Model

Polya's four-step process has provided a model for the teaching and assessing problem solving in mathematics classrooms: understanding the problem, devising a plan, carrying out the plan, and looking back.

Understanding the problem. Students must understand the meaning of a problem statement; identify the known, the unknown, and the relationship between them; and recognize all concepts that are needed for solving the problem.

Devising a plan. Students must clarify the relations between parts of the problem, combine previously learned knowledge to develop thoughts for solving a problem, and develop a plan.

Carrying out the plan. Following the planned path, students execute all calculations previously identified in the plan.

Looking back. Students examine the solution and carefully revise the course that they refer to in an attempt to see if other problem solving paths exist.

The emphasis of using this specific model was on dividing the problem solving procedure into stages and the concentration on the skill details of each stage. In order to help students cope with difficulties encountered in solving problems, several researchers have developed game-based learning tools based on Polya's problem solving strategy. As students typically find games engaging, many academics developed narrative-centered learning games that embed problem solving interactions. For example, Lester et al. [30] designed a game-based learning environment called "Crystal Island" with the dual goals of improving students' problem solving effectiveness and their problem solving efficiency. The problem solving process was designed according to Polya's problem solving strategy. The results indicated that students have had significant gains from pretest to posttest on a four-part Polya process ordering question designed to probe at problem solving skill acquisition. Similarly, Ortiz-Rojas et al. [31] analyzed the impact of gamification, using badges, on learning performance. Students involved in a gamified computer programming course were supposed to attain higher intrinsic motivation, self-efficacy, and engagement as compared to students in a control condition. The authors adapted the four-problem solving techniques proposed by Polya to a programming context: Identify the input/output variables, select or create algorithms that are needed to solve the problem, write syntactically correct code lines, and think of an alternative way besides the one already used, to write the program, using fewer lines. Research outcomes showed a significant differential impact of studying in the gamified condition in terms of engagement. Similarly, Karatas and Baka [32] stated that students' gaining of problem solving skill in school mathematics is closely related to the learning environment to be formed. They focused on helping students to develop their problem solving skills and achievement in mathematics through a learning activity designed by Polya's heuristic phases of the problem solving approach. The study revealed that while the experimental group students' achievements of problem solving increased, control group students' achievements on problem solving have not changed significantly. Recently, Yang et al. [33] examined how to foster pupils' mathematical communication abilities by using tablet PCs to support students' math creations (including mathematical representation, solution, and solution explanation of word problems) and reciprocal peer-tutoring activities. The game activity flow involved four steps: understanding the problem, drawing a representation, writing a solution, and explaining the solution. These steps were designed according to Polya's findings [29] about problem solving. The results showed that students' mathematical representations and solution explanations became more accurate after the learning activity. Combination of Polya's strategy with schema representation has also been studied [34]. In this study, the authors developed a computer-assisted problem solving system that involves the operations of addition and subtraction and focused on the stages that are problematic for students. The system was empirically demonstrated to be effective in improving the performance of students with lesser problem solving capabilities.

These previous experiments show that game-based learning designed according to Polya's problem solving strategy has a positive impact on children's problem solving abilities. Therefore, it is necessary to incorporate learning theories into the design of serious games which is suitable for the development of students' problem solving skill.

2.2 Serious Games and Word Problem Achievement

A number of issues can affect the success of students at primary education level including well-being, teacher quality, levels of poverty, and parental support. If a student fails to grasp mathematics basics at primary education level, then secondary level becomes all the more difficult. Wilson et al. [35] highlight the need for novel teaching approaches including the use of computing technology and computer games to promote engagement in primary education and thus reduce difficulties later on.

Game-based learning and problem solving skills have been gaining an enormous amount of attention from researchers. Given numerous studies support the positive effects of games on learning, a growing number of researchers are committed to developing educational games to promote students' problem solving skill development in schools. To date, a number of researches have been conducted regarding the efficiency of game-based learning in several domains such as math, business, computer science, psychology, and biology [36–38]. However, no agreement has been reached regarding the positive effect of game-based learning. For example, some studies [36, 39] pointed out that game-based learning might be better than traditional classroom instruction as it could enhance students' motivation for learning and provide them with opportunities to acquire new knowledge and skills, whereas others [40] did not find strong evidence which supports the association between game-based learning and students' high academic achievements and have raised questions regarding the methodology of studies that observe transfer [41, 42].

Michael and Chen [43] defined serious games as "games that do not have entertainment, enjoyment, or fun as their primary purpose." However, Abt [44] noted that "this does not mean that serious games are not or should not be entertaining." One of the key strengths of serious games is that they can allow students to observe, explore, recreate, manipulate variables, and receive immediate feedback about objects and events that would be too time-consuming, costly, or dangerous to experience first-hand during traditional school science lessons. In a recent systematic review of empirical evidence on serious games, Connolly et al. [37] identified 129 reports on the impact on learning and engagement. The most frequently found outcomes were related to knowledge acquisition, content understanding, and motivation.

Several studies have been done in order to analyze the impact that serious games have on learning processes and, in particular, on problem solving skills [10], as well as on areas within the school curriculum such as language, science, and mathematics [45]. For instance, Chen et al. [19] designed a digital game-based learning system for multiplication and division in basic mathematics based on iconic representation animation. The results showed that game-based instructional materials could increase students' learning achievements, better than traditional instruction with significant benefits. Similarly, Eseryel et al. [46] examined the complex interplay between learners' motivation, engagement, and complex problem solving outcomes during game-based learning. Findings of this study suggested that learners' motivation determines their engagement during gameplay, which in turn determines their development of complex problem solving competencies. They also pointed out that learner's motivation, engagement, and problem solving performance are greatly impacted by the nature and the design of game tasks. Development of serious games for mobile devices has also been studied [44, 47–49]. For example, Sanchez et al. [50] developed a problem solving collaborative game for eighth-grade science classes' curriculum. A high degree of user satisfaction with the final product was found, and results indicate that the experience is contributed to the development of the student's problem solving skills obtaining positive gains as an outcome of this experience. Lester et al. [30] presented the design of the "Crystal Island" learning environment and described its evolution through a series of pilots and field tests. Results indicated that "Crystal Island" produced significant learning gains on both science content and problem solving measures.

The use of serious games may enhance the learning of word problem solving process. Yet, the enhancement is restricted to only exploring the learning experience states. For instance, according to Liu et al. [51] in various studies such as [52–54], students' responses to such problem solving experiences were positive, and the serious games were shown to be able to motivate students to solve problems. However, little is known regarding how games may influence student acquisition of the process skills of problem solving. Thus, there is an imperative need to have a better understanding of the impacts of games on problem solving strategies. To this end, we have conducted a theoretical study to enhance the influence of serious games on problem solving in terms of word problem solving process enhancement. The aim of this study is to adopt a serious game to inculcate students how different parts of a solution effort relate to each other following the solving process proposed by Polya.

2.3 Probabilistic Assessment in Game-Based Learning

Learner's knowledge or gained knowledge through the game is assessed by various game-based learning environments [11, 55]. However, elaborating effective models of student knowledge acquisition in game-based learning environments presents significant computational challenges. First, student knowledge models must deal with the student's reviewable reasoning when attempting to solve problem. Furthermore, knowledge models must dynamically model knowledge states that change over the course of a narrative interaction. To address these challenges, many scholars developed probabilistic approaches to modeling user knowledge during interactive narrative. For example, Shute [56] uses Bayesian inference networks to carry out continuous, unobtrusive assessment in games and indicate the strength of evidence. Kickmeier-Rust et al. [57] proposed a probabilistic assessment on the basis of interpreting the learner's behavior and actions within the game in a noninvasive way. Additionally, Lester et al. [58] developed a dynamic Bayesian network approach to modeling user knowledge during interactive narrative experiences. An initial version of the model has been implemented in Crystal Island game. Conlan et al. [59] defined knowledge skills in the game for each task. Learner knowledge skills are assessed using a probabilistic embedded assessment. In the last decade, learner emotions and motivation have been subjects of increasing attention. Learner emotions such as joy or distress toward the game, admiration, or reproach toward a helping agent were assessed using a probabilistic method using learner interaction with the game and questionnaire in Conati and Maclaren [60]. The learner's achievement emotions such as anticipatory joy, hope, anxiety, anticipatory relief, and hopelessness were also detected and assessed using a probabilistic model [61].

Summary of the review above reveals that using probabilistic assessment approaches in game-based learning environments may lead to accurate knowledge state modeling.

3 System Design

Children with average and low mathematics achievement had basic learning ability and need remedial instruction. They require more practice, time, and scaffolding compared to students at other achievement levels to monitor and control their learning processes. For example, Krutetskii [62] believes that problem solving processes of students with average or low mathematics aptitude differ from those of students with high mathematics aptitude. Therefore, we choose subjects of average and weak ability. Furthermore, Ketelhut et al. [63] found that after using *River city* game, student inquiry learning was enhanced and low science achievers did nearly as well as high science achievers with appropriate pedagogical strategy being employed.

Using Polya's problem solving model, the proposed serious game is designed to guide average and low-achieving second-grade students through the parts of the problem solving process that they often ignore or fail to understand by providing steps to identify what is given and what is requested in the problem or how to organize the solving plan. We chose the second-grade level because students at this point supposedly have robust knowledge of these types of problems [64], as well as sufficient technology mature to elude rapidly any game interaction issues.

3.1 Student Profile

According to many scholars, there are two main procedural steps in problem solving: (i) transforming the problem into mathematical sentences and (ii) computation of the operation involved in the mathematical sentences. These two main procedural steps involve several other mathematical sub-skills that students must know and which lack of one of them might result in difficulties and confusion in the process of problem solving. Our serious game targets students with the following difficulties: incomplete mastery of number facts, weakness in computation, inability to connect conceptual aspects of math, difficulty to make meaningful connection among information, incompetency to transform information mathematically, incomplete mastery of mathematical terms, and incomplete understanding of mathematical language.

3.2 Game Characteristics

Academics proposed many essential game characteristics [11]. Malone [65] suggested that the four central game characteristics are challenge, fantasy, complexity, and control. Prensky [66] used the expression "game feature" and observed six key structural game features: rules; goals; outcomes and feedback; challenge, competition, and opposition; interaction; and representation or story [67]. Dynamic visuals, interaction, rules, and goals were also designated as structural parts of a game [68]. In this paper, we focused on four game characteristics: story line, challenge, immediate rewards, and integration of gameplay with learning content. A story line refers to the problem solving scenario, which describes the steps that a player needs to follow to solve successfully word problems. A challenge is defined as a mission requiring math knowledge and skills in word-based addition and subtraction questions. Immediate rewards consist in points received by a player as soon as each challenge is successfully accomplished. Thus, this current study uses a set of selected game characteristics in a word problem solving learning context to bring a positive impact on word-based addition and subtraction question performance.

3.3 Description of the Serious Game

"Tamarin" is a serious game designed to support second-grade mathematics teaching and learning activities. It was designed using sophisticated 3D graphics from the "Unity 3D" platform, where the design of some elements such as reward, actions, logic, and resources is aligned with the real-time strategy of commercial games, since this type of games is potentially a tool for developing problem solving skills [13].

This instructional software is a dynamic creation and investigation tool that enables students to explore and understand solving process of mathematical word problems. "Tamarin" was created through consideration of Polya's solving model, which helps students gain positive attitudes toward mathematics. The game starts with an enjoyable animation story, and the overall design of the software is oriented around general problem solving strategies, with interactive exercises about mathematical problems and solutions based on "real-life" action activities. Tamarin game is suitable for average and gifted students, as well as for those having difficulty with mathematics. In all stages of the solving process, there are playful activities designed so that students can discover learning principles to enhance their word problem solving skills. "Tamarin" offers four types of problems based on the classification of Vergnaud [69]: (1) put together; (2) change-get-more; (3) change-get-less; and (4) compare (see Table 1). Feedback is included at the end of each problem solving scenario, which helps students to evaluate their own knowledge learning. Furthermore, score keeping and useful information about student performance are stored for future retrieval.

Table 1 Vergnaud's classification of word problems

Addition situations	Subtraction situations
Change-get-more Missing end Ali had three marbles. Then, Omar gave him five more marbles. How many marbles does Ali have now? Missing change Ali had three marbles. Then, Omar gave him some more marbles. Now Ali has eight marbles. How many marbles did Omar give him? Missing start Ali had some marbles. Then, Omar gave him five more marbles. Now Ali has eight marbles. How many marbles did Ali have in the beginning?	**Change-get-less** Missing end Ali had eight marbles. Then, he gave five marbles to Omar. How many marbles does Ali have now? Missing change Ali had eight marbles. Then, he gave some marbles to Omar. Now Ali has three marbles. How many marbles did he give to Omar? Missing start Ali had some marbles. Then, he gave five marbles to Omar. Now Ali has three marbles. How many marbles did Ali have in the beginning?
Put together Missing all Ali has three marbles. Omar has five marbles. How many marbles do they have altogether? Missing first part Ali and Omar have eight marbles altogether. Omar has three marbles. How many marbles does Ali have? Missing second part Ali and Omar have eight marbles altogether. Ali has three marbles. How many marbles does Omar have?	**Compare** Missing difference Ali has eight marbles. Omar has five marbles. How many more marbles does Ali have than Omar? Missing big Ali has three marbles. Omar has five more marbles than Ali. How many marbles does Omar have? Missing small Ali has eight marbles. He has five more marbles than Omar. How many marbles does Omar have?

The game consists of two types of playing environments, the "virtual supermarket" and four "science fiction computation rooms." The player is characterized in the game by an avatar selected beforehand. Before entering into the supermarket, the system displays the problem statement to the player (see Fig. 1) among other problems stored in the problem solving information database. Each problem is designed to ask the player to buy food. The player serves himself directly on the shelves to take what he needs according to the needs of the problem (see Fig. 2). The problem missing part may either be requested as the first operand and the second operand, or be the result of the mathematical expression. This part of the game was developed in order to present the students with a real-life situation by stimulating them to discover the knowledge of problem solving.

Problem statement: Your mother asks you to buy a cheapest bag of sugar and a cheapest bag of coffee too. What is the total amount of expenses?

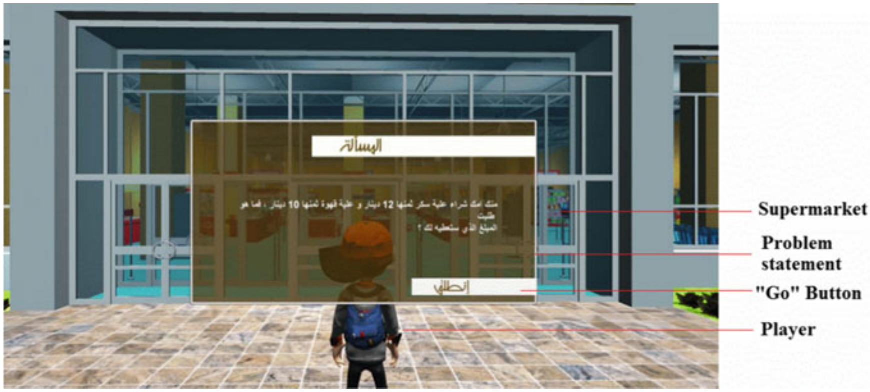


Fig. 1 Display of the problem statement

Fig. 2 Looking for asked products

3.4 The Problem Solving Process

Students' problem solving guidance process is portrayed in Fig. 3. First, the adequate problem is provided from problem solving data storage among other problems. After buying food in the supermarket, the game proposes a fun activity, in room 1, to rule on player's understanding of the problem. Room 2 enables the student to build a mathematical equation that represents the solution plan. Room 3 suggests, according to the operation type, a calculation interface that provides information about student's procedural skills. Finally, in room 4, a questionnaire form is provided to student to validate his solution. Problem information is provided at each stage of the problem-solving process, and assessment of each stage is recorded in the student-tracking database. The system displays feedback messages subsequent to problem solving completion.

Room 1: Understanding the problem stage. In this room, firstly, three arrow targets are presented to the player, and each one describes a goal of a problem, but only one

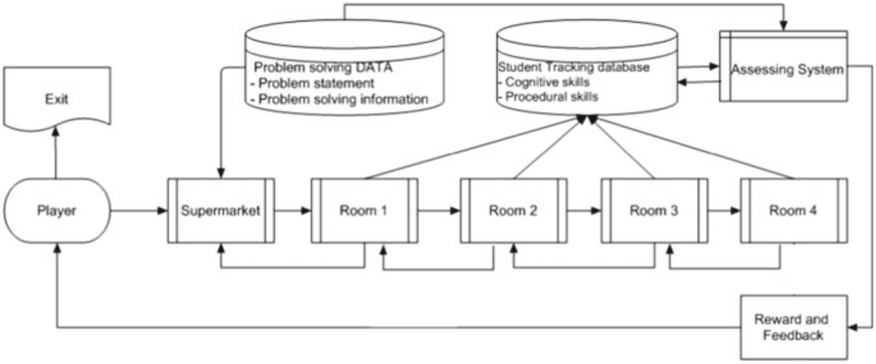


Fig. 3 Students' problem solving guidance flowchart

goal is correct as presented in Fig. 4 (i.e., goals presented for this problem are: 1—calculation of the total amount to be paid, 2—calculation of the rest amount to be paid, and 3—calculation of coffee price). The player has to shoot with an arrow the right target, if the correct goal is selected then he moves to the second step of the same stage, and otherwise a new simplified version of the problem will be displayed to the player. If one of the following attempts is successful, then the final score will consider the number of attempts. Secondly, the player has to find out the context of the problem by playing basketball as depicted in Fig. 5; the context of the problem describes the nature of the question (addition or subtraction). Moreover, students have to distinguish between what is known and what is requested in the problem by selecting appropriate responses. At the end of this stage, the door to the next room will open.

Room 2: Making a plan. Once the context is determined, students are asked, in this stage, to identify the two operands of the mathematical equation and their values among a list of operands displayed on screen. The player can scroll through this list of operands and choose the required ones (see Fig. 6). Finally, during the last step of

Fig. 4 Screenshot of identifying the problem goal (understanding of problem)

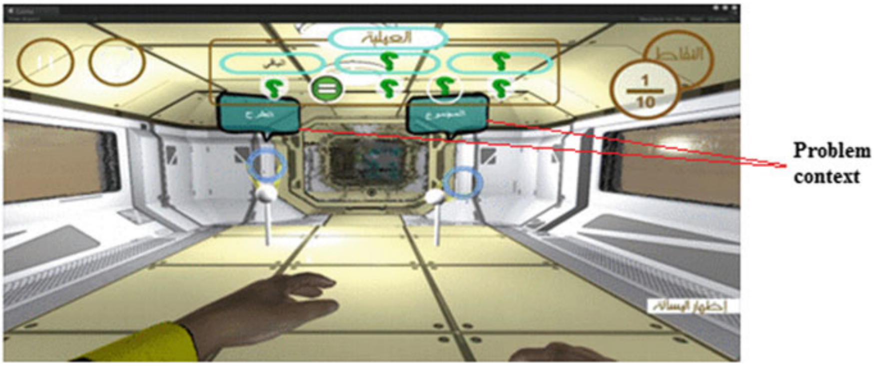


Fig. 5 Screenshot of identifying the problem context (understanding of problem)

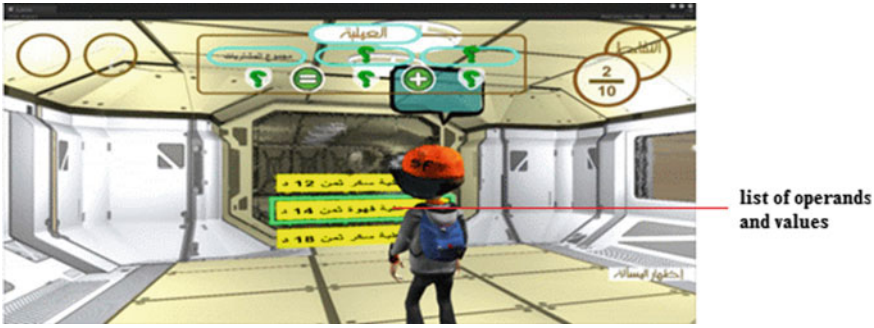


Fig. 6 Screenshot of identifying the operands and their values (making a plan)

this stage, all that is required is a result label. According to the problem missing part, a result label may either be requested as first or second operand of the equation. As the player progresses through the steps of this stage, the question marks in the equation are replaced by the selected values. Finally, the designed plan corresponds to the mathematical equation shown in the top of screen with a question mark indicating the missing part value. After the student has finished the plan design, the system compares the solution plan created by the student with the one built into the system. The serious game then produces suggestions regarding the student's problem solving, stores them in the student-tracking database, and displays after the student completes the problem.

Room 3: Executing the plan. At this stage, the game evaluates the player's procedural knowledge. As shown in Fig. 7, the game provides a graphical preview of all the addition and subtraction worksheets in a vertical problem format. The player must provide the values of both operands with a ball gun. The number of balls introduced into the holes represents, from right to left, the units and tens of the operands and result. Large holes are reserved for digits of operands, and little holes are used if

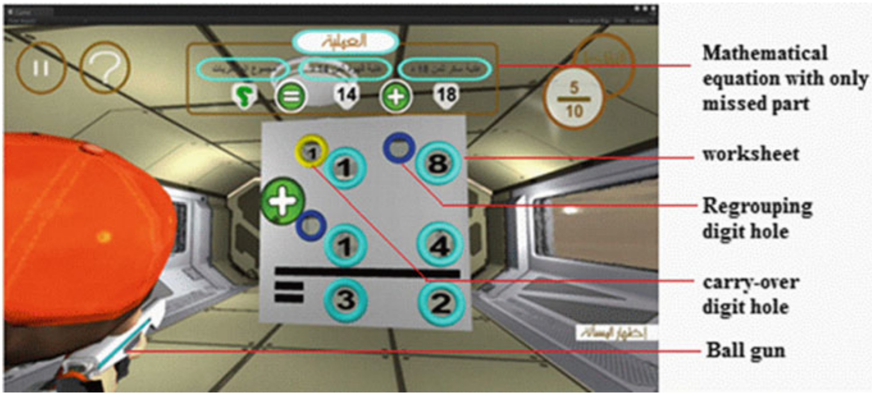


Fig. 7 Screenshot of calculation (executing the plan)

regrouping is required for subtraction (exchanging one of the tens for 10 units or one of the hundreds for 10 tens) or for addition when the process involves a carryover number. All the calculation is done in a fun way to relieve the stress of the student. After finishing the calculations of all the columns, the game assesses the student's answer and feedback is stored. A last stage is needed to validate the student's problem solution.

Room 4: Reviewing the solution. Finally, during the solution's reviewing stage, we will determine how precisely and clearly the students recognize the variants of the solved problem expressed (put together, change-get-more, change-get-less, and compare) to verify their solutions.

During this stage, the player answers questions as depicted in Fig. 8. In order to validate the solution given in the previous stage, the game proposes questions that are related to the problem and the student must answer with true or false. After

Fig. 8 Screenshot of reviewing the solution

completing this stage, the student presses the evaluation button that triggers the game to evaluate the results, and messages appear to indicate whether any mistakes were made. Additionally, the correct problem solving steps are displayed simultaneously next to the student's answers.

4 Game-Related Metrics

In order to better identify and to characterize the skills to be assessed within one-step mathematical word problem solving domain of competence, we propose to adopt a goal-oriented analysis to produce the competency model related to word problem solving domain. This analysis allows first to identify high-level skills and then after to classify sub-skills hierarchically related to these high-level skills. For each skill in the model, the goal-oriented analysis allows to detect all sub-skills that contribute positively to its accomplishment (see Fig. 9), until reaching the actions to be performed by the player via interaction with the serious game. To assess students' problem solving process ability, we consider the four main skills: understanding the problem, making the plan of resolution, executing the plan, and reviewing the solution. These competencies are considered as main skills that student must master to solve a mathematical word-based problem. Then, other skills are derived from the main ones until obtaining actions, reflecting identified skills, that student can carry out via the serious game.

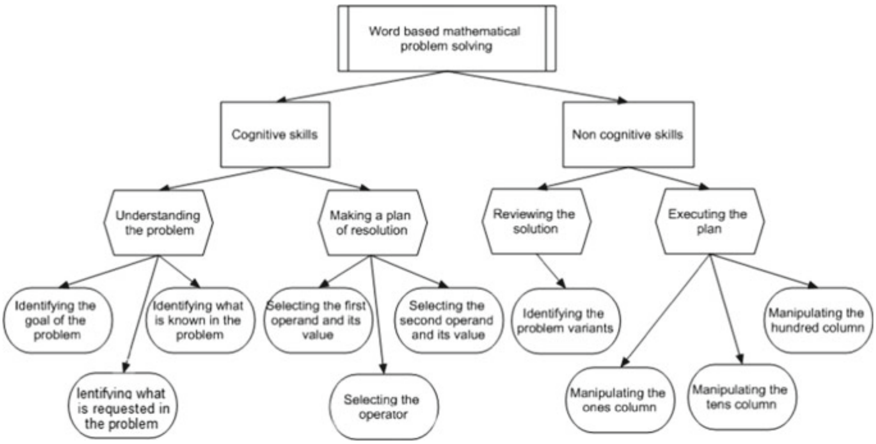


Fig. 9 Illustration of the competency model for mathematical word-based problem solving

4.1 Knowledge Representation

Probabilistic logic is a logical system whose truth values can range over the set of real between 0 and 1. The truth value associated with a formula represents the probability that this formula is true. The main advantage of this formalism, besides those inherited from logical systems, is that it makes it possible to represent the notion of uncertainty. In our case, uncertainty represents lack of information about student knowledge state.

The approach that we adopt to assess student's problem solving skills is based on a probabilistic logic system. It consists in comparing the solution of the student with that of the expert. The analysis of the possible differences between these two solutions makes it possible to identify the knowledge entities that were used by the expert and by the student and those that were used by the expert but not by the student. The optimistic diagnosis may be initially preferred. The learner is then considered to have used the concepts mastered by the expert. If subsequent interaction proves that the student does not master some of these concepts, the previous diagnosis must be changed. In this system, the expert knowledge is divided into entities. Each entity is associated with the numbers of times the student used this knowledge entity or not in each context [70]. This approach has been used in several intelligent tutor systems such as integration [70], ET [71] and FBM [72].

The reason we have chosen this formalism to represent student knowledge is that this formalism allows us to implement the student's reviewable reasoning. Using probabilistic logic, student knowledge is set up as a set of skills. Each element of this set denotes an association between a skill displayed by the student and a confidence factor for that skill. This confidence factor corresponds to the probability of belief that the student owns this competence. The calculation of this factor is given by the probability theory which states that the sum of the probabilities associated with the contradictory propositions of the language is equal to 1. In our case, a proposition denotes the association by the student of a concept used by the expert to a type of question. Its contradiction is not to use this concept for this type of question. At this level, a representation of knowledge in terms of skill controlling and skill deficit is suitable [70, 73]. The confidence factor is thus calculated by dividing the weight associated with the student's answer to the type of considered question (denoted by α) by the sum of the weights associated with the other answers already given to the same type of question as shown in Formula (1).

$$p_i^t = \frac{\alpha_i^t}{\sum_{k=1}^{n} \alpha_k^t} \tag{1}$$

where n is the number of answers already given to the same type of question and p_i^t the probability associated with the skill at time t. The assessment given this way is based on the psycho-cognitive hypothesis that states the more recent an answer is, the more it reflects the current cognitive state of the learner.

4.2 Updating Beliefs About Knowledge

Low-Level Skill Probability Adjusting. Beliefs about student's knowledge are updated after the analysis of each activity in the game. The modification consists in changing the confidence factor. This update is only done on the part of knowledge that corresponds to the last activity of the player in the serious game. The body of knowledge is thus divided into several competences, according to the activities of the game to which the knowledge refers. The revision to be made to the skill's probability, which is concerned by the activity of the player, is carried out in two phases according to the principle used in FBM [72].

Firstly, a devaluation is performed for all the weights associated with the answers already given by the student with respect to the considered skill as depicted in Formula (2):

$$\mathrm{dev}(\alpha_i) = \alpha_i * \tau \tag{2}$$

This decrease in the factors that are associated with the old answers implements the notion of temporal relativization. In our case, we will take into account only two types of student's answers (the one that corresponds to the expert's answer and the one that does not correspond to the expert's answer). In FBM, the parameter τ of the data dating system is set to 0.9. We resume this valuation. The weights associated with all answers are initialized to 0.

Secondly, the weight of the last answer corresponding to the student activity in the game for the targeted skill must be increased. For this purpose, we use a reinforcement function that is used in FBM [72] as presented by Formulas (3) and (4).

$$\mathrm{reinf}(x) = x + 1 \tag{3}$$

We then obtain:

$$\alpha_i^{t+1} = \mathrm{re\,inf}\left(\mathrm{dev}\left(\alpha_i^t\right)\right) = \mathrm{dev}\left(\alpha_i^t\right) + 1 \tag{4}$$

$$p_i^{t+1} = \frac{\alpha_i^{t+1}}{\sum_{k=1}^{n} \alpha_k^{t+1}} \tag{5}$$

In summary, the update of the confidence factors associated with answers reflecting a skill type follows the following algorithm: Let "identifying what is known in the problem" be a skill. The activity of the player in the game can be interpreted as two types of answer, a correct answer (equivalent to that of the expert) or a wrong answer (different from that of the expert). Let coeff_{R1} and coeff_{R2} be the weights of, respectively, the correct answer and the wrong answer. In addition, let $\mathrm{prob}_{\mathrm{exp}}$ and $\mathrm{prob}_{\mathrm{Nexp}}$ be, respectively, a probability that the player will give the same answer as that of the expert and a probability that the player will give a different answer than that of the expert.

Algorithm confidence factors update:

Update_belief_coeff (activity player_answer, expert_answer)
begin
If (Player_answer is_same_as Expert_ answer)
Then
$coeff_{R2} \leftarrow dev(coeff_{R2})$
$coeff_{R1} \leftarrow reinf(dev(coeff_{R1}))$
Else
$coeff_{R1} \leftarrow dev(coeff_{R1})$
$coeff_{R2} \leftarrow reinf(dev(coeff_{R2}))$
endif
$prob_{exp} = \frac{coeff_{R1}}{coeff_{R1}+coeff_{R2}}$
$prob_{Nexp} = \frac{coeff_{R2}}{coeff_{R1}+coeff_{R2}}$
end

High-Level Skill Probability Adjusting. An additional high-level model summarizes only the information on probabilities of mastery and lack of skills related to the problem solving process without details of the player's answer history. This model reports on high-level skills of word problem solving process such as understanding the problem, making a plan, executing the plan, and reviewing the solution. We point out that each of these skills consists of other sub-skills (see Fig. 9) that are associated with confidence factor probabilities and weights that indicate contribution rate to high-level skill achievement as depicted in Fig. 10.

We adopt Bayesian networks to calculate high-level skill probabilities, as they are very convenient for representing systems of probabilistic causal relationships. As such, they have remarkable properties that make them better than many traditional methods in determining the effects of many variables on an outcome. Furthermore, using Bayes' theorem, we can calculate the probability of the effects from the observation of the causes. In our case, in view of the example depicted in Fig. 10, the effects are the high-level skill "making a plan of resolution" and causes are sub-skills that make up this high-level skill. The fact sub-skills supply high-level skills may easily be modeled in the network by adding a directed arc from sub-skills to high-level skills and setting the probabilities appropriately. For example, for the high-level skill "making a plan of resolution" there are three sub-skills, namely 1—selecting the first

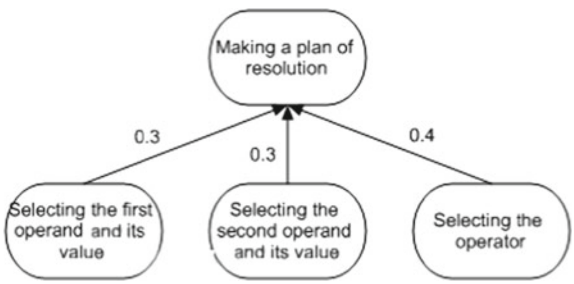


Fig. 10 Sub-skill contribution rates to high-level skill achievement

Table 2 Skill probabilities to consider in Bayesian network

Probability	Meaning
$P\ (M)$	Conditional probability: the probability of the high-level skill (to calculate)
$P\ (F)$	The probability of the "selecting the first operand and its value" sub-skill
$P\ (S)$	The probability of the "selecting the second operand and its value" sub-skill
$P\ (O)$	The probability of the "selecting the operator" sub-skill
$P\ (X \mid Y)$	Conditional probability: X to occur if Y occurs

operand and its value, 2—selecting the second operand and its value, and 3—selecting the operator. The probability of controlling this high-level skill depends on the conditional probabilities. Table 2 presents all probabilities we need to calculate the high-level skill probability, and Formula (6) gives required calculation.

We then obtain:

$$p(M) = \sum_{i=1}^{n} p\left(\frac{B}{A_i}\right) p(A_i) = p(M/F)p(F) + p(M/S)p(S) + p(M/O)p(O) \quad (6)$$

4.3 Problem Solving Competency Model

The skill descriptive model provides information to the teacher about the student's ability to solve a given type of problem. The competency is divided into two main categories (cognitive and non-cognitive); each of these categories is then divided into several skill subclasses until reaching the levels that correspond to the activities to be undertaken by the player in the game. Some skills are evaluated according to the solved problem type (i.e., put together, change-get-more, change-get-less, and compare). The competency model indicates the student's probability of skill controlling in terms of providing similar answers to those of the expert as well as the probability of skill deficiency in terms of providing different answers from those of the expert. Moreover, an additional high-level model allows the teacher to have a general idea about the player's solving skills and difficulties with respect to each step of the process (i.e., understanding the problem, making a plan, executing the plan, and reviewing the solution) (see Table 3).

Table 3 Competency model of skill controlling probability evolution

Skill_ID	Type_of_problem	
Time	Probability of correct answer	Probability of wrong answer
t_0 (initialization)	**0**	0
t_1 (correct answer)	$\frac{\text{reinf(dev(0))}}{\text{reinf(dev(0))+dev(0)}} = \frac{1}{1} = 1$	$\frac{\text{(dev(0))}}{\text{reinf(dev(0))+dev(0)}} = \frac{0}{1} = 0$
t_2 (wrong answer)	$\frac{\text{dev(1)}}{\text{dev(1)+reinf(dev(0))}} = \frac{0.9}{1.9} = 0.47$	$\frac{\text{reinf(dev(0))}}{\text{dev(1)+reinf(dev(0))}} = \frac{1}{1.9} = 0.53$
t_3 (correct answer)	$\frac{\text{reinf(dev(0.47))}}{\text{reinf(dev(0.47))+dev(0.53)}}$ $= \frac{1.423}{1.9}$ $= 0.748$	$\frac{\text{(dev(0.53))}}{\text{reinf(dev(0.47))+dev(0.53)}} = \frac{0.477}{1.9} =$ 0.251

5 Conclusion

Educators have highlighted the importance of problem solving competence. Consequently, many approaches have been proposed to enhance such competence. This paper proposes to combine Polya's problem solving model with a serious game to assist students in developing their problem solving process mastery. Such serious games address the features of problem solving through the simulation of embodied experience.

Integrating games into education is not easy to achieve. There is an attempt to integrate serious games on an ad hoc learning methodology to develop and improve problem solving skills. Therefore, adopting and evaluating a methodology to improve these skills are very important contributions to educational systems. Specifically, game designs that feature a blending of established learning theories with game design elements proven successful in the entertainment game industry are most likely to lead to effective learning [74].

In this paper, we have addressed the issue of whether students gain from a game-based learning using Polya's problem solving strategy by providing assistance at each stage to help average and low-achieving second-grade elementary students improve their abilities in solving basic word-based addition and subtraction questions, and enhance their willingness to continue. The emphasis when using this model was on dividing the problem solving procedure into stages so that students can better understand the semantics and context of mathematical word problems. Furthermore, stages at which errors occur when a student encounters difficulties may provide a valuable help for student support.

Another weighty argument in using Polya's model is that till now game-based learning studies often fail to use theoretical foundations [20, 38]. For example, Wu et al. [38] reviewed 567 published studies and found that game-based learning tended to yield positive outcomes when learning theories were incorporated into the design, but surprisingly most studies did not address learning theories. However, Qian and Clark [74] revealed in their study that 76% of 29 papers explicitly referenced at least one established learning theory in the research design or in the game design, with constructivism being the most popular one, and a variety of other learning theories. According to Young et al. [75], successful game-based learning is not simply providing students with a game and expecting increased motivation and knowledge acquisition. “Rather, educational games need to be designed and researched with careful attention to contemporary learning theories.”

Our work has been guided by an interest in developing word problem solving activities to include serious game with a logic that is closely matched to an attractive game on the market, integrating it with learning contents. Developing games for learning opens up new possibilities for understanding how teaching and learning practices are mediated by technology and under what conditions those practices actually improve learning. This gameplay can be beneficial to fostering the acquisition of mathematical competences, in particular word problem solving. This in turn requires suitable training and the presence of the teacher in the classroom to guide the activity, monitoring it, and to evaluate the students’ learning. We believe that this first experience could be improved and represent a promising future line of work. We also believe it is necessary to continue the study of learning-embedded games, meaning that a good performance of the game is possible when the contents are learned.

In addition, the ideas related to using low-level skill and high-level skill assessment, and competency model of skill controlling can help to reduce teachers’ workload in relation to supervising the students’ problem solving activity. This would allow teachers to spend more time in promoting student learning. If serious games were easy to employ and provided valuable embedded assessment tools, then teachers would more likely want to utilize them to support student learning throughout a scope of educationally valued skills. The ideas and tools within this paper are planned to help teachers facilitate learning, in a fun manner, of word problem solving valuable skills not currently supported in school.

Carrying out the experiments as well as result analysis will be the subject of our next paper. Specifically, we will compare mathematical word problem solving skills of students who will utilize the combination of the serious game with Polya’s model with students who employ the same serious game based on general strategy instruction. Our aim is to highlight the effect of the combination of serious games with Polya’s strategy on student problem solving achievement. Participants will be randomly assigned to treatment conditions.

The recommendation for future research on this topic is as follows: As advocated above, future empirical study should be carried out to investigate if the “Tamarin” serious game activities can contribute to the development of problem solving and the motivation for learning among primary education students. In addition, further research studies should be conducted to investigate teacher’s readiness, attitudes, and knowledge about serious games teaching in primary education.

Acknowledgements We gratefully thank Karim Khattou and Toufik Achir for their cooperation and contribution to the development of “Tamarin” serious game and for supplying data used in problem information database.

References

1. Roman, H. T. (2004). Why math is so important. *Tech Directions, 63*(10), 16–18.
2. Samuelsson, J. (2008). The impact of different teaching methods on students' arithmetic and self-regulated learning skill. *Educational Psychology in Practice, 24*(3), 237–250.
3. Cankoy, O., & Darbaz, S. (2010). Effect of a problem posing based problem solving instruction on understanding problem. *Hacettepe University Journal of Education, 38,* 11–24.
4. Su, C., & Cheng, C. (2013). 3D game-based learning system for improving learning achievement in software engineering curriculum. *Turkish Online Journal of Educational Technology, 12*(2), 1–12.
5. Chen, Z., Liao, C. C., Cheng, H. N., Yeh, C. Y., & Chan, T. (2012). Influence of game quests on pupils' enjoyment and goal pursuing in math learning. *Educational Technology & Society, 15*(2), 317–327.
6. Harskamp, E. G., & Suhre, C. J. M. (2006). Improving mathematical problem solving: A computerized approach. *Computers in Human Behavior, 22*(5), 801–815.
7. Stern, E., & Lehrndorfer, A. (1992). The role of situational context in solving word problems. *Cognitive Development, 7*(2), 259–268.
8. Flick, L., & Bell, R. (2000). Preparing tomorrow's science teachers to use technology: Guidelines for science educators. *Contemporary Issues in Technology and Teacher Education, 1*(1), 39–60.
9. Akpan, J. P. (2001). Issues associated with inserting computer simulations into biology instruction: A review of the literature. *Electronic Journal of Science Education, 5*(3), 1–32.
10. Klopfer, E., Yoon, S. (2005). Developing games and simulations for today and tomorrow's tech Savvy Youth. *Tech trends. Linking Research & Practice to Improve Learning, 49*(3), 33–41.
11. Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming, 33*(4), 441–467.
12. Wouters, P., van Oostendorp, H., ter Vrugte, J., Vandercruysse, S., de Jong, T., & Elen, J. (2017). The effect of surprising events in a serious game on learning mathematics. *British Journal of Educational Technology, 48*(3), 860–877.
13. Prensky, M. (2006). *Don't bother me mom, I'm learning!: How computer and digital games are preparing your kids for 21st century success and how you can help!*. St. Paul: Paragon House.
14. Shute, V. J., Rieber, L., & Van Eck, R. (2011). Games and learning. In R. Reiser & J. Dempsey (Eds.), *Trends and issues in instructional design and technology* (3rd ed., pp. 321–332). Upper Saddle River: Pearson Education.
15. Green, C. S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science, 18*(1), 88–94.
16. Brown, D. J., Ley, J., Evett, L., Standen, P. (2011). Can participating in games based learning improve mathematic skills in students with intellectual disabilities? In *1st International Conference on Serious Games and Applications for Health (SeGAH)* (pp. 1–9). IEEE.
17. Lee, J., Luchini, K., Michael, B., Norris, C., Soloway, E. (2004). More than just fun and games: Assessing the value of educational video games in the classroom. In *CHI'04 extended abstracts on Human factors in computing systems* (pp. 1375–1378). ACM.
18. Chen, M. P., Ren, H. Y. (2013). Designing a RPG game for learning of mathematic concepts. In *IIAI International Conference on Advanced Applied Informatics* (pp. 217–220). IEEE.
19. Chen, H. R., Liao, K. C., Chang, J. J. (2015). Design of digital game-based learning system for elementary mathematics problem solving. In *8th International Conference on Ubi-Media Computing (UMEDIA)* (pp. 303–307). IEEE.
20. Li, M. C., & Tsai, C. C. (2013). Game-based learning in science education: A review of relevant research. *Journal of Science Education and Technology, 22*(6), 877–898.
21. Polya, G. (1945). *How to solve it*. Princeton: Princeton University Press.
22. Hepworth, D. H., Rooney, R. H., Rooney, G. D., Strom-Gottfried, K. (2016). *Empowerment series: Direct social work practice: Theory and skills*. Nelson Education.

23. Abdullah, N., Halim, L., Zakaria, E. (2014). VStops: A thinking strategy and visual representation approach in mathematical word problem solving toward enhancing STEM literacy. *Eurasia Journal of Mathematics, Science & Technology Education*, *10*(3).
24. Berch, D. B., Mazzocco, M. M. M. (2007). Why is math so hard for some children? In *The nature and origins of mathematical learning difficulties and disabilities*. Maryland: Paul H.Brookes Publishing Co.
25. Johnson, A. (2010). *Teaching mathematics to culturally and linguistically diverse learners*. Boston: Pearson Education.
26. Montague, M., & Dietz, S. (2009). Evaluating the evidence base for cognitive strategy instruction and mathematical problem solving. *Exceptional Children, 75,* 285–302.
27. Thevenot, C., & Oakhill, J. (2008). A generalization of the representational change theory from insights to non-insight problems; the case of arithmetic's word problems. *Acta Psycologica, 129*(3), 315–324.
28. Hagit, Y., & Anat, Y. (2010). Learning using dynamic and static visualizations: Students' comprehension, prior knowledge and conceptual status of a biotechnological method. *Journal Research in Science Education, 40*(3), 375–402.
29. Polya, G. (1973). *How to solve it: A new aspect of mathematical method*. Princeton: Princeton University Press.
30. Lester, J. C., Spires, H. A., Nietfeld, J. L., Minogue, J., Mott, B. W., & Lobene, E. V. (2014). Designing game-based learning environments for elementary science education: A narrative-centered learning perspective. *Information Sciences, 264,* 4–18.
31. Ortiz-Rojas, M., Chiluiza, K., Valcke, M. (2017). Gamification in computer programming: Effects on learning, engagement, self-efficacy and intrinsic motivation. In *European Conference on Games Based Learning* (pp. 507–514). Academic Conferences International Limited.
32. Karatas, I., & Baki, A. (2017). The effect of learning environments based on problem solving on students' achievements of problem solving. *International Electronic Journal of Elementary Education, 5*(3), 249–268.
33. Yang, E. F., Chang, B., Cheng, H. N., & Chan, T. W. (2016). Improving pupils' mathematical communication abilities through computer-supported reciprocal peer tutoring. *Journal of Educational Technology & Society, 19*(3), 157–169.
34. Chadli, A., Tranvouez, E., Dahmani, Y., Bendella, F., & Belmabrouk, K. (2018). An empirical investigation into student's mathematical word-based problem-solving process: A computerized approach. *Journal of Computer Assisted Learning, 34*(6), 928–938.
35. Wilson, A., Hainey, T., & Connolly, T. M. (2013). Using Scratch with primary school children: An evaluation of games constructed to gauge understanding of programming concepts. *International Journal of Game-Based Learning, 3*(1), 93–109.
36. Boyle, E., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., et al. (2015). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers & Education, 94*(2), 178–192.
37. Connolly, T. C., Boyle, E. A., Hainey, T., Macarthur, E., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59,* 661–686.
38. Wu, W. H., Chiou, W. B., Kao, H. Y., Hu, C. H. A., & Huang, S. H. (2012). Re-exploring game-assisted learning research: The perspective of learning theoretical bases. *Computers & Education, 59*(4), 1153–1161.
39. Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of Educational Computing Research, 34*(3), 229–243.
40. Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica, 129,* 387–398.
41. Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science, 8*(4), 445–454.

42. Kristjansson, A. (2013). The case for causal influences of action videogame play upon vision and attention. *Attention, Perception, & Psychophysics, 75,* 667–672.
43. Michael, D. R., Chen, S. L. (2005). *Serious games: Games that educate, train, and inform*. Muska & Lipman/Premier-Trade.
44. Abt, C. C. (1987). *Serious games*. University Press of America.
45. Mayo, M. J. (2007). Games for science and engineering education. *Communications of the ACM, 50*(7), 31–35.
46. Eseryel, D., Law, V., Ifenthaler, D., Ge, X., & Miller, R. (2014). An Investigation of the Interrelationships between Motivation, engagement, and complex problem solving in game-based learning. *Educational Technology & Society, 17*(1), 42–53.
47. Sanchez, J., Salinas, A., Sáenz, M. (2007). Mobile game-based methodology for science learning. In *International conference on human-computer interaction* (pp. 322–331). Berlin, Heidelberg: Springer.
48. Huizenga, J., Admiraal, W., Akkerman, S., & Dam, G. T. (2009). Mobile game-based learning in secondary education: Engagement, motivation and learning in a mobile city game. *Journal of Computer Assisted learning, 25*(4), 332–344.
49. Schwabe, G., & Göth, C. (2005). Mobile learning with a mobile game: Design and motivational effects. *Journal of Computer Assisted learning, 21*(3), 204–216.
50. Sánchez, J., & Olivares, R. (2011). Problem solving and collaboration using mobile serious games. *Computers & Education, 57*(3), 1943–1952.
51. Liu, C. C., Cheng, Y. B., & Huang, C. W. (2011). The effect of simulation games on the learning of computational problem solving. *Computers & Education, 57*(3), 1907–1918.
52. Gestwicki, P. V. (2007). Computer games as motivation for design patterns. *ACM SIGCSE Bulletin, 39*(1), 233–237.
53. Bayliss, J. D. (2007). The effects of games in CS1-3. In: *Proceedings Of Microsoft Academic Days Conference on Game Development in Computer Science Education* (pp. 59–63).
54. Barnes, T., Richter, H., Powell, E., Chaffin, A., & Godwin, A. (2007). Game2Learn: Building CS1 learning games for retention. *ACM SIGCSE Bulletin, 39*(3), 121–125.
55. Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C. S., Estock, J. L., et al. (2009). Relationships between game attributes and learning outcomes. *Simulation & Gaming, 40*(2), 217–266.
56. Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. *Computer games and Instruction, 55*(2), 503–524.
57. Kickmeier-Rust, M. D., Hockemeyer, C., Albert, D., Augustin, T. (2008). Micro adaptive, non-invasive knowledge assessment in educational games. In *Second IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning* (pp. 135–137). IEEE.
58. Lester, J. C., Ha, E. Y., Lee, S. Y., Mott, B. W., Rowe, J. P., & Sabourin, J. L. (2013). Serious games get smart: Intelligent game-based learning environments. *AI Magazine, 34*(4), 31–45.
59. Conlan, O., Hampson, C., Peirce, N., Kickmeier-Rust, M. (2009). Realtime knowledge space skill assessment for personalized digital educational games. In: *Proceedings of 9th IEEE international conference on advanced learning technologies* (pp. 538–542). Riga, Latvia.
60. Conati, C., Maclaren, H. (2009). Modeling user affect from causes and effects. In: G.-J. Houben, G. McCalla, F. Pianesi, M. Zancanaro, (Eds.), *Proceedings of the 17th international conference on user modeling, adaptation, and personalization*. (LNCS, Vol. 5535, pp. 4–15). Berlin: Springer.
61. Muñoz, K., Kevitt, P. M., Lunney, T., Noguez, J., & Neri, L. (2011). An emotional student model for game-play adaptation. *Entertainment Computing, 2*(2), 133–141.
62. Krutetskii, V. A. (1969). Mathematical aptitudes. *Soviet Studies in the Psychology of Learning and Teaching mathematics, 2,* 113–128.
63. Ketelhut, D. J., Nelson, B. C., Clarke, J., & Dede, C. (2010). A multi-user virtual environment for building and assessing higher order inquiry skills in science. *British Journal of Educational Technology, 41*(1), 56–68.
64. Fuchs, L. S., & Fuchs, D. (2005). Enhancing mathematical problem solving for students with disabilities. *The Journal of Special Education, 39*(1), 45–57.

65. Malone, T. W. (2001). What makes computer games fun? *Byte, 6*(12), 258–277.
66. Prensky, M. (2001). *Digital game-based learning*. New York: McGraw-Hill.
67. Ke, F. (2011). A qualitative meta-analysis of computer games as learning tools. In *Gaming and simulations: Concepts, methodologies, tools and applications* (pp. 1619–1665). IGI Global.
68. de Felix, W., & Johnston, R. T. (1994). Learning from video games. *Computers in the Schools, 9*(2), 119–134.
69. Vergnaud, G. (1982). A classification of cognitive tasks and operations of thought involved in addition and subtraction problems. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 39–59). Hillsdale: Lawrence Erlbaum Associates.
70. Wenger, E. (1987). *Artificial intelligence and tutoring systems—Computational and cognitive approaches to the communication of knowledge*. Morgan Kauffman publishers, Inc.
71. Tasso, C., Fum, D., & Giangrandi, P. (1992). The use of explanation-based learning for modelling student behavior in foreign language tutoring. *Intelligent tutoring systems for foreign language learning* (pp. 151–170). Berlin, Heidelberg: Springer.
72. Kuzmycz, M., & Webb, G. I. (1992). Evaluation of feature based modelling in subtraction. *International conference on intelligent tutoring systems* (pp. 269–276). Berlin, Heidelberg: Springer, Heidelberg.
73. Ohlsson, S. (1994). Constraint-based student modeling. *Student modelling: The key to individualized knowledge-based instruction* (pp. 167–189). Berlin: Springer, Heidelberg.
74. Qian, M., & Clark, K. R. (2016). Game-based Learning and 21st century skills: A review of recent research. *Computers in Human Behavior, 63,* 50–58.
75. Young, M. F., Slota, S., Cutter, A. B., Jalette, G., Mullin, G., Lai, B., et al. (2012). Our princess is in another castle: A review of trends in serious gaming for education. *Review of Educational Research, 82,* 61–89.

Chapter 9
Designing a 3D Board Game on Human Internal Organs for Elementary Students

Yu-Jie Zheng, I-Ling Cheng, Sie Wai Chew and Nian-Shing Chen

Abstract The importance of learning about the human body and its internal organ is undeniable as it encourages an individual to care about their health and lifestyle. Elementary school students who are beginning to learn about the human internal organ are usually faced with the difficulty of comprehending vital information of each internal organ, such as its appearance, its position in the human body, and its role to ensure the livelihood of a human being. Facing with this difficulty, students usually could not fully understand the important roles these internal organs play, and this would be a challenge to students to pay attention in taking care of their own health. Several medical researches had ventured into enabling such learning through an experiential learning process via a board game, enabling learners to gain new knowledge of the human internal organ and understand vital roles of the internal organ, and the consequences if one does not take care of them. Enlighten by a "Human Body Model" toy manufactured by MegaHouse, this research attempted to improve the game playing process of the toy by introducing additional software and the usage of sensors in the toy, resulting in a 3D board game. Students' interaction data were collected by the reader, and instantaneous feedbacks were provided to students accordingly. The results of the research show that the design of the 3D board game had effectively improved students' learning experience and also their learning performance. Several decisions on the research design and noticeable observations during the research process were discussed in the chapter for the usage of teachers, educators, and future researchers.

Y.-J. Zheng · I.-L. Cheng · S. W. Chew
Department of Information Management, National Sun Yat-sen University, Kaohsiung, Taiwan
e-mail: zyj940328@gmail.com

I.-L. Cheng
e-mail: chengi428@gmail.com

S. W. Chew
e-mail: chewsw@mis.nsysu.edu.tw

N.-S. Chen (✉)
Department of Applied Foreign Languages, National Yunlin University of Science and Technology, Douliu, Taiwan
e-mail: nianshing@gmail.com

A. Tlili and M. Chang (eds.), *Data Analytics Approaches in Educational Games and Gamification Systems*, Smart Computing and Intelligence,
https://doi.org/10.1007/978-981-32-9335-9_9

1 Introduction

Science, Technology, Engineering, and Mathematics (STEM) are core subjects in schools and education. Although the health and physical education is not considered as one of the major academic discipline in school, health, and physical education is strongly related to a student's personal health. With the knowledge of one's health and one's own body, students would learn to take precaution in illness prevention, assist students to maintain a healthy lifestyle, and possess a strong body and mind [1]. Learning should be in touch with the realities of life experience, especially when it comes to something as personal as knowing and understanding one's own health and body [2, 3]. At the age of 7–11, children are in a concrete operational stage where they can only think about specific existence of the matter [4]. Piaget's [4] study pointed out that the cognitive development of a child aged 7–11 could not comprehend abstract concepts and knowledge, such as the idea of being and staying healthy, formation of human muscles, and the role of different human organs. Hence, it might be difficult to have elementary school students understand about their body and their internal organs, and the importance of taking care of their personal health.

In medical and health-related education, experiential learning theory is widely used to assist students in mastering the learning content. The past research had often utilized experiential learning theory with gaming, stimulation, practical exercises, and/or computer-based programs, such as online or multimedia programs, especially for medical and health-related education [5–7]. This was mainly because the learning activities designed using experiential learning theory emphasize an active knowledge construction process between the learner and its surrounding environment [2]. For example, previous studies had shown that the usage of board games in medical and health professionals' learning and training process can facilitate the abstract organ appearance into concrete appearance in order to reinforce learners learning subject through playing around the board [8]. As Treher [9] shared, using board game provides a visual metaphor to connect with information and offer young children a hands-on chance (with physical devices) to develop knowledge of the subject.

For elementary school students in Taiwan, the topic on the introduction to the human organ is part of their "health and physical" class syllabus. The knowledge about human organs is difficult for students to fully understand as they are abstract concepts which they are unable to witness firsthand to comprehend the role and importance of these organs. Hence, students could not internalize the knowledge of the human organs and the importance of taking care of them. By utilizing the experiential learning theory from previous educational board games, this research designed a 3D board game for students to learn about the human internal organs by playing a game. The objective of this research was to explore the effectiveness of using a 3D board game in learning about human internal organs for elementary school students. The research question of this research was "Comparing with a 2D board game design, how effective is a 3D board game design in teaching elementary school students about the human internal organs?"

2 Literature Review

The research explored and compared the effectiveness of elementary school students learning about the human internal organs using a 3D board game and a 2D board game. In designing the board game and the learning activities for these students to learn about the human internal organ, the research utilized the game-based learning and experiential learning pedagogies to better improve students' learning experience and learning performance.

2.1 *Game-Based Learning*

There is a significant relationship between games and learning [6, 10]. Game-based learning is a learning process of integrating playing time with learning time [11]. The past research had shown that students will learn more effectively in a game-based learning setting as students would tend to be more engaged in the game, and they would put in more effort in understanding the learning material embedded into the game in order to complete the goal of the game [12, 13]. With technology innovation, students have a lot of opportunities to touch with different types of games. Even so, studies have proved that board games can be effective motivational and learning tool for students to learn [14]. In Whittam and Chow's [8] study, they utilized a board game to enable students to learn about the importance of different degree of burn. The educational board game had a positive effect on the students' learning performance as it not only enhanced students in term of their knowledge and understanding, throughout the game playing process, but the educational board game also provided students with enjoyment and instilled their interest on the topic [8]. Learning about the human internal organ is difficult for elementary students to visualize and hard for them to understand the importance of each organ. Hence, this research introduced the concept of game-based learning by designing a 3D board game for elementary Grade 4 students to learn about the human internal organ. The research design is further discussed in the following section.

2.2 *Experiential Learning Theory*

In experiential learning theory, "learning is the process whereby knowledge is created through the transformation of experience. Knowledge results from the combination of grasping experience and transform it" [2, p. 67]. Experiential learning focuses on the process of learning instead of the results of learning. There were four modes proposed in the experiential learning theory model: (1) Concrete Experience (CE), (2) Reflective Observation (RO), (3) Abstract Conceptualization (AC), and (4) Active Experimentation (AE). Concrete Experience (CE) involved having students encountering a new experience or given the opportunity to reinterpret their existing experience. Reflective Observation (RO) is a learning mode where students could notice and evaluate the inconsistencies between the experience and their own understanding of the knowledge. Abstract Conceptualization (AC) is a reflective process on the given experience which assists students in generating a new idea or induces students in making modification of an existing understanding of a particular abstract concept. Active Experimentation (AE) is a learning mode where students apply and utilize the knowledge gained during the experience and observe the changes and difference in the results of such change [15].

Experiential learning is a process of "constructing knowledge that involves a creative tension among the four learning modes that is responsive to contextual demands" [16, p. 1216]. The past research had shown that experiential learning is effective in improving the learning process and learning experience of students which in turn would result in an improved learning performance [17–19]. Much research in the medical field had found that experiential learning played a crucial role in the learning process as it engages students in reflecting the new knowledge with reference to their own experience and understanding [5, 8]. In order to improve the learning process and experience of elementary school students in learning about abstract concepts, such as the human internal organ, this research adopted the experiential learning theory in designing and improving the learning process for students. In designing the learning process on learning about the human internal organ, all four learning modes of experiential learning were utilized in guiding the design of the learning process.

3 Research Design

This research's objective was to explore the design of using a 3D board game in teaching elementary school students about the human internal organs. In the attempt of improving students' learning process on understanding the human internal organ, this research utilized the "Human Body Model" toy which was manufactured by MegaHouse (Japan). This toy was a 3D model of a human which enables players to place the relevant 3D internal organs at it designated position. With the 3D human body toy, the research further develops an electronic board game called the "Organ Savior Game." The detailed descriptions of the development and design of the "Organ Savior Game" are as follows.

3.1 Learning Topic and Learning Content

In designing the learning content of the research, Kolb's [2] experiential learning theory was utilized with the four learning modes as reference (i.e., Concrete Experience, Reflective Observation, Abstract Conceptualization, and Active Experimentation).

This research utilized the "Human Body Model" toy manufactured by MegaHouse (Japan) as one of its core instruments in designing the electronic board game called the "Organ Savior Game." The "Human Body Model" toy consists of a 3D human body model with ten 3D models of different internal organs, such as heart, lung, liver, stomach, cholecyst (gallbladder), spleen, large intestine, small intestine, kidney, and bladder (as shown in Fig. 1). Out of the ten internal organs, eight were selected in accordance with the difficulty level suitable for elementary school students and were used as the learning topics of this research (i.e., heart, lung, liver, stomach, large intestine, small intestine, kidney, and bladder).

With the learning topic decided, the research further designed the learning content of each learning topic by referring to the textbook used by the "Health and Physical" classes of elementary school students in Taiwan [20–24]. There were two stages in the learning process of the game: (1) learning stage, and (2) challenge stage. The learning stage provided students with Concrete Experience and Reflective Observation from Kolb's [2] experiential learning theory, whereas the challenge stage provided students with Abstract Conceptualization and Active Experimentation.

In designing the learning content of the learning stage, learning materials related to each internal organ were selected from these textbooks and summarized to portray important learning information for each organ. These materials were then classified into (1) the appearance and location of the internal organs, (2) the function or role of the organs, and (3) some additional information, such as the description of the daily activities involving the usage of the organs, ways to keep these organs healthy, and preventive measures taken to ensure the organs are healthy. The prepared learning content and materials were discussed with elementary school teachers of the "Health and Physical" classes to ensure the accuracy of the prepared learning content and its appropriateness for the students.

Fig. 1 Human body model and its playing cards manufactured by MegaHouse

As an example, for the lungs, the information on the appearance and location of the lungs and its function or role was provided as follow:

> This organ is on both sides of the chest. It is like a sponge, just that for a sponge it absorbs water, and the lungs it absorbs air. How do the lungs breathe? When the air enters the trachea through the nose or mouth, and arrives to the lungs. (translated from Mandarin, see upper part of Fig. 5)

For the lungs, the additional information was provided as follows:

> Note that people with smoking habits will have a layer of tar attached to their lungs. It is very unhealthy. When you have a cold, the effect of coughing helps the lungs clear germs and prevent lung infections. However, the mucus produced by cough is pathogenic, so remember to bring a mask when you have a cold! By increasing the intake of yellow orange fruits and vegetables like carrots and bananas, it can improve immunity and reduce the incidence of lung cancer and stomach cancer. (translated from Mandarin, see lower part of Fig. 5)

The learning contents selected for this study were intended to be closely related to the students' daily life with the aim of providing students with Concrete Experience and Reflective Observation as mentioned in the experiential learning theory. Using the lungs as an example, students of this age understood the action of breathing and would have experienced having cough or a cold in the past. The learning contents for each organ were designed using context which students could have experienced in the past and provided them an opportunity to re-evaluate their understanding of their existing experience.

With the content of the learning stage completed, the items for the challenge stage were designed. There were a total of five patients' conditions provided for students to provide their diagnosis on each patient's condition. Students were provided with the conditions faced by the patients similarly with when one was to consult a doctor when they were unwell. After listening to the patient's condition, students were then required to determine the organ which was in need of attention and the suggested consultation to be provided to the patient. The health conditions of the patients were designed in accordance with the learning contents provided in the learning stage. This stage was designed in accordance with Abstract Conceptualization and Active Experimentation as mentioned in the experiential learning theory where students had to understand the conditions of the patients, identify the source of their illness, and provide them with the appropriate consultation.

3.2 System Design

With the learning topics and the learning content prepared, the research further venture into using the "Human Body Model" toy in portraying the learning activity by designing an electronic board game called the "Organ Savior Game." The "Organ Savior Game" was designed to provide students with an interactive game play where instantaneous feedback was given in accordance with the interactions that the students made with the "Human Body Model" toy. This instantaneous feedback mechanism

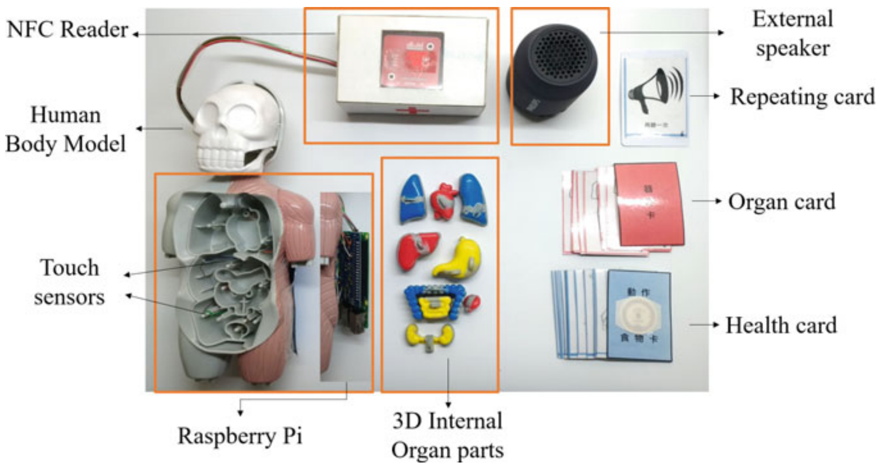


Fig. 2 Instrument setup of the "Organ Savior Game"

was enabled in the "Organ Savior Game" with the embedment of a single-board computer (i.e., Raspberry Pi) along with the support of different sensors used on each 3D internal organ parts and playing cards of the game (i.e., touch sensor, near-field communication technology: NFC reader and NFC tags) as shown in Fig. 2. In order to enable an audio feedback mechanism throughout the game play process, external speakers were used. The designed "Organ Savior Game" consisted of two stages (i.e., learning stage and challenge stage). Throughout both stages, the designed instrument setup of the "Organ Savior Game" was used.

The designed scenario of the "Organ Savior Game" was set at the Organ Savior Hospital where students were greeted by Dr. Organ, the president of the hospital. The goal of the game was to have students to learn about the eight human internal organs (i.e., learning stage) through a series of game play interaction with the instrument setup of the game. Thereafter, students were required to complete the challenge presented by Dr. Organ in determining whether learners were ready to be a doctor (i.e., challenge stage). In the following section, the instrument setup and the description of both stages (i.e., learning stage and challenge stage) of the "Organ Savior Game" are discussed.

3.2.1 Instrument Setup of the "Organ Savior Game"

The design of the "Organ Savior Game" was heavily based on the "Human Body Model" toy manufactured by MegaHouse (Japan). Hence, one of its core instruments used in the game was the "Human Body Model" toy which is a 3D anatomy of the human body with eight internal organ pieces provided. By utilizing the existing toy, the research further enriched the "Human Body Model" toy by introducing and implementing the Raspberry Pi, touch sensors, and near field communication technology (i.e., NFC reader and NFC tags) in the toy and in the game itself.

Beginning with the “Human Body Model” toy and its 3D internal organ parts (as shown in Fig. 2), inside of the human body model, there were designated piece holders for each internal organ parts. These piece holders were then covered with individual touch sensors which were used to determine whether the correct internal organ parts were placed in its correct location by the students throughout the gaming process. In order for the touch sensors to be activated, each 3D internal organ parts were coated with silver conductive adhesive paint to enable the conduction of electricity when these internal organ parts were placed into each place holders.

As the “Organ Savior Game” was designed as an electronic board game, playing cards were designed where there were the “Health cards,” the “Organ cards,” and the “Repeat card” (as shown in Fig. 2). These cards were prepared to each have an individual NFC tag embedded in them except for the “Organ cards.” The “Health cards” consisted of cards portraying different healthy lifestyles and preventive activities that students would learn about while learning about the internal organs (as shown in Fig. 3a). Throughout the game, students would utilize the “Health cards” in completing the designed tasks. The NFC reader was designed in a medical box where students were instructed to place their playing cards on to verify their answers. The “Organ cards” were used to help students in the group to determine whose turn it was to participate in the game (see Fig. 3b).

Throughout the game, in order to provide instantaneous feedback from students’ interaction data and signal received by the different sensors, a Raspberry Pi was designed to be embedded behind the human body model where it functions as a computer or processor of the “Organ Savior Game.” The designed learning contents were transformed into audio scripts which were pre-recorded and were programmed to be broadcasted in response to the actions taken by the students during the game playing process. The operating program in the Raspberry Pi was the Python which was written in Python language. The Raspberry Pi was used to read signals received from the touch sensors which were placed inside of the human body model and provide the relevant audio recording as the given response toward the placement of each internal organ pieces to its designated location in the human body model.

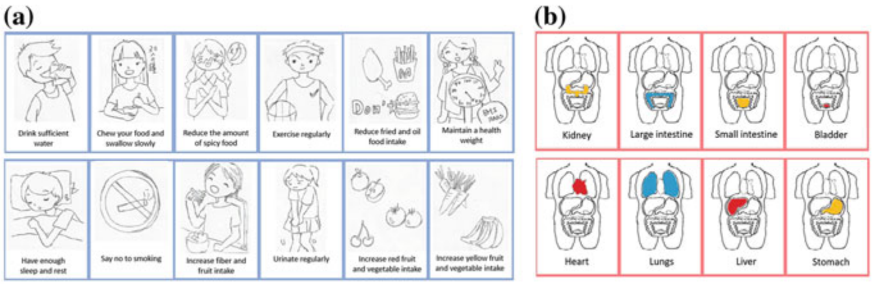


Fig. 3 Contents of **a** “Health cards” and **b** “Organ cards”

Furthermore, the Raspberry Pi was also used to communicate between the playing cards of the game which had individual NFC tags and the NFC reader with the designed program for the electronic board game. In response to the placement of each card to the NFC reader during each task, the Raspberry Pi would provide its relevant response. The audio scripts were broadcast through an attached external speaker to the Raspberry Pi. The designed program for the "Organ Savior Game" consisted of a back end platform which enabled the monitoring process of each student's live learning performance, and these results were recorded in the platform throughout the game playing process. Teachers and educators could utilize the learning process data collected to determine learning topics that students usually have difficulty in answering, learning topics that students required repetition to understand the material, and activities that students commonly answer wrongly during the learning process.

3.2.2 Stage 1: Learning Stage

Prior to the commerce of the game, the "Organ cards" and the "Health cards" were distributed among the group. The "Organ Savior Game" began with a greeting from Dr. Organ, the president of Organ Savior Hospital. The pre-recorded greeting by Dr. Organ was broadcasted to introduce students about the background of the game. Students were informed that the game was to begin with a learning stage. During this stage, Dr. Organ would provide the information on the name, the appearance, the location, and the function or role of a randomly selected internal organ to students. Students with "Organ card" referring to the provided information would be required to make the move for the group. This design of the game was to ensure that all students of the group were participating in the game. Thereafter, the students had to identify which internal organ was being referred to among the given eight 3D internal organ parts. Students were required to place the 3D internal organ parts on the designated place holders in the human body model. With each attempt of placing an internal organ into a place holder with its touch sensor, the game would revert back with an immediate feedback on whether the selected internal organ was correct, or whether the organ was placed on the correct location based on the students' interaction. The game would seek for students to continuously find the correct internal organ in reference to the given information and have them place the correct organ in the correct location in the human body model. For example, when Dr. Organ mentioned about the "heart," students with the heart "Organ card" would be required to identify the 3D internal organ part of the heart and place it in the correct location inside the human body model.

Once they had managed to find the correct internal organ and placed it in its designated location in the human body model, the game will continue on by sharing additional information on the internal organ which included ways to keep that internal organ healthy, activities to avoid damaging the performance of the internal organ, and more. After listening to the given additional information, the game would broadcast questions in regards to the given additional information on the internal organ. Students were required to answer these questions by placing the provided "Health cards" on

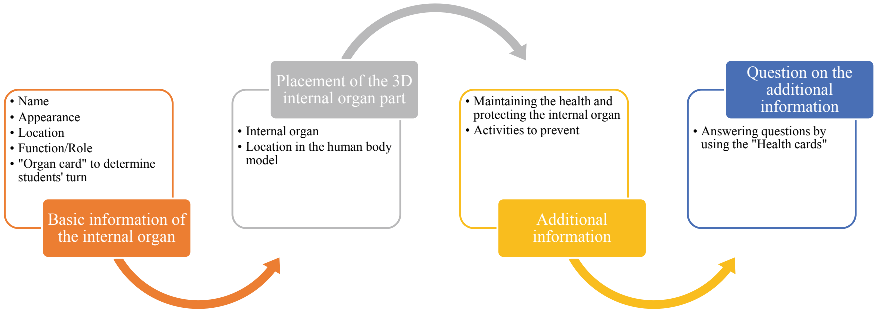


Fig. 4 Segments of the learning stage in the "Organ Savior Game" for each internal organ

the medical box (i.e., the NFC reader). On each "Health card" was information on different healthy lifestyles and preventive activities which were used to answer these questions. In reference to the different "Health cards" students placed on the medical box, the game would inform students on whether they answered the questions correctly. There were questions that had multiple different answers which were placed on different "Health cards," hence students were required to find all the correct cards in order to complete the questions. Students could utilize the "Repeat card" anytime by placing it on the medical box (i.e., the NFC reader) throughout this process to replay the information given for the particular internal organ to assist them in completing this task.

After completing the question for the particular internal organ, the game would carry on with the next randomly selected internal organ, and the learning process was repeated for each internal organ until all eight internal organs were completed by the students. The different segments of the learning stage in the "Organ Savior Game" were as shown in Fig. 4. Throughout the learning process in the game playing process, the game provided students with a new learning experience where students gained new knowledge about the human internal organs, and they had to reinterpret their existing knowledge and experience in regards to the topic. By undergoing this process, students could modify their own understanding on the human internal organ when they encounter any inconsistencies between their existing understanding and experience with the information provided during the game playing process. The learning stage was designed based on the Concrete Experience and Reflective Observation learning modes proposed by Kolb [2].

3.2.3 Stage 2: Challenge Stage

With the completion of the learning stage of the "Organ Savior Game," the game would continue with the challenge stage of the game. At the beginning of this stage, students would be greeted again by Dr. Organ where they were congratulated by Dr. Organ for completing the learning stage. Students were then informed about the scenario of the challenge stage by Dr. Organ. In the challenge stage, students would play the role of a doctor where audio recordings of patients seeking for their consultation were played. After listening to the patients' condition and health issues, students were required to utilize the knowledge that they had learnt in the learning stage in regards to each internal organ and provide each patient a diagnosis. For example, one of the patient's conditions in the game was "Doctor! I recently felt that my heart is pumping quickly, and I feel pressured around my chest."

There were a total of five patients designed in the challenge stage of the "Organ Savior Game." For each patient, students had to identify the internal organ in question by placing the 3D internal organ parts in the relevant location inside of the human body model. Similarly, the game would provide students with an instantaneous response on whether they had selected the correct internal organ for each patient. After identifying the internal organ which was related to the patient's health issue, students would then utilize the "Health cards" to provide each patient their diagnosis and relevant health recommendations. For each patient, student had one minute to identify the correct internal organ and two minutes to provide the correct diagnosis and health recommendations. The students' performance for each patient was recorded in the game.

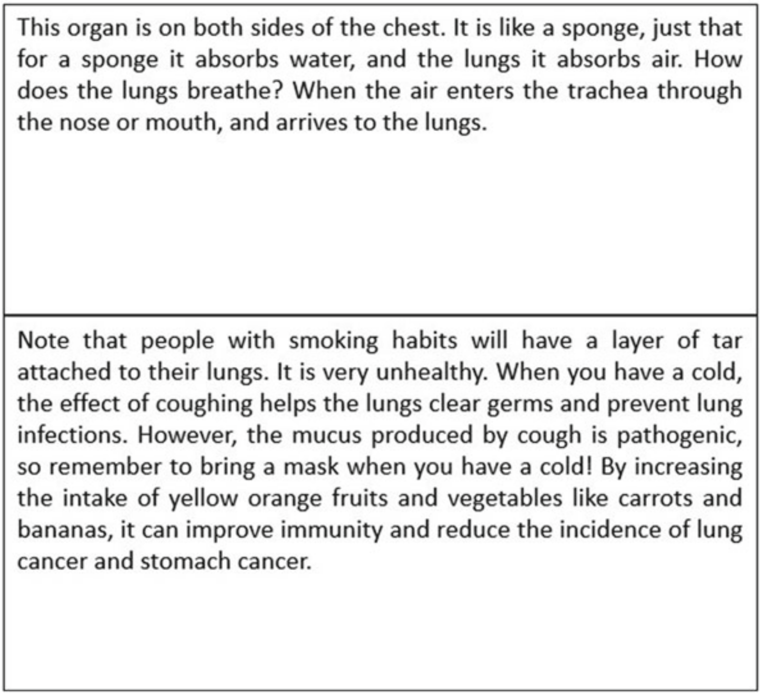
By providing students with the simulation of playing the role of a doctor and providing patients with the relevant diagnosis in accordance to their health conditions, students were given the opportunity to utilize the knowledge they had learnt in the learning stage. The challenge stage was designed to provide students with the opportunity to reflect on the knowledge that they had on human internal organ, generate new ideas and understanding on the topic, and assist students in modifying their understanding of an abstract concept. Furthermore, students could apply their knowledge on the human internal organ and verify their understanding through this game playing process. The challenge stage was designed based on the Abstract Conceptualization and Active Experimentation learning modes proposed by Kolb [2].

4 Experimental Process

The research's objective was to explore the effectiveness of using a 3D board game for elementary school students in learning about human internal organs. In order to explore the effectiveness of the designed 3D board game "Organ Savior Game," a conventional 2D board game was designed using the same learning topic and learning material as a comparison.

The conventional 2D board game was design in a manner where students were to complete the game on a playing board. The "Organ cards" were used in the conventional 2D board game in replacement of the 3D internal organ parts as shown in Fig. 3b. The conventional 2D board game was similar with the 3D board game where there were two similar stages (i.e., learning stage and challenge stage). For the learning stage, all the basic information (i.e., the appearance, the location, and the function or role of the internal organ) was displayed on the board in wordings, and audio recording was also played during the game playing session (see Fig. 5). Students would then go through the internal organ playing cards to determine which internal organ does the given information refers to by placing the card on the board. As there were a total of eight learning topics (i.e., eight different internal organs), a total of eight information boxes were prepared on the board for the learning stage.

As for the challenge stage, a different board was prepared where the patients' condition and health issues were listed in wordings and the similar audio recording was played during the game play process. On the board, a human anatomy was printed where students had to place the internal organ that they deemed to be involved in accordance with the patient's condition and health issues (see Fig. 6). For this, additional 2D human internal organ parts were prepared. Students had to identify the correct internal organ and place the 2D internal organ part on the correct location on the printed 2D human anatomy. Thereafter, on the lower portion of the board, students had to use the "Health cards" to provide patients with their diagnosis and relevant health recommendations by placing the card in place on the board. There were a total of five patients for the challenge stage, hence five information boxes were prepared on the board.


This organ is on both sides of the chest. It is like a sponge, just that for a sponge it absorbs water, and the lungs it absorbs air. How does the lungs breathe? When the air enters the trachea through the nose or mouth, and arrives to the lungs.

Note that people with smoking habits will have a layer of tar attached to their lungs. It is very unhealthy. When you have a cold, the effect of coughing helps the lungs clear germs and prevent lung infections. However, the mucus produced by cough is pathogenic, so remember to bring a mask when you have a cold! By increasing the intake of yellow orange fruits and vegetables like carrots and bananas, it can improve immunity and reduce the incidence of lung cancer and stomach cancer.

Fig. 5 Designed boards for the conventional 2D board game for the learning stage

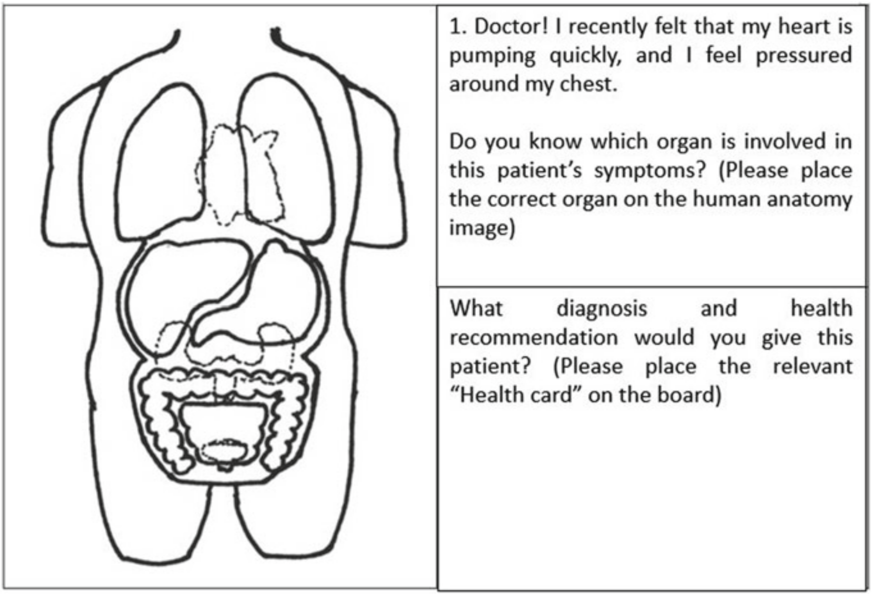


Fig. 6 Designed boards for the conventional 2D board game for the challenge stage

4.1 Participants

The learning content of this research was designed for elementary Grade 4 students to learn about the human internal organs. The participating students of this research were all Grade 4 students from an elementary school in Kaohsiung, Taiwan. A total of 74 Grade 4 students participated in the research with 36 students randomly assigned to the control group and 38 students randomly assigned to the experiment group. Students of the control and experiment group each formed individual groups of three or four which resulted in nine individual groups formed in the control group, and ten individual groups formed in the experiment group. The gender distribution of each group was shown in Table 1.

Table 1 Distribution of participants

Group	No. of individual groups	Male	Female	Total
Control	9	15	21	36
Experiment	10	16	22	38
Total	19	31	43	74

4.2 Pretest and Posttest

The questions for the pretest and posttest were questions designed in accordance with the learning topics and learning content used in the research with reference to the questions used in textbook of the “Health and Physical” classes of elementary school students in Taiwan. The designed questions were then evaluated by teachers of the “Health and Physical” classes to ensure that the questions are suitable and appropriate for the students. Both tests were prepared in a similar manner which consisted of multiple choice questions and fill in the blank questions with option of answers provided. The students completed the pretest before the research and posttest after the research whereby 15 min was provided for each session.

4.3 Experimental Procedure

The research was conducted at the classroom of the students in school. The control group completed the research using the conventional 2D board game which consists of game boards, 2D human internal organ parts, “Organ cards,” and “Health cards.” The experiment group completed the research using the 3D board game which consists of the 3D human model, 3D internal organ parts, “Organ cards,” “Health card,” “Repeat card,” “external speaker,” and a medical box (i.e., NFC reader). The learning contents designed for both the control and experiment group were similar with the difference in the game playing mechanism. Each individual group of both groups was assigned with an observing researcher. For the control group, the observing researchers would play the role of determining whether the group had answered the questions in both learning stage and challenge stage correctly, and they would record down the performance of the students in each individual group.

Prior to the commerce of the research, students completed the pretest individually. After completing the pretest, students were then randomly assigned into individual groups of three or four. For both control and experiment group, each individual group was to ensure that they had all the items required for the game. A briefing session on the goal of the game was conducted before the game began. During the briefing session, students were instructed to distribute all the playing cards among the individual group members (i.e., “Health cards” for both control and experiment group; internal organ playing cards for control group only). Thereafter, students began the playing the game among their own groups. The control group completed the game using a conventional 2D board game (see Fig. 7a), and the experiment group completed the game using a 3D board game (see Fig. 7b). After completing the game, students then completed the posttest. The total time of the research was 90 min with the experiment procedure shown in Fig. 8.

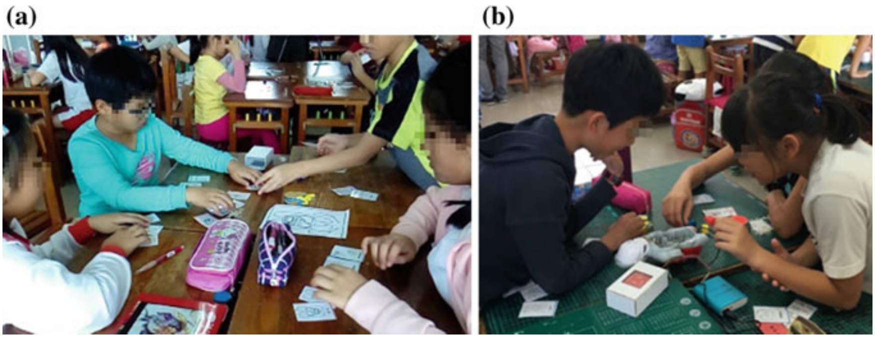


Fig. 7 **a** Control group and **b** experiment group completing the game

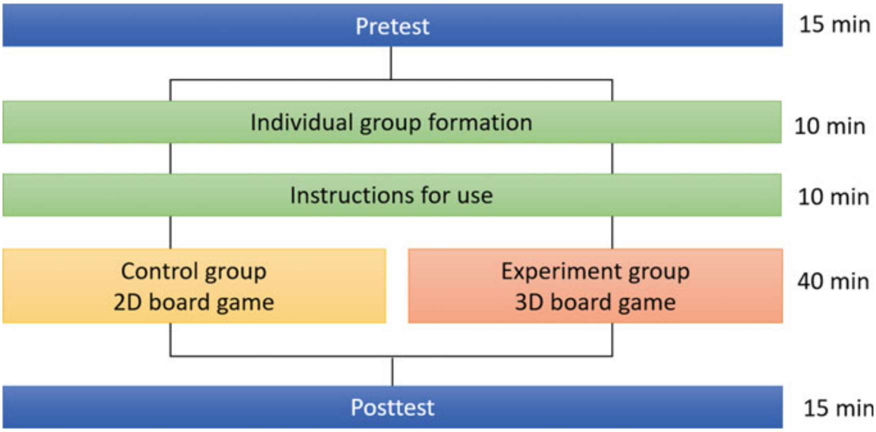


Fig. 8 Experimental procedure

5 Results

The research was conducted to explore the effectiveness of using a 3D board game for elementary school students in learning about human internal organs. For comparison, the 3D board game was compared with a 2D board game which was designed using similar game instructions and contents. A *t*-test was conducted on the pretest results which showed that there was a slight difference between the two groups where $p < 0.10$. Hence, the results of the posttest were analyzed using ANCOVA with the pretest results used as the covariate variable. Prior to the ANCOVA analysis, the homogeneity test was conducted with the results of $F = 3.23$, $p = 0.076$ (>0.05) which indicated that the results had fulfilled the assumption of homogeneity of regression.

The results of the ANCOVA analysis on the posttest results were shown in Table 2. The results indicated that there was a significant effect on the grouping of the students (i.e., control group and experiment group) with $F = 4.58$ ($p = 0.036$, $\eta^2 = 0.061$).

Table 2 ANCOVA results on students' learning performance (posttest result)

	N	Mean	SD	Adjusted mean	SE	F	η^2
Control group	36	50.58	12.52	49.54	1.91	4.58*	0.061
Experiment group	38	54.29	13.12	55.28	1.86		

Note $^{*}p < 0.05$. Dependent variable: Posttest results, Covariate variable: Pretest results

With the adjusted mean of the experiment group (=55.28) higher than the control group (=49.54) showed that the students learnt more effectively using the 3D board game as compared to the 2D board game. Based on Cohen's [25] suggestion, the effect size (η^2) for the ANCOVA analysis of the different board games was large (η^2 > 0.138)

6 Discussion

The research explored into using a 3D board game to improve the learning process of elementary Grade 4 students in learning about the human internal organs. The results of the research indicated that the design was effective in assisting students in improving their learning performance with comparison with a conventional 2D board game. It implies that by using the 3D board game designed with an immediate and interactive feedback based on students' interaction with the toy can assist in improving students' learning achievements on learning about the human internal organ and the health information compare to conventional 2D board game. One of the main contributions of this research was the introduction to a new method to better improve the learning experience and learning performance of elementary school students. By integrating a computer software into a conventional board game, students were able to enjoy the learning process through an enjoyable game playing process, with instantaneous feedback and results provided by the software in order to determine whether the moves that they had taken was correct. Furthermore, researchers, teachers, and educators were able to monitor the learning process data provided in the backend system to identify learning topics that students find difficult to understand and require further elaboration or material to facilitate their learning, learning topics that require more attention as students commonly make mistakes in completing the learning activities, etc. These data provided teachers and educators with information about the effectiveness of the designed learning topics and their students' performance during the learning process which could enable further enhancement to the designed learning activities.

Several research design decisions were made in order to improve on existing issues faced in the classroom, including (1) the usage of "organ cards" for students to take turn making a move, (2) the usage of human voice recorded narratives and matching sound effects, (3) using the system to evaluate the students' work instead of a tutor, and (4) using board games items to induce learning. During the game playing process, students in each group were given the “organ cards” at random. For each round, students who had the particular “organ card” of the aforementioned internal organ were responsible to make a move for the group when the question provided was related to that particular organ. This game play mechanism was introduced to prevent the learning session being controlled by one or two students in the group, leaving out other members of the group in the participation of the game or learning process. For this research, students in the group were working together to learn about the learning contents of the game. In the future, the game play mechanism could be modified to allow for internal competition among group members to encourage more active participation of students in the game.

Besides on the game playing mechanism during the game, throughout the game playing process, all audio files were designed to enrich students’ learning experience by using recorded audios. Instead of using computerized virtual voice, these recorded audios were recorded using an actual human voice which was filled with emotions and intonation, providing students with a better visualization of the whole situation. For example, the patient in pain would speak slowly, sound weak and some may even cough. Patients who were conveying a positive health information, a happy, and uplift tone was used as compared to a low and discouraging tone on the sharing of activities which would deteriorate one’s health. Additionally, matching sound effects of the situation was also utilized to build up the situation for students during the game play process. For example, during the “Challenge stage,” before each patient begins sharing about their health conditions, the sound effect of a door opening was played. It was observed during the research that students were excited and happy when this sound effect was played, and students were getting ready to pay attention to listen to the patient. The sound effect assisted students in better visualizing the actual situation of being a doctor in the consulting room, and as the patient was entering the room to consult them for their diagnosis. This indicated that by using recorded audios and proper sound effects of the given situation, it would enrich the learning experience for students and enable students to not only be better engaged in the game playing process, but also to provide students with the anticipation on the events that were about to occur.

Throughout the research process, there were certain observations that were distinct and worthy of mentioning and discussed in this chapter. The results of this research indicated that there was a significant difference between the control group and the experimental group (i.e., 2D board game and 3D board game, respectively). This may be contributed by the fact that in the game play mechanism for the 2D board game, an observing researcher was assigned to each group to facilitate in determining whether the group’s move (i.e., card placement) for each learning activity was correct. As a comparison, for the 3D board game, the verification mechanism was completed instantaneously by the designed software, which means that the observing researchers of the 3D board game group were not involved in the game playing process directly. This may also contribute to the lower performance of the control group as the observing researchers may had interrupted the students’ game playing process when they were providing the groups feedback on their moves. Initially, when the observing researchers were introduced in the 2D board game playing mechanism, it was thought that it would have a positive effect on the students' learning performance as the observing researchers would play the role of a "tutor" throughout the game playing process. However, it was noticed that this may not be the case as the observing researchers may be interrupting the students' game playing process. This observation may provide an important output to future research, as "tutors" or "guides" were commonly used to assist students during the learning process, especially in station games, school field trips, and classroom group learning. In designing the role of the "tutor" or "guide," it was vital that the "tutor" or "guide" would not disrupt the students' learning process or flow to ensure a fully immersed learning process.

Furthermore, it was also observed that students were curious about the designed 3D board game and was eager to begin playing the game. Before the beginning of the game, with all the instruments required for the 3D board game laid out on the table for the students, students showed signs of curiosity where they asked the observing researchers about each item on the table, they were eager to know how to play with these items and were impatient and wanted to begin playing the game. After the game playing process, students expressed their interest on the topic as they asked the observing researchers whether they could play the game again, and also asked the observing researchers for further information on a certain internal organ. These observations were mainly observed from the students of the 3D board game group, which indicated that the design of the research had improved and enrich the learning experience for students, and the introduction of different sensor and technology could assist in instilling students' interest on the learning topic.

The autonomy of the students during the game play process is important, especially when the students showed interest and eagerness to learn using the designed game. The current design of the 3D board game did not provide students with the autonomy to decide which internal organ they would like to learn about as a designated learning course was designed for the learning process. Although the "repeat card" played the role of allowing students to listen to the recording again, students could not head back to the previous internal organs to listen to the information again as a revision. In the future, the game design could consider providing students with the autonomy right to decide on parts that they would like to learn more about or to have a revision on. Likewise, as this research has shown that students were enthusiastic about learning more about the human internal organ while learning using this method of game playing process, other learning topics may also consider using such design in carrying out the learning process, providing students a game playing situation while learning about the learning topics itself in an enjoyable manner.

7 Conclusion

In designing an improved 3D board game to enable elementary school students to learn about the human internal organs in an interactive and engaging manner, several other important elements on the design of the learning activities and learning process were encountered. The introduction of complementary software along with different sensors to the existing toy "Human Body Model" manufactured by MegaHouse had shown to be effective in enhancing the learning experience in elementary school students and instill their interest on the topic itself. Although the design of the improved game play had indicated significant improvements, the research had also uncovered some other issues which would require much attention in the improvement of the future research. These include the design of the game to ensure the participation of all students, methods to enrich the learning experience throughout the game using rich media, the benefit of using new methods and mediums to convey the learning topics to students, the importance of the tutor in the learning process, and students' autonomy during the game playing process. These few elements are crucial items that were uncovered from this research which could be useful to the future research and for teachers and educators who are looking into new methods to teach their students. This research provided an alternative approach to teachers and educators in teaching about the human internal organ, and also for other learning topics by using different games or toys and incorporating these sensors to facilitate the students' "playing" process. For example, educators could utilize the shopping cart toy set to allow students to learn about mathematics by having students purchasing groceries of different prices and calculate the total. The research design could be adopted to fulfill the needs and requirements for other learning topics which could be conveyed in a game playing process enrich with an interactive software and different sensor throughout the learning process.

Acknowledgements This work was supported by the Intelligent Electronic Commerce Research Center from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan and by the Ministry of Science and Technology, Taiwan, under project numbers MOST-108-2511-H-224-008-MY3, MOST-107-2511-H-224-007-MY3, and MOST-106-2511-S-224-005-MY3.

References

1. Joseph, W., & David, P. (2016). *Adapted physical education and sport* (6th ed.). Champaign, IL: Human Kinetics.
2. Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. Upper Saddle River, NJ: FT press.
3. Peralta, L. R., Dudley, D. A., & Cotton, W. G. (2016). Teaching healthy eating to elementary school students: A scoping review of nutrition education resources. *Journal of School Health, 86*(5), 334–345.
4. Piaget, J. (1965). The stages of the intellectual development of the child. *Educational Psychology In Context: Readings For Future Teachers, 63,* 98–106.
5. Blakely, G., Skirton, H., Cooper, S., Allum, P., & Nelmes, P. (2010). Use of educational games in the health professions: A mixed-methods study of educators' perspectives in the UK. *Nursing and Health Sciences, 12*(1), 27–32.
6. Papastergiou, M. (2009). Exploring the potential of computer and video games for health and physical education: A literature review. *Computers and Education, 53*(3), 603–622.

7. Sand, J., Elison-Bowers, P., Wing, T. J., & Kendrick, L. (2014). Experiential learning and clinical education. *Academic Exchange Quarterly, 18*(4), 43–48.
8. Whittam, A. M., & Chow, W. (2017). An educational board game for learning and teaching burn care: A preliminary evaluation. *Scars, Burns and Healing, Vol. 3*. https://doi.org/10.1177/2059513117690012.
9. Treher, E. N. (2011). *Learning with board games*. Northern Minnesota, MN: The Learning Key Inc.
10. Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE Review, 41*(2), 16.
11. Tobias, S., Fletcher, J. D., & Wind, A. P. (2014). Game-based learning. *Handbook of research on educational communications and technology* (pp. 485–503). New York, NY: Springer.
12. Qian, M., & Clark, K. R. (2016). Game-based learning and 21st century skills: A review of recent research. *Computers in Human Behavior, 63,* 50–58.
13. Wu, W. H., Hsiao, H. C., Wu, P. L., Lin, C. H., & Huang, S. H. (2012). Investigating the learning-theory foundations of game-based learning: A Meta-analysis. *Journal of Computer Assisted learning, 28*(3), 265–279.
14. Chiarello, F., & Castellano, M. G. (2016). Board games and board game design as learning tools for complex scientific concepts: Some experiences. *International Journal of Game-Based Learning (IJGBL), 6*(2), 1–14.
15. Kolb, A. Y., & Kolb, D. A. (2005). Learning styles and learning spaces: Enhancing experiential learning in higher education. *Academy of Management Learning and Education, 4*(2), 193–212. https://doi.org/10.5465/amle.2005.17268566.
16. Kolb, A. Y., & Kolb, D. A. (2012). Experiential learning theory. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 1215–1219). Boston, MA: Springer.
17. Binson, B., & Lev-Wiesel, R. (2018). Promoting personal growth through experiential learning: The case of expressive arts therapy for lecturers in Thailand. *Frontiers in Psychology, 8,* 2276.
18. Hsieh, Y. H., Lin, Y. C., & Hou, H. T. (2015). Exploring elementary-school students' engagement patterns in a game-based learning environment. *Journal of Educational Technology and Society, 18*(2), 336–348.
19. Stein, M. (2018). Theories of experiential learning and the unconscious. In *Experiential Learning in Organizations* (pp. 19–36). Routledge.
20. Han, D.-K. (2008). *Wode diyitang youqu de renti changshi ke [My First interesting human body class]*. Taipei, Taiwan: Meiyi Academy.
21. Shuri, T. (2010). *Renti gouzao yu jineng de aomi (tujie ban) [The Mystery of human body structure and function (Graphic version)]*. Taichung, Taiwan: Morning Star Publication.
22. Yan, M. (2013). *Wajue qianzai de nengli: Shenti gongneng [Mining potential energy: Body function]*. Changchun, China: Jilin Publishing Group Co., Ltd.
23. Zhang, R. (2013). *Renti de bimi: Chatu ban [The Secret of the Human Body: An Illustration]*. Lanzhou, China: Gansu Science and Technology Press.
24. Han, Z. (2010). *Renti de aomi [Human mystery]*. Yinchuan, China: Ningxia Children's Publishing House.
25. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum.

Part IV
Modeling Learners and Finding Individual Differences by Educational Games and Gamification Systems

Chapter 10
Learner Modeling and Learning Analytics in Computational Thinking Games for Education

Sven Manske, Sören Werneburg and H. Ulrich Hoppe

Abstract Various approaches of game-based computational thinking (CT) environments were designed to better support the development of CT skills, such as abstraction, algorithmic thinking, generalization, or decomposition. We provide a classification of game-based environments for CT according to their characteristics such as the programming tools offered to the learners. In contrast to environments with open-ended tasks, goal-oriented learning environments have the potential to guide learners toward becoming a computational thinker. We present a framework for designing and evaluating game-based CT environments. This framework combines the use of methods of learning analytics with a suitable learning progression in order to provide appropriate dynamic guidance, scaffolds, and feedback to the learners depending on their actual state of programming. Finally, we evaluated our environment ctGameStudio with a study from the science festival "ScienceNight Ruhr 2018" using this framework.

1 Introduction

Programming can be conceived as personal and creative form of intellectual expression [45]. In this sense, code written by learners represents their cognitive state regarding their understanding and a constructive solution of a given problem. This possibly uncovers misconceptions that are not limited to syntactic or semantic mistakes. Misconceptions can have to do with the problem, or they can relate to the structure of code in a general sense. So-called code smells are subjectively defined characteristics of (correct) source code that indicate problems regarding the coding style. Examples of code smells are code redundancy or long methods, which downgrade the readability and maintainability of a program. The metaphor of a "smell" emphasizes that it defines a loss in quality on a stylistic level. Without making a statement

S. Manske (✉) · S. Werneburg · H. U. Hoppe
Faculty of Engineering, University of Duisburg-Essen, Duisburg, Germany
e-mail: manske@collide.info

H. U. Hoppe
e-mail: hoppe@collide.info

A. Tlili and M. Chang (eds.), *Data Analytics Approaches in Educational Games and Gamification Systems*, Smart Computing and Intelligence,
https://doi.org/10.1007/978-981-32-9335-9_10

about the correctness of a program, it points more to a "value system" for software development [35].

The automatic assessment of code quality measures and the definition of indicators and metrics for such characteristics of software artifacts have a long history in computer science. The first grading system for assessing students' programming has been developed by Hollingsworth [18]. The students received specific tasks, and the "automatic grader" checked the correctness of the solution and gave feedback whether the solution is correct or not. However, the system was not completely automatic. Overflow, storage selection, too much time for executing, and operator decisions needed a manual entry. While such a system can support instructors in finding correct solutions, it is not beneficial for learners who struggle with programming. In contrast to recent research in assessment tools in learning contexts, it did not provide any guidance or learner support such as hints to improve programming.

With LOGO, a popular system was created to support learners getting started in programming [43]. Research has shown that students learned programming fast and easily [11]. Although LOGO was designed specifically to benefit learners, they still had to face challenges such as syntactic constraints [32]. While syntactic mistakes usually lead to non-executable programs, it shifts the learners' focus from the semantics of a program to its syntax. Sometimes, this requires the learner to understand very technical or even obscure error messages, instead of implementing a meaningful algorithm in first instance. With the popularization of computational thinking (CT), visual block-based programming tools became a standard way for beginner courses in schools and beyond [60]. Learning environments like Scratch handle the problem of syntactic constraints. In contrast to LOGO, "the code blocks only lock in syntactically valid ways, therefore 'bugs' are always semantic errors" [32].

Brennan and Resnick [3] developed a framework for assessing projects developed with Scratch. With project portfolio analysis, artifact-based interviews, and design scenarios, they assessed computational concepts, practices, and perspectives. However, Moreno-León and Robles presented one of the first approaches of assessing Scratch projects automatically [40]. For different computational skills, they have defined metrics, which count the presence of specific computational constructs, for example, the use of a conditional statement ("if"). Depending on the complexity of the used constructs, 0–3 points are assigned to the targeted concept so that learners can get a clear and quick feedback on their programming performance. In the example of a conditional statement, the differentiation on the level of complexity is distinguished by using "if" (1 pt.), "if-else" (2 pt.), or the use of logical operations (3 pt.). To achieve a maximal score, a table of necessary programming fragments is given. If the students use these programming constructs, they can receive all points, although the programming artifact does not necessarily have a good quality on a semantic level. Such an approach results from the trade-off between authoring efforts for reference solutions and the generalizability of automatic indicators. The indicators included in the work of Dr. Scratch can be used for any Scratch project without any adaptation costs.

There are also other approaches to assess programming artifacts in the context of CT. However, the transfer to CT is often not clear, because CT contains long-term competences and it cannot be measured easily with situational computational artifacts. We elaborated a framework, which augments a pedagogical model for the level design of computational thinking games with guidance and scaffolding in order to give students the opportunity to train their CT competences and to foster CT in self-regulated learning scenarios. With this work, we classify recent learning analytics techniques and facilitate code metrics to identify this transfer to CT competences for CT games. Finally, we present an overview of our studies using these tools and give an outlook on the upcoming research. We present a study of an "hour of code"-like event during the science festival *ScienceNight Ruhr*, where students had the opportunity to explore concepts of computational thinking by using our game-based learning environments. We use activity metrics to describe their programming behavior, and we will show the advantages of learning analytics techniques to identify of students' problems while programming and how connected guidance components can support them while programming. Combining these approaches of a task design of the environments, with a guidance framework and learning analytics serves as a prototypical model for designing game-based computational thinking environments.

2 Gamification of Computational Thinking Tools

Wing introduced CT as "solving problems, designing systems, and understanding human behavior, by drawing on the concepts fundamental to computer science" [65]. CT is essentially based on "the creation of 'logical artifacts' that externalize and reify human ideas in a form that can be interpreted and 'run' on computers" [19]. In this sense, computational thinking can be seen as a process, which involves the use of computational tools methods in order to accomplish a certain task or solving a given problem.

By using text mining techniques on a corpus based on a literature review of CT definitions, Hoppe and Werneburg state out that, in this context, "games are often used as examples [for teaching] to illustrate general principles of CT" (Fig. 1). This might be justified in the favorable nature of games for educational purposes. Wang and Chang showed that games have positive effects on the flow experience of students and support the learning of CT-specific competences on a motivational level [57]. However, there is a big difference between game-based environments for CT. While some environments are intended to create games and microworlds, others are more focused on learning particular computational concepts and support the learner in a specific progression acquiring certain CT skills. In the following sections, we provide a classification of game-based learning environments and present models for learning progressions to support and foster CT competences. However, the success in acquiring CT skills is framed by different kinds of learning strategies in this context.

Fig. 1 Word cloud of CT definitions according to Hoppe and Werneburg [19]

2.1 *Classification of Computational Thinking Environments*

In the following, we present some environments, which are focused on creating logical artifacts in the sense of computational thinking. Some of the widely used environments to foster CT are used to create games or have a game-based environment. We categorize such CT environments, which have a strong relation to games or gamification, as computational thinking games. In such approaches, a game engine or a sandbox to create games is provided. The learners use their computational thinking skills in order to accomplish a task, solve a puzzle, and interact with a stage using computational methods and constructs. Stages are typical components of such environments—independent of the specific sub-type. A stage corresponds to a microworld where virtual agents can interact. The stage can be manipulated using computational constructs through a given interface, namely a programming tool. Visualizing this interaction in the microworld has the potential to help learners in debugging processes—"an essential component of both CT and programming" [26]. With appropriately adapted domain-specific languages and associated programming tools, learners can control and observe the behavior of the virtual agents in their microworld. Two different aspects can group CT environments: On the one hand, it is possible to distinguish between the programming tools provided to learners: text-based programming tools, visual block-based programming tools, and alternative visual programming tools. On the other hand, differences in the concept of the microworld are observable, for instance, open task environments and goal-oriented environments (Table 1).

The **open task environments** contain sandboxes where students can author own games or stories. These environments have a higher degree of freedom for students because they are less restricted through explicit given tasks to progress and encourage teachers to be more proactive in designing the exercises.

The **goal-oriented environments** usually involve a level system or a task that needs to be completed. Such environments follow a stricter learning progression, usually with less degrees of freedom or particular constraints.

RoboCode and RoboBuilder are environments with open sandboxes for robot fights against other players. A solution counts as "good," if the robot can defeat the enemies, which are controlled through the other players' programs. LightBot and Pirate Plunder contain level systems, which the student has to complete. The *ctGameStudio* environment takes advantage of both microworld characteristics (Fig. 2). It includes both, a level system and an open sandbox, where the learners develop their own algorithms in order to prepare for robot fights. Following this approach, ctGameStudio combines a task-oriented environment ("RoboStory") with an open sandbox ("RoboStrategist"). The users control a virtual robot in a microworld using a visual block-based programming tool [61, 63]. In order to learn the basic CT concepts and programming constructs, a scaffolded learning progression supports the learner. The system includes "RoboStory," it is a level-based environment which focuses in each unit with different sub-levels on a specific abstraction type (loop abstraction, etc.) and uses different microworld constructs (geometric concepts, etc.). Additionally, advanced learners can develop own algorithms and strategies in the open sandbox. In contrast to the pre-defined levels and goals, in this environment, the learners are motivated through a tournament mode where they compete against other learners in a virtual robot fight. The tournament is still situated in the same microworld using the same representations of source code in the visual block-based tool. In this sense, the goal-oriented environment "RoboStory" blends seamlessly

Table 1 Categorization of a selection of CT learning environments

Programming tool	Microworld concept	
	Open task	Goal-oriented
Text-based	LOGO [9] Greenfoot [30] JavaKara [16]	RoboCode [41] Karel the robot [44]
Visual block-based	Scratch [48] Alice [27] Open Roberta [25] AgentSheets [47] Snap! [38]	RoboBuilder [58] Blockly games [14] ctGameStudio [63] Pirate Plunder [50]
Visual alternatives	KidSim [55] Kodu [33] Kara [16]	LightBot [12] Program your robot [26] ctMazeStudio [62]

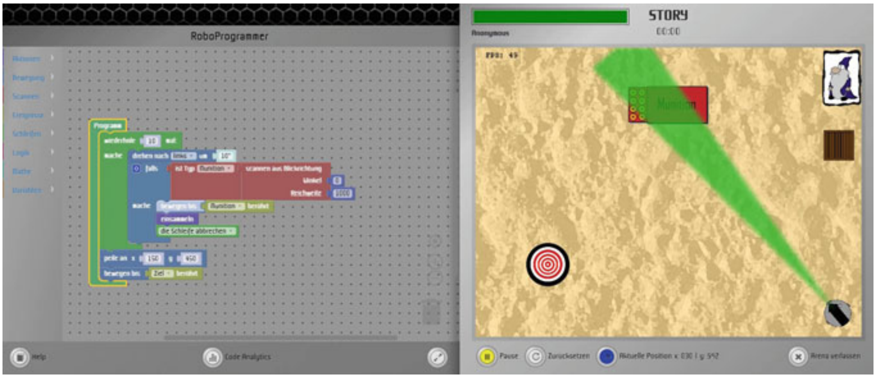


Fig. 2 Using scanning to find an object in *ctGameStudio*

into the open sandbox "RoboStrategist." It contains an open stage, where students compete in a tournament. In this mode, they can develop different algorithms to fight against other robots from other students.

Text-based environments are CT environments that primarily feature a text-based input, for example, a code editor. The logical artifacts that learners create are text-based and usually correspond to source code. Such environments are close to integrated development environments ("IDEs"), but support learners through tailored and more specific interfaces. For example, a reduced set of commands or code templates support students during their programming processes. By using these support mechanisms, the students have a low threshold to start programming, but syntactic errors are possible (and likely).

Visual block-based programming tools offer a solution to avoid the aforementioned shortcomings of text-based environments regarding the hurdle of expressing a syntactically correct solution. The logical artifacts that learners create in such environments are similar to the text-based artifacts, but are masked by a user interface, which prevents mistakes on a syntax level. Common environments (cf. 1) use a puzzle metaphor for the composition of the logical artifacts. In this sense, blocks are the puzzle pieces that stand for a certain command, for example, a statement or an expression in programming. Each puzzle pieces dragged into the editor only snaps if the combination leads to a syntactically valid composition. Additionally, in analogy to syntax highlighting in text-based editors, a color coding of blocks can be used to group according to (semantic) categories of commands. However, block-based visual programming tools are less powerful, potentially more cumbersome to use for larger projects, and inauthentic for students.

Each of the **programming tools with visual alternatives** presented in Table 1 has its own approach regarding the type of logical artifacts the learners create and the computational model they use to operate on. For example, Kara builds on finite-state automata, while KidSim and ctMazeStudio work with (visual or textual) reactive rule-based programming tools, which "use representations in the computer that are analogous to the objects being represented and to allow those representations to be directly manipulated in the process of programming" [54]. ctMazeStudio is a learning environment from the goal-oriented category following a game-based approach to create and learn about algorithms to escape mazes. The users control a virtual "Minotaur" in a microworld using different programming tools (a visual alternative with a visual reactive rule-based approach and then a visual block-based programming tool) [19]. Such an environment can be used to explore the transition between different representations.

The alternative representations of programming are mostly used to pave the way to (visual) block-based programming tools [19], because "the conceptual gap exists between the representations that people use in their minds when thinking about a problem and the representations that computers will accept when they are programmed [and] this gap is as wide as the Grand Canyon" [54].

2.2 Learnability of Computational Thinking in Game-Based Learning Environments

Learning computational thinking and acquiring all the related competences poses challenges for learners and facilitators of such environments. Therefore, we propose the use of a certain learning progression as an underlying model for the design of CT environments. For game-based environments, studies had shown the effectiveness of games in the context of CT. For example, *ctGameStudio* is a scaffolded learning environment for understanding CT concepts and abstractions in subsequent levels, which motivates the students to progress in the game [61, 63]. The concept of Scratch to author own games and to tell stories brings the learners to express their ideas in a creative way [13, 32]. LightBot offers a puzzle-like approach, which lets students reflect on their ideas and refine their solutions on a motivational level [12].

With assessment schemes like Variables-Expressions-Loops (VEL) [13] and the CT test (CTt) [10], questionnaires are given to measure the level of the learned CT competences. However, the studies of Grover and Basu [13] and Werneburg et al. [61, 63] have shown that computational constructs like variables and computational abstractions like loops are a high hurdle for students when they start programming. Therefore, it is necessary to support learners in adapting new computational models or constructs with an appropriate mechanisms of guidance and scaffolding in such learning environments.

Level systems in game-based learning environments give the opportunity to implement models like the Use-Modify-Create progression presented by Lee et al. [31]. With an adaption of this model (see Fig. 3), it is foreseen to introduce in a first level a new concept. The students can use and analyze given code (as a template) to study its behavior, which is visualized through virtual agents in the given microworld. In a next step, learners modify the given code or vary parameters to solve the first level in the progression.

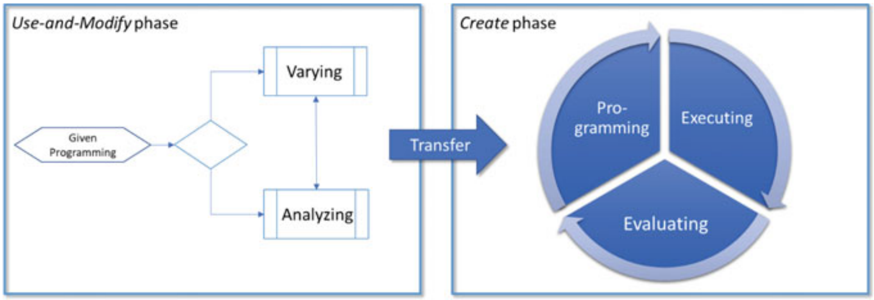


Fig. 3 Use-Modify-Create progression according to Hoppe and Werneburg [19]

This combination of using and modifying describes a customization of given programming code to become familiar with newly introduced CT concepts. In a second level, the students have to create their own programming artifacts from scratch using the respective computational concept. This deepens their knowledge about the particular concept without using a given programming code [61].

However, the transfer from using and modifying a given code to producing their own program is a big step for learners. Due to the high degree of freedom, the learners can fail because of problems that are not controllable as in the previous level, where they are pushed into the right direction by a given code template. Such a high delta in the difficulty between the two phases might cut off or prevent flow experiences on part of the learners. To address this problem, there is a need for just-in-time feedback for students and teachers to support the step-wise process of developing CT competences in such a learning progression model, because "the current goals of the user are more important than long-term interests, especially for providing just-in-time access to relevant information" [4]. Likewise, learners can become underchallenged, and this can lead to boredom and losing the pleasure of programming. To tackle this issue, we propose an extension of such a learning progression model by a "challenge" phase. Challenges and transfers to more abstract tasks are necessary to deepen the knowledge and to counteract the boredom of the underchallenged students.

In summary, we advocate for the use of a certain learning progression model for game-based learning environments. Particularly in the context of computational thinking games and playful programming, the literature has shown the usefulness of the Use-Modify-Create model. With an extension of this model through a challenge phase, we propose the 3C model ("Customize-Create-Challenge") to address particular motivational challenges. Complementary to the design of a learning progression or level system in a game-based environment, the learners need support in each step depending on their learning strategies.

2.3 Learning Strategies for Computational Thinking Games

Grover et al. identified three major strategies when students program with Blockly games: "(1) basic iterative refinement […]; (2) breaking down a problem into sub-problems and recombining the partial solutions into an overall solution; and (3) testing and debugging" [14]. Especially in the Customize-Create-Challenge model, the behavior (3) is observable by many learners in the Customize phase and primarily behavior (1) in the Create phase [61].

For behavior (2), there are two possible approaches: top-down or bottom-up. Students can define all sub-problems at the beginning and solve them after this planning phase or they start bottom-up with solving a sub-problem before they structure the following sub-problems.

Kiesmüller [28] designed a framework for Kara to give appropriate feedback (Table 2). For example, "hill climber" use behavior (1). If their code has a bad quality, they get hints to structure the code. Students using the "trial-and-error" approach (3) get an introduction to structured problem solving, and students using the "divide-and-conquer" approach (2) get hints if a branch is missing or if there is an error in the currently edited branch. Kara uses an alternative representation as a programming tool (Table 1). In the case of the presented study, providing specific tasks makes the environment goal-oriented. The visual representation as a finite-state automaton allows the easy preparation of incoming conditions in the description of the individual states at the beginning. Simple clicks allow the addition of individual new restrictions, so that learners can quickly learn the top-down method.

Table 2 Individual feedback depending on the solving strategy according to Kiesmüller [27]

Problem-solving strategy	Quality of the solution				
	Very bad	Bad	Medium	Good	Very good
Hill climbing	Hints to structure the code first		Technical error message and hints for the current sequence		
Trial and error	Hints to structured problem solving			Technical error message and hints for faulty sequence	
Top-down	Number of branches correct: notes on faulty sub-branch		Technical error message and motivating remark		
	Branch number wrong: marking the missing sub-branch				
Bottom-up	Number of branches correct: notes on faulty sub-branch		Technical error message hints for faulty sequences		
	Branch number wrong: marking the missing sub-branch				

In contrast to this, Meerbaum-Salant et al. [37] identified that learners used a bottom-up strategy while programming with Scratch. They classified this strategy as a bad habit and suggested to provide dynamic feedback while programming in order to lead students to better practices.

Werneburg et al. [63] identified two kinds of strategies applied by students using activity metrics such as the number of runs ("program executions") and the changes per run as indicators. If the students know how to solve a problem, they started with behavior (1) and solved the problem within a few runs. If they were unsure how to solve the problem, then behavior (3) was more likely to be observed and they needed more runs with fewer changes per run when they started programming. If a problem was not solved, a "third" strategy with an unstructured pattern of changes per run could be observed.

In summary, different learning strategies can result in success while programming and students prefer their own learning strategy [29]. Additionally, the use of a particular learning strategy depends on the given task and its difficulty, but learners also adapt different strategies if a new abstraction type for programming is introduced. In the presented studies, such cases lead to a less structured (hill climbing, frequent parameter variations) or even unstructured behavior (trial and error). To circumvent this, we propose to incorporate support mechanisms to force structured programming behavior and to guide learners through computational thinking processes. To implement guidance mechanisms, dynamic feedback or other interventions through a system, a precise view about CT processes and, respectively, the logical artifacts generated by the learners is necessary.

3 Automated Assessment for Computational Thinking

Computational thinking tools are environments to support the creation of logical artifacts in a given context. The field of automated assessment, particularly in the context of programming, has been framing research on computational thinking. On the one hand, it provides means to evaluate learners' artifacts during the processes that involve CT, but also enables dynamic feedback in order to improve CT competences over time. While traditional methods of assessment involve psychometric tests in order to get information about the learners' cognitive state, automated tools facilitate analytic methods, which focus on logical artifacts and programming processes captured by software systems. Román-González et al. [49] developed an assessment for research in the field of computational thinking. They identified three complementary assessment tools and mapped each tool to categories in a revised version of Bloom's taxonomy. They recommend psychometric tests like the CTt for understanding and remembering, item pools like Bebras tasks for analyzing and applying, and analytic tools like Dr. Scratch [49] for evaluating and creating.

However, the concrete tools proposed can be seen more as one example of an assessment framework for CT in the context of middle school. To develop and implement tools, which give students just-in-time feedback in learning processes involving CT, particularly when evaluating and creating code, a structured and automated analysis of their programming artifacts and processes is necessary. The automated analysis of these artifacts traditionally employs code metrics, which originates in the discipline of software science [6]. Applications in educational contexts exist, which apply metrics to identify the programming behavior by counting particular computational constructs and concepts the learners used [28, 40, 63]. Such counting metrics are a first step, but the characteristics of code are manifold. Automated assessment tools typically use a decomposition of source code into specific features that capture the respective code characteristics. Lately, the emerging field of learning analytics is framing the research in CT by combining computational methods of process and artifact analysis. The following sections give an overview about the use of code metrics in this context and present how learning analytics brought a new take on CT assessment. Finally, we present a framework for learning analytics, which combines a model for CT processes with code metrics. Such a framework can be adapted in order to improve learning by placing direct interventions or by generating feedback for learners.

3.1 Code Metrics to Analyze Students' Programming Artifacts

Traditionally, software metrics estimate the software complexity and measure costs for developing and sustaining software [15]. Ten years after the pioneering work from Halstead, there were 500 interdisciplinary references including software metrics, control structure complexity, logical complexity, and psychological complexity [56]. In the following, we focus on metrics with respect to the context of computational thinking in education. We explicitly exclude productivity-oriented measures that mainly focus on aspects of software engineering in corporate contexts.

Software Metrics are (mathematical) functions to quantify certain characteristics of source code. Some examples employ counting tokens in the source code, include program length, volume, difficulty, and language level, and combine that into an integrated and comprehensive system of formulas [15]. In addition to counting particular structural elements, such as operators, a typical metric is "lines of code," which exists in many variations. For example, such metrics can be used for the detection of code smells like long methods or empty blocks. In the context of Scratch, long methods and code duplicates decrease the comprehension and potentially complicate the modifiability of Scratch programs [17].

Control Structure Complexity was introduced by McCabe to measure the "cyclomatic complexity" of a program, because a "50 line program consisting of 25 consecutive 'IF THEN' constructs [has] as many as 33.5 million distinct control paths" [36]. For this analysis, programs are represented as control flow graphs. The cyclomatic complexity can be calculated as the number of linearly independent paths in this graph. Dealing with a lower complexity can be supportive for learners in developing their logical artifacts as it possibly prevents side effects that might occur due to high level of branching.

Logical Complexity involves data flow analysis and visualize dependencies between procedural objects and data objects [22]. These measures consider the span of dependencies of computations and the proximity of expressions dealing with the same data object. Especially in learning environments such as Scratch or Snap, hints about these complexity measures can be helpful when working with multiple sprites or first class objects, where the overview might be lost easily.

Psychological Complexity distinguishes between behavioral methodological symptoms of complexity [53] and involves the structure, indentation, choice of variable names, and documentation of programming artifacts. For example, Moreno and Robles observed bad habits in Scratch programs, for instance by not renaming sprites which resulted in difficult debugging processes in large projects, because for students it was "hard to know what object relates to a given statement" [39].

Dynamic Metrics capture the runtime behavior of a program, in contrast to static code metrics, which are based on the code structure. While static code metrics are evaluated at compile time, dynamic code metrics execute the program or approximate the runtime characteristics using heuristics. For example, an approximation of the execution can be achieved by using particular code representations such as a dynamic call graph of a program [51]. Typically, dynamic code analysis aims to forward engineer code, for example, in the case of malware detection [64], or in the optimization of runtime behavior of a program [7]. In educational contexts, other metrics and representations are used, such as the "visited Lines of Code," which quantifies—in contrast to the static lines of code metric—how many lines are visited at runtime. Such a characteristic can identify a brute-force solution [34].

3.2 Learning Analytics for Computational Thinking

Ihantola et al. presented a review of recent systems of automatic assessment of programming assignments [21]. They completed the review of Ala-Mutka [1] with tools until 2010. They followed their approach that both static and dynamic analysis to assess functionality, efficiency, and testing skills are important features for assessing systems. Thus, they categorized the systems in (1) automatic assessment systems for programming competitions and (2) automatic assessment systems for (introductory) programming education.

To categorize assessment techniques for CT game environments, we collected data by searching for phrases "Computational Thinking" AND (*specific game environment according to Table* 1) AND "automated assessment" from the conference proceedings and journals through ACM Digital Library and IEEE Xplore beginning in 2006 with the first definition of CT by Wing. We reduced the set of papers to an amount of 29 by filtering after reading the abstracts. Figure 4 shows that most papers concentrate on Scratch as a prototypical example for using visual block-based programming tools in an open-world environment. Moreno et al. contributed with Dr. Scratch a fundamental work for analyzing Scratch projects [40]. It utilizes static code metrics and the appearance of specific code constructs in order to assess computational thinking based on code artifacts. The evaluation of code artifacts poses challenges for the research in automatic assessment of CT skills and code quality. Limitations exist when it comes to assessing the correctness of a solution or the accomplishment of a task with a certain code artifact.

Therefore, we categorize a microworld concept as an open task or as goal-oriented. Apart from the microworld concept, the research in the field of CT assessment is framed by using different programming tools that depend on different representations of the code artifacts. We classified popular examples of CT learning environments depending on the microworld concept and the used programming tool according to the typology elaborated earlier on in Table 1.

It is challenging for software tools in open task environments to check the correctness of a code artifact automatically due to the degrees of freedom and the lack of a computational, closed solution. Defining milestones, requirements and conditions are helpful to create a framework for the assessment. Such requirements might include the appearance of a specific programming concept or the integration of a certain code feature. As a caveat, the assessment has to be interpreted with care because a code artifact can be "good" due to the existence of a certain code feature but the solution can be insufficient to the respective task specification, it contains semantic errors, or even unveils a low understanding on the part of the learner.

By setting concrete goals for a task that can be checked automatically, such environments enable the incorporation of correctness with respect to specific features as a characteristic to the assessment. This gives the opportunity to compare different solutions automatically and to define more specific indicators to assess the quality of a solution.

Fig. 4 Word cloud of the tf-idf-weighted words in the analyzed papers

Text-based Programming Tool Blikstein analyzed the programming behavior of students with text-based programming tools in open-ended environments [2]. He presented the behavior of students and analyzed the code events and non-code events like compilation frequency, code size, code evolution patterns, frequency of correct/incorrect compilations, etc., to find prototypical coding profiles and styles. Jamil presents in his work the MindReader system, which matches code fragments, analyzes data flow, and tests randomly the code to identify possibly valid solutions [23]. For goal-oriented environments like Karel the robot or RoboCode, there are no automated assessment tools available. These environments allow learners to create algorithms in an open sandbox, which allows generating an infinite range of possible solutions and combinations—despite having a goal-orientated microworld concept. However, it can be possible to some extent to transfer the ideas of Blikstein and Jamil with a specific focus on particular sub-goals to specify.

Visual Block-Based Programming Tool As can be seen in Fig. 4, most of the papers which are in the intersection of the categories "visual block-based programming tool" and "open task environment" were dominated by Scratch. The work of Moreno et al. [40] is a central point in this collection. They designed metrics for different computational practices, which count the use of certain syntactical programming constructs to identify the students' CT level. Ota et al. [42], Seiter, and Foreman [52] follow similar approaches. Meerbaum-Salant et al. [37] identified as bad habit "extremely fine-grained programming"—students used few scripts while programming and the authors recommend decomposition strategies. Werneburg et al. [63] presented different features for analyzing students' programming behavior in the goal-oriented environment *ctGameStudio*. Activity metrics are used to define features, for example to describe the advanced planning behavior with the feature *# changes per run* (Table 3).

Weintrop and Wilensky [59] also analyzed visual block-based programming tools for a goal-oriented environment using metrics to count blocks added to a program. A difference between successful and unsuccessful runs could be observed. In each time slice, they added 1.6 blocks to the program and increased its complexity. Between successful battles, the learners added 3.3 blocks. This example shows that such simple metrics might serve as an indicator to predict the quality of the outcome.

Table 3 Interpretation of the activity metrics according to Werneburg et al. [61]

Indicator	Interpretation
# Runs	Testing and evaluating behavior of the created programming code
# Changes per run	Trial-and-error behavior or advanced planning
# Creates	Active extensions of the programming solution
# Consecutive changes per create	Structured editing
Time spent in minutes	Measure for efficiency

Visual Alternatives In contrast to the previous categories of tools, "visual alternatives" are domain-specific programming tools that do not simply mask text-based logical artifacts. They provide a domain-specific graphical user interface far away from conventional programming, such as pictorial input methods or flowcharts. The *KidSim* environment features so-called graphical rewrite rules [55], which show the current situation ("before") and the subsequent situation ("after"). Because of the domain-specific character of such tools, it is challenging to define general methods for the analysis of such environments. However, metrics for counting like in the work of Kiesmüller [28] help to classify learners' behavior. They used activity-time analytics to identify behavioral patterns of the programming processes in order to provide appropriate feedback to the learners. Although the environment used for the evaluation is open-ended, the design for the evaluation has a strong goal orientation.

Another example of a visual alternative programming tool is the goal-oriented environment *ctMazeStudio*. In order to accomplish the given task, the learners have to escape mazes by discovering and applying maze algorithms, such as the wall following strategy. The learners create so-called reactive rules when they discover a new situation for the virtual agent in the microworld, where none of the already defined rules matches the current situation. When the learners proceed in finishing mazes, more complexity such as cycles is added to the maze. By advancing, the learners try to generalize their rule set in order to create a strategy as general as possible. Finally, it should be possible for them to solve all mazes with the same rule set that consists of four rules. In this environment, the analysis of the logical artifacts, which is the current rule set of the learners, is embedded into the design. Based on an activity log, loops in the solution are detected, which triggers dynamic feedback for the learner. The addition, modification and deletion of rules are captured data to analyze the programming behavior.

In summary, combining all of these approaches has the potential to foster CT in game-based environments. We highlight the importance of combining product- and process-oriented analyses using code and activity metrics that are as specific as necessary for the particular programming tool. With useful and supportive interventions, which facilitate the analytics, it is possible to support the development of computational thinking competences on the part of the learners. Examples of such interventions can be hints for a clear programming style, hints to check the (partial) correctness of solutions, or hints on how to structure code. We propose a framework for learning analytics in computational thinking games that facilitates the combination of such analytics with the respective thinking processes.

3.3 Framework for Learning Analytics in Computational Thinking Games

Learning analytics (LA) can be seen as the use of analytic methods on learning data targeting various stakeholders with the aim to improve learning processes and environments [8]. Typically, three different types of computational methods are used in LA according to Hoppe [20]: (1) the analysis of content, (2) the analysis of processes using methods of sequence analysis, and (3) the analysis of social network structures. The latter is of minor interest, as the scenarios and examples of computational thinking environments that have been part of this investigation do not incorporate social or collaborative aspects.

However, this model needs to be adapted to an application-specific context. Programming activities are iterative processes, where the learners create executable software artifacts in each iteration, which can lead to a complete or a partial solution for a given problem or task. This interleaves the aspects product and process of the model by Hoppe [20]. Therefore, we align this model to the process of programming according to the creation cycle in the Use-Modify-Create progression (Fig. 3) and incorporate product- and process-oriented analyses.

In the cyclic process of programming, learners create code artifacts, execute them, and evaluate the respective results corresponding to their expectation and to their respective goal orientation. If the results satisfy their reception of a goal or provide a solution to a given, closed problem, the learners might decide to end their programming. If it does not satisfy the given conditions, they might want to refine the code and re-enter the cycle.

Early approaches for the automated assessment focus on the use of static code metrics in order to analyze programming artifacts. In addition to the analysis of static code features, dynamic metrics have the potential to capture runtime information of code artifacts. Therefore, analytical models for computational thinking might comprise programming behavior of the learner (analysis of processes) and runtime behavior of a program (analysis of products). In this sense, the analysis of runtime behavior is still located in the "product" aspect of the learning analytics model; it counts as a method of content analysis using dynamic code analysis.

To improve learning through the embedding of a learning analytics framework, it is necessary to incorporate the results into the learning situation, for example, by providing direct interventions or feedback. Corresponding to the cyclic programming model that consists of the phases *programming*, *evaluating*, and *executing*, various analytics methods can affect specific phases.

Figure 5 shows the alignment of analytics methods to this model and outlines the different possibilities to provide feedback to the learner. During the execution, the learner has to realize whether he has solved the problem, for example, when he has found a (task-specific) solution. Particularly in open-ended environments, this is difficult to check. Goal-oriented environments circumvent this by offering a more structured path, for example, through a story mode or a level structure.

However, in open task environments, unit tests have the potential to support the evaluation of solutions. In addition, partial solutions could be identified by calculating similarities to reference solutions. Source code similarities have been part of the research in the field of software metrics, for example, on calculating differences on abstract syntax tree [46], but also in the field of plagiarism detection [5, 24].

Psychological aspects of coding can be beneficial for the evaluation process. Such aspects targeting the psychological complexity improve the readability and

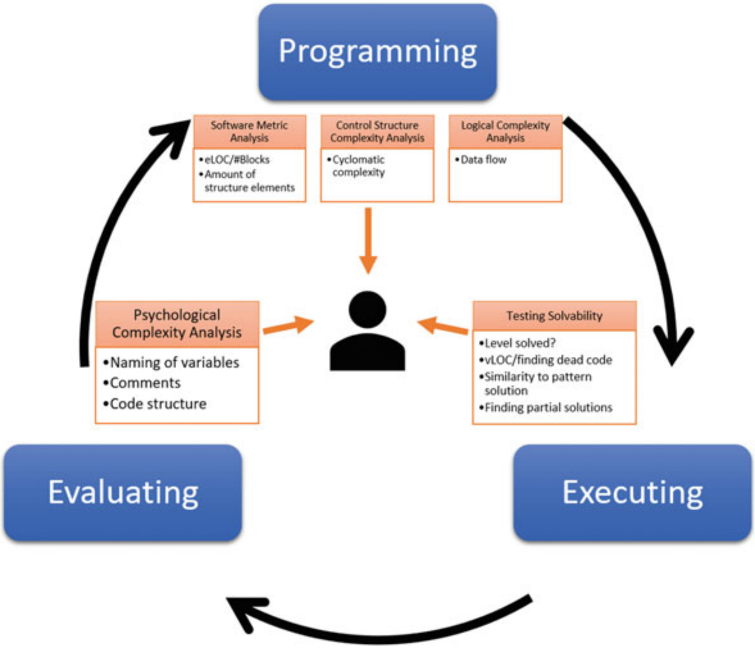


Fig. 5 Assessment framework for computational thinking games

maintainability of the code. When learners have to debug their solutions, a higher complexity limits the advances to find proper solutions. In addition to aspects like the meaningful naming of variables or the co-existence of comments, this consists of structural aspects such as redundant or dead code. The programming phase can be supported by using software metrics. Such metrics can be embedded in a way that they prompt the learners to create an awareness of possible problems or flaws. Examples can be long methods or incorrect branching in conditional statements. For such cases, typical static code metrics such as cyclomatic complexity can be facilitated. However, the specific composition of metrics and the design of interventions and feedback are highly domain-specific and need to be adapted to the particular tool.

4 Evaluation of Programming Behavior

In this work, we presented an overview of game-based computational thinking environments and tools to assess computational thinking competences to learners. We aligned cognitive models of programming to analytical methods and created a framework for learning analytics for such game-based environment. Using this framework, we evaluated an experimental study that has been rolled out at the *ScienceNight Ruhr*, one of the largest science festivals in Germany.[1] During the study, participants used the *ctGameStudio* in a self-regulated learning scenario after a brief introduction about computational thinking and the game-based environment.

4.1 Experimental Setting

The aim of this evaluation is to get a better understanding of students' learning progression to build interventions and feedback using the proposed learning analytics model. The analysis of the programming behavior is based on activity metrics, and the product analysis consults the final solutions of learners.

In contrast to a supervised setting, the students had no time constraints for the programming. For such events, it is mandatory that participants are allowed to leave at any time. During this experiment, 54 students used the *ctGameStudio* and more than a half of the participants played longer than 30 min ($M = 34.5$, $SD = 24.22$). Here, individual subjects took up to nearly two hours of time (cf. Fig. 6).

During the experiment, the participants had the opportunity to use the *RoboStory*-mode of the *ctGameStudio*, a guided learning environment with a level system focusing on several abstraction types. As described earlier, this environment consists of a microworld with a virtual agent (the robot), which can be controlled by the learner using visual block-based programming. Each level targets a specific computational construct or abstraction type such as *loops* or *object types* to be learned and applied by the user. Most of the levels have a dedicated target the robot has to reach. In the first level, the player has to use a "moveForward" command with a particular range. In further levels, the user needs to facilitate loops to move in the shape of a square, to follow another character using conditionals, or to scan objects using event listeners and distinguishing object types. A specialty of this environment is that the program created by the learner needs to provide a solution in advance. When the learner clicks the "run" button to execute the code, there is no more chance to interactively modify anything in the environment. Therefore, the whole behavior of the virtual agent needs to be programmed in advance. Consequently, a solution provided by a learner is an

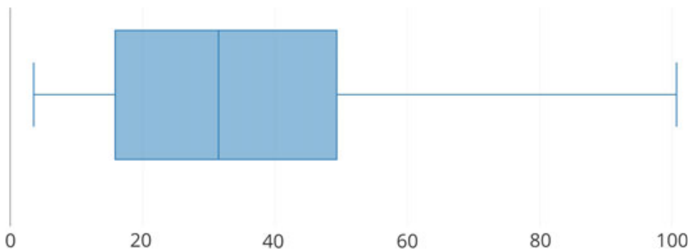


Fig. 6 Distribution of the time spent in minutes, how long the students used *ctGameStudio*

[1] *ScienceNight Ruhr*, science festival: https://www.wissensnacht.ruhr/english/ retrieved 2019-02-18.

algorithm for a particular aspect. The task design of the ctGameStudio follows the previously introduced 3C model (“Customize-Create-Challenge”). Each level represents one of these computational constructs, such as loops. Every level consists of sub-levels, where the first one follows the “customize” phase. For the example of loops, a given template with a loop is presented to the learner and the level can be solved customizing the given code. In the next sub-level, the “create” phase, the learners have to write their own code from scratch using the particular concept. With the final sub-level, the learners can master each concept by having a more difficult task around this concept in a “challenge” phase.

4.2 Analysis of the Learner Progression

During the experiment using the *ctGameStudio*, different kinds of data have been collected through the system. According to Blikstein [2], we collected coding events as well as non-coding events from the users. Examples of coding events are creating a block while programming and non-coding events are for example clicking on the “run” button to execute the code [61].

The feature *# runs* describes the testing and evaluation behavior of the learner when creating programming code. One observation of a prior study was that students needed less runs if the level was easy to solve [63]. The boxplots in Fig. 7 show the distribution of the number of runs per sub-level.

The first sub-level of each unit (1.1, 2.1, and 3.1) contains the “customize” phase, where the students use and modify given programming code. The learners should try out the given code and then modify it in order to get a correct solution. Especially in level 2.1 and 3.1, the boxplots show that much more runs were needed than in the respective subsequent sub-levels, which contain the creation part of the Customize-

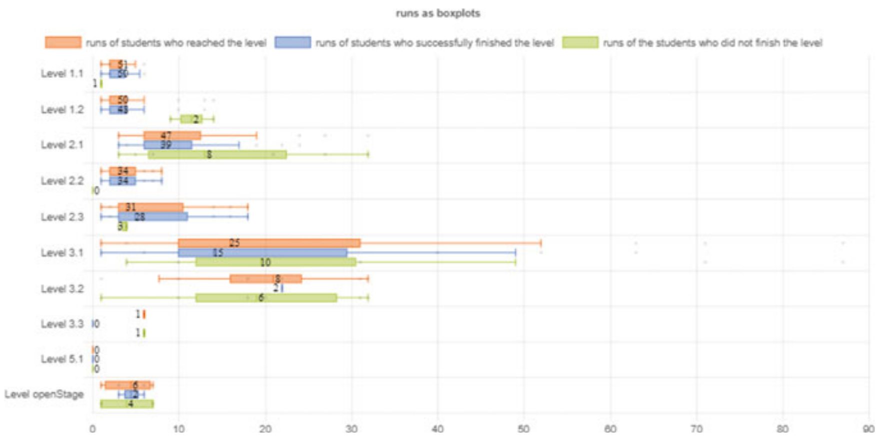


Fig. 7 Distribution of the runs of the students per sub-level in *ctGameStudio*

Create-Challenge progression model. Level 2.3 was the only challenge level, which was tried out. The students needed more runs than before, because of the increasing degree of difficulty. For example, the median of the number of runs was in level 2.1 at 8, fell down in level 2.2–3, and increased in level 2.3–5 by students who successfully finished the level. Level 2.1 was not finished by 8 of 47 participants. The median for this group was 13. To help students with more guidance, 8 runs can be a threshold for this sub-level to provide interventions after this number of runs to guide them to a correct solution. However, there were also students who left the game after 3 runs. This can be regarded as a special case of the setting, as the subjects could always cancel.

To analyze how the approach of the Customize and Create phases was adopted, we captured all changes per run and with the feature *# changes per run*, the advanced planning behavior can be analyzed. This includes creating a block, moving a block in the structure of the program, deleting a block, and varying a parameter.

Level 2.1 is a level to customize a given source code related to loops. As can be seen in Fig. 8, only a few persons did for the first run of level 2.1 many changes. In average they did 3.02 changes in their first run ($SD = 6.33$). 57% of the participants made no changes, and only 28% of the students did more than four changes. For the second run, in average 5.6 changes were done by the participants ($SD = 7.56$) and only 14% of the participants did no change for this run. In the following runs, the participants did in average 2.4 changes ($SD = 1.23$). In summary, the idea of using the given source code before modifying was applied by most of the users. They analyzed the behavior of the robot acting in the microworld and modified the given code after analyzing it for the second run. After the second run, mainly a refinement of the code could be observed.

Level 2.2 is a level to create source code from scratch without starting from a given code template. However, in the previous level, the learners explored the newly introduced abstraction type “loop.” It could be observed that the users did in average 13.8 changes per run ($SD = 11.89$) before the first run, and before each of the following runs 4.3 changes per run ($SD = 5.55$). In this case, they were able to

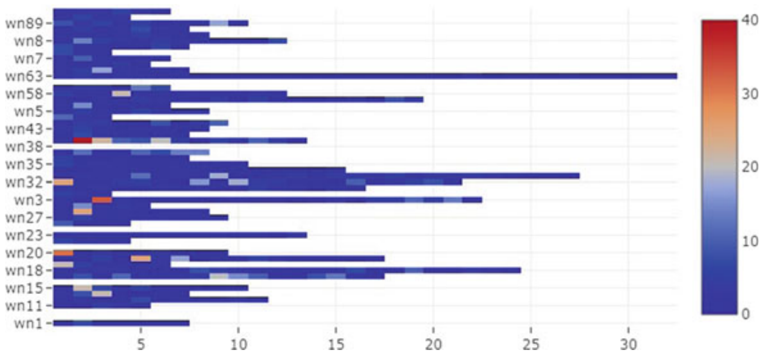


Fig. 8 Changes per run in the customize phase in level 2.1

construct their idea to solve the level at the beginning and only did some refinements of their solutions in the subsequent runs (Fig. 9).

Table 4 shows the results for all levels tested by the participants. It presents the number of runs (minimum, average, maximum) in all sub-levels as well as the changes per run as regression line.

The sub-levels of unit 1 were easy to solve for the students so that they incorporated most of the changes before the first run. Changes for the following runs can be seen primarily as refinements of the code and do not involve many changes in the semantics. The first two sub-levels of unit 2 have been discussed in detail above. However, level 2.3 was a level, which contained a "challenge" for the students. It is based on the previous level (2.2) and requires the learners to build a similar algorithm, but for a generalized problem. As expected, the students needed in average more runs and made more changes per run after the first run. In unit 3, the abstraction type "event" has been introduced. Although the learners could start with a given code template in the first sub-level, it could not be observed that this had an impact on the runs in the following levels. In this example, we can observe an abstraction gap,

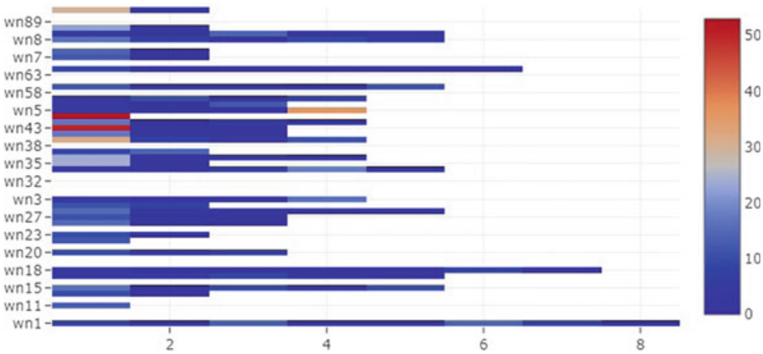


Fig. 9 Changes per run in the create phase in level 2.2

Table 4 Runs of the students using the *ctGameStudio*

Level	Runs			Regression of the changes per run
	Minimum	Average	Maximum	
1.1	2	4.30	9	$y = -1.1x + 9.9$
1.2	2	5.48	15	$y = -0.45x + 11.4$
2.1	4	11.795	35	$y = 0.03x + 7.4$
2.2	1	4.61	19	$y = -0.88x + 10.89$
2.3	1	7.06	19	$y = -0.277x + 11.46$
3.1	1	24	88	$y = 0.09x + 6.86$
3.2	2	20.5	33	$y = 0.19x + 8.22$
3.3	7	7	7	$y = 10.6x - 9.67$

which is a major hurdle for the learners. This unveils that additional mechanisms to guide the learners are necessary in order to force the understanding of the specific abstraction during the programming process. However, only 8 of the 15 students who successfully finished level 3.1 started level 3.2. To evaluate the sub-levels that are later in the learning progression, another study might be needed. As a limitation of this study—in analogy to other work in this field—computational thinking is a thought process and can only be captured and learned to some extent. In this case, the duration of the experiment was limited to 60 min. More long-term studies are needed to re-design curricula and environments in order to foster CT effectively. Otherwise drawing definitive and generalizing conclusions from such experiments needs to be handled with caution. Still, such conclusions are useful to design and re-design levels or guidance mechanisms, and to fine-tune parameters like thresholds for prompts and interventions.

5 Conclusion

In this work, we presented the state of the art of game-based environments to foster computational thinking. With a classification of the game-based CT environments related to microworld concepts as well as to the used programming tools, we identified different learning strategies throughout programming processes. Additionally, we presented in this paper various learning analytics techniques to characterize and analyze the logical artifacts created by learners and the corresponding processes these artifacts where created in. The programming process is influenced by the design of the kind of task the learner has been assigned to. Particularly for task-oriented environments, we proposed the use of a learning progression model ("Customize-Create-Challenge"). An essential part of the Customize-Create-Challenge model is the create phase, where students program, execute, and evaluate their solutions. For this step of creating, we assigned different metrics to provide tools to analyze the students' behavior and to develop appropriate feedback. As a first step, we used activity metrics and implemented features such as # runs and # changes per run in the investigated CT game ctGameStudio. In an experimental study using this environment, we have shown at which points students struggle while programming during the ScienceNight Ruhr. To circumvent this, interventions have to be placed to successfully improve learning and to observe a gain in CT. Additional metrics to analyze the programming artifacts are the next step for a refinement for dynamic and adaptive guidance, scaffolds, and feedback. The use of dynamic and static code metrics for a similar use case has been proposed in the context of creative problem solving with programming, which demands CT competences [34].

However, embedding such direct interventions that aim to interfere with the actual learning needs a concrete alignment to the pedagogical model that underlies the level and task design in addition to the guidance and scaffolding framework. Therefore, we presented a three-layered approach that combines the (1) task design, (2) guidance mechanisms, and (3) learning analytics. Following this model, the evaluation presented in this work can be used to parameterize the guidance framework of the ctGameStudio. For example, the median of the number of runs in a cluster of low-performing learners might serve as a threshold for direct interventions. In addition to the assessment of computational thinking skills to the learners, such indicators can be useful to fine-tune and tweak the guidance mechanisms for future use.

Acknowledgements We dedicate this publication to the memory of Sören Werneburg who designed and developed the ctGameStudio and ctMazeStudio environments and conducted this study.

References

1. Ala-Mutka, K. M. (2005). A survey of automated assessment approaches for programming assignments. *Computer Science Education, 15*(2), 83–102.
2. Blikstein, P. (2011). Using learning analytics to assess students' behavior in open-ended programming tasks. In: *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, (pp. 110–116). LAK '11, ACM, New York, NY, USA.
3. Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. In: *Proceedings of the 2012 Annual Meeting of the American Educational Research Association*, (Vol. 1, p. 25). Canada: Vancouver.
4. Budzik, J., & Hammond, K. J. (2000). User interactions with everyday applications as context for just-in-time information access. In: *Proceedings of the 5th International Conference On Intelligent User Interfaces*, (pp. 44–51). ACM.
5. Cosma, G., & Joy, M. (2012). An approach to source-code plagiarism detection and investigation using latent semantic analysis. *IEEE Transactions on Computers, 61*(3), 379–394.
6. Curtis, B. (1981). The measurement of software quality and complexity. *Software Metrics: An Analysis and Evaluation,* pp. 203–224.
7. Engler, D. R. (1996). Vcode: A retargetable, extensible, very fast dynamic code generation system. *SIGPLAN Notices, 31*(5), 160–170.
8. Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning, 4*(5/6), 304–317.
9. Feurzeig, W., et al. (1969). *Programming-Languages as a Conceptual Framework For Teaching Mathematics*. Final report on the first fifteen months of the logo project.
10. González, M. R (2015). Computational thinking test: Design guidelines and content validation. In: *Proceedings of EDULEARN15 Conference*, (pp. 2436–2444).
11. Gorman, H., & Bourne, L. E. (1983). Learning to think by learning logo: Rule learning in third-grade computer programmers. *Bulletin of the Psychonomic Society, 21*(3), 165–167.
12. Gouws, L. A., Bradshaw, K., & Wentworth, P. (2013). Computational thinking in educational activities: An evaluation of the educational game light-bot. In: *Proceedings of the 18th ACM Conference on Innovation and Technology in Computer Science Education*, (pp. 10–15). ITiCSE '13, ACM, New York, NY, USA.
13. Grover, S., & Basu, S. (2017). Measuring student learning in introductory block-based programming: Examining misconceptions of loops, variables, and boolean logic. In: *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, (pp. 267–272). ACM.
14. Grover, S., Bienkowski, M., Niekrasz, J., & Hauswirth, M. (2016). Assessing problem-solving process at scale. In: *Proceedings of the Third (2016) ACM Conference on Learning @ Scale*, (pp. 245–248). L@S '16, ACM, New York, NY, USA.
15. Halstead, M. H., et al. (1977). *Elements of software science (operating and programming systems series).* New York, NY: Elsevier Science Inc.

16. Hartmann, W., Nievergelt, J., & Reichert, R. (2001). Kara, finite state machines, and the case for programming as part of general education. In: *Proceedings IEEE Symposia on Human-Centric Computing Languages and Environments, 2001*, (pp. 135–141). IEEE.
17. Hermans, F., & Aivaloglou, E. (2016, May). Do code smells hamper novice programming? A controlled experiment on scratch programs. In: *2016 IEEE 24th International Conference on Program Comprehension (ICPC),* (pp. 1–10).
18. Hollingsworth, J. (1960). Automatic graders for programming classes. *Communications of the ACM, 3*(10), 528–529.
19. Hoppe, H. U., & Werneburg, S. (2018). Computational thinking—More than a variant of scientific inquiry. In: S.-C. Kong, H. A. (Ed.), *Computational Thinking Education*. Springer.
20. Hoppe, U. (2017). Computational methods for the analysis of learning and knowledge building communities. In: C. Lang, G. Siemens, A. F. Wise, & D. Gaevic, (Eds.), *The Handbook of Learning Analytics*, (1st ed., pp. 23–33). Society for Learning Analytics Research (SoLAR), Alberta, Canada. http://solaresearch.org/hla17/hla17-chapter1.
21. Ihantola, P., Ahoniemi, T., Karavirta, V., & Seppa¨l¨a, O. (2010). Review of recent systems for automatic assessment of programming assignments. In: *Proceedings of the 10th Koli Calling International Conference on Computing Education Research*, (pp. 86–93). Koli Calling '10, ACM, New York, NY, USA.
22. Iyengar, S. S., Parameswaran, N., & Fuller, J. (1982). A measure of logical complexity of programs. *Computer Languages, 7*(3–4), 147–160.
23. Jamil, H. M. (2017, July). Automated personalized assessment of computational thinking mooc assignments. In: *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*, (Vol. 00, pp. 261–263).
24. Ji, J. H., Woo, G., & Cho, H. G. (2007). A source code linearization technique for detecting plagiarized programs. In: *ACM SIGCSE Bulletin*, (Vol. 39, pp. 73–77). ACM.
25. Jost, B., Ketterl, M., Budde, R., & Leimbach, T. (2014). Graphical programming environments for educational robots: Open roberta-yet another one? In: *2014 IEEE International Symposium on Multimedia (ISM)*, (pp. 381–386). IEEE.
26. Kazimoglu, C., Kiernan, M., Bacon, L., & Mackinnon, L. (2012). A serious game for developing computational thinking and learning introductory computer programming. *Procedia-Social and Behavioural Sciences, 47,* 1991–1999.
27. Kelleher, C., Pausch, R., Pausch, R., & Kiesler, S. (2007). Storytelling alice motivates middle school girls to learn computer programming. In: *Proceedings of the SIGCHI conference on Human Factors In Computing Systems*, (pp. 1455–1464). ACM.
28. Kiesmüller, U. (2008). Automatisierte identifizierung der Problemlösestrategien von Programmieranfängern in der Sekundarstufe I. In: *DDI*, (pp. 33–43).
29. Kolb, D. A. (1976). *Learning style inventory technical manual*. MA: McBer Boston.
30. Kölling, M. (2010). The Greenfoot programming environment. *ACM Transactions on Computing Education (TOCE), 10*(4), 14.
31. Lee, I., Martin, F., Denner, J., Coulter, B., Allan, W., Erickson, J., et al. (2011). Computational thinking for youth in practice. *ACM Inroads, 2*(1), 32–37.
32. Lewis, C. M. (2010). How programming environment shapes perception, learning and goals: Logo vs. scratch. In: *Proceedings of the 41st ACM Technical Symposium On Computer Science Education*, (pp. 346–350). ACM.
33. MacLaurin, M. B. (2011). The design of Kodu: A tiny visual programming language for children on the xbox 360. In: *ACM Sigplan Notices*, (Vol. 46, pp. 241–246). ACM.
34. Manske, S., & Hoppe, H. U. (2014, July). Automated indicators to assess the creativity of solutions to programming exercises. In: *2014 IEEE 14th International Conference on Advanced Learning Technologies*, (pp. 497–501).
35. Martin, R. C. (2009). *Clean code: A handbook of agile software craftsmanship*. Pearson Education.
36. McCabe, T. J. (1976, December). A complexity measure. *IEEE Transactions on Software Engineering, SE-2*(4), 308–320.

37. Meerbaum-Salant, O., Armoni, M., & Ben-Ari, M. (2011). Habits of programming in scratch. In: *Proceedings of the 16th Annual Joint Conference on Innovation and Technology in Computer Science Education*, (pp. 168–172). ITiCSE '11, ACM, New York, NY, USA.
38. Mönig, J., & Harvey, B. (2018). *Snap!: A Visual, Drag-and-Drop Programming Language*. http://snap.berkeley.edu/snapsource/snap.html. Last accessed October 24, 2018.
39. Moreno, J., & Robles, G. (2014, October). Automatic detection of bad programming habits in scratch: A preliminary study. In: *2014 IEEE Frontiers in Education Conference (FIE) Proceedings*, (pp. 1–4).
40. Moreno-León, J., & Robles, G. et al. (2015). Analyze your scratch projects with Dr. scratch and assess your computational thinking skills. In: *Scratch Conference*, (pp. 12–15).
41. Nelson, M., & Larsen, F. N. (2001). Robocode. *IBM Advanced Technologies*.
42. Ota, G., Morimoto, Y., & Kato, H. (2016, September). Ninja code village for scratch: Function samples/function analyser and automatic assessment of computational thinking concepts. In: *2016 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, (pp. 238–239).
43. Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books, Inc.
44. Pattis, R. E. (1981). *Karel the robot: A gentle introduction to the art of programming*. Wiley & Sons, Inc.
45. Peppler, K., & Kafai, Y. (2005). Creative coding: Programming for personal expression. *Retrieved August 30*(2008), 314.
46. Peters, L. (2005). Change detection in xml trees: A survey. In: *3rd Twente Student Conference on IT*.
47. Repenning, A. (1993). Agentsheets: A tool for building domain-oriented visual programming environments. In: *Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors In Computing Systems*, (pp. 142–143). ACM.
48. Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., et al. (2009). Scratch: Programming for all. *Communications of the ACM, 52*(11), 60–67.
49. Román-González, M., Moreno-León, J., & Robles, G. (2017). Complementary tools for computational thinking assessment. In: S. C Kong, J Sheldon, & K. Y Li (Eds.), *Proceedings of International Conference on Computational Thinking Education (CTE 2017)*, (pp. 154–159). The Education University of Hong Kong.
50. Rose, S., Habgood, J., & Jay, T. (2018). Pirate plunder: Game-based computational thinking using scratch blocks. In: *Proceedings of the Academic Conferences and Publishing International Limited*.
51. Ryder, B. (1979). Constructing the call graph of a program. *IEEE Transactions on Software Engineering, 5,* 216–226.
52. Seiter, L., & Foreman, B. (2013). Modeling the learning progressions of computational thinking of primary grade students. In: *Proceedings of the Ninth Annual International ACM Conference on International Computing Education Research*, (pp. 59–66). ICER '13, ACM, New York, NY, USA.
53. Shneiderman, B. (1977). Measuring computer program quality and comprehension. *International Journal of Man-Machine Studies, 9*(4), 465–478.
54. Smith, D. C., Cypher, A., & Schmucker, K. (1996). Making programming easier for children. *Interactions, 3*(5), 58–67.
55. Smith, D. C., Cypher, A., & Spohrer, J. (1994). Kidsim: Programming agents without a programming language. *Communications of the ACM, 37*(7), 54–67.
56. Waguespack, L., Jr., & Badiani, S. (1987). Software complexity assessment: Annotated bibliography. *SIGSOFT Softw. Eng. Notes, 12*(4), 52–71.
57. Wang, L., & Chen, M. (2010). The effects of game strategy and preference matching on flow experience and programming performance in gamebased learning. *Innovations in Education and Teaching International, 47*(1), 39–52.
58. Weintrop, D., & Wilensky, U. (2012). Robobuilder: A program-to-play constructionist videogame. In: *Proceedings of the Constructionism 2012 Conference*. Athens, Greece.

59. Weintrop, D., & Wilensky, U. (2014). Program-to-play videogames: Developing computational literacy through gameplay. In: *Proceedings of the 10th Games, Learning, and Society Conference*, (pp. 264–271).
60. Weintrop, D., & Wilensky, U. (2015). To block or not to block, that is the question: Students' perceptions of blocks-based programming. In: *Proceedings of the 14th International Conference on Interaction Design and Children*, (pp. 199–208). ACM.
61. Werneburg, S., Manske, S., Feldkamp, J., & Hoppe, H. U. (2018). Improving on guidance in a gaming environment to foster computational thinking. In: *Proceedings of the 26th International Conference on Computers in Education*. Philippines.
62. Werneburg, S., Manske, S., & Hoppe, H. U. (2018). ctstudio. https://ct.collide.info. Last accessed October 24, 2018.
63. Werneburg, S., Manske, S., & Hoppe, H. U. (2018). ctGameStudio—A game-based learning environment to foster computational thinking. In: *Proceedings of the 26th International Conference on Computers in Education*. Philippines.
64. Willems, C., Holz, T., & Freiling, F. (2007). Toward automated dynamic malware analysis using cwsandbox. *IEEE Security and Privacy, 5*(2), 32.
65. Wing, J. M. (2006). Computational thinking. *Communications of the ACM, 49*(3), 33–35.

Chapter 11
Motivational Factors Through Learning Analytics in Digital Game-Based Learning

Rafael Luis Flores, Robelle Silverio, Rommel Feria and Ada Angeli Cariaga

Abstract As learning analytics is still an emerging discipline, there is a lack of a standardized method for its data collection and analysis, especially in educational games where players' data can vary greatly. This paper presents an LA model for determining students' motivation within a game-based learning environment by analyzing their in-game data. In the proposed model, three motivational factors are assessed: goal orientation, effort regulation, and self-efficacy. This paper also presents implementations of the game Fraction Hero developed using the RPG Maker MV engine as well as the Learning Analytics system and dashboard. In the experiment, thirty-one Grade 6 students from the University of the Philippines Integrated School were asked to answer a 10-item survey about their self-perceived motivation toward solving fraction problems, and afterwards play the game for data collection. Based on the results, it was revealed that the students' in-game motivation was significantly higher than their self-perceived motivation.

1 Introduction

Thanks to the increasing popularity of e-learning systems, Learning Analytics (LA) has drawn the attention of researchers for its potential for optimizing teaching methods and assessing student learning behaviors [1]. In addition, several studies have recently turned their attention to Digital Game-Based Learning (DGBL), an emerging e-learning trend which incorporates serious learning with interactive entertainment

R. L. Flores (✉) · R. Silverio · R. Feria · A. A. Cariaga
Department of Computer Science, College of Engineering, University of the Philippines, Diliman, Philippines
e-mail: rafaelluis6198@gmail.com

R. Silverio
e-mail: robelle.silverio25@gmail.com

R. Feria
e-mail: rpferia@up.edu.ph

A. A. Cariaga
e-mail: adcariaga@up.edu.ph

A. Tlili and M. Chang (eds.), *Data Analytics Approaches in Educational Games and Gamification Systems*, Smart Computing and Intelligence,
https://doi.org/10.1007/978-981-32-9335-9_11

[2]. These games are rich sources of educational data [3] and, thanks to their highly interactive nature, are ideal for capturing students' interaction data for the purposes of better understanding their learning behaviors and processes [4].

That being said, it is possible to use learning analytics to measure students' motivation toward learning, which can be done by looking at behaviors indicating either high or low motivation [5]. This paper presents motivation as a "second-order" variable [6], being dependent on three factors: goal orientation, effort regulation, and self-efficacy. The indicating behaviors defined in the LA model are classified based on the aforementioned factors.

The LA system will output scores for overall motivation, as well as for each factor based on its model. The students' self-perceived motivation scores based on their survey responses will also be evaluated using these factors, for comparison.

2 Related Works

2.1 Student Motivation

Many scholars agree that there are two distinctive types of motivation: intrinsic and extrinsic. **Intrinsic motivation** refers to the desire to do something for the natural enjoyment aroused from the involvement in the activity per se, while **extrinsic motivation** refers to the engagement in an activity as an instrumental means to an end [7]. Motivation is generally considered a crucial factor in determining students' success and achievement. A higher motivation to learn has been linked not only to better academic performance but to greater conceptual understanding as well [8]. As a psychological construct, it energizes, directs, and sustains behavior toward a certain goal [5]. In the context of learners, it is no doubt important as it is a main driver toward academic achievement. However, due to its nature, it cannot be directly observed but inferred from overt behavior of the learner [9]. As such, it is very difficult to measure, especially in online learning environments where interactions are largely virtual.

As a second-order variable, motivation can be measured based on factors such as attitudes and perceived goals. According to Richardson, there are three key motivational factors [10]: (1) **goal orientation**, which is the learner's perceived reason for pursuing their achievement [11], (2) **effort regulation**, the learner's engagement in learning activities [12], and (3) **self-efficacy**, the belief in one's own capabilities to perform a task successfully [13]. All three factors must be taken into consideration in order to accurately assess student motivation.

2.2 Digital Game-Based Learning

In recent years, numerous researchers have turned to digital games as tools for motivating students to learn certain subjects (e.g., mathematics). Digital Game-Based Learning is an e-learning trend that connects educational content and video games and makes learning difficult topics more accessible, engaging and enjoyable [6, 14]. In general, educational games have a high impact on learners' engagement and motivation due to their entertainment aspect [15]. Moreover, the emergence of learning analytics in the field of educational games has introduced an increasing demand for the collection, analysis and presentation of the in-game data in multiple ways [3].

2.3 Learning Analytics

Learning Analytics as a discipline has opened new possibilities in the field of DGBL. E-learning management systems as well as educational games can be used to generate data that can potentially be harnessed and used to better understand student learning [16]. The use of LA has implications in how educational games determine motivation, as it can be observed through learner activity within the game environment rather than using a typical questionnaire [17]. Although motivational factors have typically been evaluated using questionnaires, this method has been heavily criticized by several studies, as students' self-perceived motivation may not correspond to reality [5]. A study conducted by Cano et al. [18] introduced the GLAID (Game Learning Analytics for Intellectual Disabilities) Model, which describes "how to collect, process and analyze video-game interaction data in order to provide an overview of the user learning experience, from an individualized assessment to a collective perspective [18]." In other words, all user interactions within video games can be collected and analyzed to be used for future assessment. Signals or game observables that can give useful information about a player's learning behavior are identified, such as timestamps, level changes, achievements, fails and other user interactions. There is still a need for empirical evidence as this is only a theoretical adaptation, and the field of Game Learning Analytics is still performed mostly through ad hoc analysis, without a systematic, standardized approach [4].

3 Methodology

This section includes the system architecture for the LA system, including the game and learning analytics components. It also includes the experiment design for testing and data collection.

3.1 Fraction Hero: The Game

The game Fraction Hero was developed by the researchers using the RPG Maker MV engine on Windows. The learning content of the game is based on the **DepEd (Department of Education) K-12 Basic Education Curriculum** [19] in the Philippines. The topics covered in the game are using the K-12 curriculum targets for Grade 6 students, or students in their sixth year of primary education, focusing on the fundamental operations involving fractions.

- Adds and subtracts simple fractions and mixed numbers without or with regrouping; multiplies simple fractions and mixed fractions; divides simple fractions and mixed fractions.

The game itself is an RPG/simulation game based around answering quizzes through battles in which the player must correctly solve fraction arithmetic problems in order to attack (see Fig. 1). The player has 5 health points (HP) and 3 magic points (MP) at the start of each battle. Enemies start with 10 HP. The problems are multiple-choice (four), with an additional choice to skip the problem (which costs 1 MP). For every problem the player may choose the difficulty:

- Addition/Subtraction
  - **Easy**: fractions are similar.
  - **Medium**: fractions are dissimilar but have the same multiple.
  - **Hard**: fractions are dissimilar (may include improper fractions).
  - **Expert**: mixed numbers (fraction parts are dissimilar).
- Multiplication/Division
  - **Easy**: fractions have low number range.
  - **Medium**: fractions have moderate number range.
  - **Hard**: fractions have high number range (may include improper fractions).
  - **Expert**: mixed numbers (fraction parts have high number range).

Fig. 1 Fraction Hero: in-game screenshots

If the player answers correctly, he/she will deal damage to the enemy and gain credits based on the difficulty selected. Otherwise, the player will lose 1 HP. Players may also use items and perks to help them during battle:

- Items (bought with credits):
  - **Eraser**: restore 1 HP.
  - **Sharpener**: restore 1 MP.
- Perks (cost 1 MP each):
  - **Open Notes**: review one page from the notebook (one topic).
  - **Cramming Mode**: extend the timer by 1 min.
  - **Bonus Points**: deal extra damage on the next problem (if correct).

Battles end when either the player or enemy hits 0 HP, or after a 5-min time limit. After each battle, the player will receive a score from 0 to 10 from the in-game teacher NPC (non-player character) based on their performance. Outside of battles, players may also review notes about the topic using the Notebook, buy items with credits earned from battles, as well as look at their stats such as accuracy and average solving time in the Student Record. The game consists of five levels, one for each operation, with the last level having random operations.

3.2 Learning Analytics System

Student motivation is evaluated through assigned weights for the three motivation categories of self-efficacy, goal orientation, and effort regulation.

Figure 2 shows the LA model used for this study. It is constructed based on the three motivation categories stated. Data from the game are assessed per individual student's play-through. The following game traces evaluated in the LA are:

1. **Difficulty versus accuracy**. Difficulty and accuracy were compared to assess students' behavior; the student's choice of difficulty based on their result in the previous problem (correct, wrong or skip). Integrating the two in-game data variables yields seven possible scenarios:
   a. The result is correct and the selected difficulty level for the next problem is increased.
   b. The result is correct and the selected difficulty level for the next problem is decreased.
   c. The result is incorrect and the selected difficulty for the next problem is increased.
   d. The result is incorrect and the selected difficulty for the next problem is decreased.
   e. The result is correct and the selected difficulty level for the next problem is the same.

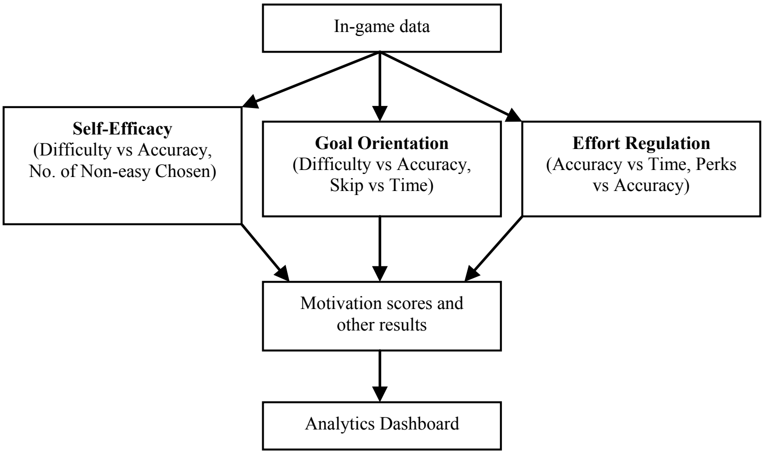


Fig. 2 Learning analytics model

   f. The result is correct and the selected difficulty for the next problem is the same (highest difficulty).
   g. The result is incorrect and the selected difficulty for the next problem is the same.

2. **Number of non-easy problems chosen**. This is the total number of selected medium, hard and expert difficulty problems.
3. **Number of non-skipped problems**. Non-skipped problems were measured to give students a reasonable score for this metric as skipping is generally considered a negative factor.
4. **Accuracy versus time**. Accuracy and time were also compared to identify students who only guess the answers. Combining these two data variables resulted in four cases:

   a. *Fast-accurate*. The student most likely solves the problem easily.
   b. *Slow-accurate*. The student is most likely able to solve the problem but needs sufficient time.
   c. *Slow-inaccurate*. The student is most likely struggling in answering the problem (but still trying).
   d. *Fast-inaccurate*. The student most likely guessed the answer.

5. **Perks versus accuracy**. Use of the Open Notes perk and accuracy were compared to examine students' engagement or mastery in solving a problem.

Calculation of the motivation score involves the assignment of proper weights for each of the three categories. The total motivation score is measured over 100 points total, composed of 35% Self-Efficacy, 35% Goal Orientation and 30% Effort Regulation as shown in Formula (1):

$$\text{Motivation Score} = (0.35 \times \text{Self Efficacy}) + (0.35 \times \text{Goal Orientation}) + (0.30 \times \text{Effort Regulation}) \tag{1}$$

The mean of all battle entries is computed after each session.

In order to calculate Self-Efficacy, Formulas (2) and (3) are added:

$$\left(\frac{\text{NonEasy Problems Chosen}}{\text{Total Problems}} \times \frac{20}{35}\right) \times 100 \tag{2}$$

$$\left(\frac{\text{DifficultyVSAccuracy}}{\text{Total Problems} - 1} \times \frac{15}{35}\right) \times 100 \tag{3}$$

DifficultyVSAccuracy is the total number of encountered cases in each battle as referenced in the *Difficulty versus accuracy* section above which are either (a), (c), (d), (f), or (g). Formulas (1) and (2) are given weights of 20 and 15, respectively, and are added for a total weight of 35 for Self-Efficacy.

Goal Orientation is the sum of Formulas (4) and (5):

$$\left(\frac{\text{NonSkipped Problems}}{\text{Total Problems}} \times \frac{15}{35}\right) \times 100 \tag{4}$$

$$\left(\frac{\text{NonEasy Problems Chosen}}{\text{Total Problems}} \times \frac{20}{35}\right) \times 100 \tag{5}$$

Formulas (4) and (5) are given weights of 15 and 20, respectively, and are added for a total weight of 35 for Goal Orientation.

Lastly, Effort Regulation is the sum of Formulas (6) and (7):

$$\left(\frac{\text{AccuracyVSTime}}{\text{Total Problems}} \times \frac{15}{30}\right) \times 100 \tag{6}$$

$$100 \times \begin{cases} \frac{15}{30}, & \text{Open Notes} \geq 3 \\ \frac{7.5}{30}, & \text{Open Notes} < 3 \text{ and accuracy} > 0.60 \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

AccuracyVSTime is the total number of encountered cases in each battle which are referenced by the *Accuracy versus time* section above as either (b), (c), or (d). Formulas (6) and (7) are given weights of 15 each and are added for a total weight of 30 for Effort Regulation.

After all the in-game data are analyzed, they are sent to the teacher's dashboard. Figure 3 displays the overall motivation for a certain class. The first graph shows the motivation scores for each student in that class. The three pie graphs in the center show the overall scores for each motivational factor. Lastly, the graph at the bottom

Fig. 3 Teacher's dashboard for overall class results

shows the overall motivation scores of the class for each topic. Figure 4 shows the analytical results for individual students' scores for each motivational factor and motivation scores for each topic. The students' *Difficulty versus Accuracy* and *Time versus accuracy* results are also displayed.

Fig. 4 Teacher's dashboard for individual student results

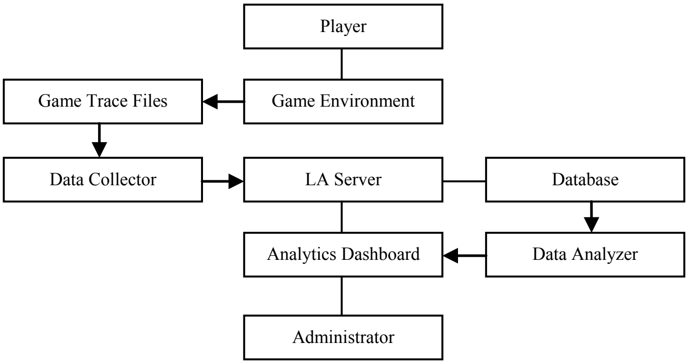


Fig. 5 LA system architecture model

3.3 System Architecture

A player interacts with the game environment which captures all his/her in-game interactions. A data collector sends the data to an LA server where it will then be stored in the database only accessible by the authorized persons. The participating administrators are assigned usernames and passwords. Every time an administrator logs in, the LA processes and analyzes the players' data so they can view the results on the dashboard. Figure 5 shows a model of the system architecture.

3.4 Experiment Implementation

A quasi-experimental design is used for the study wherein all the subjects are tested under the same conditions. A single group of respondents are chosen using purposive sampling. This is because the learning material of the game is designed for Grade 6 students. In the experiment, the group is exposed to two different stimuli to compare their outcomes. In this case, a single group of students from the same class are asked to participate in both a survey and game session, with each measuring their motivation level. The outcomes of the two stimuli are then analyzed for comparison. This experimental design was chosen due to a limitation of available participants.

The survey is intended to measure the scores of each student per motivational category: self-efficacy, goal orientation, and effort regulation. To satisfy this goal, a Scale Response type of survey is used as it categorizes each respondent's standing in a certain concern and the like [20]. Furthermore, a Likert Scale is used to rate the students' motivation level. The range of values are 1–5 with 1 being the lowest and 5 being the highest.

The survey was developed by the researchers of this study with the guidance of Dr. Belinda Cabrera Silverio, who is a research consultant, thesis/dissertation adviser, panelist, and statistician of the University of Makati and Western Colleges in Cavite.

Table 1 Survey items mapped to their motivation categories

Survey items	Motivation category
I am challenged when the fraction problem is hard or difficult	Self-efficacy
I like to solve hard or difficult fraction problems rather than easy fraction problems	Self-efficacy
I am eager to answer fraction problems which are hard or difficult	Self-efficacy
I make sure that I finish all difficult fraction problems	Self-efficacy
I want to answer math problems about fractions	Goal orientation
I don't skip any fraction problems	Goal orientation
I feel excited when I am asked to answer fraction problems	Goal orientation
I don't make a guess in answering fraction problems	Effort regulation
I enjoy answering fraction problems	Effort regulation
I don't give up when answering fraction problems	Effort regulation

To ensure validity and reliability, the research survey instrument was validated by Dr. Lucia B. Dela Cruz, who is a lecturer and thesis/dissertation adviser at the graduate school of the University of Makati. Additionally, Dr. Dela Cruz is regularly invited by educational institutions and private organizations as a resource speaker for test construction and development.

Thirty-one Grade 6 students from the University of the Philippines Integrated School participated in the experiment. The students were given the 10-item Likert Scale questionnaire that measured their self-perceived motivation toward solving fraction problems. The items in the survey were patterned with the in-game's motivation data indicators as shown in Table 1.

After answering the survey, the students played the game within a 40-min session. The students were then instructed to write a game code on their survey sheet for the researchers to map their surveys to their in-game data files. The researchers encoded and matched students' survey data with their corresponding in-game data.

4 Results

4.1 LA Motivation Results

The students' LA motivation scores are considered passing if they are greater than or equal to 60. The greater the student's score is above the passing rate, the higher that student's level of motivation is. Otherwise, the student is unmotivated. Conversely, the lower the result from the passing rate, the poorer that student's level of motivation is.

Table 2 LA mean motivation results versus self-perceived mean motivation

	Self-efficacy	Goal orientation	Effort regulation
LA mean score	62.55	78.77	89.01
Self-perceived motivation mean score	73.55	68.17	73.76
Mean difference	−10.99	10.6	15.25
z-value	2.61	3.17	3.5
α/2 (0.025)	1.96	1.96	1.96
Significantly different?	Yes	Yes	Yes

From the experiment, the students' mean motivation score was 76.94. Only two out of the thirty-one students were categorized as unmotivated. Table 2 shows the mean results per category.

The students' self-efficacy mean score in answering fraction problems was barely above the passing rate (62.55). It is possible that many of them could not trust themselves to handle the Expert difficulty problems in the game. On the other hand, their goal orientation mean score of 78.77 implied that they were motivated to answer and master solving the fraction problems. Lastly, the students were highly engaged in solving the fraction problems as indicated by their effort regulation mean score of 89.01, which had the highest score among all the categories.

4.2 *Comparison Between LA Motivation Results and Self-perceived Motivation Per Category*

The LA mean scores for each motivational category were calculated based on the formulas stated in the Learning Analytics System section. For the self-perceived motivation scores, the survey responses of all students were totaled and averaged based on the motivation category mapped to each item.

To determine whether there is a significant difference between the LA motivation score and self-perceived motivation score with the sample sizes of two groups are greater than 30, a test for Difference of Means based on Two Independent Samples is used wherein z is the test statistic. The said test statistic is computed using this formula:

$$z = \frac{((\bar{x} + \bar{y}) - d_0)}{\sqrt{\frac{S_x^2}{n_1} + \frac{S_y^2}{n_2}}}$$

where

$\bar{x}$ = the mean LA motivation score,
$\bar{y}$ = the self-perceived motivation score,
d_0 = the mean difference of population means.

S_x^2 = the variance of the LA motivation score,
S_y^2 = the variance of the self-perceived motivation score, and
n_1, n_2 = the sample sizes.

The mean difference is obtained to support the result of the hypothesis testing, such as giving a concrete value of how much higher/lower the mean self-perceived motivation score is compared to the LA motivation mean score. Table 2 shows the results.

The first column implies that the students' perceived confidence in answering hard/difficult types of fraction problems is higher than what their actual selection of difficulty in the game implied. The difference between their self-perceived motivation mean score and LA mean score in Self-Efficacy is -10.99, which means that their mean self-perceived motivation in the said category is higher by 10.99. In contrast, their LA mean score in Goal Orientation and Effort Regulation is higher than their corresponding mean self-perceived motivation score by 10.6 and 15.25, respectively. These results suggest that the students' desire to improve and engagement in answering fraction problems within the game is higher than they thought.

4.3 LA Mean Motivation Score Versus Self-perceived Motivation Score

The LA mean motivation score is computed by adding the mean scores of the three categories; the sum is then divided by 3 which results in a score of 76.74. The same method is applied to calculate the self-perceived mean motivation score which gives a result of 65. To prove that the LA mean motivation score is higher than the self-perceived mean score, it is necessary to check if they are significantly different. To do so, we use a hypothesis testing of two independent samples (greater than or equal to 30) where sigma squared in both samples are different. The two indicated independent variables are proven to be significantly different using a 0.05 level of confidence and a z-value of 3.2042. There is sufficient evidence to claim that the LA mean score is higher than self-perceived mean motivation by 11.74, which implies that their actual motivation is greater than their self-perceived motivation.

5 Conclusion

The student participants of this research experiment were mostly highly motivated in solving fraction problems based on the obtained results. By comparing the motivation scores from each test, we found that there is a significant difference between the students' overall self-perceived motivation scores and LA motivation scores in all motivation categories. These results suggest that motivation measured through a

Learning Analytics system may be higher than when using traditional instruments such as surveys or questionnaires.

References

1. Charlton, P., Mavrikis, M., & Katsifli, D. (2013). The potential of learning analytics and big data. *Ariadne* (71).
2. Prensky, M. (2003). Digital game-based learning. *Computers in Entertainment (CIE), 1*(1), 21.
3. Cariaga, A. A., & Feria, R. (2015). Learning analytics through a digital game-based learning environment. In *2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA)* (pp. 1–3). https://doi.org/10.1109/iisa.2015.7387992.
4. Alonso-Fernandez, C., Calvo, A., Freire, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2017). Systematizing game learning analytics for serious games. In *2017 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1111–1118). https://doi.org/10.1109/educon.2017.7942988.
5. Mubeen, S., & Reid, N. (2014). The measurement of motivation with science students. *European Journal of Educational Research, 3*(3), 129–144.
6. Malone, T. W. (1980). What makes things fun to learn? Heuristics for designing instructional computer games. In *Proceedings of the 3rd ACM SIGSMALL Symposium and the First SIGPC Symposium on Small Systems* (pp. 162–169). ACM.
7. Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67.
8. Usher, A., & Kober, N. (2012). *Student motivation: An overlooked piece of school reform*. Summary. Center on Education Policy. http://www.cep-dc.org.
9. Reid, N. (2006). Thoughts on attitude measurement. *Research in Science and Technological Education, 24*(1), 3–27.
10. Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353.
11. Tariq, S., Mubeen, S., & Mahmood, S. (2011). Relationship between intrinsic motivation and goal orientation among college students in Pakistani context. *Journal of Education and Practice, 2*(10).
12. Kim, C., Park, S. W., Cozart, J., & Lee, H. (2015). From motivation to engagement: The role of effort regulation of virtual high school students in mathematics courses. *Educational Technology and Society, 18*(4), 261–272.
13. Ersanlı, C. Y. (2015). The relationship between students academic self-efficacy and language learning motivation: A study of 8th graders. *Procedia-Social and Behavorial Sciences, 199,* 472–478.
14. Kafai, Y. B. (2001). The educational potential of electronic games: From games–to–teach to games–to–learn playing by the rules cultural policy center. University of Chicago, October 27, 2001.
15. Erhel, S., & Jamet, E. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers and Education, 67,* 156–167.
16. Sclater, N., Peasgood, A., & Mullan, J. (2016). *Learning analytics in higher education*. London: Jisc. Accessed February 8, 2017.
17. All, A., Castellar, E. P. N., & Van Looy, J. (2014). Measuring effectiveness in digital game-based learning: A methodological review. *International Journal of Serious Games, 1*(2).
18. Cano, A. R., Fernandez-Manjon, B., & Garcia-Tejedor, A. (2016). *Glaid: Designing a game learning analytics model to analyze the learning process in users with intellectual disabilities* (p. 176). https://doi.org/10.1007/978-3-319-51055-2_7.

19. Department of Education. Department of Education: K to 12 basic education curriculum. http://www.deped.gov.ph/k-to-12/bec-cgs/als-program.
20. Dela Cruz, L., & Silverio, B. (2019). *Practical research 2*. ISBN = 978-971-821-839-6. Mutya Publishing House, Inc. Potrero, Malabon City.

Chapter 12
Designing, Developing and Evaluating Gamification: An Overview and Conceptual Approach

Ana Carolina Tomé Klock, Isabela Gasparini and Marcelo Soares Pimenta

Abstract Gamification defines the use of some game features in contexts other than games. Because of its tendency to increase user motivation and engagement, many areas are applying gamification to improve the user experience. Combined with that, more and more positive outcomes can be found in the literature, predominating over neutral or adverse effects. However, original or reformulated concepts to which gamification is associated are introduced with every new result. This chapter aims to organize and clarify these concepts according to seven different properties: personal, functional, psychological, temporal, playful, implementable, and evaluative. This work discourses about users and their profiles; computational systems and their characteristics; the desired stimuli and incentives; the schedule of reinforcement and the player journey; the game elements; the system development process, and; the consequences and how to measure them. The main contribution of this chapter is the comprehensive view of the gamification through a user-centered approach.

1 Introduction

Gamification is the use of game elements and design for purposes unrelated to games in order to get people motivated to achieve specific goals [12]. Nick Pelling first coined the word "Gamification" in 2002, but it only started to become famous in 2010 [8]. In recent years, gamification has been applied in many different areas, and it has motivated people to change behaviors, to develop skills, and to drive innovation [50].

A. C. T. Klock (✉) · M. S. Pimenta
Institute of Informatics, Federal University of Rio Grande do Sul (UFRGS), Porto Alegre, Brazil
e-mail: actklock@inf.ufrgs.br

M. S. Pimenta
e-mail: mpimenta@inf.ufrgs.br

I. Gasparini
Computer Science Department, Santa Catarina State University (UDESC), Joinville, Brazil
e-mail: isabela.gasparini@udesc.br

A. Tlili and M. Chang (eds.), *Data Analytics Approaches in Educational Games and Gamification Systems*, Smart Computing and Intelligence,
https://doi.org/10.1007/978-981-32-9335-9_12

Nowadays, gamification is applied in several different parts of our lives, just like: shopping, hanging out, working out, recycling, and learning [13]. The e-commerce eBay, for instance, implements points to show users status and badges for the best sellers. The mobile application Swarm creates a sense of progression when the users "check-in" in a place and share their experience (represented by levels). Nike+ also is an example of a mobile application that rewards users for their training with points that unlock awards, achievements, and surprises. RecycleBank is a website that gives points to users when they use less water or energy. Another example is Duolingo, a website and mobile application that helps students learn a new language using points, levels, and rankings.

With this large-scale application by computational systems in their various contexts, everyday new or reformulated concepts from other areas are incorporated into gamification. So, this chapter aims to organize and clarify these concepts, grouping them according to seven different properties explored in the literature: personal, functional, psychological, temporal, playful, implementable, and evaluative. These groups were defined based on the chronological order of the gamification process: after identifying the user profile and goals, we can determine their tasks over the system, as well as the appropriate stimuli to each user according to the purpose of the gamification in the system. These stimuli should be reinforced from time to time with suitable game elements. With all this project in hands, the developer team can implement the gamification, which should be evaluated afterward.

For this, the chapter is structured as follows: the Sect. 2 explores the users and their profiles (Sect. 2.1), the systems and their characteristics (Sect. 2.2), the desired stimuli and incentives (Sect. 2.3), the schedule of reinforcement and the player journey (Sect. 2.4), the game elements and design (Sect. 2.5), the system development and quality control (Sect. 2.6), and the consequences and how to measure them (Sect. 2.7). Finally, Sect. 3 describes the final remarks about this work.

2 Gamification Properties

Based on peer-reviewed works about gamification wrote from 2010 onwards and published on many academic search engines (e.g., ACM Digital Library, IEEE Xplore, Science Direct, Scopus, SpringerLink, Wiley Online Library, Web of Science), the multiple concepts explored by the literature were classified according to seven groups of properties. All the procedures covered by the gamification process in the literature were extracted and later grouped by affinity. These groups were divided as follows: (i) personal, related to the user profile; (ii) functional, related to the tasks to be performed; (iii) psychological, related to stimuli to achieve the purpose of gamification; (iv) temporal, related to the interventions that occur during the interaction; (v) playful, related to the game elements and design; (vi) implementable, related to the application of gamification to the system; and (vii) evaluative, related to the analysis of the results of the gamification.

Chronologically, once the users of the system and their objectives are known (i), it is possible to identify the tasks to be performed by them (ii) and the appropriate stimuli to each profile (iii). Continuous reinforcements should be adopted to stimulate users, and the evolution of interaction should be considered (iv) to select the most suitable game elements for each characteristic (v). With the project completed, the development (vi) and the evaluation of the results obtained with the gamification (vii) begin. In this way, seven ordered groups were named as: (i) Who?; (ii) What?; (iii) Why?; (iv) When?; (v) How?; (vi) Where?; and (vii) How much?, as illustrates 1. Each group analyzes the several similar concepts that can be taken into account during the design, development, or evaluation of gamification. The following sections describe the purpose, the concepts considered, and the actors that can be involved in each group application.

2.1 Personal Properties: Who?

The purpose of the first group is to identify the users who are part of the target audience and which characteristics of these individuals interfere in gamification. As proposed by many works available at the literature, some characteristics influence users experience during their interaction with gamified systems (e.g., age, sex, motivational style, culture, and player type), as described below.

Age The age of the user is one of the characteristics that influence the gamification. The work of Attali and Arieli-Attali [5], for example, conducted two controlled experiments to analyze the effect of experience points on students' performance during an assessment of basic math concepts. The first experiment did not show a significant influence on the points regarding the accuracy of the responses. It was performed with 1218 adults aged between 18 and 74 years. However, adults who had access to experience points answered the questions more quickly compared to the control group. The second experiment conducted with 693 adolescents from the last years of elementary school found the same results regarding the accuracy and speed of responses. Still, the adolescents from the experimental group had a higher satisfaction rate than the participants in the first experiment [5]. Thus, Attali and Arieli-Attali [5] suggest that adolescents tend to be more positively affected by experience points in gamified systems than adults concerning their satisfaction.

Other works that used gamification outside the educational scope also evaluated the influence of the age of users in gamified systems. The work of Bittner and Shipper [7] and Conaway and Garay [9] aimed to improve customer loyalty through gamification and affirm that users engagement in gamified systems is inversely proportional to their age (i.e., the younger the users, the more engaging is their experience with the game elements). Concerning game elements, Bittner and Shipper [7] used customization, badges, feedback, narrative, points, relationships, and leaderboards, while Conaway and Garay [9] implemented challenges, feedback, narrative, progression, and relationships.

In general, it is not possible to identify which and how the game elements applied outside the educational context affected users. However, the work of Attali and Arieli-Attali [5] provides evidence that the use of experience points is more satisfactory for younger users.

Sex In addition to age, studies also indicate that the user's sex influences gamification. The work of Su and Cheng [47] carried out an experiment with three classes of students of 4th grade. Despite not defining which game elements were applied and not using a control group, Su and Cheng [47] identified that male students had a better performance than female students in the gamified system. Another study, conducted by Pedro et al. [39], proposes a gamified virtual learning environment with feedback, badges, experience points and leaderboards to compare the motivation and performance of students aged from 12 to 13 years. During a controlled experiment, the students were divided into two groups: the experimental one, which used the gamified version of the environment, and the control one, which used the environment without gamification. As a result, Pedro et al. [39] concluded that there was a positive motivational effect for male students who used gamification, but it was not possible to identify significant differences in the motivation of female students or the performance of students of both sexes.

Conaway and Garay [9] identified that women are more motivated to use websites if they are gamified with challenges, feedback, narrative, progression, and relationships. Whereas the work of Koivisto and Hamari [26] defined that women are more motivated to perform physical activities with the use of challenges, badges, levels, points, and relationships. The works of Conaway and Garay [9] and Koivisto and Hamari [26] have not identified the most motivating game elements for men.

In general, the work of Pedro et al. [39] suggests that the use of feedback, badges, experience points, and leaderboards will help to motivate the male audience, whereas the works of Conaway and Garay [9] and Koivisto and Hamari [26] suggest the use of challenges, badges, feedback, narrative, levels, points, progression, and relationships will motivate the female audience. However, these studies were carried out in a specific domain with only some of the game elements, and other studies should be performed to assess whether such results can be generalized.

Motivational Style The motivational style is another characteristic that influences gamification, as explained by Hakulinen and Auvinen [22]. Hakulinen and Auvinen [22] explore a concept of psychology called "Achievement Goal Theory" that characterizes the motivational style of the users according to their goals orientation and behaviors. Concerning the orientation, goals can be mastery or performance-oriented. These goals can be further subdivided by the behavior of the user to achieve them:

- Mastery-approach: users focusing on overcoming challenges, improving their competence and learning as much as possible;
- Mastery-avoidance: users avoiding the possibility of failing or doing worse than they have done before;
- Performance-approach: users focusing on demonstrate and prove their abilities to the others;

- Performance-avoidance: users avoiding looking incompetent or less able than others by pretending they are effortless achievers.

Such goals are not mutually exclusive, but a composition of goals at different intensities. As a result, Hakulinen and Auvinen [22] identify that mastery-approach, mastery-avoidance, and performance-approach predominant users tend to be more motivated with badges than performance-avoidance predominant users. This result was the only conclusion withdrawn from this study.

Culture The culture is another characteristic that appeared in the literature, although it was little explored. Although culture has many dimensions to be analyzed, only one study, conducted by Almaliki et al. [2], analyzed the influence of the geographical localization on the feedback provided about the quality of the system used. For this, the study involved users from Europe (the United Kingdom, the Netherlands, and Spain) and the Middle East (Saudi Arabia, Iran, and Egypt). As a result of quantitative and qualitative analyses, Almaliki et al. [3] identified that Middle Eastern users were more motivated by the feedback than users from Europe. Thus, Almaliki et al. [3] conclude that there are some game elements (e.g., badges, customization) that tend to motivate users in the Middle East more than users from Europe.

Player Type Player type is the most explored characteristic in the literature, grouping users according to their gaming preferences. Among the various typologies, the only one that analyzes the profile of the users of gamified systems is the one proposed by Marczewski [33]. It describes six player types according to their motivations for the use of gamified systems, being: Achievers, Disruptors, Free Spirits, Philanthropists, Players, and Socializers.

- Achievers are motivated intrinsically by competence and mastery. These players try to learn new things and improve themselves by overcoming challenges. As their primary motivation is mastery, they are not interested in showing their progress to other players. However, they often compete with others as a way to become better, treating them as challenges to be overcome in the system;
- Disruptors are motivated by change. They try to deregulate the system and force a change, either directly or through other players. This change may be negative (e.g., chasing other players or discovering system failures that spoil the experience of others) or positive (e.g., influencing other players to behave differently or improving the system by adjusting the flaws encountered);
- Free Spirits are intrinsically motivated by autonomy and self-expression. They like to explore the system unrestrictedly and build new things (e.g., customizing their environment with more extravagant avatars and creating more personal content);
- Players are extrinsically motivated by the rewards. In such cases, gamification must reward them at the same time as it attempts to motivate them intrinsically. Thus, players would have both intrinsic (e.g., developing skills, helping others) and extrinsic (i.e., rewards) motivation to use the system;
- Philanthropists are motivated intrinsically by meaning and purpose. They are altruistic players, because they like and usually help other players without expecting a

reward for it. They make the system meaningful to themselves and appreciate to be considered as part of something more significant (i.e., part of a purpose);
- Socializers are players who are intrinsically motivated by relationships. They interact with other users and aim to create social connections.

Such types are not mutually exclusive, and each user can be a combination of several types. Marczewski [33] provided a questionnaire for correct identification of the percentages of the users' player type. Some works that use this typology are Herbert et al. [24] and Gil et al. [19].

The work of Herbert et al. [24] presents a gamified virtual learning environment called "Reflex", which analyzes the variation of students motivation and their behaviors based on Marczewski's typology [33]. The Reflex system presents the content to students based on their curricular learning objectives and accompanies their interactions. Herbert et al. [24] conducted experiments with second-year undergraduate students in a Computer Science course and, based on the results of the questionnaire, correlated students' behaviors with their player types. Herbert et al. [24] suggest that missions and levels motivate Achievers; customization and content unlocking motivate Free Spirits; gifts movitate Philanthropists; badges, points, and virtual goods motivate Players; and relationships motivate Socializers. The Disruptor player type has not been evaluated.

The work of Gil et al. [19] presents a preliminary study on a gamified-based educational system that evaluates the use of game elements by the intrinsically motivated player types proposed by Marczewski [33]. For this, some game elements were implemented in the learning activities of the system and an experiment was carried out to verify the effectiveness of this implementation, and the relation between the game elements and the player types. The 40 first-year undergraduate students of the Computer Science course participated for 5 hours in the C Programming Language and Abstract Data Types disciplines. As a result, Gil et al. [19] identified that the elements used were in line with those recommended by Marczewski [33]: Achiever appreciated challenges, Free Spirits appreciated unlocking content, Philanthropists appreciated gifts, and Socializers appreciated the competition.

Other Considerations About Personal Properties Based on the previous subsections, it can be inferred that there are elements of games more suitable for each user. Thus, it is essential to identify users and their characteristics (e.g., age, gender, player type) to select the most appropriate game elements (if any) to stimulate certain behaviors. To identify users and their characteristics, some methods from the Human–Computer Interaction (HCI) area can be applied, such as data gathering methods: questionnaires, interviews, focus groups, and user observation. As a consequence, the actors involved in this group are final users, their supervisors or domain specialists, and people with knowledge in applying the methods for identifying user characteristics (e.g., HCI specialists).

The quantity and nature of these characteristics vary according to the needs of the system and may, in the future, support adaptive gamification. As the theme is still recent and few works are dealing with the different characteristics of users in the process of gamification, there may be other characteristics that have not yet been identified and explored by the literature, like personality traits [11, 23].

2.2 Functional Properties: What?

The goal of the second group is to identify the behaviors that must be performed by the target audience during interaction with the system to aid in the achievement of the purpose of the system. This group comprises the tasks available that should be performed by the users, guiding the creation of stimuli to carry them out and the inclusion of the appropriate game elements. These functional properties may vary according to the scope of the system: learning, shopping, recycling, exercising, each context will have different purposes.

When talking about learning, for instance, behaviors can be related to interaction, communication, performance, etc. Interaction, within the educational context, encompasses the student's various actions in the system (e.g., student-interface, student-content) [34]. Communication is related to the tools that support discussions between students and teachers to assist in solving exercises, and possible difficulties with the system, which may occur synchronously (e.g., online chat) or asynchronous (e.g., message board and discussion forum) [40]. Performance allows the evaluation of the student and can be carried out through exercises and tests.

It is up to the domain specialists (e.g., teachers, personal trainers) and system analysts to determine which of the available functionalities should be stimulated (i.e., the user still does not do it, but must to), discouraged (i.e., the user should stop doing it) or maintained (i.e., the user must continue to perform it).

2.3 Psychological Properties: Why?

The third group identifies the stimuli to be generated in the target audience to perform the desired behaviors. Thus, the incentives to stimulate users during the interaction with the system are defined. These incentives can persuade users to, for instance, access the system more frequently, communicate with others, and perform better.

The primary purpose of persuasion, when applied to computational products, is to change the behavior of the users [16]. For example, the system can persuade the user to access more items through visual communication and feedback. It should be remembered that strategies of persuasion, when applied in the technological area, should take into account the ethical implications related mainly to data privacy (e.g., avoiding the player's exposure in situations that may cause him/her embarrassment) [31].

When using a gamified system, the user experience involves both sensory experience (sensory-motor aspects), as well as significant experience (cognitive aspects) and emotional (motivational aspects) [32]. The sensory-motor aspect is related to the inputs and outputs of the interaction, where several emotional outputs of the users are obtained through visual, auditory and tactile stimuli inserted in the gamified system. The cognitive aspect is related to the direction and support to the user to accomplish the task, by adapting the interaction to the user profile and feedback of the relevant information (e.g., objectives, results). The motivational aspect is related to the manipulation of emotions and the use of persuasion to engage users with the purpose of the system (e.g., learning, buying). Thus, the main stimuli generated by gamification are: fun, motivation and engagement [32].

Fun can be categorized into four types according to the emotion it gives rise to: easy fun, hard fun, people fun, and serious fun. Easy fun is driven by curiosity and creativity, hard fun is triggered by the evoked emotion of triumphing over an opponent (i.e., Fiero), people fun is driven by entertainment with other users, and serious fun is driven by satisfaction in changing the way the others think, feel or act in order to achieve a higher purpose [28]. Fun can also be linked to the player types [28]. For example, people Fun is most often triggered by users who have the Socializer player type described by Marczewski [33].

Motivation consists of a set of biological and psychological mechanisms whose objective is to guide an individual to continually perform certain behaviors until a goal is reached [30]. Briefly, Ryan and Deci [44] define motivation as the stimulus that an individual receives to achieve a goal, and it can, according to Gagné and Deci [18], be divided between intrinsic, extrinsic and a motivation. Intrinsic motivation refers to personal and internal motivation, in which the individual performs the activity because he/she desires to. Extrinsic motivation refers to external motivation, in which the individual performs the activity for the tangible (e.g., money, good grades) or intangible reward (e.g., praise, admiration) he/she receives. The a motivation refers to the lack of motivation, that is, the individual does not find endogenous or exogenous benefits to perform the activity [18, 36]. The motivation can also be analyzed according to its duration (i.e., short and long-term) [44].

Engagement is an affective and cognitive state that is related to commitment to work, but does not focus on a particular object, event, individual, or behavior [45], its peak being called the “flow state”. The flow state is defined by Csikszentmihalyi [10] as a mental state in which the individual is so involved in an activity and considers it so rewarding that he realizes it even if it is difficult, costly, or dangerous. During the flow state, the concentration of the individual is totally turned to activity, the self-consciousness disappears and the perception of time becomes distorted. To achieve such a state, challenges must be compatible with the individual’s abilities. Otherwise, the experience may become tedious (when the skills are far superior to the challenges) or distressing (when the challenges are far superior to the skills).

Within the educational context, engagement can also be classified as cognitive, behavioral and emotional [17]. Cognitive engagement encompasses the psychological investment of the student in the learning process (i.e., the effort to understand the content). Behavioral engagement encompasses students’ participation in curricular and extracurricular activities. Emotional engagement, on the other hand, encompasses emotional reactions (e.g., interest, frustration) of students about the elements of the educational environment (e.g., activities, other students, teachers).

Depending on the need of the gamification project, other stimuli may be used. Among other concepts, fun, engagement, and flow state make up the so-called “player experience”, a gaming concept that describes a player’s physical, cognitive, and emotional experience during the game [6]. Therefore, it is essential to involve game designers and HCI specialists to identify which stimuli should be used to compose player and user experiences during interaction with the system.

2.4 Temporal Properties: When?

The temporal group identifies the most appropriate situations to stimulate the target audience to perform the desired behaviors. These situations can be classified in two ways: the player’s journey, which is directly related to the evolution of the user about the desired tasks, and the frequency of reinforcement, where reinforcements are applied to motivate the user and keep him/her motivated.

The player’s journey guides users through interaction with the system, indicating the behaviors to be performed and providing the feeling of progress [20]. According to the user experience with the tasks, the system must adopt more appropriate ways to drive the user. Thus, the system should support the novice user so that he feels interested in its use. By becoming more proficient in the task, the user must be surprised and rewarded to continue exploring the system and, consequently, build a habit. Finally, when gaining mastery over the tasks and contents, the user who has invested a particular time in the system (i.e., expert) should be pleased to maintain his loyalty and return occasionally. In addition to the player’s journey, it is essential to apply reinforcements derived from Behavioral Theory to keep users motivated to achieve the goals proposed by the system.

Behavioral Theory suggests that systematically applied rewards and punishments motivate people, condition their actions, and reinforce their responses in anticipation of new rewards or punishments [50]. Reinforcement stimulates the desired behaviors through benefits and can be divided between positive and negative. Positive reinforcement provides the individual with more items that he/she likes (e.g., rewards), while negative reinforcement removes items he/she does not like (e.g., the need to perform a specific task) [21]. Punishment, however, creates a series of conditions to avoid unwanted behaviors and can also be divided between positive and negative. Positive punishment provides the individual with more items that he/she does not like (e.g., reprimand) and negative punishment removes items he/she likes (e.g., freedom) [21]. These reinforcements are used to motivate both extrinsically, through rewards, and intrinsically through feedback.

The frequency of reinforcements is classified in three ways: continuous, proportional, and temporal [15]. The continuous frequency applies a reinforcement to the user to each action performed. Proportional frequency applies a reinforcement to each given number of actions, which can be a fixed number (e.g., every five activities) or a variable one (e.g., first, fifth, and seventh activity). The temporal frequency applies a boost to each given period, which may be fixed (e.g., every ten minutes) or also variable (e.g., within five, thirty, fifty minutes). Exemplifying in the educational context, it can be said that there is a continuous reinforcement if the student receives feedback for each answered exercise, a proportional reinforcement if the student receives a badge in the first and tenth publication held in the discussion forum (i.e., variable reinforcement) and a temporal reinforcement if the student receives a certain quantity of points upon accessing the system every two days (i.e., fixed reinforcement).

To identify what stage of the journey the player is in and in which situations to insert reinforcements, it is important that game designers, domain specialists, and systems analysts work together. In this way, the player's journey should be oriented to the tasks that the domain specialist proposes, while the systems analyst evaluates the viability of the implementation. The reinforcement moment is defined by the domain specialist and the game designer, identifying the tasks to be reinforced and the adequate frequency according to the importance and difficulty of the task.

2.5 Playful Properties: How?

The goal of the fifth group is to design the gamification to stimulate the desired behaviors on the target audience in certain situations. Thus, it is chosen the most appropriate game elements to apply gamification to the system based on the users, the tasks, the stimuli, and the situations.

These game elements are a series of tools that, if used correctly, generate a significant response of the players [51]. According to Werbach and Hunter [50], such elements can be divided according to the MDC model (i.e., Mechanics, Dynamics, and Components). In this model, Mechanics are processes that stimulate player action and engagement (e.g., competitions), Dynamics are managed aspects that do not belong directly to the game (e.g., relationships), and the Components are specific instances of one or more mechanics or dynamics (e.g., leaderboards). Briefly, the MDC model hierarchically organizes the game elements based on their abstraction, as detailed below.

Dynamics At the highest abstraction level of the MDC model are the Dynamics, which are aspects controlled by gamification, but which are not implemented directly. Emotions, narratives, progressions, rules, and relationships are examples of dynamics.

- *Emotions* are the perceptions of the users that directly influence their behavior [28]. Some examples of emotions that can be aroused are: curiosity, competitiveness, frustration, happiness, fear, surprise, disgust and pride;
- *Narratives (or stories)* are plots that interconnect the other game elements implemented. The narrative is an experience that can be appreciated by the player and does not necessarily present a linear story, which may be the unfolding of a sequence of events and even be altered according to the choices made by the player [46];

- *Progressions* express the player's evolution over time [50]. Progression allows players to track their development, demonstrating that each completed activity is related to new content and not just a repetition of something already seen. The progress of the player is strictly controlled by some mechanisms that block or unblock access to specific content [1];
- *Rules* impose limits on what players can and cannot do during the game. Rules are imposed characteristics (constraints or forced commitments) that players are unable to change, forcing them to find alternative ways to achieve the goal [14];
- Relationships are social iterations that generate feelings of camaraderie, status, and altruism [50]. Relationships are a way that players have to interact with others (e.g., friends, team members, and opponents).

Mechanics In the second level of abstraction are the Mechanics, which are ways to induce the player to perform certain activities within the system [50]. Challenges, chances, competitions, cooperations, customization, feedback, rewards, and win states are examples of mechanics.

- *Challenges* are puzzles or other activities that require effort to be resolved [50]. They are important for guiding novice players while they can be used to add depth and meaning to expert players [51]. Challenges are commonly used to provide a sense of progression;
- *Chances* are elements of randomness within the game. Chances serve as a variable proportional reinforcement that rewards the player after a series of activities. For example, a player has a 10% chance of receiving 50 more experience points than he/she usually gets while performing activities. This additional reward possibility keeps the activities consistent, as players increasingly perform them in the hope of receiving such a reward [48]. Chances can also be used to arouse various emotions (e.g., surprise, frustration) in users;
- *Competitions* and *Cooperations* are used to promote interaction between players [50]. In competition, players (or groups of players) compete against others, stimulating the existence of a winner and a loser. In cooperation, players work together to achieve a shared goal. Both can be used to stimulate the relationship between users and arouse emotions;
- *Customization* is the possibility to modify some of the elements available in the system, and can happen in several ways: even simpler interface elements provide an opportunity for customization (e.g., avatar, player name). For example, by changing the background color of the system, customization can add value to the player's experience [51]. Its use is mainly related to emotions;
- *Feedback* returns relevant information to the players [50]. This element is used to generate a cycle of engagement, where the player is motivated to perform a given activity, and this activity provides feedback that reinforces his/her motivation to carry out new activities. The main uses of feedback involve the reinforcement of system rules and the unfolding of narratives;
- *Rewards* are benefits given to players as a way of recognition for their efforts, such as badges that indicate their achievements and items that allow the customization of their characters. In addition to showing appreciation for the time players invest,

offering something in return recognizes their success and insight. Rewards are valuable because they create meaningful measures of progress, reinforce system rules, and help maintain user interest over time [14].

- *Win states* are goals that make a player or a group of players winners or losers. The victories are related to the results of a game that, based on rules, feedback, or rewards, defines the win state [51]. They are directly linked to relationships.

Components At the most concrete and effectively implemented level are the Components, which are specific ways of achieving the mechanics and, consequently, the dynamics. Avatars, content unlocking, emblems, gifts, leaderboards, levels, missions, points, and virtual goods are examples of components.

- *Avatars* are the visual representation of players in a virtual world. Avatars can maintain privacy and anonymity while providing a form of individuality and self-expression to the player [49]. They are commonly used as a form of customization;
- *Content unlocking* is the release of some aspect of the system conditioning the performance of a particular activity by the player. In such cases, the system disables some functionalities until the player completes specific challenges (i.e., reaches a goal) [50]. Content unlocking is usually considered a reward and provides a sense of progression;
- *Emblems* are visual representations of the player's achievements, being awarded when some goal is achieved and serving as a form of follow-up of the player's progression [51]. Emblems can be represented in a variety of ways (e.g., badges, medals, and trophies) [50]. According to Antin and Churchill [4], the emblems present five socio-psychological functions: goal setting, guidance, reputation, status, and group identification. Goal setting determines which goals the player must meet. The guidance helps the players about the possible types of activity within the system. Reputation encapsulates the interests, knowledge, and past interactions of a player. Status reports a player's achievements in their profile, without having to brag explicitly. Finally, group identification allows players to identify others with similar goals, creating a sense of a group. Emblems can be assigned to users by completing challenges as a form of reward and feedback;
- *Gifts* are possibilities to share the resources that a player has with others. According to Schell [46], the player feels satisfaction in surprising another player with a gift. This satisfaction is not only related to the fact that the other player is happy, but the player who offered the gift was responsible for that happiness. Thus, its main use is to encourage cooperation and altruism among players;
- *Leaderboards* show the achievements and progression of the player, giving meaning to the other components (e.g., points and levels), contextualizing the scores (i.e., indicating how good or bad the player is when compared to the others). Leaderboards are also used to increase interest in the game design since it provides a goal to achieve (e.g., a specific position, be better than some other player) [14]. Leaderboards, in general, encourage competition between players;
- *Levels* are markers that identify the progress of the player over time, usually based on completed missions or experience points gained. Therefore, levels are usually tied to challenges and may also appear as a form of feedback. According to

Zichermann and Cunningham [51], levels can be categorized between difficulty, game, and player. The levels of difficulty serve to indicate the effort required by the player to evolve in the game (e.g., easy, medium, and difficult). The game levels are used to indicate the evolution of the player, measured by the completed missions. Player levels indicate the player's experience, measured by the experience points won;

- *Missions* are a set of challenges with specific goals and respective rewards. Missions usually appear in the form of a task that can be achieved in the short-term (e.g., reaching a specific score, completing a certain number of tasks) for a larger goal. When completed, the missions provide a reward to the player [50];
- *Points* are numerical representations of progression. Points can be categorized, according to Zichermann and Cunningham [51], as: Experience Points, Redeemable Points, Skill Points, Karma Points, and Reputation Points. Experience Points are used to reward the player for the activities performed. Redeemable Points are used as the bargaining item. Skill Points are used to reward the player for specific activities. Karma Points are used to assist other players by encouraging altruistic behavior. Finally, Reputation Points indicate the trust between two or more players. Due to this wide variety of types, the points can be used to achieve any of the described mechanics;
- *Virtual goods* are items that exist only virtually and have a value in meaning or money [50]. Such items can be divided into three categories: collectible, consumable, and customizable. The collectible items are those that have an aesthetic purpose (e.g., virtual decoration), the consumables ones are those that can only be used a certain number of times (e.g., virtual food), and the customizable ones are those used to customize the game or the player (e.g., virtual clothing) [25]. Virtual goods can be used as forms of customization or reward, for example.

Other Considerations About Playful Properties As described in Sect. 2.1, some game elements may be more recommended for users with specific characteristics (e.g., the effectiveness of experience points in user satisfaction is inversely proportional to their age [5]). Attention should also be paid to what behaviors we wish to stimulate in users according to the purpose of gamification in the system. For example, to influence student performance improvement, game elements such as missions, challenges, and progressions can be used [50]. The stimuli that we wish to generate through the game elements are also important. For example, you can apply elements such as narratives, progressions, relationships, and emotions to awaken the easy, difficult, people, and serious fun, respectively [28]. The most appropriate time to apply each element is also evaluated. In the player's journey, for example, challenges can be taken to guide novice users during system interaction, rewards to keep habit-builder users motivated, and badges to make expert users feel special [20].

Thus, this group defines the entire design of the gamification, which includes all the game elements to be used, how they interact with each other, and how they influence each of the previously defined groups. To that end, game developers, HCI specialists, and system analysts should be involved in designing the player and user experiences, as well as the feasibility of implementing the system.

2.6 Implementable Properties: Where?

After designing the gamification to stimulate the desired behaviors on the target audience in certain situations, the process of implementing the game elements in the system begins. To apply gamification to the system, we can follow models from the HCI area (e.g., Star Model), Software Engineering area (e.g., Cascade Model) or even a mixture of both areas, depending on the skills and knowledge of the development team.

A model of the HCI area that can be adopted is the "Interaction Design Life Cycle" proposed by Rogers et al. [42], which incorporates the activities of interaction design. Interaction design activities encompass the establishment of requirements, the design of alternatives, the prototyping, and the evaluation [42]. The establishment of requirements identifies the users and the type of support that the interactive product could provide, forming the basis of the product requirements and sustaining the subsequent design and development. The design of alternatives consists of suggesting ideas to satisfy the requirements, covering the conceptual design (what the users can do in the product and which concepts should be understood for the interaction to occur) and physical design (considers product details—e.g., colors, images, sounds). Prototyping encompasses techniques that allow users to evaluate user interaction with the product, and maybe of low fidelity (e.g., paper-based prototypes) or high fidelity (e.g., functional prototypes). Finally, the evaluation of the design process can determine the usability or user experience of the product by measuring a variety of metrics or defined criteria. The results of this evaluation may require a review of the design or requirements. This life cycle generates a final evolutionary product, where the number of repetitions of the cycle is limited by the resources available, finalizing the development through the positive evaluation of the design process.

In this way, those involved in the vary according to the model adopted. As the needs and part of the project have already been surveyed during other groups (e.g., personal and functional properties), the final users may not be involved, except in cases adopting participatory design or even in cases where the user participates in the validation of the final product. Some examples of actors involved in this group are system analysts, developers, HCI specialists, and test analysts.

2.7 Evaluative Properties: How Much?

The last group suggests the evaluation of how much the gamification in the system was able to stimulate the desired behaviors on the target public in certain situations. Unlike the evaluation of the system that found in software development processes, the "How much?" group is responsible for evaluating only the effect of gamification on final users, not covering the usability and functionality tests.

The methods adopted for the evaluation vary according to the type of research [27]. Descriptive research, which is focused on describing a situation or set of events (e.g., X is happening), usually observes users, conducts field studies, focus groups, or interviews. Relational research, which identifies relationships between variables (e.g., X is Y-related), using methods such as user observation, field studies, and questionnaires. For experimental research, which identifies the causes of a situation or a set of events (e.g., X is responsible for Y), controlled experiments are preferred [27]. In general, according to the comparative study performed by Ogawa et al. [38], the most commonly adopted method to evaluate the influence of gamification on students is the controlled experiment.

The controlled experiment usually begins with the hypothesis investigation, and there must be at least one null hypothesis and an alternative hypothesis [27]. The null hypothesis generally defines that there are no differences between what is being tested (e.g., gamification does not influence student interaction), while the alternative hypothesis always determines something mutually exclusive to the null hypothesis (e.g., gamification influences student interaction). Thus, the goal of any controlled experiment is to find statistical evidence that refutes the null hypothesis to support the alternative hypothesis [43]. Also, a good research hypothesis must meet three criteria: (i) use clear and precise language; (ii) focus on the problem that must be tested by the experiment; and (iii) clearly define the dependent and independent variables [27].

Dependent variables reflect the results that should be measured (e.g., interaction) while independent variables reflect at least two conditions that affect these outcomes (e.g., whether or not gamification is used) [37]. Dependent variables are usually measured using quantitative metrics [27]. For example, to measure student interaction with content and interface, we can analyze the number of accesses to concepts and the duration of access to the system. To measure the communication, we can use the number of messages sent and the number of topics created and answered in the discussion forum. To measure performance, the students' final grades are examples of metrics that can be used. In addition to the "What?" group evaluation, we can also create hypotheses to assess the stimuli that should have been generated by gamification (i.e., "Why?" group). For instance, to measure engagement, it would be possible to analyze the amount and duration of user accesses during a period [29].

The ideal condition is that the only variation within the experimental environment is the independent variable and, to avoid external factors and remove potential biases, a protocol must be defined to guide the experiment and make it replicable [41]. Following this protocol, the participants are divided, the experiment is executed to collect the data to allow the measurement of the dependent variables, and the analysis of the results using different statistical tests of significance is made to accept or refute the hypotheses defined [27].

Regarding the division of participants, we can adopt the between-subject or within-subject approach for experiments with only one independent variable. In the between-subject approach, participants are randomly divided into groups, each group is exposed to only one condition of the independent variable and, in the end, dependent variables are compared between groups. In the within-subject approach, all participants perform the activities in all conditions, and each dependent variable is compared with all conditions of the same participant [41].

Each approach has advantages and limitations that must be analyzed before the choice. The between-subject approach avoids the learning effect, which would allow the participant to perform the task more quickly the second time, and more effectively controls confounding variables (i.e., biases), such as fatigue [27]. On the other hand, this approach requires a larger sample, it is more difficult to achieve statistically significant results, and individual differences have a more significant impact. The within-subject approach contrasts the advantages and limitations of the between-subject approach, as it is suitable for smaller samples, can adopt several statistical tests and can isolate individual differences, but is impacted by the learning effect and by confounding variables [27].

After dividing the participants according to the chosen approach, the execution of the experiment begins and, consequently, the data collection. The data collected can be quantitative and qualitative. Quantitative data are those representing numbers resulting from a count or measurement and are subdivided into: discrete and continuous. Discrete is the one that assumes values within a finite and enumerable set resulting from a count (e.g., number of hits to the system); while continuous assumes values within the set of real numbers and are the result of a measurement (e.g., the rate of exercises correctness) [35]. Qualitative data are those that categorize some aspect related to what is being observed, being subdivided in: nominal, that assume values without a predetermined ordination (e.g., male, female); and ordinals, which have an ordering (e.g., high school, college) [35].

At the end of the data collection, the statistical analysis of the data starts. According to the chosen approach, we must use adequate tests of significance that enable the refutation of the null hypothesis. All significance tests are subject to errors [27]. The errors can be classified as type 1 (or false positive) error, in which a true null hypothesis is refuted; and the error of type 2 (or false negative), in which a false null hypothesis is accepted [43]. To avoid type 1 errors, a low probability (0.05) is usually adopted so that the difference between the two groups compared (i.e., p-value) is equal or not. Thus, the probability of erroneously rejecting the null hypothesis is less than 0.05 [27]. To avoid type 2 errors, the use of a relatively large sample is suggested [41].

In this last group, the main ones involved are the final users, who effectively participate in the evaluation and generate the data to be analyzed. In addition to them, their supervisors can carry out proper accompaniment and guidance. To obtain qualitative data, it is important to involve HCI specialists and to analyze quantitative and qualitative data a specialist or at least one person who is knowledgeable in statistical techniques should be involved.

3 Final Considerations

The chapter organized and clarified many gamification concepts while described various properties that can be considered during the gamification of computational systems to assist in the proper design, development, and evaluation. These properties were divided into seven groups, and each of which encompasses several fundamental concepts for gamification.

The first group, “Who?”, is responsible for personal properties and identifies the gamification target audience. It explores some characteristics of the users that influence the gamification: age, sex, goal, culture, and player types. Since gamification is influenced by these and probably other characteristics, it is possible to infer that gamification is not suitable for all users.

The second group, “What?”, is responsible for functional properties and identifies the behaviors that should be stimulated, discouraged, or maintained so that users achieve the purpose of gamification in the system. Thus, the functionalities available in the computational system are raised.

The third group, “Why?”, is responsible for the psychological properties and identifies which stimuli the gamification should generate in users so that they perform the desired behaviors. These stimuli are directly related to gamification and its aesthetics: user and player experience, persuasion, motivation, fun, engagement, and flow state.

The fourth group, “When?”, is responsible for temporal properties and identifies the most appropriate situations for users to be encouraged to perform the desired behaviors. The situation can identify the user expertise with the tasks (player’s journey), classifying their experience as “novice”, “habit-builder” or “expert” and defining what the system should provide for each level. The situation can also identify the most appropriate moments to include reinforcements, be it reward or feedback, on the behaviors performed. Such reinforcements may be continuous, temporal, or proportional.

The fifth group, “How?”, is responsible for playful properties and identifies the game elements that should be used to encourage users to perform the desired behaviors in the given situations. From the users, the behaviors, the stimuli, and the situations defined in the previous groups, the game elements are modeled to achieve the purpose of gamification in the system.

The sixth group, “Where?”, is responsible for the implementable properties and identifies the changes that must be made to the system so that gamification can stimulate users to perform the desired behaviors in the given situations. There are several models that can be adopted to assist in this implementation, from Software Engineering or/and HCI areas. It is up to those involved in this group to define which methodology is most appropriate based on the skills of the development team.

The last group, “How much?”, is responsible for the evaluative properties and analyzes if the implementation of gamification in the system stimulated the users to perform the desired behaviors in the determined situations. This last group suggests that hypotheses are defined based on the purpose of gamification in the system, the metrics to evaluate such hypotheses, and a protocol to control the experiment. The results will allow the improvement of the gamification.

The main contribution of this chapter is the comprehensive view of personal, functional, psychological, temporal, playful, implementable, and evaluative properties of gamification, guiding its design, development, and evaluation. This grouping can be used by both educators and researchers in order to define how to gamify a

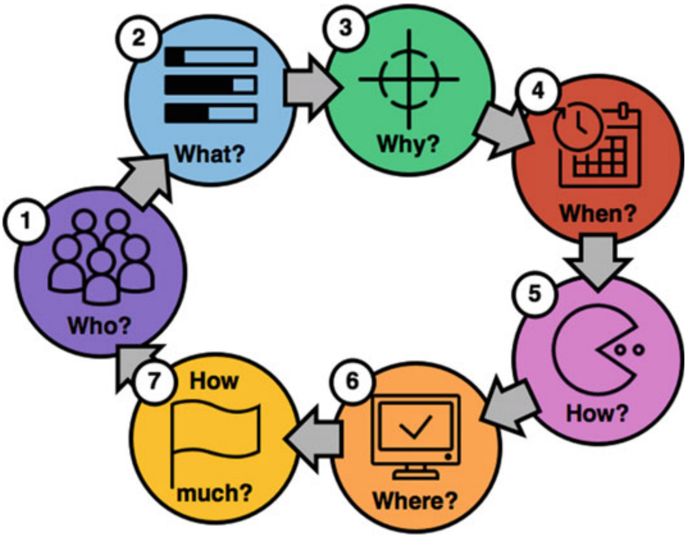


Fig. 1 Seven groups of gamification properties

computational system, following the proposed order in Fig. 1. As future work, these properties will be considered for the application of gamification in a computational system, in order to verify the efficacy and fullness of these properties.

Acknowledgements We thank the partial financial support of FAPESC, public call FAPESC/CNPq No. 06/2016 support the infrastructure of CTI for young researchers, project T.O. No.: 2017TR1755—Ambientes Inteligentes Educacionais com Integração de Técnicas de Learning Analytics e de Gamificação. We are also grateful for the financial support of CNPq National Council for Scientific and Technological Development. This study was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001.

References

1. Adams, E., & Dormans, J. (2012). *Game mechanics*. Berkeley: New Riders Games.
2. Almaliki, M., Jiang, N., Ali, R., & Dalpiaz, F. (2014). Gamified culture-aware feedback acquisition. In *Proceedings of the 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing* (pp. 624–625). Washington: IEEE Computer Society.
3. Almaliki, M., Ncube, C., & Ali, R. (2014). The design of adaptive acquisition of users feedback. In *Proceedings of the 2014 IEEE Eighth International Conference on Research Challenges in Information Science* (pp. 1–12). Marrakech: IEEE Computer Society.
4. Antin, J., & Churchill, E. F. (2011). Badges in social media. In *CHI 2011 Gamification Workshop Proceedings* (pp. 1–4). Vancouver: ACM Press.
5. Attali, Y., & Arieli-Attali, M. (2015, April). Gamification in assessment. *Computers and Education, 83*, 57–63.
6. Bernhaupt, R. (2010). *Evaluating user experience in games*. New York: Springer Publishing Company.

7. Bittner, J. V., & Shipper, J. (2014). Motivational effects and age differences of gamification in product advertising. *Journal of Consumer Marketing, 31*(5), 391–400.
8. Burke, B. (2016). *Gamify: How gamification motivates people to do extraordinary things*. Routledge.
9. Conaway, R., & Garay, M. C. (2014, December). Gamification and service marketing. *SpringerPlus, 3*(1), 1–11.
10. Csikszentmihalyi, M. (1990). *Flow*. New York: HarperCollins.
11. Denden, M., Tlili, A., Essalmi, F., & Jemni, M. (2018, July). Does personality affect students' perceived preferences for game elements in gamified learning environments? In *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)* (pp. 111–115).
12. Deterding, S., Khaled, R., Nacke, L. E., & Dixon, D. (2011). Gamification. In *CHI 2011 Gamification Workshop Proceedings* (pp. 12–15). Vancouver: ACM Press.
13. Duggan, K., & Shoup, K. (2013). *Business gamification for dummies*. Hoboken: Wiley.
14. Ferrera, J. (2012). *Playful design*. LLC, New York: Rosenfeld Media.
15. Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. East Norwalk: Appleton-Century-Crofts.
16. Fogg, B. J. (2002). *Persuasive technology*. San Francisco: Morgan Kaufmann Publishers Inc.
17. Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004, March). School engagement. *Review of Educational Research, 74*(1), 59–109.
18. Gagné, M., & Deci, E. L. (2005, January). Self-determination theory and work motivation. *Journal of Organizational Behavior, 26*(4), 331–362.
19. Gil, B., Cantador, I., & Marczewski, A. (2015). Validating gamification mechanics and player types in an e-learning environment. In G. Conole, T. Klobucar, C. Rensing, J. Konert, & É. Lavoué (Eds.), *Proceedings of the 10th European Conference on Technology Enhanced Learning* (pp. 568–572). Cham: Springer International Publishing.
20. Gilbert, S. (2015). *Designing gamified systems*. New York: CRC Press.
21. Griggs, R. A. (2008). *Psychology*. New York: Worth Publishers.
22. Hakulinen, L., & Auvinen, T. (2014). The effect of gamification on students with different achievement goal orientations. In *Proceedings of the 2014 International Conference on Teaching and Learning in Computing and Engineering* (pp. 9–16). Washington: IEEE Computer Society.
23. Hamari, J., Koivisto, J., & Sarsa, H. (2014, January). Does gamification work?—a literature review of empirical studies on gamification. In *2014 47th Hawaii International Conference on System Sciences* (pp. 3025–3034).
24. Herbert, B., Charles, D., Moore, A., & Charles, T. (2014). An investigation of gamification typologies for enhancing learner motivation. In *Proceedings of the 2014 International Conference on Interactive Technologies and Games* (pp. 71–78). Washington, IEEE Computer Society.
25. Ki, E. N. (2014). The nature of goods in virtual world. In A. Lakhani (Ed.), *Commercial transactions in the virtual world: Issues and opportunities* (pp. 103–118). Hong Kong: City University of Hong Kong Press.
26. Koivisto, J., & Hamari, J. (2014, June). Demographic differences in perceived benefits from gamification. *Computers in Human Behavior, 35*, 179–188.
27. Lazar, J., Feng, J. H., & Hochheiser, H. (2010). *Research methods in human-computer interaction*. Hoboken: Wiley.
28. Lazzaro, N. (2009). Why we play. In A. Sears, J. A. Jacko (Eds.), The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications (2nd edn, pp. 679–700). New York: CRC Press.
29. Lehmann, J., Lalmas, M., Yom-Tov, E., & Dupret, G. (2012). Models of user engagement. In *Proceedings of the 20th International Conference on User Modeling, Adaptation, and Personalization* (pp. 164–175). Berlin: Springer.
30. Lieury, A., & Fenouillet, F. (2000). *Motivação e aproveitamento escolar*. São Paulo: Loyola.
31. Llagostera, E. (2012). On gamification and persuasion. In *Anais do XI Simpósio Brasileiro de Jogos e Entretenimento Digital* (pp. 12–21). Porto Alegre: SBC.

32. Marache-Francisco, C., & Brangier, E. (2014). The gamification experience. In K. Blashki & P. Isaias (Eds.), *Emerging research and trends in interactivity and the human-computer interface* (pp. 205–223). Hersey: IGI Global.
33. Marczewski, A. (2015). *Even ninja monkeys like to play*. Charleston: CreateSpace Independent Publishing Platform.
34. Mattar, J. (2009). Interatividade e aprendizagem. In F. M. Litto & M. Formiga (Eds.), *Educação a distância: o estado da arte* (pp. 112–120). São Paulo: Pearson.
35. Morettin, P. A., de O. Bussab, W. (2004). Estatística Bsica (5th edn.). São Paulo: Editora Saraiva.
36. Mukherjee, K. (2009). *Principles of management and organizational behavior*. McGraw-Hill Education (India) Pvt Limited: New Delhi.
37. Oehlert, G. W. (2010). *A first course in design and analysis of experiments*. New York: Freeman and Company.
38. Ogawa, A. N., Klock, A. C. T., & Gasparini, I. (2016). Avaliação da gamificação na área educacional. In Anais do Simpósio Brasileiro de Informática na Educação (pp. 440–449). Porto Alegre: SBC.
39. Pedro, L. Z., Lopes, A. M. Z., Prates, B. G., Vassileva, J., & Isotani, S. (2015). Does gamification work for boys and girls? In *Proceedings of the 30th Annual ACM Symposium on Applied Computing* (pp. 214–219). New York, ACM Press.
40. Pereira, A. T. C., Schmitt, V., & Dias, M. R. A. C. (2007). Ambientes virtuais de aprendizagem. In *Ambientes Virtuais de Aprendizagem em Diferentes Contextos* (pp. 4–22). Rio de Janeiro: Ciência Moderna Ltd.
41. Purchase, H. C. (2012). *Experimental human-computer interaction*. New York: Cambridge University Press.
42. Rogers, Y., Sharp, H., & Preece, J. (2011). *Interaction design* (3rd ed.). Hoboken: Wiley.
43. Rosenthal, R., & Rosnow, R. (2008). *Essentials of behaviorial research* (3rd ed.). Boston: McGraw Hill.
44. Ryan, R. M., & Deci, E. L. (2000, January). Intrinsic and extrinsic motivations. *Contemporary Educational Psychology, 25*(1), 54–67.
45. Schaufeli, W. B., Martnez, I. M., Pinto, A. M., Salanova, M., & Bakker, A. B. (2002, September). Burnout and engagement in university students. *Journal of Cross-Cultural Psychology, 33*(5), 464–481.
46. Schell, J. (2014). *The art of game design* (2nd ed.). Boca Raton: CRC Press.
47. Su, C. H., & Cheng, C. H. (2013, November). A mobile game-based insect learning system for improving the learning achievements. *Procedia—Social and Behavioral Sciences, 103*, 42–50.
48. Sylvester, T. (2013). *Designing games*. Sebastopol: O'Reilly Media Inc.
49. Vasalou, A., & Joinson, A. N.: Me, myself and i. Computers in Human Behavior 25(2), 510–520 (mar 2009).
50. Werbach, K., & Hunter, D. (2012). *For the win*. Philadelphia: Wharton Digital Press.
51. Zichermann, G., & Cunningham, C. (2011). *Gamification by design*. Sebastopol: O'Reilly Media Inc.

Part V
Conclusion

Chapter 13
Data Analytics Approaches in Educational Games and Gamification Systems: Summary, Challenges, and Future Insights

Ahmed Tlili and Maiga Chang

Abstract This chapter summarizes the reported findings of this book to facilitate the adoption of data analytics in educational games and gamification systems. Specifically, this chapter presents the objectives of adopting data analytics which is finding individual differences; doing learning assessments and knowing more about the learners. It then presents the collected metrics and applied analytics techniques in order to achieve these objectives. Additionally, this chapter highlights several limitations reported by other authors during the adoption of learning analytics. These limitations should be considered by researchers and practitioners in their context to facilitate learning analytics adoption. Finally, this chapter provides future insights about the learning analytics field.

1 Objectives of Adopting Data Analytics

The inclusion of data analytics within educational games and gamification systems can make them smart by achieving several objectives, highlighted in this book, as follows:

- **Finding individual differences**: Traditional learner modeling instruments, such as questionnaires, have been reported to be lengthy and not motivating. With the help of data analytics, a system includes educational game is capable of modeling learners implicitly. Furthermore, individual differences like competences (computational thinking in particular) and motivation can also be found (see Chaps. 11 and 12 for more details). The modeling process is considered a crucial step in order to provide personalized and adaptive learning services based on individual differences that are further highlighted in this chapter).

A. Tlili (✉)
Smart Learning Institute of Beijing Normal University, Beijing, China
e-mail: ahmed.tlili23@yahoo.com

M. Chang
School of Computing and Information Systems, Athabasca University, Athabasca, Canada
e-mail: maiga.chang@gmail.com

A. Tlili and M. Chang (eds.), *Data Analytics Approaches in Educational Games and Gamification Systems*, Smart Computing and Intelligence,
https://doi.org/10.1007/978-981-32-9335-9_13

- **Do learning assessments**: In online learning environments like MOOCs, massively multiplayer educational games, where thousands of learners are taking their pace for studying and doing learning activities. It becomes very difficult for teachers to monitor learners' learning progress, assess learners' skills and knowledge levels, and take care of individual needs. Data analytics have been applied to solve this issue and help teachers. For instance, Tlili and his colleagues develop iMoodle that can identify at-risk students and provide them personalized learning support (see Chap. 6). With data analytics' help, the completion rate of a course can also be improved. Seaton and her colleagues provide learning analytics dashboard for the learners so they can easily keep track of their progress, realize their habits and weaknesses to help them overcome obstacles and achieve better learning outcome (see Chap. 7).
- **Know more about the learners**: Traditional educational games and gamification systems are black boxes where teachers cannot see or know, besides the final scores and levels cleared, how their learners do in the learning process and behave towards the learning goal. Data analytics approaches have been applied to overcome this limitation. Ifenthaler and Gibson explore the learning engagement and its relationship with learning performance in the context of game-based learning (see Chap. 3). Also, Shute, Rahimi, and Smith explore the importance of including learning supports and its impact on learning performance when using the Physics Playground game (see Chap. 4).

Based on the reported chapters in this book, Fig. 1 presents a generic framework of adopting data analytics in educational games and gamification systems to achieve the three objectives mentioned above. When learners use and interact with the developed educational game or gamified system, several metrics (traces) are created based on the interactions and collected into the database. The data analytics module(s) is (are) developed either as the built-in the game and system or accessories of the game and system. The module takes the collected metrics as inputs and does proper analysis and produce results as outputs for achieving a particular objective.

It should be noted that no chapter reports the use of cloud computing technology to store the collected metrics. Also, no chapter has parents, as stakeholders, for the application of data analytics in educational games and gamification systems. Therefore, further investigation is needed regarding these two matters.

2 Collectable Metrics and Traces

It has been seen that the more data is collected within educational games and gamification systems, the more possibilities we will have to enhance the learning process. Kinshuk et al. [1] highlight that to provide smart learning every bit of information that each learner comes into contact with should be collected. For example, to predict at-risk students, Tlili and his colleagues in Chap. 6 use the following five metrics, namely: (1) Number of acquired badges which highlights the number of conducted

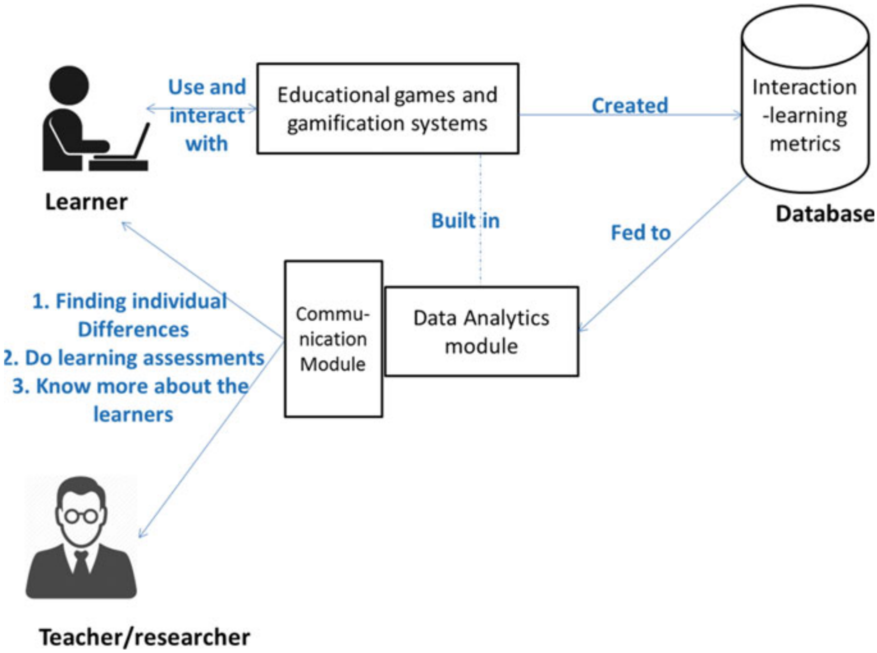


Fig. 1 Generic process of adopting data analytics in educational games and gamification systems

learning activities, since every time a student finishes a learning activity, he/she gets a badge; (2) Activities grades which refer to the value assigned by teachers to assignments and quizzes requested and delivered by students; (3) Student's rank on the leaderboard which is based on the acquired number of points (4) Course progress which can be seen in the progress bar; and, (5) Forum and chat interactions which refers to students' participation in online discussions, such as the number of posts read, posts created and replies.

To identify the motivation of students in an educational game, Flores, Silverio, Feria, and Cariaga use in Chap. 12 the following five metrics, namely: (1) difficulty versus accuracy: compared to assess students' behavior; the student's choice of difficulty based on their result in the previous problem (correct, wrong or skip); (2) number of non-easy problems chosen: the total number of selected medium, hard and expert difficulty problems; (3) number of non-skipped problems: measured to give students a reasonable score for this metric as skipping is generally considered a negative factor; (4) accuracy versus time: were also compared to identify students who only guess the answers; and, (5) perks versus accuracy: were compared to examine students' engagement or mastery in solving a problem.

To assess computational thinking skills, Montaño, Mondragón, Tobar-Muñoz, and Orozco use in Chap. 5 the following seven metrics, namely, (1) abstraction and pattern recognition: focus on not having unused code, the use of functions in the code, and the use of clones of blocks of code (a specific functionality of the Scratch environment);

(2) flow control: assessment of the correct use of every control instruction (such as if and for statements), and also the adequate use in nesting those statements; (3) input control: assessment of the adequate use of statements designed to capture user input into the code, the naming of variables, and the use of non-user-defined variables; (4) data analysis: assessment of the treatment and transformation of the data through the use of data transformation blocks or statements, and also their adequate nesting if necessary; (5) parallelism and threading: assessment of the adequate use of threading and multi-tasking enabling blocks; (6) problem-solving: assessment of the student's ability to decompose a problem into multiple smaller ones in order to address them more easily; and, (7) algorithmic thinking: assessment of the student's ability to develop sequences of tasks, that would be translated into blocks of code, in order to solve a problem.

While Ghergulescu and Muntean [2] mentioned that little is known about the collected traces and used metrics in game-based learning environments, it has been seen that different types of metrics could be collected by asking three questions as Fig. 2 shows:

(1) *What types of metrics should be collected?* Two types of metrics (traces) can be collected, namely, (1) generic metrics which can be found in most educational games and gamification systems, such as the number of signing into an educational game or gamification system, and the time spent on the game or system; and, (2) specific metrics which are defined based on the designed learning envi-

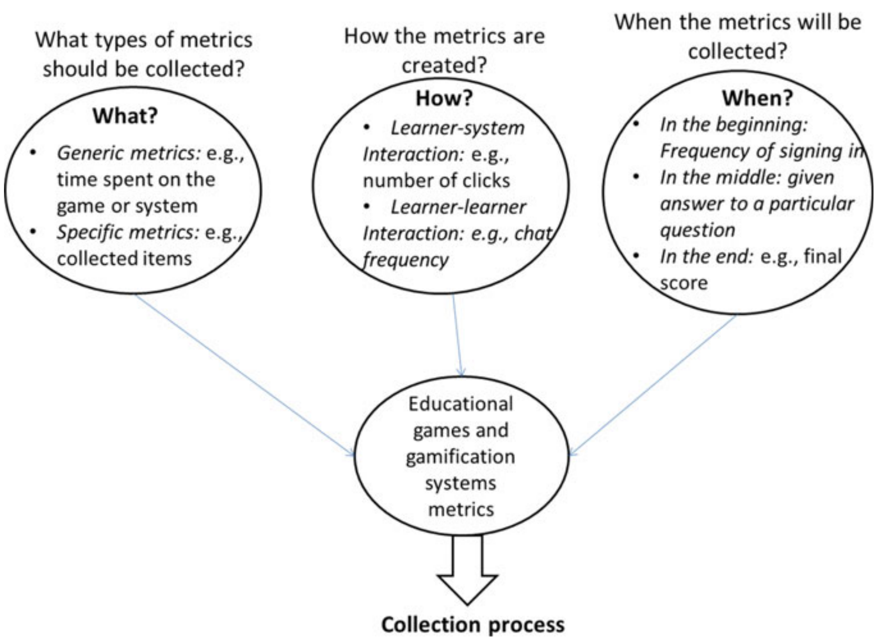


Fig. 2 Types of collected metrics in educational games and gamification systems

ronment, such as number of collected items (badges, points, coins, etc.) and the selected game path.

(2) *How the metrics are created?* One kind of metrics (traces) is created when the learners interact with the educational game or the gamification system, while the others can be created when the learners are interacting with others within the game or the system. For instance, chat frequency is created when learners start to chat together within the game or the system to solve a particular learning activity.

(3) *When the metrics will be collected?* Metrics (traces) can be collected at the beginning, in the middle or at the end of the game-play or the usage of the gamification system. For instance, the final score is collected at the end while the number of times to sign in the game or the system is collected at the beginning.

To extract useful information from the collected metrics (discussed above), different analytics techniques are applied, as discussed in the next section.

3 Analytics Techniques

Based on the chapters included in this book, three analytics techniques are usually adopted in educational games and gamification systems:

- *Data Visualization*: It uses visualization such as pie charts and histograms to represent data. This can help to communicate information clearly and efficiently to stakeholders (e.g., teachers, students, etc.). For instance, the authors in Chaps. 6 and 7 all adopt data visualization techniques to create dashboards for both teachers and students.
- *Data Mining*: It aims to discover hidden information and meaningful patterns from massive data. In this context, several algorithms are adopted. For example, Tlili and his colleagues in Chap. 6 adopt association rules mining and Apriori Algorithm to predict at-risk students.
- *Sequential Analysis*: It allows exploring, summarizing, and statistically test cross-dependencies between behaviors that occur in interactive sequences. For example, Moon and Liu in Chap. 2 conduct a systematic literature review on 102 articles that work on sequential data analytics (SDA) in game-based learning.

Several studies also reported the abovementioned analytics techniques are commonly adopted in games [3–5]. Several challenges, on the other hand, are reported by the authors, in their chapters, which might hinder the adoption of data analytics in educational games and gamification systems. These challenges are discussed in the next section.

4 Challenges

Based on the chapters included in this book, several challenges are reported by the authors. These challenges should be considered by researchers and practitioners in their context to enhance the adoption of data analytics in educational games and gamification systems.

Moon and Liu in Chap. 2 highlight two limitations while adopting data analytics approaches, specifically sequential analysis, in educational games and gamification systems, namely: (1) the need for high computational power in order to collect and analyze big data; and, (2) sequential analysis is often performed as post hoc analysis. Therefore, it is challenging to ensure the validity of the results without cross-validating with the participants. In addition, the participants may not even recall some certain behaviors because the data is captured at a fine granularity. Another issue with post hoc analysis is if the scope of the study is biased, data collection will be biased which in turn leads to an invalidated biased results.

Ifenthaler and Gibson in Chap. 3 and Montaño, Mondragón, Tobar-Muñoz, and Orozco in Chap. 5 report that one of the challenges is collecting large enough data so the applied data analytics approach within their gamified systems can be more accurate. Tlili and his colleagues in Chap. 6 highlight the challenge of protecting learners' privacy while applying educational games and gamification systems. They also discuss the importance of having a predefined time of keeping the learners' stored data.

5 Conclusion

Game-based learning environments and learning analytics are gaining increasing attention from researchers and educators since they both can enhance learning outcomes. Therefore, this book covered a hot topic which is the application of data analytics approaches and research on human behavior analysis in game-based learning environments, namely educational games and gamification systems, to provide smart learning. Specifically, this book discussed the purposes, advantages, and limitations of applying theses analytics approaches in these environments. Additionally, this book helped readers, through various presented smart game-based learning environments, integrate learning analytics in their educational games and gamification systems to, for instance, assess and model students (e.g., their computational thinking) or enhance the learning process for better outcomes. Finally, this book presented general guidelines, from different perspectives, namely, collected metrics, applied algorithms and the encountered challenges during the application of data analytics approaches, which facilitate incorporating learning analytics in educational games and gamification systems.

Future directions for readers to consider and focus could be: (1) investigating the use of data analytics in educational games and gamification systems for health education in particular; and, (2) investigating how Internet of Things (IoT), which is a new technology that is gaining increasing attention from researchers and practitioners, could help the application of data analytics in educational games and gamification systems.

References

1. Kinshuk., Chen, N. S., Cheng, I. L., & Chew, S. W. (2016). Evolution is not enough: Revolutionizing current learning environments to smart learning environments. *International Journal of Artificial Intelligence in Education, 26*(2), 561–581.
2. Ghergulescu, I., & Muntean, C. H. (2012). Measurement and analysis of learner's motivation in game-based e-learning. In *Assessment in game-based learning* (pp. 355–378). New York: Springer.
3. Chaudy, Y., Connolly, T., & Hainey, T. (2014). Learning analytics in serious games: A review of the literature. In *European Conference in the Applications of Enabling Technologies (ECAET)*, Glasgow, (2014).
4. Siemens, G. (2010). *What are learning analytics*? Retrieved August 12, 2016, from http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/.
5. Kim, S., Song, K., Lockee, B., & Burton, J. (2018). What is gamification in learning and education? In *Gamification in learning and education* (pp. 25–38). Cham: Springer.