Designing an Educational Game Evaluation Framework Based on Game Mechanics

Satrio Adi Rukmono
School of Electrical Engineering and Informatics, Institut Teknologi Bandung
Bandung, Indonesia
Email: sar@itb.ac.id

Fais Zharfan Azif
School of Electrical Engineering and Informatics, Institut Teknologi Bandung
Bandung, Indonesia
Email: 23518037@std.stei.itb.ac.id

Muhammad Zuhri Catur Candra
School of Electrical Engineering and Informatics, Institut Teknologi Bandung
Bandung, Indonesia
Email: m.candra@itb.ac.id

***Abstract*—Children in everyday life are increasingly using educational games. However, the quality of each of the many educational games available varies. Some evaluation frameworks exist, but most are prone to the evaluator's subjectivity, which cannot be compared objectively. This study aims to formulate a framework that evaluates the quality of educational games objectively based on the game mechanics used. The framework is built upon Bloom's taxonomy as the basis to ascertain the academic side and MDA (Mechanics-Dynamics-Aesthetics) Framework to distinguish the game side. Then, it assesses each educational mechanic based on a standard in the evaluation framework to obtain an accurate, quantifiable score as a measure. Validation of the framework involves using the framework to evaluate existing educational games and comparing the results with expert reviews. With this framework, an educational game quality can be measured objectively and quantitatively based on the technical and fundamental elements that exist in each game.**

***Index Terms*—*educational game, Bloom taxonomy, game evaluation framework, game* mechanics**

I. Introduction

In recent years, technological developments have experienced rapid growth in all fields. Of the many areas, one in high demand is the entertainment sector, especially games. Because more and more games are being played, people start using games as a means for education. However, while tons of educational games have emerged, only a few are of good quality [1]. To find effective games, they need to be evaluated before being used.

Fortunately, several frameworks are available for evaluating educational games. The Game-Based Learning Evaluation Framework [2] and Triadic Game Evaluation [3] offer a way to assess an educational game from a subjective perspective. Therefore, the frameworks are still subjected heavily to the evaluator's opinion, making it difficult to compare with another evaluator's evaluation objectively.

For that, a more objective evaluation standard is needed to assess the various kinds of educational games in circulation so that learners and teachers can determine which games are effective. This study aims to build a game evaluation framework that objectively assesses an educational game to be comparable even among different evaluators.

There are four sections in this paper. Related works are discussed in Section II. Section III presents the framework, while Section IV describes the evaluation process. Section V concludes the discussion and offers suggestions for future works.

II. Related Works

A. Game-Based Learning Evaluation Framework

Games-Based Learning (GBL) Evaluation Framework [2] aims to provide a general framework for evaluating game-based learning, including educational games. This framework is based on a literature study containing empirical evaluation and the methods and metrics used. This framework provides a way to evaluate specific aspects of an educational game, including the game concept. Connolly et al. formulated an effective Game-Based Learning with six components plus one optional component, *Collaboration*. The six components are Learner *Performance*, *Motivation*, *Perceptions*, *Attitudes*, *Preferences*, and *Environment*. This framework relies on an evaluator's expert opinion that may differ from person to person. The framework we propose in this paper aims to cover the subjectivity problem by making objective judgments so that assessments between games can be compared even when performed by different evaluators.

B. Triadic Game Evaluation

Triadic Game Evaluation (TGE) is a framework for evaluating serious games based on the Triadic Game Design (TGD) framework [3]. TGE adopts the philosophy from TGD, where a game must be designed by considering three important aspects; *Reality*, *Meaning*, and *Play*. *Reality* considers the real-world aspects modelled in the game. *Meaning* is the game's goal or purpose, while *Play* is the experience the player feels when playing the game. These three components are considered essential to make an excellent serious game. TGE intends to evaluate the game based on these three aspects by examining the criteria that make up those aspects. These criteria are *flexibility*, *validity*, and *fidelity* for aspects of *Reality*; *motivation*, *relevance*, and *transfer* for aspects of *Meaning*; and *engagement*, *immersion*, and *fun* for *Play*. This framework has shortcomings, mainly in the interpretation of its aspects by evaluators. Different evaluators may have a diverse understanding of an aspect, underlining the framework's reliance on subjective analysis. The proposed framework in

978-1-6654-0446-4/2021/$31.00 ©2021

this research also tries to alleviate this by evaluating a more fundamental component of games: game mechanics.

C. The MDA Framework

The Mechanics-Dynamics-Aesthetics (MDA) Framework is a framework to analyze games [4]. It considers three game components that players encounter, provides precise formal definitions for the components, and explains how they relate to one another to shape the player's experience. The MDA Framework describes *Mechanics* to include the game's rules, the basic actions a player can take, and the algorithms and data structures involved in the game's implementation. *Dynamics* are the behaviours resulted from the interaction of mechanics and player input during run-time, while *Aesthetics* are defined as the player's emotional responses while playing the game. The MDA Framework is shown in Fig. 1.

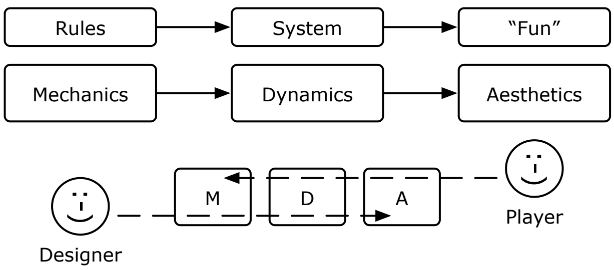

Fig. 1: The MDA framework.

For game designers, game mechanics generate dynamics, which in turn create aesthetics. Meanwhile, players experience a game through its aesthetics, provided by game dynamics, which emerged from the mechanics.

The MDA Framework attempted to provide a formal basis for game analysis. However, it is particularly not suitable for gamified content, and by extension, educational games. The framework proposed in our study adopts some concepts introduced in the MDA Framework.

III. The EGM Framework

Our framework, dubbed the EGM Framework for *Educational Game Mechanics*, derives its concepts of educational targets based on Bloom's taxonomy, which is then combined with the aesthetics and dynamics of the MDA framework to determine which main and supporting mechanics should be used in the game. Data regarding the list of primary and supporting mechanics is obtained from the analysis of various existing educational games. Fig. 2 shows a model of the designed framework.

The mechanics are separated into two parts. The first part concerns the *educational* side of the mechanics, which at its core provides the main mechanics based on educational targets, and the second part offers *supporting* mechanics based on aesthetics and dynamics. The main mechanics were identified from the research of Arnab et al., who mapped learning mechanics with game mechanics [5].

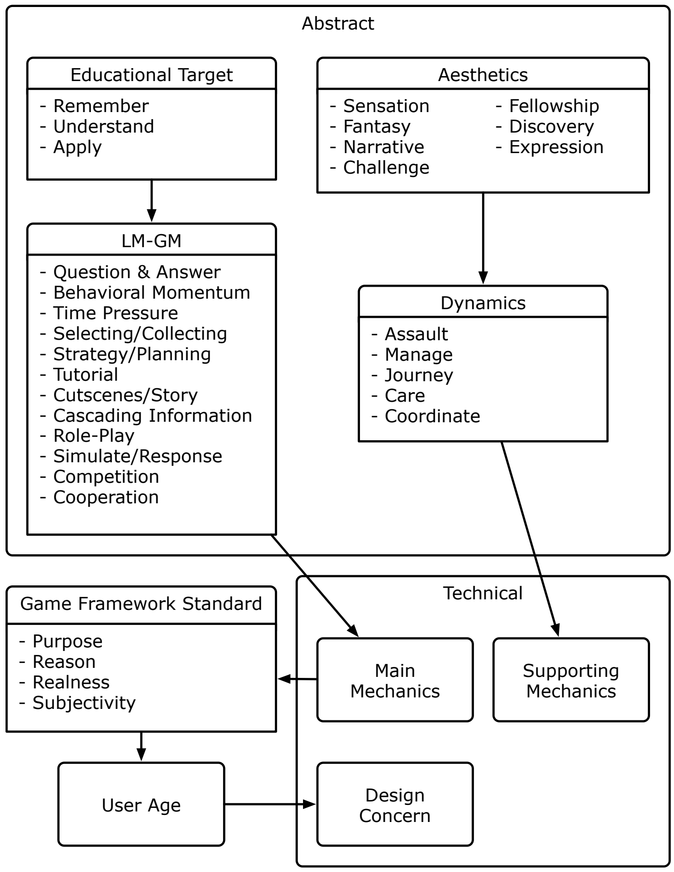

Fig. 2: The model for our framework.

A. Evaluation Standard

Based on the components in existing educational game evaluation frameworks, four elements must be evaluated in an educational game: *Purpose*, *Reason*, *Realness*, and *Subjectivity*.

The first element, *Purpose*, is the educational goal offered by games and how games transfer knowledge to players. This element is supported by GBL Evaluation Framework's *Deployment* component, which focuses on implementing the game into a learning environment, and the *Goals* component, focusing on how game goals help achieve educational goals. The *Purpose* is also supported by tacit knowledge, which suggests that education goals should be reflected implicitly by in-game goals according to Game Object Model (GOM) version II [6], and the transfer element from Meaning in TGE, which suggests how games transfer knowledge to players. We agree that in-game goals must implicitly reflect education goals. By completing these goals, players indirectly acquire an understanding that aligns with their education goals and makes the transfer of knowledge more accessible.

To ensure the effectiveness of educational content delivery, players need to engage with an educational game for a certain period. For players to be interested in continuing to play the game, they need a *Reason*. This element is supported by the *Motivation* component, one of the main elements of the GBL

Evaluation Framework, and the *Play* aspect of the TGE. GOM version II also suggests contents that can make the game feel more attractive, such as emotion, backstory, and plot.

Another vital element is *Realness* or the likeness of the game to reality. *Perceptions* from the GBL Evaluation Framework, which discusses the views of instructors or learners about how realistic games need to be, support this element. GOM version II also discusses the relationship with reality with several additional elements, including *authenticity*, *role models*, *social space*, and *gender*. In addition, TGE has one significant aspect, *Reality*, that includes the validity of realism presented by the game. Based on these points, *Realness* is deemed necessary because the closer the educational game is to the real world, the more seamless it will be for players to apply their in-game knowledge to real life.

The final element to consider is the *Subjectivity* of the evaluator. Each evaluator has different standards, which are the primary considerations in the GBL Evaluation Framework on four of its elements. *Subjectivity* is critical because educational games have specific target users and varied learning materials, so the framework must discern differences in player demographics and learning materials. The preferences, perceptions, and the attitude of the instructor/learner must be considered as they are the ones that know best about the game's requirements, such as the depth of the material needed and the theme raised.

B. Evaluation Metrics

Each mechanic identified in the model is given a metric according to the standard evaluation framework formulated.

Question-and-Answer as a mechanic is more suitable for children who have entered school age because this mechanic has relatively strict rules where players must answer questions. This rule is certainly not appropriate for toddlers who cannot read yet and preschoolers who are not used to answering questions. *THE GROWTH* game [7] uses Question and Answer as one of its mechanics, where players are rewarded for correct answers and punished for wrong answers. Several solutions that can have degrees of correctness provide nuanced rewards, making players more entranced in playing. Questions of the same topic can be grouped to compel a better understanding of the topic.

Behavioral Momentum is a mechanic that encourages a shift in player behaviour through motivation or punishment aversion. Hervas et al. formulated a taxonomy that can be used to support changes in player behaviour [8] and believed that Behavioural Momentum could be modelled in the form of gradual goals to learn from easy to complex so that the motivation is given following the player's ability, or is designed to influence how players perform through punishment. Social status can also be a motivation in ranking, leaderboards, or sharing. In addition, actions in the form of countdowns or routine schedules can shape disciplined behaviour in players. Finally, there is scoring that can provide an assessment or feedback for player performance.

Time Pressure is a mechanic that can be used to accommodate different learning speeds. The optimal Time Pressure should emulate real-life conditions, neither too accelerated nor slowed down [9]. However, Time Pressure applied inconsequently can reduce immersion and break the game's flow due to unruly time constraints [10].

In **Selecting/Collecting**, the player's primary interaction with the game is choosing an object from a more extensive assortment of things. Klopfer et al. used Selecting/Collecting mechanics with many ways to collect clues to increase engagement in the game *Mystery at the Museum* [11]. *EmoJump* provides a list of objects that must be collected so that players can make object selections according to the goal [12]. Selecting/Collecting can also be used in the form of gathering resources needed to create something.

Strategy/Planning focuses on understanding the task given by the game and building solutions to these problems. According to Bottino et al., to help players comprehend the situation, direct feedback is needed to stipulate input to the players' actions [13]. Backtracking can help players see a hypothesized solution from the start and replay the action if considered less than optimal. Providing tips for optimal strategies can also be done to motivate players to think deeper. Some games, like *Tetris*, hint at the next part of the puzzle for the player to anticipate. A view of previous players' actions can also be provided to reflect on the steps they took. Difficulty levels can be made gradual from less complex with little variability to a more complex level. For a higher level of education, complexity can attract players who like strategy games. Hints can also be used if a player is stuck in a level.

Tutorials that appear as needed and according to the player's current impediment provide more benefits to the player than tutorials that appear in the game's beginning [14]. In addition, sound-based tutorials are more favourable because they do not cause players to lose immersion mid-game. It can be more motivating for players when tutorials appear in the appropriate context. Better tutorials are less time-consuming and do not require idle time for players to read tutorials in visual tutorials. White also explained that tutorials could be given more frequently for novice players and decrease according to the player's skill increase.

The **Cutscenes** mechanic helps the game tell its story to the player. According to Dubbelman, narratives in a game can be made realistic by the choices made by players based on moral dilemmas or something that often exists in real life [15]. The suspense in anticipating what will happen can motivate players to keep playing. Moffat et al. demonstrate Holocaust history through a game by giving players the perspective of a Holocaust survivor. By providing an authentic story and character background and an appropriate environment, this game conveys the history and cruelty of the Holocaust well [16].

Information and Instructions are mechanics that are distinctive from tutorials. While tutorials tell players how to play and offer tips for playing effectively, Information and Instructions aim to increase the player's knowledge necessary in completing the game and provide instructions as goals or objectives that must be achieved to end the game. The Player Interaction Framework [17] divides information and instructions into four types. The first is *presentation objects* for information that appears immediately, e.g., a box containing instruction text; the second is *background objects* where players must actively search for concealed information; the third is *person-to-character/object* where the game highlights things or characters that are important to interact with, possibly using a glowing effect on objects or "!" above the head of a Non-Playable Character (NPC); and finally, *storage objects* where players can view the collected items and read descriptions, combine, or analyze these items for new information. The selection of the type of information must be tailored to the game's purpose. From the point of view of motivation, instructions neither increase the player's motivation nor reduce it [18].

Role Play is a mechanic where players can take a particular role in the game. The conflicts between existing roles can increase player engagement to try different roles [19]. The possibility of having player-set objectives also increases the creativity and control of the player. Customs in the real world can be adopted in-game when teaching social norms to make role play feel authentic, including the relationship between roles in the game and those roles in the real world.

Simulate/Response is a mechanic that is used to allow players to experiment freely. This mechanic relies on the principle of cause and effect [20]. Pirker, in her research, stated that interaction is an essential component in simulation games [21]. Interactions can be interactive challenges such as missions that give a sense of purpose to the simulation while still allowing the player to experiment on achieving these goals. Players can also be motivated to do various experiments by providing limited information such as text labels in several places. The game can also provide features that make experiments more interesting such as graphics, effects, and animation since a game's reaction to a player's actions often stimulate their cognition [22]. However, simulations are more suitable for inexperienced players because experienced players are not interested in experimenting in known fields. Giving more things or variables for the player to control also yields positive results. Limiting available actions should be done with care, for irrelevant things or variables can also make the player lose focus. In addition, the representation of an object in various forms such as gas and liquid can help players better understand the thing. Finally, the simulation must have believability, concurring with real-life conditions or corresponding to what players already know.

Competition is a mechanic that allows players to compete for higher status than other players. This status can be obtained by defeating other players, getting a higher score than other players, or other means. Burguillo et al. conducted research related to the application of Competition-based Learning in games. They found that friendly competition, where the extra points obtained are rewards but not punishment, will motivate players more, especially those less reliable in playing the game [23]. Competition between groups is also beneficial because it encourages deeper collaboration within the group due to the desire to beat other groups. The final idea is to create a competitive environment that students desire or according to the educational content being carried. Cooperation is a game where players work together to attain the same goal, like puzzles that must be solved together or resources to be shared. Based on the results of their research, Seif El-Nasr et al. found that an effective cooperation mechanic is to invent mutual goals for players, shared obstacles, similar objects/items, and different but complementary roles for each player [24]. In addition, they also found that games with camera models that follow the main player make players often wait for other players so that it can affect the play experience.

Table I shows an example of a guide to measure the effectiveness of the mechanic Selecting/Collecting. A game will be given one point for each compliance with a metric in the appropriate age group.

TABLE I:
METRICS FOR THE MAIN MECHANIC
SELECTING/COLLECTING

Age Group	**Purpose**	**Reason**	**Realness**
Toddler	Only collect one object for each type	Only one way to pick objects	Realistic objects
Pre-school	Lists the types and number of items to collect	Only one way to pick objects	Realistic objects
Primary school	Lists the types and number of items to collect	Varied means to pick objects	Realistic objects
Tween	Lists the types and number of items to collect	Varied means to pick objects	Realistic objects

IV. EVALUATION

A. Framework Applicability

The goal of this evaluation is to show the framework's applicability in a real-world scenario. Three educational games around periodic table material are assessed in this section. The games are *Periodic Table Quiz* by Sean Burnham, *Periodic Table – Game* by Chernykh [1], and *Periodic Table Game* by Melissa Marinus. These games are evaluated with the tween age group because periodic table material is widely explored in junior high school. In addition, the weight scores for each mechanic are generalized due to the lack of data to be able to give different weights to mechanics. Scores are obtained from the assessment of both main mechanics and supporting mechanics; the latter is obtained using a mechanics recommendation tool based on machine learning. Scores for Question-and-Answer Mechanic are shown in Table II, Scores for Time Pressure Mechanic in Table III, and Scores for Selecting/Collecting Mechanic in Table IV. In the cumulative

[1] www.chernykh.tech

scores, as shown in Table V, both Chernykh and Marinus's table-themed game achieved higher scores, thus are deemed to be better at educating children about the periodic table.

TABLE II:
SCORES FOR THE MAIN MECHANIC **QUESTION-AND-ANSWER**

Standard	Metrics	Burnham	Cherynkh	Marinus
Purpose	Questions are grouped by topics	1	1	0
	Feedback if wrong answer	1	1	1
Reason	Reward	0	1	0
	Reward depends on correctness	0	1	0
Realness	Punishment	0	1	0
	Randomly generated questions	1	1	0
Total		3/6	6/6	1/6

TABLE III:
SCORES FOR THE MAIN MECHANIC **TIME PRESSURE**

Standard	Metrics	Burnham	Cherynkh	Marinus
Purpose	Gradually increasing difficulty	N/A	1	N/A
Reason	Ability to choose stricter time limit	N/A	0	N/A
Realness	Neither too fast nor too slow	N/A	1	N/A
Total			2/3	

TABLE IV:
SCORES FOR THE MAIN MECHANIC **SELECTING/COLLECTING**

Standard	Metrics	Burnham	Cherynkh	Marinus
Purpose	Provides list of items to collect	N/A	N/A	0
Reason	Provides alternative means to pick objects	N/A	N/A	0
Realness	Realistic objects	N/A	N/A	1
Total				1/3

TABLE V:
FINAL SCORES

Score	Burnham	Cherynkh	Marinus
Main mechanics	0.50	0.83	0.23
Supporting mechanics	0.30	0.00	0.60
Final score	0.80	0.83	0.83

B. Expert Review

Model validation was carried out by interviewing three experts from two major game companies in Indonesia. The experts are the Vice President of Gamification from Agate, Chief Executive Officer of Maulidan Games, and Chief Operating Officer of Maulidan Games. According to one expert, our EGM model can be considered complete because it is derived from the widely used Bloom's Taxonomy. Game components derived from aesthetics and dynamics are also regarded as sufficient in game design to be used as a basis for evaluating a game. He commented on the use of Bloom's Taxonomy which only considers the bottom three levels. However, from the educational perspective, children aged 12 years and under are not expected to perform analysis and more complex tasks, which belongs to Bloom's Taxonomy level 4+.

Meanwhile, other experts commented on a complete framework model, with the educational side derived from Bloom's Taxonomy and linked to game mechanics through LM-GM. One expert's attention is more focused on the absence of demographic aspects of the user that should influence the user experience and the effectiveness of mechanics towards learning objectives. He also agrees that the demographic factors of the user, especially the target age of the user, must be considered because, for example, games for toddlers who cannot read are undoubtedly different from grade 5 elementary school children. The motivations of the two users to play educational games also differ, so evaluating educational games should consider the age aspect of the target user.

V. CONCLUSIONS

By utilizing the framework built in this study, the usefulness of an educational game can be quantitatively measured so that it can be assessed objectively based on the suitability of game mechanics compared to other games that are already on the market and through existing standards for educational games. The evaluation results can also be used to enhance the game's next iteration.

For future works, one weakness of this framework is that evaluators need to identify mechanics present in an educational game before using this framework. For further research, the identification process can be simplified so that evaluators can identify mechanics in an education game more quickly, making the framework more accessible to people. Another topic that warrants future study is the weighting for each game mechanic, which involves finding the effect of one mechanic on another.

Acknowledgment

The authors would like to express our gratitude to the three experts for their contribution to this research. Without their valuable input, this research can never be complete.

References

[1] S. Papadakis and M. Kalogiannakis, "Mobile educational applications for children: what educators and parents need to know," *International Journal of Mobile Learning and Organisation*, vol. 11, no. 3, pp. 256–277, 2017.

[2] T. Connolly, M. H. Stansfield, and T. Hainey, "Development of a general framework for evaluating games-based learning," in *Proceedings of the 2nd European conference on games-based learning*. Universitat Oberta de Catalunya Barcelona, Spain, 2008, pp. 105–114.

[3] C. Harteveld, "Triadic game evaluation: A framework for assessing games with a serious purpose," in *Workshop of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, 2010.

[4] R. Hunicke, M. LeBlanc, and R. Zubek, "Mda: A formal approach to game design and game research," in *Proceedings of the AAAI Workshop on Challenges in Game AI*, vol. 4, no. 1. San Jose, CA, 2004, p. 1722.

[5] S. Arnab, T. Lim, M. B. Carvalho, F. Bellotti, S. De Freitas, S. Louchart, N. Suttie, R. Berta, and A. De Gloria, "Mapping learning and game mechanics for serious games analysis," *British Journal of Educational Technology*, vol. 46, no. 2, pp. 391–411, 2015.

[6] A. Amory, "Game object model version ii: a theoretical framework for educational game development," *Educational Technology Research and Development*, vol. 55, no. 1, pp. 51–77, 2007.

[7] P. Lameras, C. Pisithpunth, and P. Petridis, "A quiz-based game for addressing growing population issues: Linking learning mechanics to the growth serious game," in *2014 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL2014)*. IEEE, 2014, pp. 344–349.

[8] R. Hervas, D. Ruiz-Carrasco, T. Mondejar, and J. Bravo, "Gamification mechanics for behavioral change: a systematic review and proposed taxonomy," in *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 2017, pp. 395–404.

[9] S. de Freitas and M. Ott, *New pedagogical approaches in game enhanced learning: Curriculum integration*. IGI Global, 2013.

[10] J. Denholm and L. Lee-Davies, *Enhancing education and training initiatives through serious games*. IGI Global, 2017.

[11] E. Klopfer, J. Perry, K. Squire, M.-F. Jan, and C. Steinkuehler, "Mystery at the museum–a collaborative game for museum education," in *Computer Supported Collaborative Learning 2005: The Next 10 Years!* Routledge, 2017, pp. 316–320.

[12] J. Leyrer, "Advancing assessment and promotion of social cognition and emotion understanding in children," Ph.D. dissertation, uniwien, 2019.

[13] R. M. Bottino, L. Ferlino, M. Ott, and M. Tavella, "Developing strategic and reasoning abilities with computer games at primary school level," *Computers & Education*, vol. 49, no. 4, pp. 1272–1286, 2007.

[14] M. M. White, "L2p noob: Examining tutorials in digital games," *Loading...*, vol. 6, no. 10, 2012.

[15] T. Dubbelman, "Narrative game mechanics," in *International Conference on Interactive Digital Storytelling*. Springer, 2016, pp. 39–50.

[16] D. C. Moffat and A. Shapiro, "Serious games for interactive stories about emotionally challenging heritage," in *2015 Digital Heritage*, vol. 2. IEEE, 2015, pp. 709–712.

[17] R. Wainess, A. Koenig, and D. Kerr, "Aligning instruction and assessment with game and simulation design. cresst report 780." *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*, 2011.

[18] S. Erhel and E. Jamet, "Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness," *Computers & education*, vol. 67, pp. 156–167, 2013.

[19] R. Linser, N. R. Lindstad, and T. Vold, "The magic circle-game design principles and online role-play simulations," in *EdMedia+ Innovate Learning*. Association for the Advancement of Computing in Education (AACE), 2008, pp. 5290–5297.

[20] A. Patino, M. Romero, and J.-N. Proulx, "Analysis of game and learning mechanics according to the learning theories," in *2016 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*. IEEE, 2016, pp. 1–4.

[21] J. Pirker and C. Gütl, "Educational gamified science simulations," in *Gamification in education and business*. Springer, 2015, pp. 253–275.

[22] W. K. Adams, S. Reid, R. LeMaster, S. B. McKagan, K. K. Perkins, M. Dubson, and C. E. Wieman, "A study of educational simulations part i-engagement and learning." *Journal of Interactive Learning Research*, vol. 19, no. 3, pp. 397–419, 2008.

[23] J. C. Burguillo, "Using game theory and competition-based learning to stimulate student motivation and performance," *Computers & education*, vol. 55, no. 2, pp. 566–575, 2010.

[24] M. Seif El-Nasr, B. Aghabeigi, D. Milam, M. Erfani, B. Lameman, H. Maygoli, and S. Mah, "Understanding and evaluating cooperative games," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2010, pp. 253–262.