International Journal of Serious Games

ISSN: 2384-8766
https://journal.seriousgamessociety.org/

Article

A Linear Programming Methodology for Evaluating Game Attributes in Serious Games

Ernesto Pacheco-Velazquez[1], Virginia Rodes-Paragarino[1], Andre Bester[2] and Lucia Rabago-Mayer[2]

[1]*Institute for the Future of Education, Tecnologico de Monterrey, Monterrey, Mexico;* [2]*BMS Lab, University of Twente, Enschede, Netherlands*
epacheco@tec.mx ; virginia.rodes@tec.mx ; a.n.bester@utwente.nl ; l.m.rabago-mayer@tec.mx

Keywords:

Serious Games
Digital Education
Higher Education
Educational Innovation
Decision Making
Linear Programming

Received: May 2024
Accepted: October 2024
Published: November 2024
DOI: 10.17083/ijsg.v11i4.802

Abstract

Serious games have emerged as powerful tools for enhancing learning experiences, particularly in educational settings where engagement, interactivity, and problem-solving are key. While the effectiveness of serious games is widely recognized, it remains challenging to find a standardized approach to evaluate their key attributes. This study presents a novel methodology based on linear programming to assess the relative importance of game attributes such as concentration, clarity of objectives, feedback, challenge, autonomy, immersion, social interaction, and knowledge enhancement. By applying this methodology, we aim to quantify the most appropriate weight that this set of attributes should have to explain the game's overall rating. The findings provide a structured framework for game developers and educators to optimize serious games, ensuring they align with user preferences and educational outcomes. This methodological approach offers valuable insights for improving game design and increasing the impact of serious games in both academic and professional training environments.

1. Introduction

Information and communication technologies (ICTs) have brought about a paradigm shift in people's perception, interpretation, and learning of the world. These technologies cause fundamental changes in the workplace and, as a result, higher education institutions need to modify their practices and curricula. Although this evolving landscape presents us with novel challenges, it is crucial to understand that it also offers a range of innovative tools and opportunities to enrich student learning experiences.

Within these opportunities, simulations, serious games (SGs), or simulation games offer us the chance to create environments where students and academics can solve real-world problems in a risk-free setting. The objective of these tools is to provide students with a better understanding of complex phenomena, expand their knowledge base, encourage the use of technological tools, and facilitate collaborative work. This way, the overall learning experience for students can be more enjoyable.

Over the years, SGs have gained significant attention due to their unique ability to engage individuals in learning and training activities [1], [2]. Learning games have been specifically designed to offer an enjoyable and captivating way for individuals to acquire knowledge and skills, while also promoting the retention and application of learning. These games are developed with the aim of facilitating learning in an interactive and entertaining manner, with a strong emphasis on enhancing engagement and motivation [3], [4], [5]. Furthermore, some authors argue that SGs also contribute to the development of critical thinking skills in students [6], improve information retention [7], stimulate creativity [8], and foster the development of problem-solving and decision-making abilities [9].

In learning games, customization to meet the individual needs and preferences of learners is also possible. The concept of individualization, which involves tailoring the game to align with the learner's abilities, preferences, and learning style, has been identified as a vehicle to enhance engagement and motivation, ultimately leading to more effective learning outcomes [10], [11]. This personalized approach can help address the challenges associated with catering to the diverse needs of learners who possess varying levels of knowledge and skills [12]. When developing a SG, it is vital to consider the individual differences among players. People have varying learning speeds and preferences, and they engage with distinct learning styles. Additionally, players may use diverse strategies or consider different elements to reach similar goals, due to having different skills or abilities [13] [14]. It is crucial to acknowledge that the player's interaction with the game is dynamic, which can give rise to challenges such as loss of motivation or predictability. Consequently, the individual characteristics and learning styles of the players can significantly impact how the game is perceived in terms of its effectiveness [15].

The effectiveness of SGs can be attributed to a wide range of attributes. One study identifies eighteen of them, including aesthetics and design, interface, gameplay, learnability, among others [16]. Additionally, research has highlighted the importance of feedback, challenge, interactivity, and immersion as key components for SG effectiveness [17], [18]. Furthermore, the significance of attributes such as the setting, integration of the game with learning, interaction, feedback, and game design has been emphasized in determining the effectiveness of these games [19], [20]. Another study points out the existence of seven primary attributes and a group of 29 sub-attributes that comprise them, categorizing them into three dimensions: effectiveness, efficiency, and satisfaction [21].

However, there is no methodology that helps determine the importance or weight of each of these attributes when users tend to evaluate a SGs. Some researchers claim that the assessment of these components is misleading [22] or that it is biased [23]. One of the major challenges in evaluating SGs is that their nature involves experts from various fields.

This article presents a methodology to determine the most appropriate weight that a set of attributes should have in explaining the overall game rating, based on the opinions of a group of users. The intent of this methodology is to determine which features have the greatest influence on the overall evaluation of a SG. The methodology is crucial because it allows us to determine the importance of various facets in the evaluation of a simulation game by players. In other words, the methodology allows us to differentiate which elements have the greatest importance for a specific simulation game within a set of attributes. It helps us to understand whether user evaluations are consistent across different options or topics. By employing a methodology, we can obtain valuable insights and make informed decisions regarding the design, improvement, effectiveness, and challenges of SGs.

In the following section, we present the theoretical framework, explaining the foundations of the decision-making models and linear programming. Next, we conduct a literature review, highlighting relevant studies on the evaluation of serious games and the methodologies used for assessing their effectiveness. Section 4 describes the background of the study and outlines the key objectives related to evaluating attributes in serious games. In the methodology section, we describe the construction of the linear programming model, followed by the results obtained from applying this model to the evaluation of game attributes. Finally, the paper concludes with a discussion of the findings, the limitations of the approach, and suggestions for future research.

2. Theoretical Framework

Decision Theory has the fundamental goal of understanding and improving the decision-making process in situations where individuals or organizations face multiple alternatives and must select the optimal option. This science aims to develop models and methods that allow for the structured evaluation of available alternatives, considering factors such as uncertainty, individual preferences, risks, and the rewards associated with each choice [24] [25]. Essentially, Decision Theory has two main branches:

- Normative Theory: This branch focuses on how rational decisions should be made. It develops mathematical models and algorithms that maximize outcomes, such as utility or benefit, under conditions of certainty or uncertainty. An example includes the use of techniques such as linear programming or expected utility theory to find the best decision based on defined objectives [24] [26].
- Descriptive Theory: This branch centers on how people actually make decisions in practice. It examines human behavior, which includes emotional factors, cognitive biases, and informational limitations. In this branch, models such as heuristics or elimination by aspects help explain how people simplify complex decisions [27] [28]. The Prospect Theory is another key contribution, showing how people make decisions under risk, often deviating from rational principles [29].

2.1 Normative Theory

Normative Theory in the context of Decision Theory focuses on how decisions should be made to be considered optimal or rational. Instead of studying how people make decisions in practice, as descriptive theory does, normative theory establishes principles and mathematical rules that decision-makers should follow to maximize outcomes, minimize errors, and optimize resources. In this area, mathematical optimization methods become fundamental. Linear Programming (LP) is a mathematical optimization technique that is widely used to optimize linear functions, which are subject to a set of linear constraints. The primary goal of linear programming is to find the optimal values of the variables that maximize or minimize the objective function, while satisfying all the constraints imposed by the problem [30]. A fundamental aspect of linear programming is that both the constraints and the objective function must be linear, meaning they cannot include quadratic or nonlinear terms. This characteristic simplifies calculations and allows the use of efficient algorithms such as the Simplex Method to solve large-scale problems [31]. Given its effectiveness in identifying optimal solutions within a set of defined objectives and constraints, linear programming is often associated with decision theory, whose primary goal is to understand and enhance the decision-making process when individuals or organizations face multiple alternatives and must choose the optimal option [32]. Typically, this entails individuals assessing different factors known as "attributes". These attributes often hold varying levels of importance for each individual, which explains why different people arrive at different evaluations when considering the same set of alternatives [24].

Decisions related to the choice or evaluation of a specific object or activity can be considered as a complex task [25]. These activities, even when performed automatically, are often regarded as demanding and challenging. One of the main reasons is that the elements being chosen or evaluated typically consist of a set of attributes that are difficult to compare [33]. Furthermore, when the evaluation involves multiple non-binary numerical attributes, the complexity increases significantly [25]. This type of problem is known as multi-attribute decision making (MADM). In this field, linear programming is employed to analyze how individuals assign weights to different attributes when making decisions. This is useful for modeling complex choices, such as purchasing a product, choosing a career, or selecting a supplier, where multiple factors must be considered simultaneously. By optimizing the combination of these factors, linear programming allows for predicting which attributes are most important to decision-makers [24]. A key work in multi-criteria decision theory introduces a systematic approach to decision-making in scenarios with multiple objectives or criteria. One of its most significant contributions is the multi-attribute utility model, which evaluates complex decisions by assigning a numerical value (utility) to each alternative based on its attributes. These attributes are weighted according to their importance, and the weighted sum of these utility values allows decision-makers to compare and select the best option [24].

Several authors have sought to combine the quantitative approach of decision analysis with empirical research on human behavior [34]. In this work, the authors examine how decision-makers assign weights to various attributes under evaluation, acknowledging that these weights are not fixed but fluctuate depending on the context and the individual's subjective preferences. A crucial aspect of decision analysis is identifying relevant criteria or attributes and subsequently weighting them to reflect their relative importance. One of the key contributions of this work is its focus on how decision-makers allocate weights to the attributes they deem important, ultimately selecting the option that maximizes their utility or overall satisfaction. The authors also argue that individuals' preferences, and thus the weights assigned to each attribute, are not static; they can shift based on the decision context or personal priorities.

Other studies also highlight the importance of tools like linear programming in decision-making [35]. However, the authors of this work argue that decision-makers do not consistently follow a single, rational approach; instead, they adapt their strategies based on task complexity, available time, and informational resources. They suggest that the decision-making process is neither static nor universal. Rather, individuals assess their decision environment—including time constraints, the amount of available information, and the significance of the decision—and adjust their strategies accordingly. This adaptive approach is crucial for understanding human flexibility when confronted with different decision challenges.

In recent years, the Best-Worst Method (BWM) has established itself as a powerful tool within multi-criteria decision theory, allowing decision-makers to efficiently structure and weight criteria [36]. Its simplicity and ability to improve consistency make it an ideal methodology for complex scenarios where multiple criteria need to be compared. The methodology has been successfully applied in areas such as supply chain management, public policy evaluation, and project selection, providing reliable and efficient results. BWM is particularly useful in situations where decision-makers must evaluate alternatives based on multiple, often conflicting, criteria.

2.2 Descriptive Theory

Descriptive Theory within the field of Decision Theory focuses on how people actually make decisions in practice. This theory is based on the observation and analysis of human behavior, recognizing that individuals often do not follow a strictly rational approach due to cognitive limitations, emotions, and biases. The goal of Descriptive Theory is to explain and predict real decision-making patterns, rather than necessarily optimize them [29]. People often rely on their own judgment to determine the importance of each attribute. This internal process is known as "importance judgment", and the weight assigned to these attributes is used to carry out the evaluation process [37]. This evaluation process is often based on common sense, and to carry it out, a set of strategies known as “heuristics” is used [38]. Although heuristics are often useful, they can sometimes lead to significant errors, commonly known as biases [39].

Heuristics in decision-making refer to mental shortcuts or rules of thumb that individuals use to simplify intricate decision problems. One issue with heuristics is that they are often susceptible to our own cognitive biases [40]. This susceptibility has contributed to heuristics being perceived as inferior techniques for decision-making, often associated with irrational decision behaviour. However, recent studies by decision-making researchers have shown that certain heuristics can be remarkably effective and even rival complex decision models in specific domains [41]. Heuristics are highly valuable in circumstances characterized by time, resource, or information constraints. They are often based on past experiences or common sense, and they allow individuals to make decisions quickly without engaging in exhaustive analysis or considering all available information [42]. While heuristics can be beneficial in simplifying decision processes and saving cognitive effort, it is also true that they can lead to biases and errors [36]. The theoretical principles of decision analysis explain a theory based on weighted additive utility [24]. This theory states that when humans face the decision to evaluate a set of attributes, people often rely on a series of heuristics that simplify the calculation. Some theories suggest that the evaluation can be solely based on the attributes that participants consider most important [28] [43]. These heuristics simplify the evaluation algorithm by considering only a small subset of attributes in the decision, instead of weighting all the attributes [44].

Among the most important works in this area is the one by Tversky [27], which explains how people simplify complex decisions by eliminating alternatives based on key attributes, without evaluating them exhaustively. Other authors analyze how people dynamically weight different attributes depending on the context and the information available [45]. More recent studies assert that people adapt their decision strategies according to the context and available resources [35]. One of the most notable works [28] focuses on intuitive decisions and the use of heuristics, examining how people make quick decisions based on a few key attributes.

Both areas recognize that one of the most important steps in formulating a problem lies in the weighting of attributes [36]. Most existing weighting methods are based on expert judgments and often involve a set of cognitive biases. For example, one of these biases is the splitting bias, which means that by presenting an attribute in greater detail, experts tend to increase the relative weight assigned to that element [46] [47]. There is also the equalizing bias, which represents the tendency of decision-makers to assign equal weights to different attributes [48] [49]. A third type of bias is called the anchoring bias, where initial values are provided and influence the experts' perception [50]. Five different methods for calculating the weights of these attributes have been identified, including simple multi-attribute rating technique (SMART), Swing, Point allocation (PA), analytic hierarchy process (AHP), and best-worst method (BWM) [33]. When using these methods, one observation is that the weights are based on the opinions of a limited set of experts, and these opinions may differ from those of non-expert users.

3. Literature Review

The evaluation of serious games is a key area of research due to their growing use in educational contexts, professional training, and the entertainment industry. Serious games are developed with purposes beyond entertainment, such as education, training, or raising awareness on social and scientific issues. Their evaluation is essential to understand their effectiveness in achieving the proposed pedagogical or training objectives, as well as to improve their design, usability, and the impact of the game on users [51] [52].

Evaluating these games allows developers and educators to measure whether the learning or training objectives are being met and to understand the user experience. Through rigorous evaluation, it is possible to identify which aspects of the game are effective for learning, motivation, and knowledge retention. Furthermore, in contexts such as health or military training, where these games are used, their evaluation is crucial to ensure that players develop the necessary skills efficiently and safely [53] [54].

The evaluation of serious games encompasses several key aspects. One of the main challenges is measuring how these games impact learning. While numerous studies have shown that serious games can enhance knowledge retention and practical skills development [53], isolating the effects of the game from other external factors remains complex. Additionally, the tools and scales used to assess learning vary greatly between studies, making direct comparisons of results difficult. Another important element is the motivation and engagement that games generate in participants. Although some studies suggest that the level of fun or engagement does not always lead to greater learning [54], many others highlight that serious games have a positive impact on both motivation and learning outcomes [55] [53] [56] [57]. A third commonly discussed aspect is skill transfer. In professional contexts such as military and medical training, it is crucial that the knowledge acquired in simulation environments effectively translates to real-world scenarios [58] [59]. The challenge here lies in designing evaluations that not only assess learning within the game but also how it translates into decision-making and practical skill development in real-life situations.

However, evaluating the methodologies used to measure the effectiveness of serious games remains an underexplored area. Some studies point to the lack of consistent evaluation standards and acknowledge the need to develop tools that adapt to the context in which games are used [60]. Although various evaluation methodologies have been proposed, such as the Game Engagement Questionnaire [61], EGameFlow [62], and the Serious Game Quality Model [63], there is still no consensus on the most suitable approach to comprehensively evaluate the diverse aspects of a serious game. Despite the growing interest in serious games, there is still a lack of systematic tools to assess the quality of their design and their impact on players [64].

4. Background and Objectives of the Study

In 2015, a logistics-focused game was developed, simulating the operations of a small company involved in the production and sale of balls. This game provides a simulation of a compact supply chain, incorporating suppliers, factories, and stores, and encompasses all the logistics-related decisions that a company would face. The game has gained popularity among students from various universities and countries, as well as small business owners and logistics practitioners. It is accessible online through the "GOAL Project" portal (https://goalproject.co).

To measure different aspects of the game, an adaptation of the EGameFlow scale [42] was applied to some of the users. The original scale consists of 56 questions and measures eight different attributes of the game: Concentration (C), Goal Clarity (G), Feedback (F), Challenge (CH), Autonomy (A), Immersion (I), Social Interaction (S), and Knowledge Improvement (K).

Before participating in the study, all participants were informed about the objectives and procedures of the experiment. They were provided with an informed consent form that explained their rights, the confidentiality of the data, and the option to withdraw from the study at any time without any consequences. Participation was entirely voluntary. The survey administered to the game participants, as described in Table 1, closely follows the proposed scale, with seven questions removed due to being considered repetitive (for example, "*The game grabs my attention*" and "*The game provides content that stimulates my attention*"), or because they were not applicable to the logistics simulator (for example, the presentation of intermediate goals when the game does not include such goals; or an increase in the difficulty of new challenges, when the simulator does not have this feature). In summary, two questions were removed from the attribute called Concentration, two questions from the attribute 'Goal Clarity', and three questions from the attribute Challenge. Since the majority of sub-attributes in these three attributes were retained, the authors consider that the survey results were not significantly affected by this modification. As a result, a total of 49 questions were administered to the participants, along with their game ratings. The survey was completed by 255 participants. For answering the questions related to each evaluated attribute, a 7-point Likert scale was used (where 1 represents "strongly disagree" and 7 represents "strongly agree").

The overall game evaluation was carried out using a scale from 1 to 10, where 1 means 'the game is terrible' and 10 means 'the game is excellent.' Each student answered this question after completing the EGameFlow scale. The results of the overall game evaluation are presented in the last row of Table 1, which explains why the evaluation score exceeds 8.

The average obtained for each question, as well as the average game evaluation and the standard deviation associated with each item, are also described in Table 1.

The main objective of this research is to present a new methodology associated with the measurement of the importance that users give to each of the attributes related to the game.

There are four secondary questions that we aim to answer with this exercise:

1) Do participants assign different weights to each attribute, reflecting varying levels of importance in the game evaluation?
2) Are the weights assigned to different attributes similar, suggesting that users give equal importance to each of the evaluated attributes?
3) Does the number of questions about a certain attribute influence people's perception of the importance of that attribute?
4) When using this method, do people tend to choose only a reduced set of attributes that they consider important?

5. Methodology and Results

The problem is to determine the weight each attribute holds in the participants' overall game rating. In other words, it involves finding the value of eight variables (F1 to F8), one for each attribute, where each variable represents the weight assigned by participants. The overall game rating is then calculated by multiplying the score for each attribute by its corresponding variable.

Table 1. Summary of basic survey statistics

Sub-attribute	Item	Mean	St. Dev.
Concentration			
C1	The game grabs my attention	6.019	1.106
C2	Most of the gaming activities are related to the learning task	6.075	1.157
C3	No distraction from the task is highlighted	5.151	1.540
C4	Generally, I can remain concentrated in the game	5.509	1.311
C5	I am not distracted from tasks that the player should concentrate on	5.491	1.317
C6	I am not burdened with tasks that seem unrelated	5.528	1.438
Goal Clarity			
G1	Overall game goals were presented in the beginning of the game	6.358	1.239
G2	Overall game goals were presented clearly	5.906	1.301
G3	I understand the learning goals through the game	6.151	1.111
Feedback			
F1	I receive feedback on my progress in the game	5.057	1.809
F2	I receive immediate feedback on my actions	5.189	1.743
F3	I am notified of new tasks immediately	5.453	1.732
F4	I am notified of new events immediately	5.377	1.690
F5	I receive information on my success (or failure) of intermediate goals immediately	5.491	1.655
F6	I receive information on my status, such as score or level	6.132	1.328
Challenge			
CH1	I enjoy the game without feeling bored or anxious	5.340	1.545
CH2	The challenge is adequate, neither too difficult nor too easy	5.811	1.260

CH3	The game provides "hints" in text that help me overcome the challenges	5.434	1.736
CH4	The game provides "online support" that helps me overcome the challenges	5.094	1.662
CH5	The game provides video or audio auxiliaries that help me overcome the challenges	5.906	1.537
CH6	My skill gradually improves through the course of overcoming the challenges	6.094	1.197
CH7	I am encouraged by the improvement of my skills	6.245	1.247
Autonomy			
A1	I feel a sense of control the menu (such as start, stop, save, etc.)	6.264	1.068
A2	I feel a sense of control over actions of roles or objects	6.245	0.958
A3	I feel a sense of control over interactions between roles or objects	6.208	0.911
A4	The game does not allow players to make errors to a degree that they cannot progress in the game	5.434	1.837
A5	The game supports my recovery from errors	5.623	1.377
A6	I feel that I can use strategies freely	6.283	1.076
A7	I feel a sense of control and impact over the game	6.132	0.958
A8	I know next step in the game	6.151	1.111
A9	I feel a sense of control over the game	5.717	1.199
Immersion			
I1	I forget about time passing while playing the game	5.264	1.679
I2	I become unaware of my surroundings while playing the game	5.226	1.854
I3	I temporarily forget worries about everyday life while playing the game	5.019	1.996
I4	I experience an altered sense of time	5.358	1.841
I5	I can become involved in the game	5.642	1.370
I6	I feel emotionally involved in the game	5.717	1.424
I7	I feel viscerally involved in the game	5.528	1.537
Social Interaction			
S1	I feel cooperative toward other classmates	5.736	1.569
S2	I strongly collaborate with other classmates	5.736	1.655
S3	The cooperation in the game is helpful to the learning	6.057	1.349
S4	The game supports social interaction between players (chat, etc)	5.057	1.666
S5	The game supports communities within the game	5.057	1.716
S6	The game supports communities outside the game	5.283	1.669
Knowledge Improvement			
K1	The game increases my knowledge	6.113	1.328
K2	I catch the basic ideas of the knowledge taught	6.245	1.178
K3	I try to apply the knowledge in the game	6.189	1.159
K4	The game motivates the player to integrate the knowledge taught	6.151	1.243
K5	I want to know more about the knowledge taught	6.019	1.280
OVERALL RATING OF THE GAME		8.661	1.707

For example, observe Table 2 and suppose that a participant assigns the following ratings to each of the attributes. Let's assume that this player assigns a high weight to the "Immersion" attribute. In this case, the game's overall evaluation (GOE) should be low, meaning that the game receives a poor rating from this participant. On the other hand, if the player assigns a very high weight to the "Feedback" attributes in the evaluation, the GOE is favourable.

Table 2. Example of players' evaluation for each attribute of the game

Attribute	Player Evaluation
Concentration (C)	7.35
Goal Clarity (G)	9.16
Feedback (F)	9.88
Challenge (CH)	8.00
Autonomy (A)	5.40
Immersion (I)	3.51
Social Interaction (S)	6.22
Knowledge Improvement (K)	8.42

In general, regardless of the evaluation each player gives to each attribute, the GOE can be determined using the formula described in Equation 1:

$$GOE = F1 * C + F2 * G + F3 * F + F4 * CH + F5 * A + F6 * I + F7 * S + F8 * K \quad (1)$$

In this example, specific values for the weight of each attribute have not been assigned. However, it is possible to conduct experiments with different values to obtain these weights. It is to be expected that assigned weights may vary depending on the objectives and characteristics of the game.

If the weights assigned to the attributes are indeed close to those that the players have considered, the overall evaluation of the game calculated by Equation 1 is close to what the participant reported.

Now consider Equation 2:

$$GOE = F1 * C + F2 * G + F3 * F + F4 * CH + F5 * A + F6 * I + F7 * S + F8 * K + L - E \quad (2)$$

In this case, the variable L can be interpreted as a shortfall, while the variable E can be associated with an excess. If the weights assigned to the attributes are appropriate, it is expected that the variables E and L have smaller values. The smaller these values are, the more suitable the assignment of weights is considered.

If the variable Xi is considered as the evaluation of attribute X made by individual i, then Equation 3 is obtained:

$$GOE_i = F1 * C_i + F2 * G_i + F3 * F_i + F4 * CH_i + F5 * A_i + F6 * I_i + F7 * S_i + F8 * K_i + L_i - E_i \quad (3)$$

It is expected that the more accurate the weights assigned to each attribute, the smaller the shortfalls or excesses are. Therefore, our objective is to minimize the sum of shortfalls or excesses among the individuals who have participated in this evaluation.

If we consider a sample of 100 individuals, we can formulate the problem using the following equations:

$$Min\, Z = \sum_{i=1}^{100} (Li + Ei) \quad (4)$$

subject to

$$GOE_i = F1 * Ci + F2 * Gi + F3 * Fi + F4 * CHi + F5 * Ai + F6 * Ii + F7 * Si + F8 * Ki + Li - Ei \quad (5)$$

$$i = 1, 2, \ldots, 100$$

In this case, Equation 4 represents the objective function, while Equation 5 represents a set of 100 constraints. A fundamental characteristic of both equations is that they are linear functions, and therefore, it is possible to use linear programming to solve the problem.

Additionally, since the sum of the attribute's weights must equal 100%, we have an additional constraint (see Equation 6).

$$F1 + F2 + F3 + F4 + F5 + F6 + F7 + F8 = 1 \quad (6)$$

The first step in performing the initial calculations involved selecting 100 random individuals (out of the 255 who completed the surveys) and using these 100 people to generate the constraints for this problem. Since the rating of attributes is calculated on a 7-point Likert scale, and the GOE is on a 1 to 10 scale, the first step was to divide the responses evaluating the attributes by 0.7. Next, the coefficients for each attribute were obtained. For example, the "Concentration" attribute consists of six associated questions, so the mean of the responses associated with this attribute was calculated and then divided by 0.7. This provides a measurement on a 1 to 10 scale.

Based on these responses, the objective function and constraints associated with each individual were constructed, and the GAMS software was used to obtain the solution to this problem (since it actually reduces to a linear programming problem). The results obtained are presented in Table 3.

Table 3. Weighted weights obtained by solving the linear programming problem.

Attribute	Weighted (%)
Concentration (C)	17.82
Goal Clarity (G)	0.00
Feedback (F)	3.34
Challenge (Ch)	0.00
Autonomy (A)	0.00
Immersion (I)	0.00
Social Interaction (S)	13.23
Knowledge Improvement (K)	65.62

Based on this data, it was decided to increase some constraints that would limit the value of the weight of each attribute. Specifically, it was requested that the maximum value for the weight of each attribute be 50%, and the minimum value for the weight of each attribute be 3%. This involved generating 16 new constraints, eight of them to limit the maximum value and another eight to ensure the minimum value of the weight of each attribute.

After incorporating these constraints, the model was run 15 times, always choosing 100 data items at random. The results obtained are shown in Table 4.

As can be observed, there is a significant difference between the results obtained from the first and second models. One of the reasons is that when there is a attribute that dominates the overall evaluation, it is advisable to limit the weight that participants give to this attribute. In this case, the importance that players attribute to Knowledge Improvement explains the rating given to the game. However, at the same time, this prevents the importance of other attributes in this evaluation from being observed. When such a dominant attribute appears in a model, constraining the importance of that attribute allows the importance of other attributes to be recognized. This is one of the reasons that explain the difference in results between the first and second models.

5.1 Methodological Discussion of the Results

A brief analysis of the results reveals that the attribute that appears to carry the most weight in users' evaluations is the Knowledge Improvement (F8). This attribute consistently receives the highest weight in 14 out of the 15 solutions examined.

Table 4. Results obtained after generating 15 models using random data.

	Weighted Value of Attribute							
Model	F1	F2	F3	F4	F5	F6	F7	F8
1	0.0574	0.0300	0.1008	0.0300	0.3457	0.0300	0.0407	0.3654
2	0.2117	0.0300	0.0300	0.0300	0.0300	0.1383	0.0300	0.5000
3	0.3095	0.0300	0.0300	0.0300	0.0300	0.0300	0.0939	0.4467
4	0.1710	0.1281	0.0368	0.1293	0.3057	0.0300	0.1101	0.0889
5	0.3100	0.0431	0.1022	0.0300	0.0300	0.0331	0.0858	0.3659
6	0.1710	0.0300	0.0858	0.0300	0.0300	0.0934	0.0599	0.5000
7	0.3508	0.0300	0.0952	0.0300	0.0300	0.0300	0.0740	0.3600
8	0.0488	0.0444	0.0300	0.2325	0.0300	0.0843	0.0300	0.5000
9	0.0467	0.0300	0.0849	0.1556	0.0300	0.1228	0.0300	0.5000
10	0.1648	0.0525	0.0300	0.0300	0.0300	0.0300	0.1627	0.5000
11	0.0700	0.1642	0.1283	0.1045	0.1760	0.0300	0.0300	0.2970
12	0.0300	0.0300	0.0300	0.3230	0.0300	0.1214	0.0300	0.4056
13	0.1257	0.1049	0.1232	0.0300	0.0300	0.0300	0.0563	0.5000
14	0.2624	0.0300	0.0300	0.0300	0.0300	0.0663	0.0513	0.5000
15	0.1504	0.0913	0.0300	0.0300	0.0300	0.1383	0.0300	0.5000
Mean	0.1653	0.0579	0.0645	0.0830	0.0792	0.0672	0.0610	0.4220
St. Dev.	0.1056	0.0433	0.0389	0.0910	0.1072	0.0447	0.0387	0.1154

Following closely behind is the attribute of Concentration (F1), which consistently ranks second in eight out of the fifteen models. In third place, we find the Challenge attribute (F4), closely followed by Autonomy (F5).

Across all the different models, it is evident that certain attributes consistently have a weight of 3% (which is the minimum requirement within the set of restrictions). This strongly suggests that these attributes had no significant weight in the actual problem, indicating that users tended to disregard these variables when evaluating the game. This finding aligns with the heuristics that people tend to employ in order to simplify their decision-making process by minimizing the number of variables they consider when evaluating a problem. It reflects a common cognitive strategy of focusing on the most salient and relevant attributes while disregarding less influential ones.

Another notable observation is that in the majority of models, a single variable carries a weight of at least 50% of the overall evaluation. Additionally, the combined weighting of two variables accounts for more than 65% of the evaluation. This pattern can be interpreted as a heuristic where individuals tend to focus on a single, highly significant variable or a very small set of such variables when making their evaluations. It suggests a tendency to prioritize and simplify decision-making by emphasizing the most influential attributes.

With the results obtained from these models, it is indeed possible to calculate a confidence interval for the average weight of these attributes. This confidence interval provides a measure of uncertainty and can help provide a more comprehensive understanding of the variability in the weights assigned to the attributes.

5.2 Discussion for Serious Games

In terms of the applicability of this methodology to the field of serious games, it is important to highlight the ease of constructing such a model and conducting evaluations of game attributes. However, it is crucial to acknowledge that the results obtained from this particular game cannot be generalized to all games. Firstly, the game assessed in this study was specifically designed for academic purposes, and not all games share the same objectives. Additionally, it should be noted that while the scale used is quite comprehensive, certain attributes were not considered, such as graphics and visual design, sound and music, user interface, among others. Therefore, it is suggested to develop a more comprehensive scale to enable a more thorough evaluation of games.

One additional comment that should be addressed is the importance of establishing a clear definition of the attributes that can be improved in a game and understanding the consequences of these attributes. For instance, the narrative of the game and the level of challenge it presents to the players are attributes that can be enhanced through investment in story creation, challenge design, and the inclusion of diverse obstacles. On the other hand, overall engagement is a consequence of the game's attributes. This does not imply that measuring engagement is incorrect, but it is essential to consider it as an intermediate goal rather than an attribute itself.

By distinguishing between attributes and their associated consequences, game developers and evaluators can better identify areas for improvement and make informed decisions about enhancing specific aspects of the game.

6. Conclusions, Limitations and Future Works

6.1 Conclusions

This study shows that knowledge improvement (F8) consistently ranks as the most critical attribute in the evaluation of serious games within an academic context, indicating its significant impact on user satisfaction. Concentration (F1) follows closely, emphasizing the importance of player engagement and focus. These findings highlight the importance of a small set of attributes that influence the overall game rating, suggesting that game designers should prioritize these elements to enhance the user experience.

Additionally, the results reveal that participants frequently employ heuristics by assigning the minimum weight to attributes they perceive as less relevant. Attributes that consistently receive the minimum required weight (3%) indicate that users tend to overlook them in their evaluation. This observation suggests that players naturally simplify their decision-making process, focusing only on the most prominent factors when rating the game.

In different samples, one attribute (knowledge improvement) dominates the overall evaluation and contributes more than 50% to the total game score. This points to a potential over-reliance on a single factor, which could overshadow the influence of other important game attributes. To mitigate this issue, the methodology adjusted the constraints, reducing the dominance of any one attribute and ensuring a more balanced overall evaluation.

By calculating confidence intervals for the attribute weights, the study provides a more nuanced understanding of the variability in user evaluations. This measure adds robustness to the methodology, offering insights into how consistent or variable user preferences are across different models and samples.

6.2 Limitations and Future Works

One of the primary limitations of this methodology lies in its reliance on a predefined set of eight attributes. While this approach encourages a more thorough and structured evaluation, it constrains the participants' ability to omit attributes that may be less relevant to them personally.

Another key limitation is that this study was conducted using only one serious game, which restricts the generalizability of the findings. To strengthen the validity of the methodology, it is recommended that future studies apply it to multiple games with similar learning objectives to determine whether the attribute weights remain consistent across different contexts.

Additionally, future research should explore the application of this methodology in varied contexts and consider the influence of different participant profiles. For example, serious games with academic objectives might yield different attribute weights when evaluated by students, professors, or professionals. Alternatively, users may prioritize different attributes in games without academic goals. This would provide valuable insights for game creators and designers on how different user groups prioritize attributes.

Finally, future studies should investigate methods that allow participants to dynamically adjust the attributes they consider important. This would provide more accurate and personalized evaluations by allowing users to include or exclude attributes based on their preferences, better reflecting their decision-making processes.

Conflicts of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

[1] F. Bellotti, R. Berta, and A. D. Gloria, “Designing Effective Serious Games: Opportunities and Challenges for Research,” International Journal of Emerging Technologies in Learning (iJET), vol. 5, pp. 22–35, Nov. 2010, doi: 10.3991/ijet.v5iSI3.1500.

[2] E. Pacheco-Velazquez, M. S. Ramirez Montoya, and D. Salinas-Navarro, “Serious Games and Experiential Learning: Options for Engineering Education,” IJSG, vol. 10, no. 3, pp. 3–21, Sep. 2023, doi: 10.17083/ijsg.v10i3.593.

[3] M. Cheng, W. Huang, and M. Hsu, “Does emotion matter? An investigation into the relationship between emotions and science learning outcomes in a game-based learning environment,” Br. J. Educ. Technol., vol. 51, no. 6, pp. 2233–2251, Nov. 2020, doi: 10.1111/bjet.12896.

[4] H. Meij, S. Veldkamp, and H. Leemkuil, “Effects of scripting on dialogues, motivation and learning outcomes in serious games,” Br J Educ Technol, vol. 51, no. 2, pp. 459–472, Mar. 2020, doi: 10.1111/bjet.12851.

[5] E. Naul and M. Liu, “Why Story Matters: A Review of Narrative in Serious Games,” Journal of Educational Computing Research, vol. 58, no. 3, pp. 687–707, Jun. 2020, doi: 10.1177/0735633119859904.

[6] K. Madani, T. W. Pierce, and A. Mirchi, “Serious games on environmental management,” Sustainable Cities and Society, vol. 29, pp. 1–11, Feb. 2017, doi: 10.1016/j.scs.2016.11.007.

[7] B. P. Bergeron, “Learning & retention in adaptive serious games,” Stud Health Technol Inform, vol. 132, pp. 26–30, 2008.

[8] J. C. K. H. Riedel and J. B. Hauge, “State of the art of serious games for business and industry,” in 2011 17th International Conference on Concurrent Enterprising, Jun. 2011, pp. 1–8.

[9] E. Pacheco-Velazquez, “Effects of the use of simulators and an online platform in logistics education,” Int J Interact Des Manuf, vol. 16, no. 2, pp. 439–457, Jun. 2022, doi: 10.1007/s12008-021-00791-z.

[10] J. Hamari and J. Tuunanen, “Player Types: A Meta-synthesis,” ToDIGRA, vol. 1, no. 2, Mar. 2014, doi: 10.26503/todigra.v1i2.13.

[11] A. Dobrovsky, U. Borghoff, and M. Hofmann, “Improving Adaptive Gameplay in Serious Games Through Interactive Deep Reinforcement Learning,” 2019, pp. 411–432. doi: 10.1007/978-3-319-95996-2_19.

[12] E. A. Pacheco-Velazquez and S. M. Viscarra-Campos, “Exploring critical factors related to reflection, engagement and self-directed learning,” in 2019 IEEE Frontiers in Education Conference (FIE), Covington, KY, USA: IEEE, Oct. 2019, pp. 1–8. doi: 10.1109/FIE43999.2019.9028546.

[13] M. Xu et al., “Personalized training through Kinect-based games for physical education,” Journal of Visual Communication and Image Representation, vol. 62, pp. 394–401, Jul. 2019, doi: 10.1016/j.jvcir.2019.05.007.

[14] R. A. Sottilare and S. Gilbert, “Considerations for Adaptive Tutoring Within Serious Games: Authoring Cognitive Models and game Interfaces:,” Defense Technical Information Center, Fort Belvoir, VA, Jun. 2011. doi: 10.21236/ADA558687.

[15] R. S. N. Lindberg and T. H. Laine, “Formative evaluation of an adaptive game for engaging learners of programming concepts in K-12,” International Journal of Serious Games, vol. 5, no. 2, Art. no. 2, Jun. 2018, doi: 10.17083/ijsg.v5i2.220.

[16] A. Calderón and M. Ruiz, “A systematic literature review on serious games evaluation: An application to software project management,” Computers & Education, vol. 87, pp. 396–422, Sep. 2015, doi: 10.1016/j.compedu.2015.07.011.

[17] M. Soflano, T. M. Connolly, and T. Hainey, “An application of adaptive games-based learning based on learning style to teach SQL,” Computers & Education, vol. 86, pp. 192–211, Aug. 2015, doi: 10.1016/j.compedu.2015.03.015.

[18] E. A. Boyle et al., “An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games,” Computers & Education, vol. 94, pp. 178–192, Mar. 2016, doi: 10.1016/j.compedu.2015.11.003.

[19] N. Faizan, A. Löffler, R. Heininger, M. Utesch, and H. Krcmar, Classification of Evaluation Methods for the Effective Assessment of Simulation Games: Results from a Literature Review. International Association of Online Engineering, 2019, pp. 19–33. Accessed: May 06, 2023. [Online]. Available: https://www.learntechlib.org/p/207576/

[20] T. Marsh, “Serious games continuum: Between games for purpose and experiential environments for purpose,” Entertainment Computing, vol. 2, no. 2, pp. 61–68, Jan. 2011, doi: 10.1016/j.entcom.2010.12.004.

[21] J. L. Gonzalez, “Jugabilidad: Caracterización de la Experiencia del Jugador en Videojuegos.,” PhD Thesis, Universidad de Grandada, Granada, España, 2018.

[22] Y. Zhonggen, “A Meta-Analysis of Use of Serious Games in Education over a Decade,” International Journal of Computer Games Technology, vol. 2019, pp. 1–8, Feb. 2019, doi: 10.1155/2019/4797032.

[23] T. M. Connolly, E. A. Boyle, E. MacArthur, T. Hainey, and J. M. Boyle, “A systematic literature review of empirical evidence on computer games and serious games,” Computers & Education, vol. 59, no. 2, pp. 661–686, Sep. 2012, doi: 10.1016/j.compedu.2012.03.004.

[24] R. L. Keeney and H. Raiffa, Decisions with multiple objectives: preferences and value tradeoffs. Cambridge [England] ; New York, NY, USA: Cambridge University Press, 1993.

[25] H. A. Simon, “A Behavioral Model of Rational Choice,” The Quarterly Journal of Economics, vol. 69, no. 1, p. 99, Feb. 1955, doi: 10.2307/1884852.

[26] J. Von Neumann and O. Morgenstern, Theory of Games and Economic Behavior (60th Anniversary Commemorative Edition): Princeton University Press, 2007. doi: 10.1515/9781400829460.

[27] A. Tversky, “Elimination by aspects: A theory of choice.,” Psychological Review, vol. 79, no. 4, pp. 281–299, Jul. 1972, doi: 10.1037/h0032955.

[28] G. Gigerenzer, Gut feelings: the intelligence of the unconscious. New York: Viking, 2007.

[29] D. Kahneman and A. Tversky, “Prospect Theory: An Analysis of Decision Under Risk,” in World Scientific Handbook in Financial Economics Series, vol. 4, WORLD SCIENTIFIC, 2013, pp. 99–127. doi: 10.1142/9789814417358_0006.

[30] F. S. Hillier and G. J. Lieberman, Introduction to operations research, Eleventh edition, International student edition. New York, NY: McGraw-Hill Education, 2021.

[31] M. S. Bazaraa, J. J. Jarvis, and H. D. Sherali, Linear programming and network flows, 4. ed. Hoboken, NJ: Wiley, 2010.

[32] M. H. Bazerman and D. A. Moore, Judgment in managerial decision making, 8th ed. New York: Wiley, 2013.

[33] J. Rezaei, A. Arab, and M. Mehregan, “Equalizing bias in eliciting attribute weights in multiattribute decision-making: experimental research,” Behavioral Decision Making, vol. 35, no. 2, Apr. 2022, doi: 10.1002/bdm.2262.

[34] D. von Winterfeldt and W. Edwards, Decision analysis and behavioral research, Repr. Cambridge: Cambridge Univ. Press, 1993.

[35] J. W. Payne, J. R. Bettman, and E. J. Johnson, The Adaptive Decision Maker, 1st ed. Cambridge University Press, 1993. doi: 10.1017/CBO9781139173933.

[36] J. Rezaei, “Best-worst multi-criteria decision-making method,” Omega, vol. 53, pp. 49–57, Jun. 2015, doi: 10.1016/j.omega.2014.11.009.

[37] T. Pajala, P. Korhonen, and J. Wallenius, “Judgments of importance revisited: What do they mean?,” Journal of the Operational Research Society, vol. 70, no. 7, pp. 1140–1148, Jul. 2019, doi: 10.1080/01605682.2018.1489346.

[38] A. Tversky and D. Kahneman, “Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty.,” Science, vol. 185, no. 4157, pp. 1124–1131, Sep. 1974, doi: 10.1126/science.185.4157.1124.

[39] M. Ahmad, S. Z. A. Shah, and Y. Abbass, “The role of heuristic-driven biases in entrepreneurial strategic decision-making: evidence from an emerging economy,” MD, vol. 59, no. 3, pp. 669–691, May 2021, doi: 10.1108/MD-09-2019-1231.

[40] T. M. Hughes, L. A. Dossett, S. T. Hawley, and D. A. Telem, “Recognizing Heuristics and Bias in Clinical Decision-making,” Annals of Surgery, vol. 271, no. 5, pp. 813–814, May 2020, doi: 10.1097/SLA.0000000000003699.

[41] N. Harvey, “Use of heuristics: Insights from forecasting research,” in Thinking & Reasoning, 1st ed., D. Lagnado and D. Read, Eds., Psychology Press, 2020, pp. 5–24. doi: 10.4324/9781003059745-2.

[42] J. N. Marewski and G. Gigerenzer, “Heuristic decision making in medicine,” Dialogues in Clinical Neuroscience, vol. 14, no. 1, pp. 77–89, Mar. 2012, doi: 10.31887/DCNS.2012.14.1/jmarewski.

[43] C. Krist, C. V. Schwarz, and B. J. Reiser, “Identifying Essential Epistemic Heuristics for Guiding Mechanistic Reasoning in Science Learning,” Journal of the Learning Sciences, vol. 28, no. 2, pp. 160–205, Mar. 2019, doi: 10.1080/10508406.2018.1510404.

[44] S. Bhatia and N. Stewart, “Naturalistic multiattribute choice,” Cognition, vol. 179, pp. 71–88, Oct. 2018, doi: 10.1016/j.cognition.2018.05.025.

[45] P. Slovic, “Choice between equally valued alternatives.,” Journal of Experimental Psychology: Human Perception and Performance, vol. 1, no. 3, pp. 280–287, Aug. 1975, doi: 10.1037/0096-1523.1.3.280.

[46] R. P. Hämäläinen and S. Alaja, “The threat of weighting biases in environmental decision analysis,” Ecological Economics, vol. 68, no. 1–2, pp. 556–569, Dec. 2008, doi: 10.1016/j.ecolecon.2008.05.025.

[47] S. K. Jacobi and B. F. Hobbs, “Quantifying and Mitigating the Splitting Bias and Other Value Tree-Induced Weighting Biases,” Decision Analysis, vol. 4, no. 4, pp. 194–210, Dec. 2007, doi: 10.1287/deca.1070.0100.

[48] G. Montibeller, “Behavioral Challenges in Policy Analysis with Conflicting Objectives,” in Recent Advances in Optimization and Modeling of Contemporary Problems, E. Gel, L. Ntaimo, D. Shier, and H. J. Greenberg, Eds., INFORMS, 2018, pp. 85–108. doi: 10.1287/educ.2018.0182.

[49] G. Montibeller and D. Von Winterfeldt, “Biases and Debiasing in Multi-criteria Decision Analysis,” in 2015 48th Hawaii International Conference on System Sciences, Kauai, HI: IEEE, Jan. 2015, pp. 1218–1226. doi: 10.1109/HICSS.2015.148.

[50] J. Rezaei, “Anchoring bias in eliciting attribute weights and values in multi-attribute decision-making,” Journal of Decision Systems, vol. 30, no. 1, pp. 72–96, Jan. 2021, doi: 10.1080/12460125.2020.1840705.

[51] F. Almeida and J. Simoes, “The Role of Serious Games, Gamification and Industry 4.0 Tools in the Education 4.0 Paradigm,” CONT ED TECHNOLOGY, vol. 10, no. 2, pp. 120–136, Apr. 2019, doi: 10.30935/cet.554469.

[52] R. L. Lamb, L. Annetta, J. Firestone, and E. Etopio, “A meta-analysis with examination of moderators of student cognition, affect, and learning outcomes while using serious educational games, serious games, and simulations,” Computers in Human Behavior, vol. 80, pp. 158–167, Mar. 2018, doi: 10.1016/j.chb.2017.10.040.

[53] P. Wouters, C. Van Nimwegen, H. Van Oostendorp, and E. D. Van Der Spek, “A meta-analysis of the cognitive and motivational effects of serious games.,” Journal of Educational Psychology, vol. 105, no. 2, pp. 249–265, May 2013, doi: 10.1037/a0031311.

[54] N. Imlig-Iten and D. Petko, “Comparing Serious Games and Educational Simulations: Effects on Enjoyment, Deep Thinking, Interest and Cognitive Learning Gains,” Simulation & Gaming, vol. 49, no. 4, pp. 401–422, Aug. 2018, doi: 10.1177/1046878118779088.

[55] J. P. Gee, “Learning by Design: Good Video Games as Learning Machines,” E-Learning and Digital Media, vol. 2, no. 1, pp. 5–16, Mar. 2005, doi: 10.2304/elea.2005.2.1.5.

[56] W. El Machtani El Idrissi, G. Chemsi, K. El Kababi, and M. Radid, “The Impact of Serious Game on the Nursing Students’ Learning, Behavioral Engagement, and Motivation,” Int. J. Emerg. Technol. Learn., vol. 17, no. 01, pp. 18–35, Jan. 2022, doi: 10.3991/ijet.v17i01.26857.

[57] S. Rosenthal and R. A. Ratan, “Balancing learning and enjoyment in serious games: Kerbal Space Program and the communication mediation model,” Computers & Education, vol. 182, p. 104480, Jun. 2022, doi: 10.1016/j.compedu.2022.104480.

[58] S. Zhu et al., “A Serious Game for Enhancing Rescue Reasoning Skills in Tactical Combat Casualty Care: Development and Deployment Study,” JMIR Form Res, vol. 8, p. e50817, Aug. 2024, doi: 10.2196/50817.

[59] T. J. Olgers, A. A. Bij De Weg, and J. C. Ter Maaten, “Serious Games for Improving Technical Skills in Medicine: Scoping Review,” JMIR Serious Games, vol. 9, no. 1, p. e24093, Jan. 2021, doi: 10.2196/24093.

[60] S. De Freitas and M. Oliver, “How can exploratory learning with games and simulations within the curriculum be most effectively evaluated?,” Computers & Education, vol. 46, no. 3, pp. 249–264, Apr. 2006, doi: 10.1016/j.compedu.2005.11.007.

[61] J. H. Brockmyer, C. M. Fox, K. A. Curtiss, E. McBroom, K. M. Burkhart, and J. N. Pidruzny, “The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing,” Journal of Experimental Social Psychology, vol. 45, no. 4, pp. 624–634, Jul. 2009, doi: 10.1016/j.jesp.2009.02.016.

[62] F.-L. Fu, R.-C. Su, and S.-C. Yu, “EGameFlow: A scale to measure learners’ enjoyment of e-learning games,” Computers & Education, vol. 52, no. 1, pp. 101–112, Jan. 2009, doi: 10.1016/j.compedu.2008.07.004.
[63] L. García-Mundo, M. Genero, and M. Piattini, “Applying a Serious Game Quality Model,” in Serious Games, Interaction, and Simulation, vol. 161, C. Vaz De Carvalho, P. Escudeiro, and A. Coelho, Eds., in Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 161. , Cham: Springer International Publishing, 2016, pp. 12–20. doi: 10.1007/978-3-319-29060-7_4.
[64] K. Mitgutsch and N. Alvarado, “Purposeful by design?: a serious game design assessment framework,” in Proceedings of the International Conference on the Foundations of Digital Games, Raleigh North Carolina: ACM, May 2012, pp. 121–128. doi: 10.1145/2282338.2282364.