Journal on Interactive Systems, 2024, 15:1, doi: 10.5753/jis.2024.4136

Games With Safety Training Purposes in the Industry: Game Design Method and Its Demonstration

Roberto Rufino Júnior [**Federal University of the State of Rio de Janeiro (UNIRIO)** | ***roberto.rufino@edu.unirio.br***]
Tadeu Moreira de Classe [**Federal University of the State of Rio de Janeiro (UNIRIO)** | ***tadeu.classe@uniriotec.br***]
Camila de Castro Lima [**Federal Institute of Education, Science and Technology of Southeast Minas Gerais (IF Sudeste MG) – Juiz de Fora Campus** | ***camilac.lima@hotmail.com***]

Research Group on Games to Complex Contexts (JOCCOM) – Graduate Program in Informatics (PPGI) – Federal University of the State of Rio de Janeiro (UNIRIO), Av. Pasteur, 458, Urca, Rio de Janeiro, RJ, 22290-255, Brazil.

Received: 16 February 2024 • **Accepted:** 15 May 2024 • **Published:** 30 May 2024

Abstract Work accidents are a concern for the industry as they can generate human, material, and economic losses. One way to mitigate them is to conduct efficient safety training, but they cannot always arouse the necessary interest and engagement of those trained. Games with a purpose are tools widely used for training in different contexts. They are identified as innovative, immersive, and attractive approaches, but they are complex to develop. This article presents SafetyPlay Game Design (SpGD), a structured method for designing safety training games with well-defined steps that support translating training elements into an immersive and playful game environment. Since we seek to build an artifact as an intervention to a problem in a real context, we based the research methodology on Action Research to conduct the investigations. Therefore, the intervention in the context of security training continued with creating and demonstrating the SpGD method based on the construction of a digital game for training. In this way, the evaluation undertaken in this research considered two moments: i) the ability of SpGD to develop the game, considering when we evaluated it by volunteers regarding game experience and perception of safety training, and ii) the validated method through interviews with game designers, experts in designing training games. The results point to evidence that the SpGD method supported the design of a game that allowed a positive gaming experience considered by game designers to be viable and useful for this purpose, in addition to allowing people to learn training concepts and risks. Therefore, this research contributes to the game design field and the industry, providing opportunities for creating interactive training that can complement traditional security training.

Keywords: Safety Training, Industry, Games With a Purpose, Game Design Method, SafetyPlay Game Design, SpGD.

1 Introduction

Year after year, work accidents claim the lives of thousands of people (TRT04, 2022). According to the United Nations (ONU, 2022), between 2012 and 2021, more than 20 thousand workers lost their lives as a result of an accident involving their work activity (TRT12, 2023).

An efficient way to provide workers with the skills and knowledge necessary to develop work activities safely is by implementing safety training (Volpe and Lorusso, 2009; Venturi *et al*., 2021). In this context, safety training incapable of transmitting insight into the dangers present in the work environment to workers may incorporate risks to a given organization, such as serious injuries, permanent disabilities, and even loss of life, as well as possible damage to property and/or the environment (Correa and Cardoso Junior, 2007).

However, safety training professionals (such as: managers and trainers) report low participant interest and engagement. They attribute this problem to factors such as the passivity of the training (e.g. slides, videos, handouts), which fail to arouse a sense of belonging in the people being trained (Rufino Júnior *et al*., 2023). As a result, there is a demand for more interesting and interactive safety training sessions (Gallerati *et al*., 2017).

In this context, purposeful games emerge as an interesting approach to meet this demand (Rufino Júnior *et al*., 2022). They can present a dynamic training environment using a simulated risk scenario without the risk being present (Mayer *et al*., 2013), providing learning and engagement of those trained (Bruzzone *et al*., 2013; Chittaro and Ranon, 2009).

Nevertheless, even though purposeful games can benefit safety training sessions, designing them is not a simple task (Forbes, 2022; Wolf *et al*., 2022). Even though industry managers are interested in such approaches, there is concern about the predictability of investing time and money to design them (Rufino Júnior *et al*., 2022). Therefore, creating a methodology capable of making the game development process for security training more systematic and simple, with less risk of mistakes in understanding and transmitting training information, is necessary.

With this in mind, this paper aims to present the proposal for the **SafetyPlay Game Design (SpGD) method and its demonstration**, explicitly conceived for designing digital games for safety training in the industry. SpGD focuses on adding agility and traceability to developing games for safety training. We based it on consolidated concepts and techniques relating to training evaluation and the development of digital games to align the essential

characteristics of safety training with game elements.

Since the artifacts of this research (SpGD method and training games) seek to solve a problem in an organizational context, changing the work environment, the research methodology chosen was Action Research (AR) (Davison *et al*., 2004). Thus, given the diagnosis of the need for safety training games, the SpGD method is designed to intervene in the organizational context, with both the training games designed using the technique and the SpGD itself being evaluated and its results leading to reflection on changes in day-to-day organizational life.

A digital game was developed for safety training using the steps provided in SpGD to demonstrate the feasibility of using the method. We evaluated the game in a quasi-experimental study to verify whether its design, with support from SpGD, provided players with a gaming experience and perception of safety training. Finally, we evaluated the SpGD through interviews with game designers to elucidate its strongness, possible improvements, and feasibility. The interviews followed the Underlying Discourse Explanation Method (Nicolaci-da Costa, 2007), and we analyzed the responses based on qualitative discourse analysis techniques grounded on the work of Strauss and Corbin (1990).

We organized the article as follows: Section 2 presents the backgrounds and related works. Section 3 introduces Action Research as the research methodology used in this article. Section 4 shows the SpGD method, and next, Section 5 demonstrates it from the design of a digital game for training purposes. Section 6 presents the game evaluation and the validation of SpGD with game design specialists. We briefly reflected on our results in Section 7. Finally, in Section 8, we presented our final remarks.

2 Backgrounds

This section introduces the conceptual basis of our research, aiming to better understand the context, concepts, and approaches we used to design the research. We also present some related works that are directly associated with our study.

2.1 Safety Training in the Industry

Industrial activities refer to the various operations and processes performed in the industrial sector to produce, transform, and manufacture goods. These activities encompass a wide range of sectors, such as manufacturing, energy production, construction, mining, chemistry, food, and beverages, among others, which present a series of risks and challenges to the professionals involved, such as exposure to dangerous substances, use of complex equipment and operations with high physical demand (Rodrigues and Santana, 2010).

In this context, it is essential to implement effective training programs that enable employees to deal with risky situations and adopt appropriate preventive measures (Martins, 2021). One of the main objectives of safety training in the industry is to provide workers with the necessary knowledge to identify and assess the risks in the work environment and the skills to implement best safety practices (Venturi *et al*., 2021).

With the advancement of technology, new training approaches have emerged in the industry. One of these innovative approaches is using digital games as a training tool (Vigoroso *et al*., 2021). Digital games offer an immersive and interactive experience, allowing realistic simulations of risky situations and the practice of specific skills. This innovative approach has effectively engaged participants, facilitating learning and retention of security-related knowledge (Bruzzone *et al*., 2013; Lu *et al*., 2022).

2.2 Training Assessment

Training assessment is an essential step in measuring the effectiveness and impact of training activities on the development of skills and performance of individuals (Lacerda and Abbad, 2003).

Several methods were proposed to evaluate training, each with specific approaches and focus. One of the widely recognized methods is that proposed by Baldwin and Ford (1988), which emphasizes results-based evaluation, seeking to identify and measure observable changes in the behavior and performance of individuals after training. Another relevant method is the one developed by Kraiger *et al*. (1993), which focuses on evaluating reactions, learning, transfer, and training results.

However, the Kirkpatrick assessment model is one of the most well-known and widely used methods (Kirkpatrick and Kirkpatrick, 2016). This model establishes four levels of training evaluation (Figure 1): reaction, learning, behavior, and results.

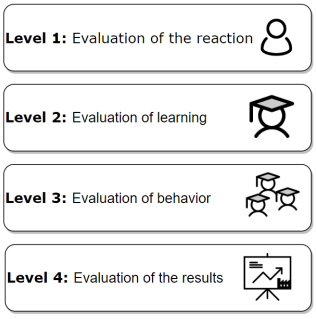

Figure 1. The Kirkpatrick model.

The reaction level evaluates participants' satisfaction and perception of the training. The learning level verifies the degree of acquisition of knowledge and skills during training. The behavior level assesses the application of learning in the work environment. Finally, the results level evaluates the impact of training on organizational results.

Although each method has its advantages and specific focus, the Kirkpatrick model is one of the most robust for training evaluation. It covers different assessment levels and allows a complete analysis of the effects of training in multiple dimensions (Agarwal *et al*., 2019).

2.3 MDA Framework

There are many different frameworks and ways to design digital games (Winn, 2009; Rogers, 2010; Schell, 2019; Classe *et al*., 2019). Still, for this research, we choose to adopt the MDA framework (Mechanics, Dynamics, and Aesthetics) because it provides a simple way to analyze games, divided into clear elements.

The MDA framework (Mechanics, Dynamics, and Aesthetics) is a conceptual model used in game design to describe and analyze the fundamental elements that make up a game (Hunicke *et al*., 2004). MDA divides the game elements into three layers: Mechanics, Dynamics, and Aesthetics.

The Mechanics layer refers to the rules, systems, and components that structure the game. It covers the most tangible elements, such as movement, environmental interaction, combat, puzzles, and other game mechanics. This layer describes the game's basic features and how players interact with them.

The Dynamic layer deals with interactions between players and the game system. It describes how mechanics and systems interact and combine to create specific game experiences. Dynamics encompasses how players' choices and actions affect the state of the game, generating challenges, rewards, and consequences. In this layer, the players' complex behaviors and strategies emerge.

Finally, the Aesthetic layer refers to the player's emotional and subjective experience when playing the game. It involves elements such as narrative, atmosphere, visual and sound aesthetics, themes, and sensations transmitted by the game. This layer creates an emotional connection between the player and the game, providing pleasure, immersion, and engagement.

The MDA framework enables a structured approach to game design, helping developers understand and balance the elements that compose the gaming experience. By using MDA, designers can describe and analyze a game's mechanical, dynamic, and aesthetic aspects in a separate and integrated way, which can lead to a deeper understanding of the desired gaming experience and assist in making more informed design decisions (Hunicke *et al*., 2004).

2.4 Related Works

In previous studies, we performed some literature reviews to understand how the industry has used games for training. Our investigations led us to perceived use contexts, technologies, and concepts that allowed us to propose our method. For instance, Rufino Júnior *et al*. (2022) investigated trainers' perception of safety training in the industry. The research results show problems such as low levels of engagement in training that used traditional methods (such as videos and presentations) and point to evidence that participatory approaches and scenario-based training can increase the effectiveness of security training. Additionally, the work conducted a literature study identifying contexts and techniques on how the industry has been using games for training situations.

In another research, Rufino Júnior *et al*. (2023) looked into the game and information systems literature to explore information about the game design process, elements, and coasts, considering the context of safety training. In that work, the authors identified types of "games for training" based on augmented reality, virtual reality, and 3D environments.

At last, Serra and Classe (2023) performed a systematic mapping of the literature to explore how the industry has been using gameplay data originating from games for training. They identified some uses of those data, such as game analytics, data visualization, and game balance. All those applications were responsible for giving training information to the organizational manager.

The search for methods that can support the creation of games for organizational training found research byClasse *et al*. (2019), which proposes the **Play Your Process (PYP) method** for designing games based on business processes and the research by Lopes *et al*. (2022), which presents **PYP4Training**, intended to guide the design of games for business process training.

Based on related work, it is possible to observe that although some studies investigate the use of games as training tools and there are some methods to design them, it is not possible to identify specific techniques aimed at safety training or risk mitigation in the industry. Therefore, evidence points to an opportunity for research into

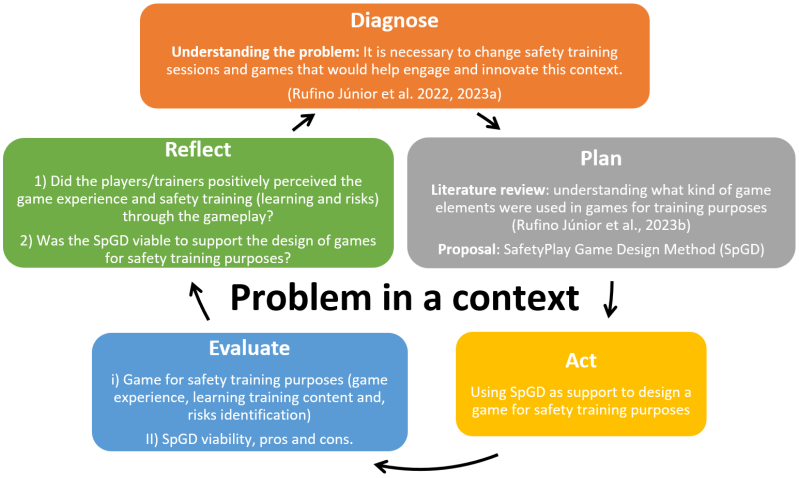

Figure 2. Instantiated research based on Action Research steps.

developing specific methods that meet the demands of safety training.

3 Action Research

Action Research is a flexible method suitable for scenarios where you want to obtain evidence about using an artifact applied in practice under a real problem (Santos and Travassos, 2011). Such an artifact is then planned, developed, and used as an intervention to the problem. By assessing data, the results lead to reflection on the issue, potentially giving rise to a new investigative cycle (Figure 3) (Davison *et al*., 2004).

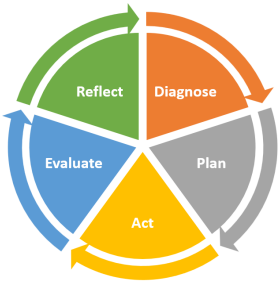

Figure 3. Action Research steps (Adapted from Davison *et al*. (2004)).

Figure 2 depicts the stages of this research concerning AR. Thus, the main challenge is making safety training sections more motivating and engaging, with the proposal of using digital games for this purpose being a problem **diagnosed** in Rufino Júnior *et al*. (2022, 2023) with industry professionals. At the same time, the research problem arises: **how to support the design of safety training games in the industry**, mitigating the erroneous transmission of training and ensuring that players/trained playfully and engagingly can learn from the training and reflect on the risks linked to work performance.

As an investigation **plan**, we conducted a literature study to identify how organizations have used purposeful games in training contexts and what elements are necessary for this, correlating them with training evaluation proposals (Rufino Júnior *et al*., 2023). Still, during planning we proposed the SpGD method to systematize these games' designs.

As **intervention/action**, we used the SpGD method to construct a digital safety training game to solve the research problem. In this way, we carried out the **evaluation** in two moments: i) the game, which was used to demonstrate the feasibility of the SpGD method, in addition to the analysis of experience, learning, and identification of security risks by the players/trained, and ii) the method itself, submitted to evaluation by game designers through the exemplification of its application in the creation of the game, collecting perceptions of the feasibility of use, strongness and weaknesses on the part of game designers.

Finally, the assessment results led to **reflection** on the research artifact about solving the problem. Our research presented two reflection questions in line with the evaluation moments: 1) Did the players/trained positively perceive the game experience and safety training (learning and risks) through the gameplay? 2) Was the SpGD viable to support the game's design for safety training purposes?

According to Carr (2007), AR is a process that involves action-reflection cycles in which researchers and research participants collaborate to identify a problem, plan and implement actions to resolve it, evaluate the results, and reflect on learned lessons. These repeated cycles allow for the adaptation and continuous improvement of approaches and interventions over time, promoting the progressive enhancement of practices and the production of knowledge relevant to solving real problems.

4 Plan: SafetyPlay Game Design Method

This research proposes the Safetyplay Game Design method to improve safety training in the industry by making it more playful, engaging, interactive, and motivating. We designed the SpGD method to guide game development teams for security training games. To do so, its stages encompass assimilating the fundamental concepts of training, translating them into game elements, going through validation and assessment cycles, and finalizing the game's production and implementation as a training alternative in industrial environments.

The SpGD method has well-defined and structured stages (Figure 4), which aim to guarantee the integration of safety training characteristics with game design elements, namely: 1) Understanding the training, 2) Mapping the training to game elements, 3) Performing brainstorming, 4) Developing the game, 5) Validating the game, 6) Assessing the game and 7) Delivering the game (Rufino Júnior and Classe, 2024). Although it is not well-represented in Figure 4, it is essential to mention that all process actors (game design team, managers, and trained people) are involved in all steps of SpGD. The Figure represents the main actor for each task.

The first stage (**Understanding the training**), conducted by the game design team, process managers, and trainers, consists of understanding the training. Its objective is to analyze the details of the training as it is currently performed so that the gameplay reflects it. To this end, Kirkpatrick's assessment model is used to identify information in the training to fill in the questions related to its four levels. With this, we expected a general and organized view of the training.

The second stage (**Mapping the training to game elements**), also conducted by game designers, process managers, and trained personnel, corresponds to the translation of training information into game elements. We did this through a semantic alignment between Kirkpatrick's four training assessment levels and the MDA framework elements (Figure 6). This step aims to mitigate problems in understanding the aspects of safety training, which could lead to players' misunderstanding during gameplay.

It is important to highlight that it is possible to associate any elements of Kirkpatrick's model with those of MDA (red dashed lines). However, when semantically analyzing

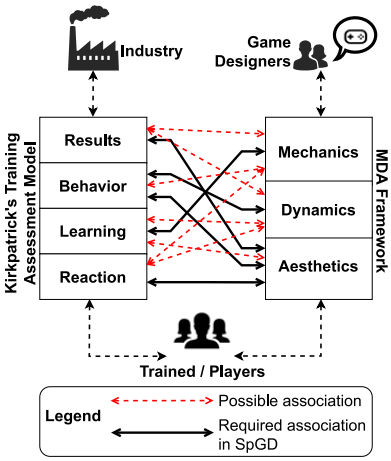

Figure 6. Mapping the training to game elements.

the meaning of the elements (considering Bloom's taxonomy (Ferraz and Belhot, 2010), procedural rhetorics (Bogost, 2008) and other theories), some relationships are more strongly indicated than others, for example, Mechanics are suggested based on information coming from Kirkpatrick's second level (learning), which contains the main actions, rules and elements that those trained must learn.

Furthermore, learning is strongly related to mechanics when thinking about procedural rhetorics. Dynamics benefit from information collected at Kirkpatrick's third level (behavior). At this level, we analyzed situations where the trained can use acquired knowledge. Finally, the game's Aesthetics are fed by references from Kirkpatrick's levels 1 (reaction), 3 (behavior), and 4 (results), benefiting from the feelings intended for those trained, the aspects of the workplace in which they can apply the knowledge and business indices that, transported into a game, could be affected by the actions of the trained.

In the third stage, **Performing a brainstorming**, game designers and process managers are trained to share ideas and thoughts about how the game should be designed based on the mapping from the previous stage. The objective of this stage is to create a documentary view of the game project (Game Design Document - GDD), which will assist developers and artists in the production of the game, with a detailed description of the game's levels, history, mechanics, rules, sounds and interfaces.

In the fourth stage, (**Developing the game**), the game is developed based on the GDD established in the previous stage. After the game's development, the fifth stage (**Validating the game**) occurs, where the game is validated by process managers and trained using information from the Kirkpatrick assessment model defined in stage 1. Such validation is important because it verifies whether the game is aligned with training objectives and conveys correct information to players. If it is not validated, we can adopt

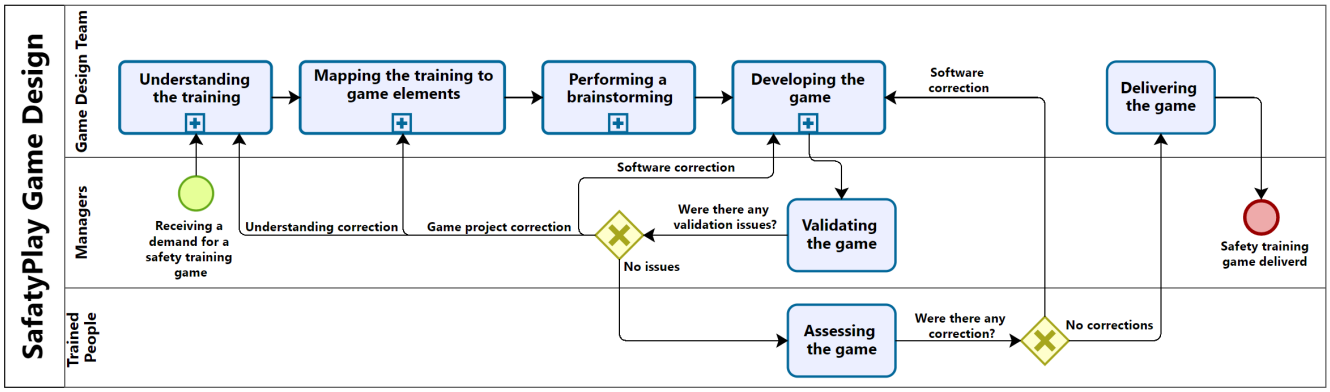

Figure 4. SafetyPlay Game Design (SpGD) method.

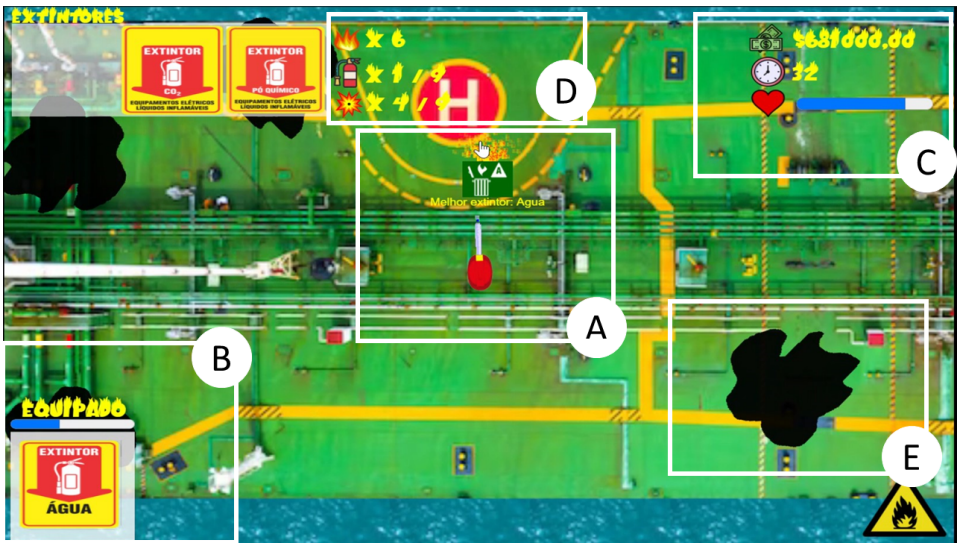

Figure 5. Game "Bob Ruff in Deck is on Fire" overview: safety training elements mapped to game elements.

three decisions: 1) the training context was poorly understood, making it necessary to return to the understanding training stage; ii) the mapping between training and game elements has gaps, so it is required to go back to the second stage of SpGD; or iii) there is an error in the game software so, you must return to the developing game stage. If approved, the game moves on to the sixth stage (**Assessing the game**) to evaluate the game experience, immersion, and engagement in the training process, besides using the training evaluation with Kirkpatrick's four levels to measure whether training was able to add the desired knowledge and experiences and ensure that those trained were aware of the risks of the presented situation.

Finally, the seventh stage is **Delivering the game** for use in safety training in the industry. At this stage, the game is made available to those who are trained and will benefit from the immersive and interactive learning experiences offered by the game.

5 Act: Game for Fire Training

We demonstrated The SpGD method to exemplify its application in creating a safety training game in the industry. In this demonstration, the "*Bob Ruff in Deck is on Fire*"[1] game was developed for fictional firefighting training, teaching about the different types of extinguishers used in these cases (Figure 5).

Firstly, those involved in designing the game (game designers and managers) realized the stage of understanding the proposed training, considering the levels in Kirkpatrick's assessment model. At each level of the model, we expected to answer some pre-determined questions. As this is a relatively simple training, Table 1[2] presents each level of Kirkpatrick, highlighting the questions and the responses of the managers of this training so that it is considered adequate in the organization.

At SpGD, that spreadsheet is developed jointly between managers and game designers so that the team can understand the context of what they will transform into a game from the beginning of the process. This approach aims to ensure a certain coherence between what organizations expect from training and which elements should be present during gameplay[3].

In the second stage, Mapping the training to game elements, the responses obtained and organized based on Kirkpatrick's levels were aligned with the MDA framework elements (Figure 7). The level of reaction, where we expect the trained to be motivated and react promptly to a possible fire situation, is connected to the game's aesthetics through fire outbreaks (Figures 5A and 7A) which, if not extinguished in time, evolve into explosions (Figures 5E and 7E). Learning the types of fire and correct extinguishers and possible consequences are linked to the mechanics of obtaining information about them (Figures 5A and 7A) and choosing the appropriate extinguisher (Figures 5B and 7B) for success in the game.

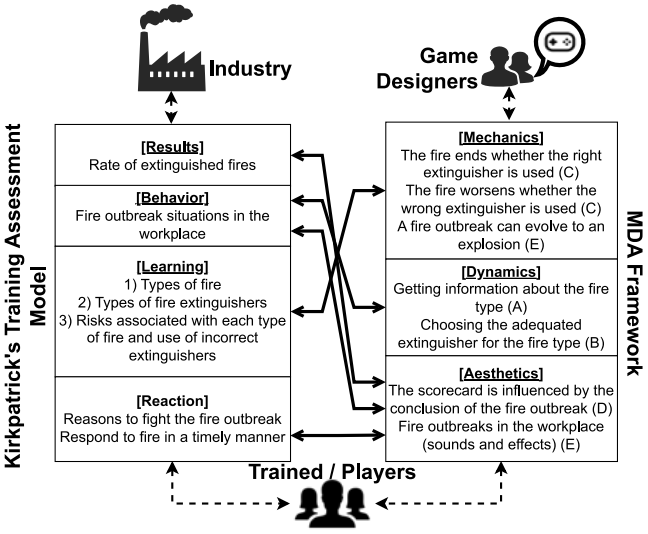

Figure 7. Demonstration of mapping the training to game elements.

Situations where the worker can demonstrate a behavior change, that is, fire situations in the workplace encourage dynamics (identifying information on the type of fire or choosing the best extinguisher for that situation) (Figures 5 C and 7C) and aesthetics (fire spots in places inspired by the work environment) (Figures 7A and E). Finally, the result, a company metric (fire completion rate), is transported to the game's scorecard (Figures 5C-D and 7C-D).

In the third stage, Performing brainstorming, the game designers and managers fed a game design document with ideas inspired by the alignment information between the training elements and the MDA, organizing the ideas around the level design (the game occurs on the deck of an oil tanker), storytelling (the player takes on the role of a member of the company's fire brigade) and the already established mechanics, dynamics and aesthetics.

With this document, a technical team (e.g. artists, programmers) developed the game. It is possible to see that the planned mechanics (obtaining information about fire and suitable extinguisher) in Figures 5A e 5B. Figure 5C represents the dynamics intended for correct or incorrect firefighting. The aesthetics (scorecard, fire source, and explosion) can be seen in Figures 5A, 3D e 3E. Therefore, we affirm the existence of an alignment between the main elements of safety training and the game's elements.

When identifying a fire source, the player can move close to it, and when hovering the mouse over the fire animation, information about which extinguisher is suitable is displayed (Figure 5B). Having identified the type of fire, the player must select the correct extinguisher, position it on the "Equip" field, and click on the flame. If the extinguisher is suitable, it will extinguish the fire. If not, the flame increases and could lead to an explosion. The player also has information such as the monetary value of the damage caused by the fires, time, and health bar (Figure 5C).

[1]Game: `https://joccom.uniriotec.br/games/deckfire` | INPI nº: BR512023001633-0

[2]Letters in Figure 5, Table 1, and Figure 7 were prepositionally allocated to reference the game design traceability. It is possible to trace the answers we got in SpGD step 1, which we used to map game elements in step 2 and develop the game in step 3.

[3]Meetings document: `https://bit.ly/BobRuff_DesignMeet22-23`

Table 1. Data from Understanding the training step.

Kirpatrick's dimensions	Questions	Answers
4 – Results	What are the tangible and measured results of the safety training?	*"It is possible to measure this safety training from the rate of identified fire outbreaks and extinguished fire outbreaks" (D).*
	Does the organization use any indexes or metrics to measure this safety training?	*"Yes! Rate of extinguished fire outbreaks." (D)*
3 – Behavior	Which situations (real or hypothetical) could the trained person use the knowledge learned in this training?	*"Trained people could use the knowledge of this safety training in practical situations or simulations of fire fighting". (A)*
		"In safety training sessions for this context, we usually use random and different fire outbreak situations, such as a fire in diverse materials. We expect that the trained person could extinguish the fire outbreak correctly." (B - E)
	What is the expected behavior from the trained person in this safety training?	*"We expected the trained person to be calm even in a stressful situation." (C)*
		"Safety training participants must reflect on this complex situation and its risks." (C)
		"They must choose the better option (extinguisher) to extinguish the fire outbreak." (B)
2 – Learning	What is the main safety information that they must learn?	*"Type of fire extinguishers and what kind of fire they must be used." (A - B)*
	How are safety training participants assessed in traditional training?	*"They do a text exam."*
		"We observe if the participant can fight against a fire outbreak in a practical simulation." (A)
1 – Reaction	How would you expect the safety training participants to react to this training?	*"We would like them to pay attention to the type of fire and what kind of extinguished they should use." (A-B)*
		"We expect them to immerse in the training and react to the fire outbreak in time." (C)
		"We want them motivated to fight against the fire, considering all the risks involved in this context, and learn with this training." (C-E)

Thus, this aimed to demonstrate the viability of the SpGD method, considering the coherence of the steps that lead to the creation of the training game conducted by game designers, managers, and trained. We conducted the validation steps with the managers who supported this game's initial stages. If they considered the result valid, we evaluated the game with the target audience: industry professionals who require fire training. Therefore, we present the evaluation with training participants in the evaluation section of this research article.

6 Evaluate

We defined two assessment steps to guide this evaluation: a quasi-experimental evaluation using the game for fire training and; semi-structured interviews with the support of the game designer. At least, we planned a reflection/communication step providing a synthesized discussion of the results of two evaluations. A summary of this process can be seen in Figure 8.

6.1 Game for Training Purpose Evaluation

At this stage of the research, we based the study design on an empirical evaluation along the lines of a quasi-experimental study (Campbell and Stanley, 2015), following the steps: 1) definition; 2) planning; 3) execution; 4) analysis and; 5) results and conclusions. According to these authors, quasi-experiments are less controlled empirical studies than traditional experiments, with no need for random selection of participants. Hence, we selected this research method because the survey's audience comprises volunteer workers from an industry.

6.1.1 Definition

The **study definition** presents the research objectives within a specific context. Then, we decided to follow the GQM (Goal-Question-Metric) (Basili, 1992) to make this work clearer and more organized. With this, we describe the objective (goal) as **Analyzing** the BobRuff in Deck is on Fire digital game developed with SpGD; **for the purpose of** evaluation; **concerning the** perception of 1) game experience and 2) safety training (learning and risks); **from the point of view of** industrial workers; **in the context** of fire safety training.

The assessment organized by GQM seeks to answer questions which, in this research, can be stated as: (Q1) Did the participants have a positive perception of the game experience? (Q2) Did the participants positively perceive the safety training in learning and identifying risks?

We based the metrics used to answer the questions on the dimensions of game experience and MEEGA+ security training (Petri *et al.*, 2019), the latter dimension covering learning variables and identification of training risks. This questionnaire consists of statements in which participants

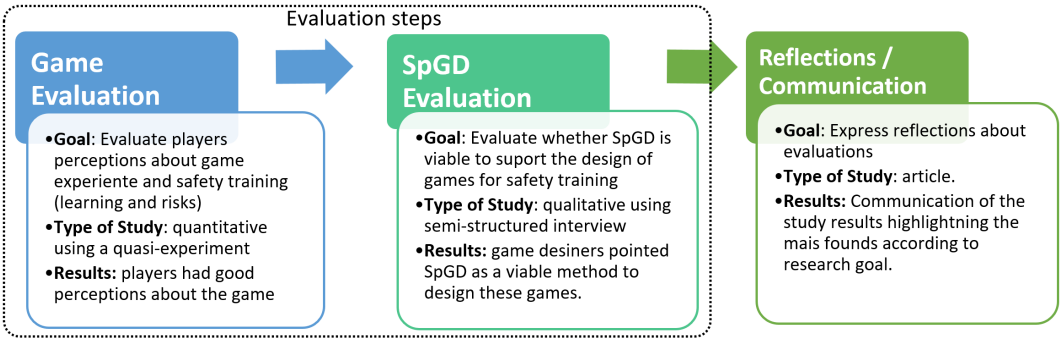

Figure 8. Evaluation steps.

indicate their degree of agreement with the statement using a Likert scale. This ordinal/interval scale allows descriptive, correlational, and inferential statistical analyses (Coelho *et al*., 2020).

6.1.2 Planning

When planning scientific evaluations, it is necessary to detail the context of the study so that it can be understood, its limitations mitigated, and academic peers can replicate it. It is essential to explain who the participants were, the study steps, what will be the instrumentation, and how the data were collected, processed, and analyzed.

The **study participants** were volunteer workers from a Brazilian industry who had already undergone fire training within their respective work environments at some point in their professional lives. We conducted the study by collecting the opinions of these professionals and accordingly in the form of an opinion survey. No sensitive information identifying the worker or company was requested, collected, analyzed, or revealed. We presented all information about the conduct and objective of the study to participants in a free and informed consent form (FICF), and their participation was recorded in the participation questionnaire.

It is worth noting that, according to CNS (Brazilian National Health Council) resolution 510/2016[4], in its Art.1, studies that are characterized as opinion research, where there is no identification of participants, do not need to be submitted and evaluated by a research ethics committee. Furthermore, all data collected follows the guidelines of Brazilian data protection law[5].

Participants performed the **research tasks** at three different times. Initially, participants had access to the electronic form with the FICF and questions related to their profile (we defined a time limit of 10 minutes for this). They were then given access to the game and asked to play it for at least 15 minutes, considering it a safety training section. Finally, participants answered their perceptions based on statements from the adapted MEEGA+ questionnaire (with a maximum time of 15 minutes).

We based the **instrumentation** on an adaptation of MEEGA+. Thus, it consisted of 52 items: 8 questions related to the participant's profile and 44 statements related to the game and its relationship to safety training (Table 2): 29 about game experience and 15 about safety training. The 44 statements were composed using a Likert scale, ranging from -2 (totally disagree) to +2 (totally agree). **Data collection**[6] was undertaken using an electronic Google form. The participants' responses were collected independently, although the classes allowed interaction and collaboration.

Rigorous empirical studies require identifying and addressing **validity threats** to the study. Table 3 presents the main threats and the treatments to mitigate them.

Data analysis used quantitative approaches, descriptive statistics (percentage, mean, median, mode, standard deviation), and inferential statistics (normality analysis, correlation, and hypothesis). The analysis software used was R Statistics 4.1.3. We defined the significance level for inferential statistics as 95% (alpha=0.05). We used tables and figures to summarize the quantitative analyses and the results.

6.1.3 Execution and Analysis

The **execution** of the evaluation occurred online from September 29, 2023, to October 13, 2023. It took place individually, in which **21 people** voluntarily participated in the study, as stated in the questionnaire's free and informed consent form. None of them attended the study more than once.

6.1.3.1 Participant's Profile

Regarding the profile of the 21 participants, the predominantly observed characteristic was the majority were men (95%), aged between 40 and 49 years old (45%), with an educational degree (41%), all workers in the oil and gas sector and who have already participated in a fire safety training session, and who rarely use games (45%).

Although the majority of participants in this study are men, they are not unique participants in the training process. Many women have been working in training contexts and hazardous situations. They will be the focus of future investigations related to this research context.

6.1.3.2 Reliability and Validity Questionnaire

To validate the reliability of the items in the MEEGA+ questionnaire, we used a statistical method, which, although the literature indicates limitations (Yang and Green, 2011), is still widely used for this purpose: the analysis of Cronbach's alpha coefficient. When interpreting the alpha value, one must observe the values (Cronbach, 1951) and interpret them as >0.9 excellent; >0.8 good; >0.7 acceptable; >0.6 questionable; >0.5 poor; and $<= 0.5$ unacceptable. According to Freitas and Rodrigues (2005), in exploratory studies, Cronbach's alpha values between 0.6 and 0.7 can be considered acceptable, and this argument is used in the reliability analysis presented in this paper.

Table 4 shows the Cronbach's alpha coefficient results applied to the MEEGA+ items. As can be seen, considering the items and their scales as a single questionnaire, the general alpha (Benreal Reliability) obtained a value of 0.91, which indicates excellent reliability in the participants' responses. Individually, the categories of gaming experience and training perception (learning and risks) were obtained, respectively, with alpha values of 0.88 (good) and 0.73 (acceptable), which corroborates the reliability and general validity of the responses.

When observing some of the variables individually, it is clear that some values demonstrate unacceptable reliability,

[4]CNS 510/2016: https://bvsms.saude.gov.br/bvs/saudelegis/cns/2016/res0510_07_04_2016.html

[5]LGPD: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm

[6]Data: https://doi.org/10.5281/zenodo.10666422

such as aesthetics (alpha = 0.40) and trust (alpha = 0.32). Regarding aesthetics, we understand that the prototype game may not have a very attractive design or the best fonts and color palettes. Therefore, we consider that this result may have influenced this value. Concerning the trust variable, we noticed that, at first, there is a lot of information for the player, which may have influenced an initial perception of a lack of trust and organization of the content perceived by the players. Consequently, both dimensions need attention in the future design process of these types of games.

However, we considered the participants' responses reliable and valid data for analyses, reflections, and considerations on the research questions defined for the study.

6.1.3.3 Participant's Perceptions and Category Correlations

When observing the percentage of responses (percentage column) and, mainly, the mode statistics in Table 4, it is observed that many of the items have values above 0 (some level of agreement with the statement - with the only exception of ACE01). Such a previous conclusion is easy to see when analyzing Figure 9.

Still, mode statistics observations cannot affirm positive answers to questions Q1 and Q2. Confirming the participants' perception, we made statistical inferences considering questions Q1 and Q2 as alternative hypotheses and, as neutral hypotheses, whether participants had negative or neutral perceptions related to them.

Table 2. Questionnaire items

Category			Code	Affirmation
Game Experience	Usability	Aesthetics (EST)	EST01	The game's design is attractive.
			EST02	The texts, colors and fonts match and are consistent.
		Learnability (APR)	APR01	I needed only a few things to start playing.
			APR02	Learning to play this game was easy for me.
			APR03	I think most people would learn to play this game pretty quickly.
		Operability (OPE)	OPE01	I consider the game to be easy to play.
			OPE02	The rules are clear and understandable.
		Accessibility (ACE)	ACE01	The fonts (size and style) used in the game are legible.
			ACE02	The colors used in the game are understandable.
		Error Protection (PTE)	PTE01	The game protects me from making mistakes.
			PTE02	When I make a mistake, it is easy to recover quickly.
	Trust (CONF)		CONF01	When I first looked at the game, I had the impression that it would be easy.
			CONF02	The content's organization helped me be confident that I would learn from this game
	Challenge (DES)		DES01	This game is suitably challenging for me.
			DES02	The game offers new challenges (offers new obstacles, situations or variations) at a good pace.
			DES03	The game does not become monotonous in its tasks (repetitive or tedious tasks).
	Satisfaction (SAT)		SAT01	Completing the game's tasks gave me a sense of accomplishment.
			SAT02	It is due to my effort that I can advance in the game.
			SAT03	I feel satisfied with the things I learned in the game.
			SAT04	I would recommend this game to my colleagues
	Fun (DIV)		DIV01	I had fun with the game.
			DIV02	Something that happened during the game (game elements, competition, etc.) made me smile.
	Focused Attention (AF)		AF01	Something interesting at the start of the game captured my attention.
			AF02	I was so involved in the game that I lost track of time.
			AF03	I forgot about my surroundings while playing this game.
	Relevance (REL)		REL01	The game content is relevant.
			REL02	It is clear how the game's content relates to real situations.
			REL03	Play is a suitable teaching method for learning about real-world situations.
			REL04	I would rather learn with this game than another way (another method).
Safety Training Perception	Short-term Learning (APC)		APC01	The game contributed to my learning about prevarication crime.
			ACP02	The game was efficient for my learning compared to other information and news (newspapers, social media, etc.).
	Learning Goals (OBA)		OBA01	It is possible to understand what is a fire type A playing the game.
			OBA02	It is possible to understand what is a fire type B playing the game.
			OBA03	It is possible to understand what is a fire type C playing the game.
			OBA04	Using the game, I understand what is the best extinguisher to fire type A.
			OBA05	Using the game, I understand what is the best extinguisher to fire type B.
			OBA06	Using the game, I understand what is the best extinguisher to fire type C.
			OBA07	Using the game, I understand that an extinguisher reaches the end.
	Risks (RISC)		RISC01	The game presented the number of fires that occurred and the total number of fires extinguished, which serves as a firefighting metric in the company.
			RISC02	The game showed that if I choose the wrong extinguisher I could cause malfunctions, explosions or other risks to the work environment.
			RISC03	The game showed a financial decrease, suggesting that poorly fought fires bring losses to the company.
			RISC04	The game showed that the longer fires took to be fought, the more outbreaks could emerge.
			RISC05	The game showed that if there is a delay in fighting fires, they could explode, even leading to the total compromise of the workstation.
			RISC06	The game shows that your health depends on you extinguishing the fire quickly but safely.

Table 3. Threats to validity

Type	Threat	Description	Treatment
Conclusion	The statistical power of the analysis method	It is related to using statistical methods that could reach wrong conclusions.	Scales and statistical methods more consistent with the metrics were selected and applied
	Violation of assumptions of statistical methods	Wrong use of the statistical tests to data that could not be evaluated with them	Use of statistical methods consistent with the scale and characteristics of the data sample.
	Bias in data selection	Favoring of data by researcher	The data used have been published so that others can repeat the analysis.
Internal	Lack of training	It happens when participants do not know how to operate the study object	To lessen the threat, participants could answer at their homes.
	Participant wear	This influences the participant behavior in studies that demand hard effort	Assessments are designed to be completed in 50 minutes.
Design	Research expectations	Influence of researchers above participants in a conscious way or not.	We planned the study so that the researcher did not communicate with participants, except in issues related to analysis objectives.
	Instrumentation	It happens when the study instruments are not adequate.	We used game evaluation questionnaires that were already used in previous research. Additionally, we made a reliability verification as the first analysis to guarantee the reliability and validity of participants' answers.
External	Planning	Use of reproducible methods.	To lessen this threat, the assessment was planned considering the design definitions of quasi-experimental studies (Campbell and Stanley, 2015)
	Generalization	It consists of generalizing the study results to a population bigger than the study participants.	We treated it using distinct participants' profiles. Each one was a worker in a Brazilian industry that already had safety training sessions in their jobs.

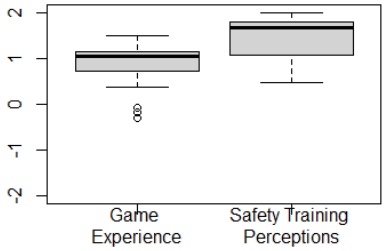

Figure 9. Participants' perceptions.

Hence, the first test was normality to determine which hypothesis test would be most suitable for sampling. In this study, as there were fewer than 30 participants, the most recommended test was the Shapiro-Wilk test (Asadoorian and Kantarelis, 2005). In addition to the mean and standard deviation values for gaming experience and perception of safety training, Table 5 also shows the results of the normality and inference tests. About the normality test, neither category follows normal behavior. In other words, their p-values are less than 0.05.

Table 5. Participants' perception analysis

Category	Weighted Average	Standard Deviation	Shapiro-Wilk (Normality test)		Wilcoxon (Hypothesis test)	
Game Experience	0.84	0.51	0.021	Non-Normal	1.51E-04	Accepted
Safety Training Perceptions	1.48	0.46	0.015	Non-Normal	6.29E-05	Accepted

By analyzing the Wilcoxon hypothesis test values, it is possible to state that all the values are below 0.05. As a result, it is possible to say with at least 95% certainty that the alternative hypotheses can be accepted. Consequently, it is possible to answer affirmatively that in Q1, the participants had a positive gaming experience. In Q2, the participants had a positive perception regarding learning and identifying training risks in the game.

The correlation analysis between the categories is vital because it identifies how each relates to the other. For this, we executed correlation tests among them. Since we perceived that the category data does not follow normal behavior (Table 5), the most indicated correlation test is the Pearson test (Benesty *et al*., 2009). According to Gasparin *et al*. (2010), we can interpret the value resulting from Pearson's correlation analysis as >0.5 large correlation; >0.3 medium correlation; >0.1 small correlation and <0.1 no correlation. This range of values can vary both negatively and positively.

Figure 10 presents the correlation measures between the game experience and safety training perceptions; we observe a strong correlation (0.63) between these dimensions. When detailing the security training perception category, it is still possible to notice strong correlations between gaming experience and risk identification (0.63), gaming and learning experience (0.53), and risk identification and learning (0.58).

Therefore, it is possible to conclude that there is evidence that participants perceived the BobRuff in Deck is on Fire game as positive in the study, with the game experience strongly correlated with learning from the training section and identification of risks of the context presented in the game. Consequently, once we developed the game with the support of SpGD, it is possible to say that there is evidence that the method can support the creation of digital games for training purposes.

Table 4. Questionnaire Reliability Analysis

Category	Variable		Itens	Percentage					Descriptive Statistics	Reliability
				P1	P2	P3	P4	P5	Mode	Cronbach's Alpha
Game Experience (Alpha = 0.88)	Usability	Aesthetics (EST)	EST01	10%	19%	29%	38%	5%	1	0.40
			EST02	10%	14%	33%	38%	5%	1	
		Learnability (APR)	APR01	0%	5%	19%	43%	33%	1	0.60
			APR02	0%	5%	19%	33%	43%	2	
			APR03	5%	14%	14%	43%	24%	1	
		Operability (OPE)	OPE01	0%	24%	29%	24%	24%	0	0.61
			OPE02	0%	10%	14%	24%	52%	2	
		Accessibility (ACE)	ACE01	14%	29%	24%	14%	19%	-1	0.72
			ACE02	0%	14%	14%	29%	43%	2	
		Error Protection (PTE)	PTE01	14%	10%	19%	33%	24%	1	0.76
			PTE02	24%	10%	29%	24%	14%	0	
	Trust (CONF)		CONF01	0%	5%	33%	29%	33%	0	0.32
			CONF02	0%	5%	19%	67%	10%	1	
	Challenge (DES)		DES01	0%	10%	19%	38%	33%	1	0.73
			DES02	14%	14%	19%	33%	19%	1	
			DES03	10%	10%	10%	52%	19%	1	
	Satisfaction (SAT)		SAT01	0%	10%	24%	29%	38%	2	0.86
			SAT02	0%	5%	10%	38%	48%	2	
			SAT03	0%	5%	5%	29%	62%	2	
			SAT04	5%	5%	5%	29%	57%	2	
	Fun (DIV)		DIV01	0%	10%	29%	14%	48%	2	0.91
			DIV02	14%	0%	24%	14%	48%	2	
	Focused Attention (AF)		AF01	0%	5%	14%	48%	33%	1	0.78
			AF02	14%	10%	24%	19%	33%	2	
			AF03	10%	5%	33%	38%	14%	1	
	Relevance (REL)		REL01	0%	5%	5%	38%	52%	2	0.74
			REL02	0%	0%	5%	5%	90%	2	
			REL03	0%	5%	5%	14%	76%	2	
			REL04	5%	0%	10%	33%	52%	2	
Safety Training Perception (Alpha = 0.73)	Short-term Learning (APC)		APC01	0%	0%	0%	33%	67%	2	0.57
			ACP02	0%	10%	0%	24%	67%	2	
	Learning Goals (OBA)		OBA01	0%	5%	0%	19%	76%	2	0.77
			OBA02	0%	5%	0%	19%	76%	2	
			OBA03	0%	0%	5%	19%	76%	2	
			OBA04	10%	0%	24%	10%	57%	2	
			OBA05	5%	0%	14%	19%	62%	2	
			OBA06	5%	0%	19%	14%	62%	2	
			OBA07	5%	5%	5%	10%	76%	2	
	Risks (RISC)		RISC01	10%	5%	19%	33%	33%	1	0.62
			RISC02	0%	0%	5%	33%	62%	2	
			RISC03	5%	0%	5%	29%	62%	2	
			RISC04	0%	0%	5%	10%	86%	2	
			RISC05	0%	0%	0%	5%	95%	2	
			RISC06	5%	10%	0%	14%	71%	2	
General Reliability										0.91

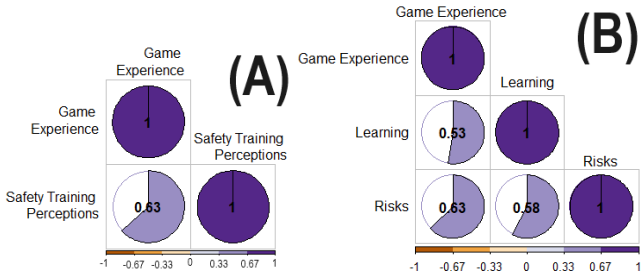

Figure 10. A) General correlation. B) Detailed safety training correlation.

6.2 SpGD Method Evaluation

6.2.1 Definition and Planning

We based the protocol that guided this part of the study on the Underlying Discourse Explanation Method (UDEM) (Nicolaci-da Costa, 2007) and included five stages: 1) selection of the sample of participants, 2) construction of the interview script, 3) conducting the interviews, 4) transcribing the statements and 5) analyzing and communicating the results.

Initially, the sample was selected (step 1), defining the desired profile of the participants, that is, game designers with experience in creating purposeful games. Next, we prepared a structured interview guide (step 2) containing specific questions that served as a guide during the individual interviews. The objective of this stage was to identify the perceptions, opinions, and suggestions of game designers regarding the SpGD method's strength, use feasibility, and improvements.

Having selected the participants according to the profile defined in the sample selection stage and having the structured script in hand, we conducted the interviews individually (stage 3), allowing direct interaction between the interviewer and the participant. The UDEM method provides that semi-structured interviews must have a pre-defined duration. For this reason, each interview lasted a maximum of 40 minutes, where the SpGD method was presented, with a detailed explanation of each of its steps and the demonstration of the technique, as well as a demonstration of the game developed using the method. Subsequently, we transcribed the interviews (step 4) to be analyzed and described in the present study (step 5).

We conducted a qualitative analysis, using Grounded Theory (GT) procedures as a basis (Strauss and Corbin, 1990) to analyze the interviews, which consists of data analysis in three phases: open coding, axial coding, and selective coding. In open coding, researchers separate data, conceptualize, and categorize. In axial coding, researchers identify possible relationships between categories. Finally, in selective coding, researchers create the central idea of the study.

Yet, since the study included interviews and qualitative data analysis, we must consider threats to validity. According to Pinto and Santos (2012), researchers in this type of research can inadvertently transmit personal bias in their interpretations since a speech coding stage was carried out. A group of researchers performed the qualitative analysis and, when discussing their interpretations, reached a consensus on the results to mitigate this threat.

The sample size may also threaten the generalization of this study. However, the UDEM method indicates the quality of the sample by considering the saturation of the statements given in the interviews. Such saturation was observed in the interviews performed in this study. Even so, specific contextual limitations must be considered in future studies to ensure the validity of the results.

6.2.2 Execution and Analysis

Game designers with academic and professional experience designing purposeful games, identified in lists from gaming communities, were invited to conduct the research. Five participants responded with their availability to participate in the study. The interviews happened between the 6th and 9th of June 2023, with participants interviewed in individual online meetings lasting 40 minutes, where the statements were recorded and transcribed. The interview report can include the questions that guided the interviews and the quotes from the participants used for this analysis (Table 6).

Table 6. Questions of interviews

ID	Question
Q1	The SpGD method uses Kirkpatrick's model as the base for its steps. How do you perceive utilizing this model to extract essential information from safety training and organize game design elements?
Q2	The second step of the SpGD method consists of mapping elements from Kirkpatricks' model to the MDA framework. How do you evaluate this mapping concerning representing safety training adequately in terms of game design?
Q3	The third step of the SpGD method aims to discuss and organize the safety training and MDA elements in a GDD. How do you perceive this creativity step concerning the game design process?
Q4	The SpGD method predicts a safety game validation with managers and an assessment of the trained to analyze if the game is aligned with the training context. How do you see the relevance of this step to the game design process?
Q5	Based on your game designer experience, what challenges do you perceive about designing games with safety training purposes? How could the SpGD method help it?
Q6	Would you consider using the SpGD method if you need to design a game for safety training purposes? Justify.
Q7	What suggestions or criticisms do you have concerning the SpGD that you would want to see in future versions?

We conducted the data analysis procedure using qualitative questions about the SpGD method. Although the GT mentions three types of coding, according to Strauss and Corbin (1990), it is up to the researcher to define which of these codings will be appropriate for the study. Therefore, this research adopted open and axial coding. Since the central idea of the study is unique and well-defined, the researchers did not consider it necessary to apply selective coding.

We read the data obtained in the interviews to implement the coding with the help of the Atlas.ti software, highlighting the important excerpts for the research, and assigning codes according to the content (open coding). Subsequently, we made the axial coding, where we grouped the codes into categories, which are sets of codes related to each other at a high degree of abstraction.

The analysis of the data obtained from the interviews resulted in identifying 14 different codes for the main contributions of the statements, grouped into three categories: strongness, improvements, and feasibility of the method. In Figure 11, it is possible to view the relationships among the codes and their respective categories.

6.2.2.1 Strongness

The main **strongness** cited by participants were concentrated on using Kirkpatrick's training assessment and the association of this information with the elements of the MDA framework. According to Participants #2 and #4, well-established theories must support the method to better understand safety training and its consequent translation into game elements. We can see this aspect in: *"It is important to be based on something already consolidated…It is a very safe way to invest"* [Participant #2]. And also that, subsequently, such mapping would help validate the result between training and game, for instance:

> I would say that it helps to understand better and define training. Before translating it into MDA mechanics... Having done it [mapping], initially, it will be more viable to use Kirkpatrick to validate the results of the training and the game. [Participant #4]

Participants #3, #4, and #5 considered that, besides the mapping being valid from their points of view, this association between Kirkpatrick's four levels of training evaluation and the elements of the MDA framework is the great highlight of the SpGD Game Design method as there is evidence that this process can make the creation of a game for security training more agile and reliable. According to Participant #5, *"This is the great thing about [the work]. As I mentioned, this is exactly the relationship between the two [Kirkpatrick and MDA]"*.

Still, on the eminent strongnesses, participants #2, #3, and #5 considered it valid to use a GDD to organize the ideas proposed in the brainstorming phase, citing that this is a good device to act as a meeting point for the game elements raised. Fragments of the interviews that illustrate these points can be seen in "*For me, yes, [the use of a GDD in the third stage] is appropriate"* [Participant #3], and "*I'm trying to see if I can think of other possibilities now off the top of my head, but I believe that Yes. I think it served [the use of GDD] very well"* [Participant #5].

The interviewees' attention was drawn to the stage where we evaluated the game developed in the fourth stage of the SpGD method with managers, trained, and the target audience. According to participants #1 and #5, this is an essential step capable of determining whether or not the process meets the objectives set for the project. Participant #5 reported that this stage:

> is just as important as all the previous development [...] If we didn't have this assessment so aligned and so refined, we could end up with the same problem initially mentioned [training disconnected from reality]

6.2.2.2 Weaknesses or Improvements

Participants suggested **improvements** in the first three stages of the SpGD method. For the first step, understanding the training, Participant #1 indicated that the technique could benefit from the addition of the ARCS (Attention, Relevance, Confidence, and Satisfaction) motivational model (Keller, 2009), which aims to improve motivation for the assessment of the generated artifact, that is, a game:

> [...] make a complement to what already exists [in the method], suddenly. If you will see how you measure engagement, speaking of Kirkpatrick, you may need to use Keller's theory. I perhaps see him [Keller] as one more complement, perhaps not exclusive" [Participant #1].

Participants #4 and #5, although considering the associations made between Kirkpatrick's levels and the elements of the MDA framework to be valid, reported the importance of allowing users of the method to make other associations, should they deem it necessary. This can be seen in *"About how this mapping was done, I would say that perhaps the reaction should also influence the game's dynamics"* [Participant #4) and:

> I don't think it's good to link [Kirkpatrick's levels with the MDA] either. I think it's very good to have this suggestion for people who like to run the model. [Participant #5]

Participant 1 suggested using a visual collaborative model to organize the brainstorming around the GDD,

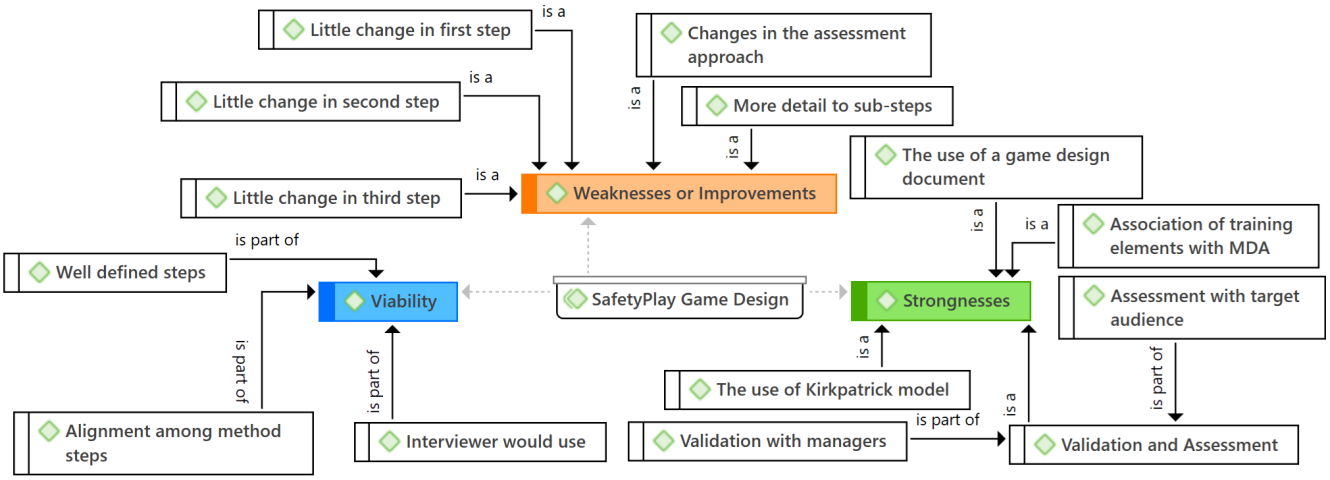

Figure 11. Codes and relationships associated with the SpGD method.

similar to the Miro [7] platform, to optimize the registration of ideas:

> The idea of [using] Canva would be interesting; it's a cool idea. Even to organize the ideas [in the brainstorming to make the GDD] ... Maybe use a structure similar to Miro's. [Participant #1]

Participant #3 contributed to the game assessment model contained in the SpGD method. According to this participant, it is important to define who will be the professional acting in the role of manager in the stages of the process, highlighting that a possible solution would be to change the term from manager to stakeholder (that is, an individual who has an interest in or influence on the project).

6.2.2.3 Viability

The indications regarding the **method's viability** are based on the statements where the participants respond that they would use SpGD if they needed to produce a game for safety training in the industry.

The simplicity of the method and the possibility of validation with managers, trainers, and the target audience were cited as points that favored its use, in addition to the well-defined steps and functions. It is worth highlighting the comment by Participant #2, who suggested that the method could be expanded to various professional training, not just those focused on safety.

Participant #2 comments that support the method's viability: *"I believe it helps. The steps in the method are very specific. Each step is well described. I believe so [the method would help overcome challenges]." [Participant #2]*. Also, Participants #4 and #1, answered about the viability:

> [I would use it because the method allows] understanding the training, using Kirkpatrick to get an idea of how the training could be evaluated. And validating with people who are there to minimize the game risk of not doing what it should. [Participant #4]

> [I would use] yes, it makes sense. It even makes sense to expand a little. It's more about the training line, in general. Not specifically, I haven't seen [another similar method] yet. [Participant #2]

7 Reflect

Based on the need to search for more modern and immersive alternatives to innovate in safety training sessions, this research presents the SpGD method as a proposed intervention in this problem context. The method aims to support the game design process for safety training, guaranteeing the traceability of training and game design elements throughout the competition process. For this, we built a training game prototype, which allowed reflections on the perception of the game experience and safety training and the possibility of interviewing game designers concerning the feasibility and positive and negative points of the method.

Did the players/trained positively perceive the game experience and safety training (learning and risks) through the gameplay?

Some authors argue that purposeful digital games can be useful for training sessions in risky situations. They allow the active involvement of trained in complex situations, immersing them in problem-solving as long as such games provide a good gaming experience (Kwegyir-Afful and Kantola, 2020). In this study, based on the game assessment by the trained, enabled us to perceive indications that the game experience contributed positively to immersion, attention, perception of the relevance of the training, and even to the fun and satisfaction of the trained (Table 4). These are indicative variables of the gameflow state (Berube, 2021), as participants' engagement and focus.

This immersion is ideal for learning during purposeful gameplay (Teichmann *et al*., 2020). We evidenced immersion in this study by identifying a strong correlation between the game experience and training perceptions, involving learning specific content related to the safety training contexts and identifying the risks involved in the presented situations. These observed values align with research by Mayer *et al*. (2013). In their study, the game is stressed by simulating risk situations and their consequences, leading people involved in the training to an experience of learning and reflection.

Therefore, it is possible to affirmatively answer that there is identification that the trained had a positive gameplay experience with the digital game for safety training purposes and that the game provided learning and risk identification within its training context. Since it was developed based on the SpGD steps, the method supported the game design process by monitoring and tracking the training in the game elements.

Was the SpGD viable to support the game's design for safety training purposes?

Although the benefits of games for training purposes are known, building them is a challenging and often unfeasible task due to the complexity involved when trying to translate the context of security training into the elements of digital games (Forbes, 2022; Wolf *et al*., 2022; Rufino Júnior *et al*., 2022). When conducting interviews with game designers experienced in building purposeful games, they pointed out that a method for building games for safety training is welcome, as long as they have well-defined steps that help understand the training.

However, a significant concern when building these games is precisely presenting the training without information being misunderstood or interpreted (Rufino Júnior *et al*., 2023). It is one of the main reasons SpGD supported training assessment models and aligned

[7] https://miro.com/pt/

them with models or frameworks for game elements, such as the Kirkpatric X MDA association. As pointed out by the interviewees in validating the method, this type of association allows not only support for game design but also to track training throughout the game creation process and, finally, validate the game results through the training evaluation model itself.

Moreover, some interviewees said that, although very useful in creating games for security training, attention is needed when carrying out Kirkpatric X MDA mapping. This suggestion, although it guarantees training traceability, can limit creativity and influence the composition of game mechanics.

In general, although there are suggestions for improvements to the method, the game designers considered it viable due to its simplicity, well-defined steps, and the involvement of managers, trainers, and game designers. Although existing, the risk of including poorly conveyed training information in the game is minimized.

In summary, there is evidence that the SpGD method is viable. It supports the design of games for safety training purposes, providing a good gaming experience that culminates in learning about training and identifying the risks involved in this context. Thus, it gives a viable option for contributing to safety training sections.

8 Final Remarks

This research aims to create a more systematic and simple game design for safety training in the industry. It presents the SafetyPlay Game Design method, which offers a structured and targeted approach to the context of risk situations in the industry.

We demonstrated the method by developing a game where fire safety training was translated into game elements, showing that designing this game using its proposed steps is possible. The game assessment was done by industry volunteers focusing on the perception of the game experience and safety training when learning concepts and identifying risks. After that, we assessed the method with game designers who attested to the technique's feasibility, all of whom highlighted that they would use it if necessary due to the simplicity and the presented theoretical-technical basis. Additionally, the same professionals contributed by highlighting strongness and suggesting improvements for future method versions.

In this sense, when analyzing the results of the evaluations, it is understood that there is evidence that the SpGD method is viable to support the process of building games for safety training. The research objective of **supporting the industry's design of safety training games** was achieved.

However, we can point out some limitations of the study. We recognize that we conducted quantitative and qualitative assessments with few participants in a controlled setting. In the quantitative assessment of the game, although the participants were volunteers with positions in the industry, it would be necessary to evaluate in the real context of a training section so that the results could be better generalized. We can say the same about the method in the interview. The game designers had contact with SpGD. They did not use it in practice in a game project for safety training purposes. Thus, the limitations and insights obtained from the assessment served as learned lessons and improvements for the subsequent investigation cycle.

In future work, we suggest that the variables investigated in the game and which had negative perceptions and the statements about suggestions for improvements to SpGD be analyzed, thought about, and implemented in the method to attest whether the changes added positive points to the process.

We also recommend that new games for safety training that use the SpGD method go through validation steps with managers and evaluations with the target audience for these types of safety training. Furthermore, the development of a computational tool is being studied to help systematize the alignments between the elements and in the other steps of the method.

With participants, we can validate the game by separating it into two experimental groups: one with prior knowledge about the training context and the other without. It will be considered in future assessments because we can understand the level of learning of the participants in each game session.

We have opportunities to examine the continued evolution of these games. In this context, it is interesting that we evaluate the method in the long term because we can constantly improve people's formations. The continuous formation takes into account the people's technological boundaries. The industrial staff can not know about the technology being used.

Finally, we understand that the study's contributions go beyond the academic field, offering real opportunities for the gaming industry and for startups and/or game houses to get involved in developing innovative training solutions in the industry, leveraging this promising market niche.

Acknowledgements

Authors thank FAPERJ (procs. E-26/210.231/2021) for partially funding this research. This paper is also an extended version of the best paper entitled *"SafetyPlay Game Design - Projetando Jogos Com Propósito de Treinamentos de Segurança na Indústria"*, published in the Proceedings of the XXII Brazilian Symposium on Games and Digital Entertainment (SBGames 2023).

References

Agarwal, N., Pande, N., and Ahuja, V. (2019). Expanding the kirkpatrick evaluation model-towards more efficient training in the it sector. In *Human Performance Technology: Concepts, Methodologies, Tools, and Applications*, pages 1092–1109. IGI Global.

Asadoorian, M. O. and Kantarelis, D. (2005). *Essentials of inferential statistics*. University Press of America.

Baldwin, T. T. and Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel psychology*, 41(1):63–105.

Basili, V. R. (1992). Software modeling and measurement: the goal/question/metric paradigm. *Technical Report: University of Maryland*, (CS-TR-2956, UMIACS-TR-92-9).

Benesty, J., Chen, J., Huang, Y., and Cohen, I. (2009). Pearson correlation coefficient. In *Noise reduction in speech processing*, pages 1–4. Springer.

Berube, D. (2021). The flow theory applied to game design. *Think Game Design*, 11.

Bogost, I. (2008). *The rhetoric of video games*. MacArthur Foundation Digital Media and Learning Initiative.

Bruzzone, A. G., Massei, M., Solis, A. O., Poggi, S., Bartolucci, C., and Capponi, L. D. (2013). Serious games as enablers for training and education on operations on ships and off-shore platforms. In *Proceedings of the 2013 summer computer simulation conference*, pages 1–8.

Campbell, D. T. and Stanley, J. C. (2015). *Experimental and quasi-experimental designs for research*. Ravenio Books, Stanford.

Carr, W. (2007). Philosophy, methodology and action research. In *The Quality of Practitioner Research*, pages 29–42. Brill.

Chittaro, L. and Ranon, R. (2009). Serious games for training occupants of a building in personal fire safety skills. In *2009 Conference in Games and Virtual Worlds for Serious Applications*, pages 76–83. IEEE.

Classe, T. M., De Araujo, R. M., Xexéo, G. B., and Siqueira, S. (2019). The play your process method for business process-based digital game design. *International Journal of Serious Games*, 6(1):27–48.

Coelho, J. A., Souza, G. H., and Albuquerque, J. (2020). Desenvolvimento de questionários e aplicação na pesquisa em informática na educação. *Metodologia de Pesquisa em Informática na Educação: Abordagem Quantitativa de Pesquisa*, 2.

Correa, C. R. P. and Cardoso Junior, M. M. (2007). Análise e classificação dos fatores humanos nos acidentes industriais. *Production*, 17:186–198.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, 16(3):297–334.

Davison, R., Martinsons, M. G., and Kock, N. (2004). Principles of canonical action research. *Information systems journal*, 14(1):65–86.

Ferraz, A. P. d. C. M. and Belhot, R. V. (2010). Bloom's taxonomy: Theoretical review and presentation of instrumental adequations for institutional goals definition (in portuguese: Taxonomia de bloom: revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais). *Gestão & produção*, 17:421–431.

Forbes (2022). Futuro do trabalho: seu treinamento será um game. Disponível em: `https://forbes.com.br/carreira/2022/05/games-invadem-treinamentos-corporativos/`. Acesso em: 19 Junho 2023.

Freitas, A. and Rodrigues, S. (2005). A avaliação da confiabilidade de questionários: uma análise utilizando o coeficiente alfa de cronbach. In *XII SIMPEP*, volume 1, pages 1–15.

Gallerati, P., Bagnato, S., Casciaro, D., Conte, A., and Maisano, M. (2017). Use of serious gaming and virtual reality applications improves students' learning retention and reduces safety risks and costs associated with training activities. In *Offshore mediterranean conference and exhibition*.

Gasparin, M., Menegotto, I. H., and da Cunha, C. S. (2010). Psychometric properties of the international otcome inventory for hearing aids. *Brazilian journal of otorhinolaryngology*, 76(1):85–90.

Hunicke, R., LeBlanc, M., Zubek, R., *et al*. (2004). Mda: A formal approach to game design and game research. In *Proceedings of the AAAI Workshop on Challenges in Game AI*, volume 4, page 1722. San Jose, CA.

Keller, J. (2009). Motivational design for learning and performance: the arcs model approach: Springer science & business media.

Kirkpatrick, J. D. and Kirkpatrick, W. K. (2016). *Kirkpatrick's four levels of training evaluation*. Association for Talent Development.

Kraiger, K., Ford, J. K., and Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of applied psychology*, 78(2):311.

Kwegyir-Afful, E. and Kantola, J. (2020). Simulation-based safety training for plant maintenance in virtual reality. In *International Conference on Applied Human Factors and Ergonomics*, pages 167–173. Springer.

Lacerda, É. R. M. and Abbad, G. (2003). Impacto do treinamento no trabalho: investigando variáveis motivacionais e organizacionais como suas preditoras. *Revista de Administração contemporânea*, 7:77–96.

Lopes, T. N., Mendes de Araujo, R., Moreira de Classe, T., and Gomes, T. (2022). Pyp4training-ludifying business process training. In *International Conference on Business Process Management*, pages 167–178. Springer.

Lu, S., Wang, F., Li, X., and Shen, Q. (2022). Development and validation of a confined space rescue training prototype based on an immersive virtual reality serious game. *Advanced Engineering Informatics*, 51:101520.

Martins, P. T. R. B. C. (2021). *Perceção dos riscos ocupacionais nos trabalhadores da indústria dos espetáculos e eventos ao vivo*. PhD thesis, Instituto Politécnico de Setúbal.

Mayer, I., Wolff, A., and Wenzler, I. (2013). Learning efficacy of the 'hazard recognition'serious game. In *International Conference on Serious Games Development and Applications*, pages 118–129. Springer.

Nicolaci-da Costa, A. M. (2007). O campo da pesquisa qualitativa e o método de explicitação do discurso subjacente (meds). *Psicologia: reflexão e crítica*, 20:65–73.

ONU (2022). Acidentes de trabalho e mortes acidentais crescem no brasil em 2021. *ONU News*. Disponível em: `https://news.un.org/pt/story/2022/04/1787092#`. Acesso em: 02 Junho 2022.

Petri, G., von Wangenheim, C. G., and Borgatto, A. F. (2019). Meega+: Um modelo para a avaliação de jogos educacionais para o ensino de computação. *Revista*

Brasileira de Informática na Educação, 27(03):52–81.

Pinto, M. d. R. and Santos, L. L. d. S. (2012). A grounded theory como abordagem metodológica: relatos de uma experiência de campo. *Organizações & Sociedade*, 19:417–436.

Rodrigues, L. B. and Santana, N. B. (2010). Identificação de riscos ocupacionais em uma indústria de sorvetes. *Journal of Health Sciences*, 12(3):31–38.

Rogers, S. (2010). *Level up! the guide to great video game design*. Wiley. OCLC: ocn475441192.

Rufino Júnior, R., Classe, T. M. d., and Siqueira, S. W. M. (2023). Games with training purpose for hazard situations in the industry - systematic mapping of the literature. In *Proceedings of the XIX Brazilian Symposium on Information Systems*, page 181–188, New York, NY, USA. Association for Computing Machinery.

Rufino Júnior, R., de Classe, T. M., and dos Santos, R. P. (2022). Jogos digitais para treinamento de situaçoes de risco na industria-rapid review. *Anais Estendidos do XXI Simpósio Brasileiro de Jogos e Entretenimento Digital*, pages 1157–1166.

Rufino Júnior, R., de Classe, T. M., dos Santos, R. P., and Siqueira, S. W. M. (2023). Current risk situation training in industry, and games as a strategy for playful, engaging and motivating training. *Journal on Interactive Systems*, 14(1):138–156.

Rufino Júnior, R., Classe, T., and Santos, R. (2022). Jogos digitais para treinamento de situações de risco na indústria - rapid review. In *Anais Estendidos do XXI Simpósio Brasileiro de Jogos e Entretenimento Digital*, pages 1157–1166, Porto Alegre, RS, Brasil. SBC. DOI: 10.5753/sbgames_estendido.2022.225970.

Rufino Júnior, R. and Classe, T. M. d. (2024). Safetyplay game design: Método para o design de jogos de treinamento de risco. *RelaTe-DIA*, 17(1).

Santos, P. S. M. and Travassos, G. H. (2011). Action research can swing the balance in experimental software engineering. In *Advances in computers*, volume 83, pages 205–276. Elsevier.

Schell, J. (2019). *The art of game design: a book of lenses*. Taylor & Francis, a CRC title, part of the Taylor & Francis imprint, a member of the Taylor & Francis Group, the academic division of T&F Informa, plc, Boca Raton - FL, third edition edition.

Serra, C. and Classe, T. (2023). Análise e visualização de dados em jogos de treinamento de situações de risco na indústria - um estudo em mapeamento sistemático. In *Anais Estendidos do XXII Simpósio Brasileiro de Jogos e Entretenimento Digital*, pages 1126–1137, Porto Alegre, RS, Brasil. SBC. DOI: 10.5753/sbgames_estendido.2023.233551.

Strauss, A. and Corbin, J. (1990). *Basics of qualitative research*. Sage publications.

Teichmann, M., Ullrich, A., Knost, D., and Gronau, N. (2020). Serious games in learning factories: perpetuating knowledge in learning loops by game-based learning. *Procedia Manufacturing*, 45:259–264.

TRT04 (2022). Abril verde: Brasil registrou crescimento de 30% em óbitos e acidentes de trabalho em 2021 na comparação com o ano anterior. *Tribunal Regional do Trabalho - 04ª Região*. Disponível em: https://www.trt4.jus.br/portais/trt4/modulos/noticias/501143. Acesso em: 19 Junho 2022.

TRT12 (2023). Cinco adolescentes são vítimas de acidente de trabalho por dia no brasil. *Tribunal Rerional do Trabalho - 12ª Região*. Disponível em: https://portal.trt12.jus.br/index.php/noticias/cinco-adolescentes-sao-vitimas-de-acidente-de-trabalho-por-dia-no-brasil. Acesso em: 20 Junho 2023.

Venturi, D., Konell, A. E., and Giovanela, A. (2021). Treinamento: importância e benefícios da disponibilização de treinamento nas organizações. *REVISTA CIENTÍFICA FAMAP*, 1(01).

Vigoroso, L., Caffaro, F., Micheletti Cremasco, M., and Cavallo, E. (2021). Innovating occupational safety training: a scoping review on digital games and possible applications in agriculture. *International Journal of Environmental Research and Public Health*, 18(4):1868.

Volpe, R. A. and Lorusso, C. B. (2009). A importância do treinamento para o desenvolvimento do trabalho. *Psicologia Online*, 1(1):01–08.

Winn, B. M. (2009). The design, play, and experience framework. In *Handbook of research on effective electronic gaming in education*, pages 1010–1024. IGI Global.

Wolf, M., Teizer, J., Wolf, B., Bükrü, S., and Solberg, A. (2022). Investigating hazard recognition in augmented virtuality for personalized feedback in construction safety education and training. *Advanced Engineering Informatics*, 51:101469.

Yang, Y. and Green, S. B. (2011). Coefficient alpha: A reliability coefficient for the 21st century? *Journal of psychoeducational assessment*, 29(4):377–392.