# The Adoption of a Framework to Support the Evaluation of Gamification Strategies in Software Engineering Education

Rodrigo Henrique Barbosa Monteiro[1] [a], Maurício Ronny De Almeida Souza[2] [b],
Sandro Ronaldo Bezerra Oliveira[1] [c] and Elziane Monteiro Soares[1] [d]

[1]*Graduate Program in Computer Science (PPGCC), Federal University of Pará (UFPA), Belém, Pará, Brazil*

[2]*Graduate Program in Computer Science (PPGCC), Federal University of Lavras (UFLA), Lavras, Minas Gerais, Brazil*

Keywords: Software Engineering Education, Gamification, Framework, Evaluation, Case Study.

Abstract: Context: gamification has been largely used to increase the engagement and motivation of students and professionals in their organizations, with a variety of models/frameworks for developing gamified approaches. Problem: the empirical data published so far are not sufficient to elucidate the phenomena resulting from the use of gamification, as there is no standardization in the specification of evaluation strategies, methods of analysis and reporting of results. Objective: therefore, the objective of this study is to present and discuss the use of a framework for the evaluation of gamification in the context of software engineering education and training. Method: for this, we executed a case study, in which the framework was used to support the design of an evaluation study for a gamification case in a software process improvement research group in a public university. Results: We report the main findings from observations and reports from the applicator of the case study, and 11 recommendations for the design of evaluation studies supported by the framework. Our main findings are: providing examples of usage of the framework improves its understanding, the framework helped the applicator in understanding that qualitative and quantitative data could be use in compliment to each other, and it helped streamlining the design of the evaluation study, considering the consistency between data to be collected, evaluation questions, and the goals of the evaluation study.

## 1 INTRODUCTION

The use of gamification – the adoption of game elements in non-game contexts (Deterding et al., 2011) – is a recurring theme in software engineering education literature in the last decade (Souza et al., 2018). There are many models, frameworks and processes for supporting the design of gamification strategies in diverse areas such as software engineering, business, health, crowdsourcing, and education. Yet, the literature on how to evaluate the effect of gamification in software engineering education is rare (Monteiro et al., 2021a).

Studies report positive and negative outcomes from the adoption of gamification (Hamari et al., 2014; Klock et al., 2018). However, generalization is difficult because: the evaluation data are heavily cou-

pled to context and individual characteristics of participants (Hamari et al., 2014); there is often insufficient detail on the design of evaluation procedures, hampering replication (Klock et al., 2018); there is lack of significant statistical data and imprecision of qualitative data (Bai et al., 2020); and there are no standard models in use to support the evaluation of gamification, hence there is no standardization in design, data analysis and report of gamification evaluation (Monteiro et al., 2021a).

Therefore, this paper goal is to report and evaluate the use of an evaluation framework for gamification. To achieve this goal, we describe a case study in which we observe the use of the framework to support the design and execution of the evaluation of a gamification case in the context of software engineering education and training. The gamification case was designed to engage participants in the proposal of solutions for problems related to software process improvements (SPI) in the context of a research group in a public university. As results, we provide a set of recommendations for the design and report of evalu-

---

[a] https://orcid.org/0000-0002-6142-8727
[b] https://orcid.org/0000-0002-0182-8016
[c] https://orcid.org/0000-0002-8929-5145
[d] https://orcid.org/0000-0002-3408-8640

Table 1: Phases and entities of the framework (Monteiro et al., 2021b).

| Evaluation Phases | Evaluation Entities |
| --- | --- |
| Contextualization of gamification | *Gamification* – game elements, context description, models and concepts used, goal of the gamification approach, methods. |
| Contextualization of the evaluation | *Evaluation* – scientific research method, duration, population; *Goals*, *Criteria*, and *Evaluation questions*. |
| Definition of methods | *Metrics*, *Indicators*, and *data collection and analysis instruments*. |
| Summarization of results | *Rounds* – description and duration; *samples* – demographics and size; and *data colletion* (for each metric and sample). |
| Analysis of Results | *Data analysis* – results for evaluation questions; and *Findings*. |

ation studies on the use of gamification for software engineering education, based on the experience of using the framework.

The remainder of this paper is organized as follows. Section 2 presents the framework for evaluation of gamification. Section 3 presents related work. Section 4 presents the study design and methods. Section 5 reports the execution of the case study. Section 6 presents and discusses the evaluation of the case study, and recommendations for the design of evaluation studies for gamification approaches in software engineering. Section 7 presents possible threats to validity and mitigation strategies. Section 8 presents our concluding remarks, contribution, limitations, and future work.

## 2 FRAMEWORK FOR EVALUATION OF GAMIFICATION

Monteiro et al. (2021b) presented a conceptual framework for evaluation of gamification in the context of software engineering education and practice. The goal of the framework is to provide a standard structure for the design of evaluation studies for gamification cases. The structure considers the planning, execution, analysis and report of results. As an expected contribution, the framework is intended to support the production of empirical data that could be more easily compared.

The design of the framework considers results from a literature review (Monteiro et al., 2021a) on the evaluation strategies used in gamification studies, and refinements on previous study (Monteiro et al., 2021b). Its structure is based on the GQIM (Goal-Question-Indicator-Metric) model – a model that drives the design of evaluation metrics from the top-down analysis of organizational goals (Park et al., 1996).

The framework is organized in evaluation phases and evaluation entities. The evaluation phases de-

scribe a sequence of decisions that drive the revision of the gamification design, and drive its designer to reflect on evaluation goals, criteria, questions, required data and data analysis procedures. The evaluation entities are sets of these data, and their relationship, that need to be documented for the evaluation.

Table 1 shows the framework components, mapping the evaluation phases to their respective entities.

## 3 RELATED WORK

To the best of our efforts, we found three studies that propose gamification models or frameworks, in the context of software engineering, with evaluation steps (Ren et al., 2020; Gasca-Hurtado et al., 2019; Dal Sasso et al., 2017), as described in (Monteiro et al., 2021a). However, we did not find any evidence of the adoption of these models to evaluate gamification cases. Nevertheless, the framework used in the current work (see Section 2) is inspired by elements of the three models (Ren et al., 2020; Gasca-Hurtado et al., 2019; Dal Sasso et al., 2017).

In addition to these models, we found related work on the proposal of evaluation frameworks in different contexts. For instance, Petri et al. (2019) present a framework for the evaluation of serious games in the context of computing education: MEEGA+ (Model for the Evaluation of Educational GAmes). In the study, the models are valuated using case studies and surveys. Therefore, this study provided insights for the design of the case study described in the current paper.

## 4 METHODS

This Section presents the study design, describing its goals, methods, actors and instruments. The goal of this study is to evaluate the adoption of a framework for the evaluation of gamification in the context of software engineering education and practice. To achieve this goal, we executed a case study, consist-

ing of the design of the evaluation of a gamification project in the context of software process improvement research group. A case study is an empirical method focusing on the investigation of a contemporary phenomenon in its real-life context, using multiple data collection methods, and without direct intervention (or active role) of the investigator in the case Wohlin (2021).

The following subsections describe the actors involved in the study (Section 4.1), the study design (Section 4.2), and instruments (Section 4.3).

## 4.1 Actors

There are three roles involved in this study: *researcher*, *observer* and *applicator*. The *researcher* role is performed by three software engineering researchers (two PhD professors and a graduate student) who were actively involved in the design of the framework. *Researchers* are responsible for planning the case study and for the analysis of results from the observation of the case study. The *observer* role is performed by one the researchers (the graduate student), responsible for supporting the *applicator* in understanding the framework and to collect data on the use of the framework. Finally, the *applicator* is a graduate student who designed a gamification strategy and applies the proposed framework to plan an evaluation study. She has 3 years of professional experience in software engineering.

## 4.2 Study Design

The execution of the case study was performed in four steps: *planning*, *meetings*, *evaluation*, and *framework usage experience analysis*.

**Planning.** This step consists in the planning of the case study, resulting in an initial schedule of meetings, design and selection of instruments, and design of support materials.

**Meetings.** The *applicator* and the *observer* held eight virtual meetings. During these meetings, the former used the framework to design an evaluation study for his gamification case, and the latter provided assistance for the understanding of the framework (on demand), and documented (text, audio and video) what happened in each meeting.

**Evaluation.** The *observer* interviewed the *applicator* (semi-structured interview) for collecting impressions on the use of the framework.

**Data Analysis.** The researchers analysed the data from the meetings records and from the evaluation phase and drew findings and insights for improvement of the framework.

## 4.3 Instruments

Considering restrictions imposed by the COVID-19 pandemic, the study was executed remotely. The observer provided the applicator with the previous work (Monteiro et al., 2021b) describing examples of use of the framework, and an *evaluation sheet* – an electronic sheet structured for documenting the entities of the framework. These instruments were intended to support the understanding of the framework, and to support its usage.

## 5 CASE STUDY

This section describes the execution of the case study. Subsection 5.1 describes the gamification case considered for the case study, and Subsection 5.2 reports the execution and results of each phase of the framework.

## 5.1 Case Study Object

The target of the application of the framework is a gamification approach to support the engagement of participants in proposing solutions to software process improvement (SPI) problems.

The gamification approach is the result of previous studies of problems related to SPI (Soares and Oliveira, 2020), and the adaptation of the Octalysis framework to compose a gamification strategy to overcome these problems (Soares and Oliveira, 2021). In this gamification approach, each participant assumes the role of a super hero, engaging in missions that relate to SPI activities. In each mission, participants propose, discuss and evaluate SPI actions. These SPI action are, then, classified according to "Customer" and "Market" competence dimensions from MOSE certification model (Rouiller, 2017). The development of these competences aim to improve the relationship between an organization and its internal and external customers.

The gamification approach was introduced to the context of a research group in a public university, for the purpose of learning about SPI and practicing the proposal of solutions to SPI problems. Seven graduate and undergraduate students of the research group participated in the gamification case, for six weeks. The designer of the gamification strategy (the *applicator*) used our framework to propose an evaluation study to assess the impact of their gamification case.

## 5.2 Case Study Report

The case study was executed in eight meetings. The following subsections reports the execution and results of each phase of the framework, as performed by the applicator.

### 5.2.1 Contextualization of the Gamification

In the first meeting, the observer presented the framework to the applicator, and answered questions regarding its use. In sequence, the applicator documented the context of his research and the goal of gamification. They listed the game mechanics and dynamics used. The applicator used an evaluation sheet, provided by the observer, to support the description of the gamification design. The *Contextualization of the Gamification* phase of the framework took two meetings. Section 5.1 presents the data documented.

### 5.2.2 Contextualization of the Evaluation

The applicator had previously designed some elements of the evaluation, prior to the use of the framework. It included: research method (case study), population sampling (the research group members), the duration of the study (six weeks), and a set of metrics. However, it was noticed that the applicator skipped important information, that the framework helped them to reflect and to review their evaluation design. For instance, there was no goal clearly stated for the evaluation study. Hence, the applicator had to propose a goal statement for the evaluation study, and realized they had to review the elements previously designed to ensure adequacy to the goal. The objective of the evaluation was to "check/analyze whether the gamified approach helped in solving SPI problems".

### 5.2.3 Methods Definition

For the definition of evaluation methods, the applicator designed evaluation criteria, questions, indicators and metrics, and documented them in electronic sheets. The evaluation considered 7 criteria: Awareness, Performance, Engagement, Learning, Positive Involvement, Participation and Satisfaction.

For those dimensions, the applicator proposed 72 questions (several questions for each SPI problem being investigated). To answer these questions, 46 indicators were designed, related to 9 metrics.

Each of these elements were documented and further detailed with data collection and analysis procedures. For instance, the applicator defined five data collection instruments, and quantitative and qualitative data analysis procedures for the evaluation questions – 11 questions with quantitative procedures, 27 with qualitative procedures, and 34 with both qualitative and quantitative procedures. The result of this phase was documented in an electronic sheet [1].

It is important to notice that, these numbers relfect the final state of the evaluation designed. However, the evaluation study was started with only a few of these elements carefully designed and documented. As a consequence, the evaluation methods had to be refactored during the study, which impacted in missed data.

### 5.2.4 Summarization of Results

This phase took place during the six weeks devoted to the execution of the evaluation study, in which the applicator collected data for computing the metrics defined in the previous phase. For each round of the evaluation study, the applicator documented all the collected data using the same evaluation sheet used for the definition of evaluation methods. However, some metrics and evaluation questions were refactored after the start of the evaluation study, and some data could not be collected.

Table 2 presents information on each round of the evaluation study. The rounds were designed as "Missions", in which the participants had to fulfil specific activities (described in the column "Description"). Table 3 presents a sample of the results as documented in the evaluation sheet, for the first round.

Table 2: Summarization of results - Rounds.

| Round | Description | Duration |
|---|---|---|
| Mission 1 | Presentation | 52hrs |
| Mission 2-5 | Dynamics | 113.5hrs |
| Mission 6 | Feedback from participants | 3.5hrs |

Table 3: Sumarization of results - Partial sample of data from Round 2.

| Participant | Metrics | | | | |
|---|---|---|---|---|---|
| | M01 | M02 | M03 | ... | M46 |
| H01 | 30 | 86 | 5 | ... | 5 |

### 5.2.5 Analysis of Results

The goal of this phase was to obtain answers for each evaluation questions, based on the analysis of their respective indicators and metrics. In the object study, the applicator used both qualitative and quantitative methods for the data analysis, following a deductive

---

[1]https://doi.org/10.5281/zenodo.5731447

analysis approach.

The applicator drew 50 analysis related to game elements, and 20 analysis related to SPI problems. In addition to the analysis of the evaluation questions, the applicator also performed an overall analysis of the evaluation goal. The applicator chose not to use the entity "Findings" of the framework to draw conclusions. Only two evaluation questions were not answered with the collected data.

# 6 EVALUATION OF THE FRAMEWORK USAGE

This Section describes the analysis of the case study reported in Section 5. After the conclusion of the case study, we performed an interview (Section 6.1) with the applicator. Finally, we draw a set of recommendations (Section 6.2) for the use of the framework, based on the analysis of the case study.

## 6.1 Interview

After the conclusion of the evaluation study, the observer interviewed the applicator in order to collect his impressions on the use of the framework. The interview focused on main problems observed during the case study.

> **Observer:** *"Did you have difficulties during the evaluation stages?"*
>
> **Applicator:** *"I did. My main difficulty was related to the phases of "Definition of Methods" and "Summarization of Results". It was largely caused by the large quantity of evaluation items. I could not visualize the framework as a whole. (...) Regarding the "Analysis of Results" phase, the collected data were well structured, therefore the execution of the analysis was fast."*

From this answer and from the observations made, we can infer that the high level of details of the evaluation study caused an elevated number of evaluation items. The applicator opted to analyze (i.e. define evaluation questions) for each gamification element and SPI problem investigated in the study, individually. It caused redundancy in the design of metrics and collected data. As a consequence, the evaluation sheet was hard to maintain.

> **Observer:** *What could facilitate the process of documenting results?*
>
> **Applicator:** *In retrospect, I believe some evaluation questions could be answered exclu-*

*sively by the analysis of metrics. However, the analysis relied too much on qualitative data.*

> **Observer:** *Whats the proportion of evaluation questions not dependant on qualitative data?*
>
> **Applicator:** *I believe they are not so representative, corresponding to approximately 30% of the evaluation questions.*

A key problem observed in the study case was that, previous to the use of the framework, the applicator had partially designed an evaluation evaluation study. Some evaluation items were only reviewed and refactored (with the support of the framework) after the start of the evaluation study. It caused problems in the collection of data.

> **Applicator:** *"Maybe it has been my biggest problem. If I had designed the evaluation study in accordance to the framework from the beginning, I could have collected more precise data. I had a lot of difficulties in some evaluation questions, because I had not planned them properly, and there was no time left to run new evaluation rounds to collect data."*

On the use of the *evaluation sheet* instrument, the applicator mentioned its usage was relevant for the phases of "Contextualization of Gamification" and "Contextualization of the Evaluation". However, the applicator found difficult to use the instrument for the definition of methods. Overall, the instrument was useful for the documentation of the evaluation study design.

> **Observer:** *"How was your experience using the evaluation sheet provided?"*
>
> **Applicator:** *"After reading the paper provided (Monteiro et al., 2021b), I understood that the basic information of the evaluation are requirements for the proper definition of methods. Therefore, I believe it makes sense to use the evaluation sheet in these phases ('Contextualization of Gamification' and 'Contextualization of the Evaluation'). It was difficult to execute the phase 'Definition of Methods' using the evaluation sheet, because each metric had to be structured. [...] It was difficult to get the full picture. "*

### 6.1.1 Findings

From the observations and the results of the interview, we list the following findings related to the use of the framework.

1. Providing examples: Providing examples of usage of the framework improves its understanding.

2. Using of qualitative and quantitative data: The initial evaluation study designed by the applicator considered only quantitative data for analysis. The framework helped the applicator in understanding that qualitative and quantitative data could be use in compliment to each other.

3. Using Evaluation sheet: Using the evaluation sheet as an instrument for guiding the use of the framework and documenting design decisions was positive. For the initial stages of the framework ("Contextualization of Gamification" and "Contextualization of the Evaluation") the sheet was sufficient and relevant. For the later phases, the complexity of the evaluation (number of evaluation questions, indicators, metrics, evaluation rounds, and other) may lead to increased complexity in the maintenance and analysis of the sheet.

4. Streamlining evaluation design: The framework helped streamlining the design of the evaluation study. It provided a sequence of decision making actions that enforced the deign of an evaluation study considering the consistency between data to be collected and evaluation questions, and between evaluation questions and the goals of the evaluation study.

## 6.2 Recommendations

From the use of the framework, we propose a set of recommendations for the design and report of evaluation studies in the context of gamification in software engineering education. The recommendations are organized in accordance to the phases of the framework, in the following subsections.

### 6.2.1 Contextualization of Gamification

1. **Document the Gamification Design.** In order to understand what should be investigated in an evaluation study, it is important to clearly document the gamification approach, its goals, the game elements used and their respective purpose. This information is important for the proper planning and understanding of the scope of an evaluation study.

### 6.2.2 Contextualization of the Evaluation

1. **Clearly Define the Scope of the Evaluation.** Clearly state the goal of the evaluation and what aspects are considered for investigation. From the use of the framework, the "Goal" and "Questions", from GQIM model, is particularly useful

for the definition of the scope of the evaluation. In addition to that, the framework entity "evaluation criteria" helps framing which aspects (e.g. performance, satisfaction, fun) are addressed by each evaluation question. From our observation, skipping this step proved troublesome, because the planning of the data collection procedures is heavily impacted by this scope.

2. **Be Careful with the Extension of the Evaluation.** Carefully planning the scope of the evaluation study is key to understand its viability. In our case study, the applicator opted for proposing an elevated number of evaluation questions (in part due to the attachment to their original evaluation design, prior to the use of the framework). This design strategy may lead to a very hard to maintain dataset, complex data analysis, greater risk of misinterpretation of data, and greater effort and time consumed for the analysis. Moreover, the data collection procedures are more prone to lose precious data.

3. **Consider Examples.** Read previous studies to understand how other researchers designed their evaluation studies. Regarding the use of the framework, Monteiro et al. (2021a) provides examples from real studies that were useful for the applicator understanding of its use. The framework can also be used for reverse-engineering previous studies for the revision of their design.

### 6.2.3 Definition of Methods

1. **Define Evaluation Methods in Accordance to the Evaluation Scope.** It is important to ensure the conformity of the definition of methods (research method, population, data collection and analysis procedures) to the scope of the evaluation. The framework helps to maintain the coupling of evaluation scope and methods by using the GQIM concept of mapping indicators and metrics (that are tied to definition of methods) to evaluation questions. Failing to this recommendation may lead to unaddressed evaluation questions, collection of unnecessary data or deviation to the scope of the evaluation.

2. **Mix Methods.** Plan the methods for each evaluation question individually, and consider using the most adequate methods for addressing each one. This may require the use of both quantitative and qualitative approaches. In addition, consider that using varied methods and data for addressing a given evaluation question may provide complimentary inputs for analysis.

3. **Plan the Maintenance of the Dataset.** During the definition of methods, plan how to store data in a opportune manner for data analysis. This may involve structuring sheets or databases that support easier retrieval of data and computation of metrics and indicators. Also consider that different evaluation questions may share the same metrics. Therefore, it is wise to decouple the metrics from the evaluation questions in the data storage structure.

4. **Run Pilot Studies.** Consider running pilot studies on smaller and (more) controlled environments/samples to validate the evaluation study design. Is the collected data sufficient for addressing the evaluation questions? Are the evaluation questions relevant for the evaluation goal? Does the evaluation method seem appropriate? Based on these questions it is possible to refine the scope and design of the evaluation study in advance, preventing risky refactoring actions in execution time.

### 6.2.4 Summarization of Results

1. **Preserve the Raw Data.** Store and report the raw data obtained from the data collection procedures and the computed metrics. These data should be available, independently, for analysis of indicators, addressing research questions and future analysis or audits.

### 6.2.5 Analysis of Results

1. **Focus on the Evaluation Questions.** The data analysis should focus on addressing the evaluation questions and, consequently, the evaluation goal. The fulfilment of the evaluation goal must be assessed based on the results of the evaluation goals.

2. **Address the Evaluation Questions based on Evidences.** The answering of each evaluation question should rely on the analysis of indicators, based on available data and the results of qualitative and quantitative methods. The framework organize the data in several entities that may help in the identification of patterns.

## 7 THREATS TO VALIDITY

This section describes the possible threats to validity of our study. We consider the categories proposed by Wohlin et al. (2012): conclusion, internal, external, and construct.

**Construct Validity.** The planning phase is threaten by construct validity bias related to the design of an initial script for the instantiation of the framework. It is possible that not every evaluation item planned was actually relevant for the design of the evaluation study. To mitigate this threat, we used the proposed framework, that was previously validated by specialists and considered elements and good practices from the state-of-the-art literature on gamification in software engineering education.

**Internal Validity.** The framework applicator could have instantiated the framework incorrectly, which could make the summarization and analysis of results unfeasible. To mitigate this, researchers reviewed the documented data and their interpretation of it, and requested revisions when necessary).

**Conclusion Validity.** This solution could have generated a threat of completion if there were suggestions from the observer, interfering with the collected results. However, this was mitigated because the researcher with the role of observer only spoke during the meetings to expose the possibilities of instantiating the framework, and the other researchers only spoke with the observer again after the completion of the evaluation stage, to give their opinion on the assessment. In addition, to mitigate the threats to conclusion related to analysis bias, we present the case study report (Section 5) and the interview performed (Section 6).

## 8 CONCLUSION

This study presented a report on the use of the framework for the evaluation of gamification in the context of software engineering education and training. From the analysis of the experience using the framework, it is possible to see that the case study applicator could improve the design of an evaluation study, and achieve positive results from the use of the framework. From this experience, it was possible to extract a set of recommendations and examples of instantiation for the framework, which can help future works in gamification and software engineering in the design of an evaluative approach suited to its context of study.

A limitation of this study is the lack of assessment of each evaluation item and instruments defined during the study case. Furthermore, the recommendations were extracted from only one experience of using the framework, and therefore cannot be generalized. For future work, it is intended to prepare and evaluate an instruction sheet for the use of the framework, recommendations and instantiation examples.

In addition, it is intended to validate the framework from new analysis of gamification evaluations, which can be performed using this instruction sheet.

# ACKNOWLEDGEMENTS

# REFERENCES

Bai, S., Hew, K. F., and Huang, B. (2020). Does gamification improve student learning outcome? evidence from a meta-analysis and synthesis of qualitative data in educational contexts. *Educational Research Review*, 30:100322.

Dal Sasso, T., Mocci, A., Lanza, M., and Mastrodicasa, E. (2017). How to gamify software engineering. In *2017 IEEE 24th International Conference on Software Analysis, Evolution and Reengineering (SANER)*, pages 261–271.

Deterding, S., Dixon, D., Khaled, R., and Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification". In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, MindTrek '11, page 9–15, New York, NY, USA. Association for Computing Machinery.

Gasca-Hurtado, G. P., Gómez-Álvarez, M. C., Muñoz, M., and Mejía, J. (2019). Proposal of an assessment framework for gamified environments: a case study. *IET Software*, 13(2):122–128.

Hamari, J., Koivisto, J., and Sarsa, H. (2014). Does gamification work? – a literature review of empirical studies on gamification. In *2014 47th Hawaii International Conference on System Sciences*, pages 3025–3034.

Klock, A. C. T., Ogawa, A. N., Gasparini, I., and Pimenta, M. S. (2018). Does gamification matter? a systematic mapping about the evaluation of gamification in educational environments. In *Proceedings of the 33rd Annual ACM Symposium on Applied Computing*, SAC '18, page 2006–2012, New York, NY, USA. Association for Computing Machinery.

Monteiro, R. H. B., de Almeida Souza, M. R., Oliveira, S. R. B., dos Santos Portela, C., and de Cristo Lobato, C. E. (2021a). The diversity of gamification evaluation in the software engineering education and industry: Trends, comparisons and gaps. In *2021 IEEE/ACM 43rd International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET)*, pages 154–164. IEEE.

Monteiro, R. H. B., Oliveira, S. R. B., and De Almeida Souza, M. R. (2021b). A standard framework for gamification evaluation in education and training of software engineering: an evaluation from a proof of concept. In *2021 IEEE Frontiers in Education Conference (FIE)*, pages 1–7.

Park, R. E., Goethert, W. B., and Florac, W. A. (1996). *Goal-Driven Software Measurement - A Guidebook*. Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA 15213.

Petri, G., Wangenheim, C. G. V., 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.

Ren, W., Barrett, S., and Das, S. (2020). Toward gamification to software engineering and contribution of software engineer. In *Proceedings of the 2020 4th International Conference on Management Engineering, Software Engineering and Service Sciences*, ICMSS 2020, page 1–5, New York, NY, USA. Association for Computing Machinery.

Rouiller, A. C. (2017). *MOSE®: Base de Competências*. Editora Pé Livre Ltda.

Soares, E. and Oliveira, S. (2021). An analysis of gamification elements for a solving proposal of software process improvement problems. In *Proceedings of the 16th International Conference on Software Technologies - ICSOFT,*, pages 294–301. INSTICC, SciTePress.

Soares, E. M. and Oliveira, S. R. B. (2020). An Analysis of Problems in the Implementation of Software Process Improvement: a Literature Review and Survey. In *CONTECSI*.

Souza, M. R. d. A., Veado, L., Moreira, R. T., Figueiredo, E., and Costa, H. (2018). A systematic mapping study on game-related methods for software engineering education. *Information and software technology*, 95:201–218.

Wohlin, C. (2021). Case study research in software engineering—it is a case, and it is a study, but is it a case study? *Information and Software Technology*, 133:106514.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M., Regnell, B., and Wesslén, A. (2012). *Experimentation in Software Engineering*. Springer.