IET Software

Special Issue: Gamification and Persuasive Games for Software

IET Journals
The Institution of Engineering and Technology

Gamification for software process improvement: a practical approach

ISSN 1751-8806
Received on 17th April 2018
Revised 18th June 2018
Accepted on 16th July 2018
E-First on 22nd August 2018
doi: 10.1049/iet-sen.2018.5120
www.ietdl.org

Eduardo Herranz[1] ✉, Javier García Guzmán[1], Antonio de Amescua-Seco[1], Xabier Larrucea[2]

[1]Computer Science Department, Universidad Carlos III de Madrid, Av. Universidad 30, Leganés, 28911, Madrid, Spain
[2]TECNALIA, Bizkaia, Spain
✉ *E-mail: eduardo.herranz@uc3m.es*

Abstract: Gamification is a research field that is intended to increase motivation, so it is especially indicated in human capital intensive environments such as the software industry. Within Software Engineering, one of the main issues regarding software process improvement (SPI) is personnel motivation in specific SPI initiatives. These issues are stronger in small and medium software development companies where employees have to deal with the pressure of deadlines and occasional work overload. To address the adoption of SPI initiatives, the researchers implemented a defined gamification framework for deployment in SPI efforts in order to increase motivation among software workers and to enhance SPI results. The framework was rolled out in a small Spanish software development organisation, which is conducting internal SPI initiatives. To validate the effectiveness of the implemented framework, a controlled experiment was carried out in which an experimental group adopted SPI improvements using a gamification approach. The implementation results show that the application of the framework does not increase personnel motivation in SPI tasks although it contributes to enhancing the SPI tasks performance. This study discusses the limitations and recommendations to implement appropriately the SPI-gamification framework in the scope of small and medium software development companies.

1 Introduction

Software process improvement (SPI) is seen as the dominant approach to improving software products in software development organisations [1]. SPI has become the primary approach to improving software quality and reliability, employee and customer satisfaction, and return on investment [2]. Personnel factors are one of the elements that can have an impact on the productivity of software teams and the effectiveness of the whole software process [3]. More specifically, in SPI people can be seen as the main factor that needs to be encouraged and supported in an organisation [4]. Among the personnel factors, motivation is very important for workers in general and, in the case of software workers, it is one of the most frequently cited causes of software development project failures [5]. SPI is also a fertile field of study with regards to the importance of motivation as a soft factor. There are recent and important works devoted to investigating motivation in SPI initiatives [6]. These studies agree on the importance of counting on new ways to both support implementation and motivate workers.

In recent years, gamification has been pointed out as powerful motivational tools in the working environment. Gamification can be defined as the use of game elements in non-game contexts to modify and influence the behaviour of people [7]. Gamification has its impact in the broad field of software engineering.

One of the expected (and documented) consequences of gamification in people's behaviour is the increase in motivation and engagement [8]. Given that motivation is an important factor for software practitioners, gamification initiatives are beginning to be implemented in several software fields of which SPI is one of the most fertile areas [9, 10]. Several of these initiatives are related to educational games for requirements engineering [11]. In recent years, several initiatives in the field have been reported regarding the use of gamification in SPICE [5, 12], organisational change management initiatives within SPI [13] or specific tools to support the gamification process [14, 15].

In spite of the benefits of gamification as a discipline, gamified applications are not easy to implement and practitioners must take into account some aspects [16] in order to avoid a poor conceptualisation of the environment that could lead to failure. This has led to the definition of a framework for gamification efforts in SPI environments that were introduced [9, 13, 17] and validated by 29 international SPI and gamification experts [17]. From the findings of the investigation, it can be stated the validity of the gamification framework, the principles that sustain it and its relevance and applicability in SPI initiatives [17]. However, to confirm the external validity of the SPI gamification framework and its practical feasibility, an empirical validation is needed [17]. Aligned with this, some authors [11, 18] reviewed the current literature on the topic, and confirmed the regular use of simple gamification mechanics and the lack of empirical evidence of the impact of gamification, turning this into one of the most relevant research issues [19].

This paper aims to bridge the gap between gamification in SPI and empirical evidence by presenting an effort to deploy the SPI-gamification framework devoted to SPI in real settings. This framework was implemented in a small Spanish software development organisation to validate its effectiveness. The specific hypotheses addressed for this experimental validation were related to the effectiveness of the framework in increasing staff motivation in SPI initiatives and the software process performance.

This paper is structured as follows: In Section 2, the methodological framework is introduced. Section 3 describes the research methodology to empirically validate the aforementioned framework. Section 4 shows the results of that empirical validation while Section 5 discusses the results. Section 6 presents the conclusions and implications of the research.

2 Gamification framework

The complexity of SPI and the intrinsic convolution of gamification lead to the need to develop a specific framework that must entail personnel and organisational aspects and define a set of processes to guide it. As stated earlier, this framework was presented in previous works [9, 13, 17]. In the last version of the framework, the lean Startup (build-measure-learn) was adopted. Fig. 1 depicts the phases of the framework (named in this research as G-SPI framework) that are also explained below:

1. *Feasibility*: The first step is to determine the viability of the organisation in adopting the framework. Several aspects are assessed, including availability of necessary resources such as time, commitment of top managers and SPI infrastructure.
2. *Business and SPI goals*: The next step is to define business objectives and their associated key performance indicators (KPIs). These aims are defined using the SMART criteria [20], meaning objectives must be specific, measurable, achievable, relevant and time-bound. Finally, SPI objectives are defined cascading them from the ones defined for business; for these objectives, metrics must be defined.
3. Activities and behaviours to enhance: In this step, SPI activities and practitioners' behaviours to be enhanced are identified. The level of intrinsic motivation of the SPI activity is
4. analysed by means of the intrinsic motivation inventory (IMI) test [21]. Finally, practitioners' behaviours and attitudes to be enhanced are identified and analysed.
5. *Player definition and motivations*: In this step, the motivational factors [22, 23] for each of the software professionals groups or SPI roles are analysed. It is also desirable to identify each group of professionals or SPI roles with some type of players' classification (for instance by means of the taxonomy presented in [24]) for the gamification proposal.
6. *Gamification elements*: For each SPI activity to be improved, it is necessary to sketch first the dynamics, mechanics and game elements defined as described in [7] and refined by means of the Octalysis framework [25] as explained in previous works [17] and second, metrics for each game element previously defined. Then, the feedback process system has to be defined in order to capture and store information on the practitioner's activity in real time. Aspects such as resistance to change and commitment must be observed and recorded.
7. *Implementation*: In this step, the gamification proposal defined in phase 5 is executed and implemented. Nevertheless, before implementing the gamification proposal, all stakeholders must be aware of the gamification project. The aim of this communication is to guarantee all parties understand and adopt the process as a critical aspect in SPI initiatives [26]. In order to support the process in an easy and affordable way, a SaaS open source tool (called Gamiware) was developed. Although initial results on its implementation show remarkable success, there are some limitations due to the lack of integration in the software engineers' workflow [14].
8. *Measurement*: The KPIs of the SPI initiative, motivation metrics and the defined game elements are collected, measured and analysed.
9. *Learning*: Results are assessed and the main inferences for future iterations are analysed and documented.

Refine loop: In order to introduce an improvement aspect, taking into account the information obtained in the process defined above, the necessary adjustments are made in stages 2–5 in order to improve process and results.

To carry out this process, apart from the different SPI roles involved in the initiative, it is necessary to ensure the commitment of key members of the organisation's top management. Without this commitment, it is not feasible to carry out the SPI initiative [6, 27–30], also in small and medium enterprise (SME) organisations [31].

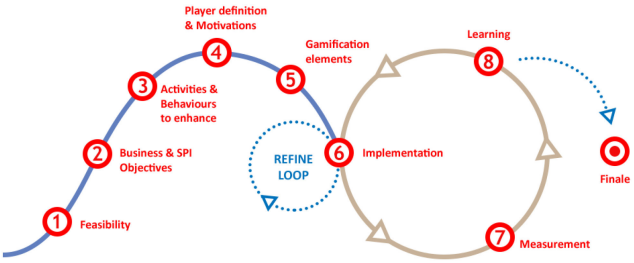

Fig. 1 *Phases of the G-SPI framework*

In addition, it is appropriate to seek the advice of a gamification expert to prepare the gamified proposal.

Fig. 2 shows the artefacts and desired results during the phases considered for the G-SPI framework.

3 Experimental validation

3.1 Experimental objectives

The main goals stated for the experimental validation to determine the G-SPI framework capability in order to increase motivation among software workers and to enhance SPI results in small and medium software development companies that have tried ineffective SPI initiatives previously.

According to these objectives, the specific hypotheses addressed for this experimental validation were:

(a) The effective implementation of the G-SPI framework contributes to increasing staff motivation in SPI initiatives implemented in a software development company.

(b) The effective implementation of the G-SPI framework contributes to increasing the software process performance in a software development company.

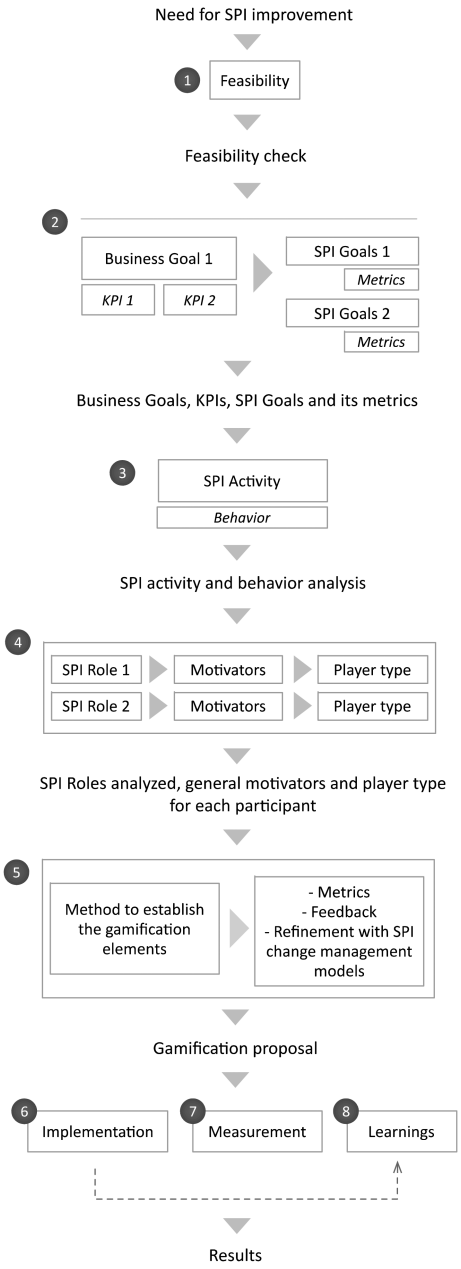

Fig. 2 *Phases traceability summary of the G-SPI framework*

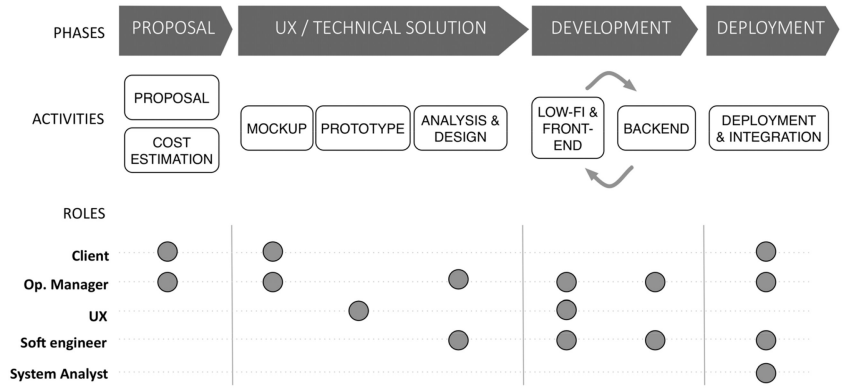

Fig. 3 *SOLUSOFT software process*

Table 1 Business and SPI goals

Business objectives	SPI objectives
BO1: improve the personnel capability to identify the tasks to complete during a week.	*SPIO1*: improve the ability to identify task focused on developing specific artefacts.
BO2: improve the personnel capability to estimate effort required for the task.	*SPIO2*: define heuristics or criteria to facilitate the task effort estimation in hours.
	SPIO3: define a data repository to store historical data on effort required to develop different types of activities in a software project.

3.2 Experimental context

SOLUSOFT is a small software development company (http://www.solusoft.es/) that provides different types of information systems: mobile apps, geographic information systems, enterprise information systems, integration of information systems, web content management systems and software maintenance. These projects are developed using Microsoft, Java (including Android) and Apple IOS. Normally, these projects have a 4-month duration on average.

At the time of the experiment execution, SOLUSOFT had 45 employees and a yearly turnover of 1.7M Euros. Based on these figures and the EU definition for small and medium companies, SOLUSOFT can be considered a small company.

The global software process used in SOLUSOFT is summarised in Fig. 3.

SOLUSOFT had started two previous SPI initiatives consisted of obtaining ISO/IEC 90003:2014 certification in 2003 and, in 2016, another SPI initiative to improve software testing processes. None of them finished due to a lack of availability and motivation of the personnel in charge of implementing the improvements designed. Finally, in 2017, SOLUSOFT started a new SPI initiative to improve project estimation and planning processes that was implemented through the application of the G-SPI framework.

The business goals stated for this SPI initiative were:

1. Improve the personnel capability to identify the tasks to complete over a week. The key issue to address is the accurate identification of tasks to be included in the project plan and its relation to actual artefacts to create during the development process. The KPI assigned to this goal is:

   KPI 1: Ratio of tasks that do not need to be done over a week.
2. Improve the personnel capability to estimate effort required for the task. The key issue to address is the accurate estimation of effort required to develop the projects. The KPIs assigned to this goal are:

   KPI 2: Ratio of tasks that fulfil its estimation
   KPI 3: Average difference between estimated and actual hours spent on a project
   KPI 4: Ratio of billing deadlines properly achieved.

The relationship between business and SPI goals are summarised in Table 1.

The activities affected by the previously mentioned SPI objectives were 'Project Planning' and 'Estimation tasks and follow-up'. The first activity was broken down into these three tasks:

Task 1: Project plan definition at highest level of detail considering the following factors: level of detail, sequentiality and dependency.
Task 2: Milestones identification related to billing milestones.
Task 3: Monitor progress of project plan.

The second activity (estimation tasks and follow-up) was broken down into the following tasks:

Task 1: Identify atomic weekly tasks.
Task 2: Estimate tasks effort.
Task 3: Register actual effort to complete a task.

The specific improvements to introduce in these activities consisted of (i) encouraging specific tasks identification and the recording of unplanned tasks at the end of a week and (ii) improving accuracy in estimating the effort of planned tasks.

3.3 Participants background

Six software engineers participated in the gamification initiative as well as the SOLUSOFT operations manager, who only acted as referee assigning gamification points to the engineers based on the SPI attributes of the planning and estimation artefacts done by the participants.

Staff taking part in the gamification companies has 1–8 years' experience as software developers in design, implementation and testing processes, but very little experience in planning and follow-up activities such as selected tasks for process improvement through gamification initiatives.

Before defining the specific gamification proposal, the participants were characterised by means of a Bartle Test [24]. The results obtained, divided by participant's teams are explained in Section 3.4, are presented in Table 2.

As stated in phase 4 of the G-SPI framework, in addition to these profiles, it is important to bear in mind that the main motivators of the gamification participants are participating in bottom-up initiatives, obtaining continuous feedback and avoiding bureaucracy [22]. It is also necessary to indicate that the main demotivators are associated with reduced creativity, lack of commitment from top management, excessive workload, time

Table 2 Software engineers (SE) participants profile

Team	Participant	Killer, %	Socialiser, %	Explorer, %	Achiever, %
1	SE1	40	40	80	40
	SE 2	27	60	73	40
2	SE 3	67	60	20	53
	SE 4	0	67	80	53
3	SE 5	53	47	60	40
	SE 6	47	40	80	33
	Mean	39	52.3	65.6	43.2

Table 3 Rewarding scheme (per person per week)

Block	Action	Points
project planning (35% of total rewarding – total points: 48)	1.1 (first week only) appropriate tasks identification	12
	1.2 (first week only) appropriate milestones identification	12
	1.3 (every week from the second week) project plan is appropriately updated to include information regarding current state of activities in progress – 8 points per week	24
project estimation and follow up (65% of total rewarding – total points: 88)	2.1 (every week from the first week) weekly tasks identified with the appropriate level of detail – 10 points per week	40
	2.2 (from first week, each week) effort estimation appropriately completed – 10 points per week	40
	2.3 (from first week, each week) registration of tasks actual effort – 2 points per week	8
	total points	136

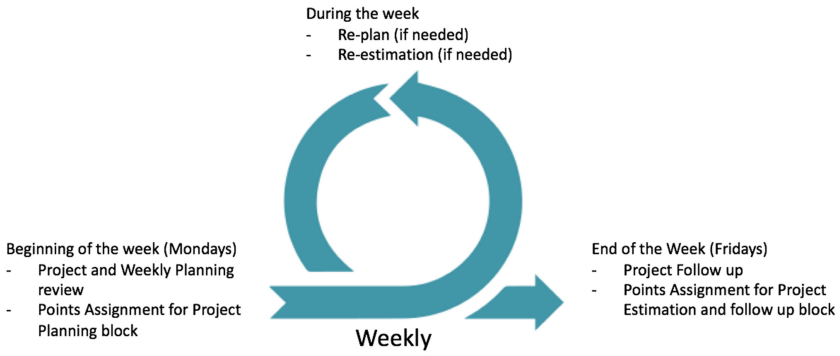

Fig. 4 *Implementation of gamification initiative*

pressures and restrictions, lack of support, cumbersome processes and inadequate communication [23].

The objectives of the projects developed by each team were:

1. The project developed by first team consisted of developing a web application using .NET framework to manage the operations central Madrid market.
2. The project assigned to team 2 had as main goal to develop a web application to manage the commercial operations using .NET framework of a new housing marketer company.
3. Finally, project by team 3 consisted of developing a social network for the users of Madrid Public Buses company to enhance the communication among the users and the company. This application is developed in iOS and Android.

3.4 Experiment definition

The experiment carried out consisted of a gamification approach using the G-SPI gamification framework to improve the task planning and estimation activities performed in three SOLUSOFT software projects.

In accordance with the gamification participants' profiles, and rolling out phase 5 of the G-SPI framework (see Section 2), the best gamification approach was exploratory. Nevertheless, due to the nature of the software development activities to improve (related to planning and estimation) and the high complexity of creating or developing an exploratory game, this approach was discarded by the operations manager. As a result, the researchers decided to create a gamification approach based on a competition, in which the participants were organised into teams (the second best approach, see Table 2 above), that is aligned with the socialiser profiles represented in the experiment participants profile.

During the implementation of the gamification initiative, the game was played weekly in the following way: at the beginning of the week (usually on Mondays) the operations manager revised the project plan upgrades and the weekly plan prepared by the software engineers, assigning the points corresponding to project planning. The actions of the rewarding scheme done at this moment are 1.1, 1.2, 1.3 and 1.4 (see Table 3). During the week, the software engineers involved in the gamification initiative upgraded the tasks information, identifying new tasks if needed, re-estimating the effort assigned to tasks and registering the actual effort spent on the already planned tasks. The action of the rewarding scheme done at this moment is 2.2 (see Table 3). Finally, this information was revised at the end of the week (usually on Fridays) by the operations manager who assigned the points corresponding to the project follow-up block. The action of the rewarding scheme done at this moment is 2.3 (see Table 3).

This sequence of tasks is presented in Fig. 4.

During the implementation of the gamification initiative the project planning, estimation and follow-up information were registered using a spreadsheet in Google Drive.

3.5 Data gathering and analysis

According to the goals stated for the validation of the gamification framework, the data to gather are:

1. Increase staff motivation in the SPI initiatives started. The researchers applied the IMI [21] before and after the implementation of the gamification initiative.

   The IMI is a multidimensional measurement device intended to assess participants' subjective experience related to a target activity in experiments. It has been used in several experiments in intrinsic motivation and self-regulation [21]. The instrument assesses participants' interest/enjoyment, perceived competence, effort, value/usefulness, pressure felt and tension, and perceived choice while performing a given activity, thus yielding six subscale scores.

   This provides relevant information on the increase in staff engagement in the project planning, estimation and follow-up that is under the scope of this gamification initiative.
2. The increase in software process performance. Two types of data were considered: gamification points, following the scheme presented in Table 3. According to the level of improvement, six levels of expertise were defined using a points-based logarithmic scale: Level 1 (Newbie): 0 points; Level 2 (Beginner): 20 points; Level 3 (Intermediate): 40 points; Level 4 (Advanced): 65 points; Level 5 (Professional): 90 points; and Level 6 (Expert): 125 point.

Metrics for the specific SPI goals assigned to the gamification initiative are shown in Table 4.

3.6 Threats to validity

In order to analyse the validity of the results obtained in this experimental work, several threats were considered prior to its execution. On the one hand, internal validity is the extent to which a causal conclusion based on the experiments defined is appropriate, preventing the introduction of systematic errors in the data used to determine the results and conclusions. On the other hand, external validity is the extent to which the results of a study can be generalised to other situations.

3.6.1 Internal validity: In this research, the factors contributing to internal validity are the personnel involved in the gamification initiative, the experiment implementation and the data used to analyse the results.

SOLUSOFT is a small software development company with the same characteristics of SMEs in this sector that experience some difficulty in implementing SPI to manage the software process effectively.

The type of personnel involved is representative of the professionals working in the software industry in SMEs, including junior and senior software engineers working with different technologies in the same organisation. What is more, engineers having different personality profiles were involved.

The data required for evaluating the increase staff motivation in SPI initiatives were IMI questionnaire and supporting interviews. The use of this IMI questionnaire is supported by pre-existing gamification research [14]. Also, complementary interviews mechanisms were used to obtain qualitative information in order to enrich the IMI quantitative data obtained [32].

The results were analysed qualitatively to identify the benefits and limitations of using gamification frameworks to introduce an SPI. The quantitative information was used to determine the degree of motivation and improvement effectiveness in this organisation. At this scale, the quantitative evidence obtained is representative of the organisational level. Moreover, the results obtained were analysed from different complementary perspectives such as competence acquisition using points obtained and specific process improvements, considering different related metrics.

3.6.2 External validity: In this research, the factors contributing to external validity are related to the type of gamification initiative selected and the type of improvements to introduce in the organisation.

The improvement of planning, estimation and follow-up activities in software projects is a managerial activity. Therefore, the conclusions obtained from this research work may not be valid when adopting gamification to improve more creative practices such as software design in software engineering.

The gamification approach selected was a team competition, so the validity of the conclusion is focused on this type of games. There are other gamification approaches such as exploratory game mechanics to introduce the types of improvements considered in the experimental validation. Nevertheless, the organisational difficulties in exploratory game mechanics prevent the use of this kind of approach in small or medium software development units.

The experiment implementation consisted of the application of the same gamification initiative to introduce several practices to software teams that had not been assigned in the past. To analyse several cases, three different teams implemented the initiative at the same time. Several limitations regarding the number of participants, number and type of improvements introduced were envisaged. In order to address these issues, the empirical study followed an orthogonal fractional design, where the participants were observed during the gamification implementation over a period of six weeks.

The authors followed an orthogonal fractional design, where each of the six participants where organised in three teams working in three different software projects having different requirements and developed in different technologies. Those teams applied the improvements designed adapting them to the specific software projects.

3.6.3 Conclusions regarding validity: The conclusion of the experimental approach validity can be discussed through its credibility, transferability and confirmability [33]. It can be stated that:

a. The use of triangulation research techniques (metrics gathering, questionnaires, interviews and observation) contributes to assure the credibility of the results obtained.
b. The results stated can be transfer to similar small and medium software development units because the organisation's structure, personnel involved and projects developed are similar to the most part of software companies having analogous characteristics.
c. To assure the confirmability of the results obtained, the artefacts and metrics produced by each team were accessible to the research team, the observations were registered in notes shared among the researchers and the communications were done through slack providing audit capabilities. Moreover, the conclusions were proposed by each researcher. These conclusions were agreed during several experimental validation meetings.

Table 4 SPI objectives with their metrics

SPI objective	Metric
improve the ability to accurately identify tasks on developing specific artefacts.	ratio of project activities correctly identified without needing to be broken down.
define heuristics or criteria to facilitate the task effort estimation in hours.	ratio of project activities having a variation against the estimation time higher that 10%.
define a data repository to store historical data on effort required to develop different types of activities in a software project.	ratio of project activities of actual and estimated time effort.

Table 5 Participants' intrinsic motivation profile evolution

Subscales	Participant 1		Participant 2		Participant 3		Participant 4		Participant 5		Participant 6	
	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After
interest/enjoyment	3.4	3.6	5	5	3.8	3.8	4	4.2	3.4	2.2	4.6	3.2
perceived competence	4.2	4	3.4	4.2	2.6	2.6	2.4	2.4	4.2	3.4	3.8	3.4
perceived choice	3	3	3.7	3.7	3.7	3.7	3.3	3	3	2.3	5	4
pressure/tension	1.4	1.6	2.6	2.6	1	1.4	2.2	2.4	2.6	2.2	1	1.8
value/usefulness	5	5	5	5	1.7	4.4	4.3	4.3	5	3	4.3	3.7

Table 6 Participants' competence evolution

	Week 1		Week 2		Week 3		Week 4	
	Points	Level	Points	Level	Points	Level	Points	Level
expected	46	intermediate	76	advanced	106	professional	136	expert
Team 1	43	intermediate	63	intermediate	85	advanced	103	professional
Team 2	44	intermediate	66	advanced	81	advanced	100	professional
Team 3	44	intermediate	71	advanced	93	professional	113	professional

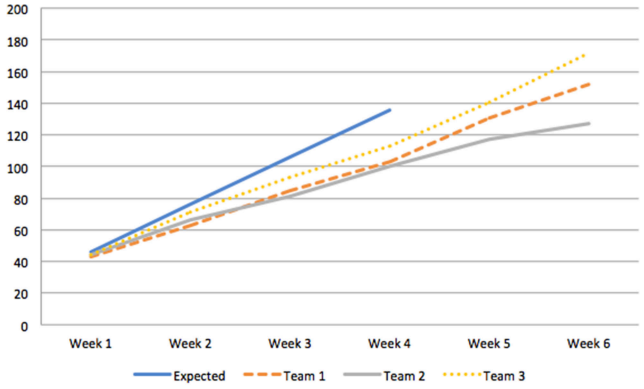

Fig. 5 *Participants' competence evolution*

4 Results

This section presents the results obtained in the experimental validation based on the objectives previously stated.

4.1 Increase staff motivation in SPI initiatives

Table 5 compares the participants' intrinsic motivation and self-regulation before and after implementing the gamification initiative. Intrinsic motivation is analysed according to the following subscales: interest/enjoyment, perceived competence, perceived choice, pressure/tension and value/usefulness. These subscales are considered in a 22-item questionnaire [21]. The values presented for each category contain the mean of the answers obtained for the items considered in each subscale.

The subscales are measured on a Likert scale of 1–5, indicating more presence of the subscale in a participant's intrinsic motivation profile.

The questionnaire was complemented with open questions based on the feedback obtained from the referee and the problems found in the gamification implementation.

In general terms, the gamification initiative implemented did not contribute significantly to increasing the participants' intrinsic motivation. Considering the data presented in Table 5, the researchers found that in several cases:

- Due to the nature of the software process being gamified (project planning, estimation and follow-up), several participants enjoyed it less than expected, possibly because the activity under gamification is not so creative.
- Participants realised they had less competency in performing project planning, estimation and follow-up than expected at the beginning of the gamification initiative. Nevertheless, others considered having less initial competence increased their motivation due to the increase in the perceived competence.
- In several cases, the pressure to perform their tasks properly increased. This increase was due to the introduction of new management responsibilities for the less experienced participants or to the continuous scoring of the job in this area.
- In the open questions, personnel whose motivation did not increase reported several difficulties, delays and subjective perception problems in the points assignment and feedback activities.

The factors influencing the evolution of the participants' intrinsic motivation profile are analysed in more detail in Section 5.

4.2 Increase the software process performance

Table 6 presents the expected and actual evolution of the participants' competence regarding project planning, estimation and follow-up.

The first row corresponds to the expected points to be achieved by the teams if they acquire the gamified capabilities as planned. Moreover, it indicates the expected path to advance through the levels of experience established for the game. The following rows correspond to the teams' actual evolution bearing in mind the points obtained during the weeks planned for the gamification initiative. As a result, the teams increased continuously their skills in project planning, estimation and follow-up but slower than expected.

As the teams did not achieve the expected competence in the weeks planned for the gamification initiative, the operations manager and research team agreed to continue the gamification initiative for two more weeks. The actual competence evolution over 6 weeks is presented in Fig. 5. At the end of the fifth week, team 3 reached the expert level. Team 1 achieved this level at the end of the sixth week. It is also remarkable that team 1 was unable to achieve the expected competence level.

Table 7 Ratio of activities correctly identified without needing to be broken down

Team	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
Team 1	0.00	0.00	0.00	0.00	0.00	0.00
Team 2	0.00	0.08	0.17	0.20	0.10	0.11
Team 3	0.58	0.17	0.33	0.18	0.00	0.17

Table 8 Ratio of activities having a variation against the estimation higher that 10%

Team	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
Team 1	0.67	0.50	0.57	0.62	0.78	0.33
Team 2	0.33	0.69	0.45	0.20	0.53	0.52
Team 3	0.33	0.33	0.50	0.31	0.29	0.40

Table 9 Ratio of activities' actual and estimated effort

Team	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
Team 1	0.24	0.58	0.35	0.21	0.25	0.20
Team 2	0.23	0.43	0.18	0.06	0.50	0.50
Team 3	0.30	0.23	0.31	0.33	0.18	0.22

This difference between the actual and the expected evolution was due to several issues such as frequency and nature of the feedback provided by the operations manager and the poor communication with the operations manager to solve relevant questions to achieve the skills and competences with respect to project planning, estimation and follow up. Additionally, the points assignment included several subjective characteristics that could have influenced the actual evolution registered. These issues are discussed in Section 5.

Regarding the first SPI objective (improve the ability to identify task focused on developing specific artefacts), the ratio of activities correctly identified without needing to be broken down was analysed (see Table 7).

It is important to mention that team 3 achieved the best performance, planning properly the activities at the beginning of the week and upgrading the plan if a task that was not considered in the initial plan had to be added.

Team 1 did not upgrade the plan during the gamification initiative. In the evaluation phase, it was found that the re-planning effort was not necessary in several of the weeks, but in other weeks such as 1 and 5 a re-plan that was not done was needed.

Regarding team 2, it should be noted that one of the members (participant 5) did not complete plans, estimations and follow-up registers from the second week of the gamification initiative because his motivation decreased during the gamification initiative. However, this team 2 participant included all the non-planned tasks in a single task called 'Non-planned tasks'. This decision did not enhance the planning process because it did not help the software engineer to identify the types of tasks missing in previous plans that should have been included in the following ones.

Regarding the second SPI objective (define heuristics or criteria to facilitate the task effort estimation in hours), the ratio of activities having a variation against the estimation higher that 10% was analysed (see Table 8).

It is important to note that in the case of team 3 there was a slight performance improvement. Regarding team 3, there were two exceptional weeks (3 and 6) when the SPI indicator did not attain a good value. The other teams did not have a noticeable improvement.

This was because several tasks with an expected effort of 1 h had an actual effort of 1.5. In percentage, the actual variation was considered 50% (and included in this ratio), but its variation in effective effort was small in relation to the total project effort. This situation can be analysed better considering a ratio of overall activities' actual and estimated effort. This metric is presented in Table 8.

Regarding the third SPI objective (define a data repository to store historical data on effort required to develop different types of activities in a software project), the ratio of activities' actual and estimated effort was analysed (see Table 9).

It is necessary to indicate that the performance achieved in the overall precision of the estimations provided was good, improving in the case of teams 1 and 3 as their software engineers used knowledge registered from similar activities previously carried out. In the case of team 2, there was no clear improvement trend. In this case, the project context underwent many delays and pressure, thus preventing the participants from dedicating the appropriate time to perform estimation activities properly.

Finally, it is important to indicate that the gamification points assigned to each team corresponded to the overall team performance, but there is no clear relationship between the SPI metrics and the gamification points trend. This is due to a bias introduced in the point assignment done each week.

5 Discussion

The discussion is based on the three stages of the experiment: before, during and after.

5.1 Gamification initiative preparation

During the steps before implementing the experiment, a number of aspects should be highlighted. On the one hand, when putting forwarding the gamification proposal (see Section 2) and based on the results of the user analysis (see Section 2), it was suggested that a majority of the participants be Explorers in the exploratory game mechanics. However, in a first presentation of the proposal to SOLUSOFT, the operations manager showed the complexity of applying this proposal on a day-to-day basis. Consequently, to ensure the feasibility of the experiment, it was decided to simplify the gamification proposal as much as possible, while maintaining its effectiveness. To this end, we opted for a gamification proposal that combined competitive and social game mechanics as a suboptimal alternative aligned with the Bartle profiles of the participants (see Table 2). This selection of the suboptimal proposal could imply a supposed decrease in the performance of the initiative. Regarding the mechanics and components of gamification used, these were basically from the PBL type [8, 34, 35], whose effectiveness has been confirmed in the short-medium term, but ineffective in the long term [7, 16, 25]. Finally, the use of Gamiware was valued as a support tool [14], but it was ruled out by adding a new barrier in the launch phase of the initiative.

Another aspect to be highlighted was the decision to include exclusively intrinsic motivators in the gamification proposal, avoiding, therefore, any type of extrinsic motivator. This design decision is justified on the basis of the Overjustification effect [36], which shows that an unbalanced introduction of extrinsic motivators can cancel any type of intrinsic motivation. However, well balanced these motivators are, the result can be more positive. Based on this effect and erring on the side of caution, it was

decided to minimise the risk of cancelling intrinsic motivation, assuming in return possible mitigation in the effectiveness of the proposed gamification [17, 34].

Finally, it is important to note that all the organisations invited to participate in the experiment were very receptive at the beginning. In the first steps, top managers showed their commitment to implementing the proposal. This commitment is critical, both in SPI [6, 26, 28, 37, 38] and gamification [13] initiatives. However, despite this, even when they had committed themselves, it was difficult to start the initiative because there was no real mobilisation of the necessary resources.

5.2 Gamification initiative implementation

Throughout the experiment, within the 5–8 phases loop of the framework (see Section 2) new limitations to the implementation and recommendations of the learning phase were identified based on measurements and observations.

First of all, it should be noted that the frequency of participants' feedback was very low. Multiple investigations indicate that, in order to obtain adequate performance in a gamification initiative, real-time feedback is necessary [7, 16, 17, 25]. However, given the impossibility of a daily manual points allocation by the organisation, a feedback of 3 days per week was agreed with the top manager. Although this frequency in itself limits the effectiveness of the proposal, the periodicity established in the feedback system was not respected, being reduced to a 1.3 feedback/week throughout the experiment. This fact affected even more the break of the motivation loop (action – feedback – motivation), which had been considered very important to maintain motivation in gamification activities [8]. For this reason, in successive implementations, it is considered critical to integrate a feedback at least daily and, if possible, automated in real time, integrated into the workflow of the participants.

Secondly, communication with the top manager was not fluid, so that all aspects of implementation included in the 'Refine Loop' (see Section 2) could not be executed as often as necessary. From this fact it can be deduced that this initiative in the organisation is not a priority. In relation to this, the lack of sense of urgency is one of the factors indicated in the gamification framework [17] and this has been pointed out by renowned experts in gamification as one of the main factors to be taken into account in any change management [37, 39, 40].

In addition, the experiment had to be restarted once. Therefore, it is necessary to highlight the sensitivity of the framework to changes in the organisation and in the teams involved. Since the gamification framework establishes a gamification proposal oriented to the profiles of the participants, any rotation of personnel can affect the development of the proposal. It must be taken into account that, if any member is replaced during the experiment, it will be necessary to perform the Bartle [24] and IMI tests [21] again as well as to review the gamification proposal to see how it affects the inclusion of the new Bartle profile.

Finally, it was detected, and later confirmed with the top manager, that the allocation of points was too subjective. Despite having agreed on the incentive scheme with the organisation, during the development of the experiment it was found that, at the time of scoring, additional factors that were not visible or previously contemplated were being taken into account. For example, if one of the participants had a heavy workload and did not record the evolution of their participation in the experiment, some points were added as a consolation. To avoid these distortions and biases, it is necessary to establish a clear objective points assignment [7, 16, 25, 38, 41], with fully objective criteria and, if possible, automated through an automatic mechanism [14] where the top manager does not have to assign points.

5.3 Gamification initiative evaluation

The discussion of what happened after the end of the experiment is divided into two blocks. Each one of the blocks corresponds to the main objectives of the research, that is the increase in personnel motivation of the SPI activities and the increase in the software project effectiveness.

5.3.1 Increase in staff motivation with SPI activities: Therefore, regarding the increase in motivation, it should be noted that the gamification proposal has not contributed to increasing the intrinsic motivation of the participants. Although the participants consider the gamification initiative 'very useful' (4.33 out of 5), the subscales of the IMI tests carried out (see Table 5) that are directly related to intrinsic motivation have different results and this prevents confirmation of the hypothesis at this point.

This non-increase in motivation can be due to multiple reasons. The first is the multitude of limitations that were identified. Regarding the motivation results, it is likely that the introduction of competitive game mechanics introduced tension among the participants and this fact may decrease the motivation and fun, as indicated by previous studies, although this is mitigated by the use of collaborative mechanics [42]. Additionally, in relation to the SPI task, several aspects that are worth emphasising were detected. On the one hand, the level of usefulness of the feedback received from the participants has been considered, on average, 'indifferent' (3.5 out of 5). On the other hand, the task of tracking software projects is, intrinsically, a boring task and the results of IMI tests show that the gamification proposal was unable to make it more fun. The difficulty in making a boring task fun has been pointed out in some studies [7, 25, 35, 43], and precisely the subscale of interest/fun in the IMI test [21] is one of the largest factors correlated with intrinsic motivation. Also, it seems that this task of tracking software projects was perceived by the participants as a control of their professional work rather than part of an SPI initiative to achieve other objectives beyond their control.

Finally, regarding the novelty of the SPI tasks, it should be noted that the organisation had selected SPI tasks that were not usually carried out among the participants. In later interviews with the operations manager, it was confirmed that the participants in the experiment had not previously carried out the follow-up software projects tasks. Moreover, two of the participants had never done such tasks in their professional lives. This implies greater resistance to change [28, 39, 40, 44, 45]. Therefore, to minimise the resistance to change, it seems more convenient to apply the gamification proposal on tasks that have already been performed and whose performance need to be improved.

5.3.2 Increase in the efficiency of the software project: Secondly, the goal of increasing the effectiveness of the process has been more than achieved. The interview with the top manager confirmed that the performance had improved greatly. This improvement was perceived in the increase in autonomy on the part of the participants and a better order when considering the tasks. In addition, a very important objective was achieved: the visibility of the tasks of each project, its breakdown and progress. Before the experiment, there was no visibility of these tasks and there were many business decisions that could not be supported by data.

However, after the implementation of the experiment, the performance of follow-up tasks did not last over time. Within a month of finishing the gamification proposal, only 60% of the participants were still doing the tasks while, a month later, none of the participants continued with them. This fact was foreseeable given that several critical factors converge. On the one hand, not only have the limitations mentioned been maintained, but the lack of feedback has been accentuated and the obligation to participate in the research initiative has disappeared. The participants did not continue on their own in the development of tasks, especially if intrinsic motivation had not increased. Besides, for a gamification initiative to last over time, it is necessary to modify periodically the proposed game dynamics [8, 34, 35].

Additionally, it is worth mentioning that the results indicate the perception of the participants' competence in the IMI tests (see Table 5) had reduced slightly. This result implies that, after the experiment, the participants considered that they were less competent than they initially were. However, although it may seem paradoxical a priori, it is a reasonable result since it was confirmed that previously they had not carried out the follow-up task. Therefore, their initial assessments were based on estimates on

something they had not done, and the gamification initiative seems to have increased awareness of their real competence in this SPI task.

Finally, the collected SPI metrics indicate that billing milestones had increased by 10% compared to the rest of the projects of the organisation that had not been conducted under the gamification initiative. In reviewing the data, there seems to be a correlation with the evolution of the three teams (each team with a different project) within the gamification initiative and the subsequent fulfilment of the milestones of the project. Thus, the project of the winning team and the one in the second position were invoiced satisfactorily. However, the project of the losing team was not billed. Therefore, it can be affirmed that 66% of the projects reached their objective.

6 Conclusions

The implementation of the gamification framework within the organisation can be considered successful, but only partially and it includes a set of relevant limitations. Although the participants have mostly considered the gamification initiative 'very useful' (4.33 out of 5), the research objectives have not been fully met. Regarding the first experimental objective (*increase in staff motivation in the SPI initiatives started*), a significant increase in intrinsic motivation has not been observed among the participants in the experiment. On the other hand, regarding the second research objective (*increase the software process performance*), some improvements in planning, estimating and follow-up tasks have been observed, although it should be noted that the positive effect of the improvements has not lasted significantly over time.

Nevertheless, it is necessary to highlight that researchers found several limitations in the introduction of the gamification. These limitations were:

a. Regarding the first experimental objective (increase in staff motivation in the SPI initiatives started), a significant increase in intrinsic motivation has not been observed among the participants in the experiment.
b. Regarding the second research objective (increase the software process performance), it should be noted that the positive effect of the improvements has not lasted significantly over time. The projects developed by the company remained similar in scope and technology; however, during the months following the gamification initiative, the availability of resources decreased and the pressure to achieve billing objectives increased. According to the opinions of the participants, these were the reasons why the improvements were not lasted as expected.
c. In several cases, the initial expectations of the gamification participants were not fully achieved. This was due to the type of game implemented (teams competition) was the second best, because the best alternative (exploratory game) was not feasible for the considered company due to the effort required to implement this alternative.

For future implementations, the gamification framework should be launched, trying to avoid the set of limitations found. In addition, new feasibility factors must be included in phase 1 of the gamification framework. These factors correspond to the need to size the effort associated with the gamification proposal and analyse the feasibility of its implementation according to the availability of the team to be involved. Also, it is necessary to transmit more effectively the sense of urgency in the introduction of the improvements required when the gamification initiative is launched.

Due to the restrictions in the use of human resources available in software engineering organisations, the systematic deployment of this type of gamification proposals is questionable, especially in SMEs. These issues are greater in small and medium software development companies where managers and software engineers have to deal with deadlines and occasional work overload.

In future research lines, new empirical evaluations should be carried out considering the limitations and critical recommendations resulting from this research. Likewise, it is also necessary to develop more empirical evaluations of the use of gamification for improving more creative software engineering processes such as design. It is also necessary to consider other non-competitive and more exploratory gamification alternatives. As with SPI initiatives [46–49], gamification implementation seems to be more feasible in large organisations, although this must be supported by new empirical research.

7 References

[1] Shih, C.-C., Huang, S.-J.: 'Exploring the relationship between organizational culture and software process improvement deployment', *Inf. Manage.*, 2010, **47**, (5–6), pp. 271–281
[2] Mathiassen, L., Ngwenyama, O.K., Aaen, I.: 'Managing change in software process improvement', *IEEE Softw.*, 2005, **22**, (6), pp. 84–91
[3] Colomo-Palacios, R., Soto-Acosta, P., García-Peñalvo, F.J., *et al.*: 'A study of the impact of global software development in packaged software release planning', *J. Univers. Comput. Sci.*, 2012, **18**, (19), pp. 2646–2668
[4] O'Connor, R., Basri, S.: 'The effect of team dynamics on software development process improvement', *Int. J. Hum. Cap. Inf. Technol. Prof.*, 2012, **3**, (3), pp. 13–26
[5] Dorling, A., McCaffery, F.: 'The gamification of SPICE', *Softw. Process Improv. Capab. Determ.*, 2012, **290**, pp. 295–301
[6] Niazi, M.: 'A comparative study of software process improvement implementation success factors', *J. Softw. Evol. Process*, 2015, **27**, (9), pp. 700–722
[7] Werbach, K., Hunter, D.: '*For the win: how game thinking can revolutionize your business*' (Wharton Digital Press, Philadelphia, 2012)
[8] Feng, Y., Ye, H.J., Yu, Y., *et al.*: 'Gamification artifacts and crowdsourcing participation: examining the mediating role of intrinsic motivations', *Comput. Human Behav.*, 2018, **81**, pp. 124–136
[9] Herranz, E., Colomo-Palacios, R., de Amescua Seco, A., *et al.*: 'Gamification as a disruptive factor in software process improvement initiatives', *J. Univers. Comput. Sci.*, 2014, **20**, (6), pp. 885–906
[10] Gómez-Álvarez, M.C., Gasca-Hurtado, G.P., Hincapié, J.A.: 'Gamification as strategy for software process improvement: a systematic mapping'. 2017 12th Iberian Conf. Information Systems and Technologies (CISTI), 2017, pp. 1–7
[11] Dalpiaz, F., Ali, R., Brinkkemper, S.: 'Special section on gamification and software engineering' *Inf. Softw. Technol.*, 2018, **95**, pp. 177–178
[12] Yilmaz, M., O'Connor, R.: 'A Scrumban integrated gamification approach to guide software process improvement: a Turkish case study', *Teh. Vjesn. – Tech. Gaz.*, 2016, **23**, (1), pp. 237–245
[13] Herranz, E., Colomo-Palacios, R., Amescua-Seco, A.: 'Towards a New approach to supporting Top managers in SPI organizational change management', *Procedia Technol.*, 2013, **9**, pp. 129–138
[14] Herranz, E., Colomo-Palacios, R., de Amescua Seco, A.: 'Gamiware: a gamification platform for software process improvement', in O'Connor, V.R., Umay Akkaya, M., Kemaneci, K., Yilmaz, M., Poth, A., Messnarz, R. (Eds.): '*Systems, software and services process improvement*' (Springer International Publishing, New York, NY, USA, 2015), pp. 127–139
[15] Ruiz, M., Trinidad, M., Calderón, A.: 'Gamification and functional prototyping to support motivation towards software process improvement'. Product-Focused Software Process Improvement, 2016, pp. 697–704
[16] Hamari, J., Koivisto, J., Sarsa, H.: 'Does gamification work? – a literature review of empirical studies on gamification'. 2014 47th Hawaii Int. Conf. System Sciences, 2014, pp. 3025–3034
[17] Herranz, E., Colomo–Palacios, R., de Amescua Seco, A., *et al.*: 'Towards a gamification framework for software process improvement initiatives: construction and validation', *J. Univers. Comput. Sci.*, no. Improving Social Aspects of the Software Development Process: Games, Gamification and Related Approaches, 2016, **22**, pp. 1509–1532
[18] Pedreira, O., García, F., Brisaboa, N., *et al.*: 'Gamification in software engineering – a systematic mapping', *Inf. Softw. Technol.*, 2015, **57**, pp. 157–168
[19] Kasurinen, J., Knutas, A.: 'Publication trends in gamification: a systematic mapping study', *Comput. Sci. Rev.*, 2018, **27**, pp. 33–44
[20] Doran, G.T.: 'There's a SMART way to write management's goals and objectives', *Manage. Rev.*, 1981, **70**, (11), pp. 35–36
[21] Ryan, R.M., Koestner, R., Deci, E.L.: 'Ego-involved persistence: when free-choice behavior is not intrinsically motivated', *Motiv. Emot.*, 1991, **15**, (3), pp. 185–205
[22] Baddoo, N., Hall, T.: 'Motivators of software process improvement: an analysis of practitioners' views', *J. Syst. Softw.*, 2002, **62**, (2), pp. 85–96
[23] Baddoo, N., Hall, T.: 'De-motivators for software process improvement: an analysis of practitioners' views', *J. Syst. Softw.*, 2003, **66**, (1), pp. 23–33
[24] Bartle, R.: 'Hearts, clubs, diamonds, spades: players who suit MUDs', *J. MUD Res.*, 1996, **1**, (1), pp. 1–25
[25] Chou, Y.K., Fuqua, J., Yuan, W.: '*Actionable gamification: beyond points, badges, and leaderboards*' (CreateSpace Independent Publishing Platform, 2015)
[26] Pries-Heje, J., Johansen, J.: Software Process Improvement eurospi.net, Alcala, Spain (2010). http://www.iscn.com/Images/SPI_Manifesto_A.1.2.2010.pdf, accessed January 2017
[27] Montoni, M., Rocha, A.R.: 'A methodology for identifying critical success factors that influence software process improvement initiatives: an application in the Brazilian software industry', in '*Software process improvement*' (Springer, Berlin, Heidelberg, 2007), pp. 175–186

[28] Narciso, H., Allison, I.: 'Overcoming structural resistance in SPI with change management'. 2014 9th Int. Conf. Quality of Information and Communications Technology, 2014, pp. 8–17
[29] Niazi, M.: 'Software process improvement: a road to success', in Münch, J., Vierimaa, M. (Eds.): '*Product-Focused Software Process Improvement*' (Springer, 2006), pp. 395–401
[30] Stelzer, D., Mellis, W.: 'Success factors of organizational change in software process improvement', *Softw. Process Improv. Pract.*, 1998, **4**, (4), pp. 227–250
[31] Sulayman, M., Urquhart, C., Mendes, E.*, et al.*: 'Software process improvement success factors for small and medium web companies: a qualitative study', *Inf. Softw. Technol.*, 2012, **54**, (5), pp. 479–500
[32] Easterbrook, S., Singer, J., Storey, M.-A.*, et al.*: 'Selecting empirical methods for software engineering research', in '*Guide to advanced empirical software engineering*' (Springer, London, UK, 2008), pp. 285–311
[33] Lincoln, Y.S., Lynham, S.A., Guba, E.G.: 'Paradigmatic controversies, contradictions, and emerging confluences, revisited', *Sage Handb. Qual. Res.*, 2011, **4**, pp. 97–128
[34] Lee, J.-C., Shiue, Y.-C., Chen, C.-Y.: 'Examining the impacts of organizational culture and top management support of knowledge sharing on the success of software process improvement', *Comput. Human Behav.*, 2016, **54**, pp. 462–474
[35] Zichermann, G., Linder, J.: '*The gamification revolution: how leaders leverage game mechanics to crush the competition*' (McGraw-Hill Education, New York, NY, USA, 2013)
[36] Tang, S.-H., Hall, V.C.: 'The overjustification effect: a meta-analysis', *Appl. Cogn. Psychol.*, 1995, **9**, (5), pp. 365–404
[37] Johansen, J., Pries-Heje, J.: 'Success with improvement – requires the right roles to be enacted – in symbiosis', *Softw. Process Improv. Pract.*, 2007, **12**, (6), pp. 529–539
[38] Hamari, J., Huotari, K., Tolvanen, J.: 'Gamification and economics 5', Gameful World Approaches, Issues, Appl., 2015, **139**, p. 139
[39] Ferreira, M.G., Wazlawick, R.S.: 'Software process improvement: a organizational change that need to be managed and motivated', *Management*, vol. **6**, p. 19, 2011
[40] Kotter, J.: 'Leading change: why transformation efforts fail', *IEEE Eng. Manag. Rev.*, 2009, **37**, (3), pp. 42–48
[41] Pink, D.H.: '*Drive: the surprising truth about what motivates us*' (Penguin Publishing Group, London, UK, 2011)
[42] Hanus, M.D., Fox, J.: 'Assessing the effects of gamification in the classroom: a longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance', *Comput. Educ.*, 2015, **80**, pp. 152–161
[43] Andonova, Y.: 'Loyalty 3.0: how big data and gamification are revolutionizing customer and employee engagement', *J. Mark. Anal.*, 2013, **1**, pp. 234–236
[44] Conradi, H., Fuggetta, A.: 'Improving software process improvement', *Software IEEE*, 2002, **19**, (4), pp. 92–99
[45] Allison, I., Narciso, H.: 'Managing change in SPI: a framework for addressing resistance', *Softw. Qual. Prof.*, 2015, **17**, (4), pp. 4–14
[46] Clarke, P., O'Connor, R.V.: 'The influence of SPI on business success in software SMEs: an empirical study', *J. Syst. Softw.*, 2012, **85**, (10), pp. 2356–2367
[47] Clarke, P., O'Connor, R.V.: 'An empirical examination of the extent of software process improvement in software SMEs', *J. Softw. Evol. Process*, 2013, **25**, (9), pp. 981–998
[48] Garcia, I., Pacheco, C., Cruz, D.*, et al.*: 'Implementing the modeling-based approach for supporting the software process assessment in SPI initiatives inside a small software company', in Lee, R. (Ed.), *Software Engineering Research, Management and Applications 2011'* (Springer, 2012), pp. 1–13
[49] Mishra, D., Mishra, A.: 'Software process improvement in SMEs: a comparative view', *Comput. Sci. Inf. Syst.*, 2009, **6**, (1), pp. 111–140