Using Gamification for improving Scrum adoption

Francisco Alexandre Borges Aurindo Barros
Instituto Superior Técnico,
Universidade de Lisboa, Portugal

ABSTRACT
Adopting an agile framework such as Scrum can help software development teams meet time, budget, and scope requirements. However, practitioners often lack enough motivation to implement agile practices. This paper explores how gamification can be used to make Scrum practices more fun and engaging. We analyze the impact of a gamification app integrated with Jira Software on two Scrum teams in different organizations. The results obtained suggest that the solution impacted positively both teams. The feedback received from the practitioners during the study was essential to improve the proposed app, for example to simplify the complex notifications approach. Some challenges, such as the Scrum Master's inflated score, still remain. In the future, the app should be tested with more diverse teams to evaluate the impact in real-world Scrum teams.

Author Keywords
Gamification; software development; software process; agile; Scrum; motivation.

INTRODUCTION
The Standish Group's CHAOS Report found that only 36% of the projects studied between 2013 and 2017 succeeded in meeting time, budget, and scope requirements [30]. Also, there is evidence that agile projects are twice more likely to succeed and one-third less likely to fail when compared against a traditional methodology. Agile teams are intended to produce working software during short iterations [2]. In Scrum, the most adopted of the agile frameworks [20], practitioners are organized in small teams, plan and track their work iteration ("Sprint") based on Scrum artefacts, implement user stories (translating the system's requirements), and communicate in Scrum events [27].

Despite the proved agile methods and techniques' benefits, like improved product quality and customer satisfaction, agile teams are still facing challenges, mostly related to human factors [5,12]. Such challenges, like improper communication, are partially explained by software development practitioners' lack of motivation to apply agile techniques in practice [5,7]. Further in this paper, the term "practitioner" will be used to refer to any of the roles a software engineer can have (like analysts or programmers).

Gamification is a recent but popular approach to make processes related to non-gaming contexts more fun, which can boost motivation [9]. Besides being successfully applied in fields like education [1] and health [15], there is evidence that gamification applied to the workplace can lead to significant improvements in employee engagement [24]. Particularly, researchers and practitioners have been applying gamification in software development to increase team motivation and help practitioners to define better goals and to focus on them [33]. Yet, very few gamification proposals were evaluated with Scrum teams in practice, despite the importance of validating these initiatives' impact on people's motivation.

This paper explores how gamification can be used to make Scrum practices more fun and engaging. In a previous work, we developed a gamification app integrated with Jira Software and analyzed its impact on a Scrum team at a Portuguese software development company. We have now studied the impact of gamification on a different Scrum team at a Portuguese small consultancy group both with the original solution and with a solution improved based on feedback received from both teams. In this paper we describe these results and compare the impact of gamification on the two teams. To the best of our knowledge, this study is one of a very few evaluating a gamification solution for Scrum with real teams.

We start the paper with a review of works implementing gamification in software development, including our previous work. Next we present the results of a task analysis with the studied Scrum team, followed by the method used. Then we present and discuss the results. The paper closes with a discussion of the results and some conclusions and future work.

RELATED WORK
Gamification uses game elements and game design to make tasks related to real-world problems and goals rewarding for themselves, thus creating incentives without incurring into high costs. Despite being related to gaming, gamified systems are not full-fledged; they just use parts of games (i.e., game elements) in an already existing process [8].

Because software development processes are brain- and collaborative-intensive, comprising some tedious activities, gamification can help making such activities more fun and attractive [22]. In the next subsection, we present and discuss relevant works related to this theme. Then, we discuss our previous work, where a gamification app was developed and evaluated with a Scrum team.

However, we must first describe concepts linked to motivation, which plays an important role in gamification initiatives. Motivation can be intrinsic (the activity is inherently interesting) or extrinsic (the activity leads to a separable outcome) [26]. The Self-Determination Theory (SDT) focuses on intrinsic motivation and is based on three

psychological needs: competence (sense of controlling the outcome and feeling mastery; related to the concept of flow, a state of mind when one is totally absorbed on an activity [6]), relatedness (sense of interacting and be connected with others), and autonomy (sense of being the source of one's own behavior) [25]. This theory also states that extrinsic motivators are effective if they promote these needs. Overall, mostly important is to understand the different motivators and how to foster them.

Gamification in Software Development

Gamification adds game elements and game design to non-game processes to engage and motivate people to adopt new behaviors [32].This approach aims at making activities related to real-world problems and goals rewarding for themselves, thus creating incentives without incurring into high costs. Despite being related to gaming, gamified systems are not full-fledged; they just use parts of games (i.e., game elements) in an already existing process [8].

Because software development processes are brain- and collaborative-intensive, comprising some tedious activities, gamification can help making such activities more fun and attractive [22]. Some research has already been conducted in this field, with different focus. Some authors tackle the main subject by discussing and proposing methodologies to apply gamification in software engineering [11,21] and to foster software process improvement initiatives [10,13,14]. Additionally, the authors of a literature review about the use of gamification in software development processes concluded that, despite the many gamified tools to support different activities, no tool supports the whole process [16]. Closely related, a framework mapping collaboration issues affecting software development teams with target behaviors and game elements to mitigate those issues was proposed [3]. Moreover, some works are targeted to specific software development processes, like the gamification of a version control system to encourage Computer Science students to commit more frequently [28] and a reputation system to improve the quality of collaboratively written code through documentation [23]. Finally, some studies directly address the application of gamification in agile development and Scrum. McClean added a lottery element to the agile process: practitioners could win a reward, and their chances increased with the number of tasks completed [19]. Yilmaz and O'Connor proposed an integrated gamification approach for Scrumban, where practitioners received points and badges for finishing tasks and helping each other [33]. Loriggio presented a methodology for teaching Scrum, supported by gamification and other theories [17]. Češka prototyped a gamified app to support Scrum development, composed by game elements like points, badges, and progression [4]. Scrum Hero is a gamification framework to support Scrum software development projects' management, based on game elements like narrative, quests, and rewards [29].

While only the last group of works directly address gamification for agile/Scrum, we can learn from them all, as they target the same players (i.e., practitioners). Overall, these works lack a proper empirical validation in the industry: some are just conceptual proposals; others are only evaluated with qualitative methods; and others are evaluated in alternative contexts (like education) using samples too small and time frames too short to support important conclusions. Furthermore, proposed solutions did not go far beyond the simplest elements (like points and badges) and are not integrated in the tools' practitioners use daily. All in all, while gamification studies to increase Scrum adoption are emerging, there is much room for growth, namely regarding diversity of game elements and empirical validations in industry.

Apart from research, some commercial tools like Jiraffe [34] exist, but there are no studies publicly available evaluating the application of these tools in industry.

Previous Research

In a previous work, we presented and evaluated a gamification solution aiming at making Scrum techniques more fun and engaging for practitioners [18]. This solution was designed using the 6D Framework, an iterative game design process composed by six steps.

In the first step, we defined the solution's objectives. The main goal was to increase practitioners' motivation to apply Scrum techniques, and from here we derived more concrete goals. In the second step, we selected behaviors that translated Scrum practices (i.e. those we want players to perform) and defined metrics for measuring them in sprint and practitioner contexts. Next, in the third step we describe what motivates target players, including factors like communication, contribute to the project's success, and receive feedback and incentives. Then, we described the cycles that will engage players based on their actions and the solution's feedback. Some nudges [31] were defined to cue users to desired behaviors. These include reminders for relevant events or rewards, immediate feedback after specific behaviors or events, rewards for performing some behaviors, and a dashboard with project's information. Fun elements were included in the fifth step so that players are likely to engage with the solution. Creating a social component to visualize each other's milestones (as practitioners like do communicate) and team achievements to help them developing a sense of belonging to something greater were some of the chosen elements. Finally, in the sixth step we selected the game elements and software used to build the gamification solution, thus promoting defined behaviors. The solution was implemented as an app for a software management tool, Jira Software [35], so that practitioners do not need to use a separate tool for gamification. This tool supports the development process and is based on "issues" i.e., problem that needs to be resolved (e.g. a bug).

In this solution, users unlock eXperience Points (XP) for actions such as resolving an issue, which allows them to progress through levels. Challenging achievements, such as resolving all issues in a sprint, are awarded with XP and rewards of two types: badges and gems (a virtual currency). Achievements can be either individual or collective, to boost cooperation. Users get positive feedback through pop-up notifications to inform them if they are performing the target behaviors or guide them otherwise.

These game elements were then implemented in Jira Software. As this tool not support Scrum events, we also created issue types for each of the four events. The Dashboard in Figure 1. displays information and statistics regarding the user's role, both general and regarding the user. A small profile shows some information like profile picture, XP earned, current level, and four featured badges that the user selected to be showcased. Below, an activity feed lists all project's events. On the right, the user can consult the rewards (s)he is closer to win and check four project's statistics that provide feedback regarding Scrum practices.

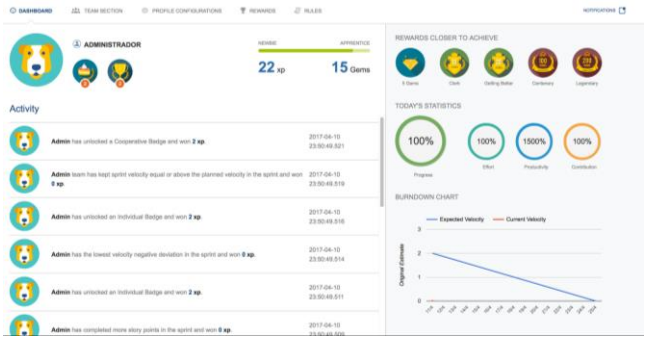

Figure 1. Project Dashboard of the gamification solution.

A similar screen provides similar information but focusing on multiple projects, and a Team screen enables the user to meet other team members by seeing their profile. The Rewards' screen displays available and unlocked rewards, and the Rules screen describes the app features.

The solution was evaluated with a Scrum team, and an analysis conducted based on quantitative data extracted from the project's Jira Software instance and interviews with team members. Results suggest that the team's Scrum practices somewhat improved after using the app, but some challenges remain. Participants liked the pop-up notifications but believe there are too many notifications launched simultaneously. They also think that a default selection should be implemented for selecting the featured badges and projects, as they forget to do this. The team did not use the Scrum events functionality, as they already have a process to register them, which they prefer. Overall, participants liked the experience but believe the app should be improved to be more challenging, for example by creating more and variate rewards. Additionally, we understood that the gamification data recorded in the database did not allow us to conduct a deeper data analysis, such as the XP evolution during the sprints.

TASK ANALYSIS

The study was conducted in a consultant group with around 70 workers providing mainly software services, but also some products. The organization started using Jira Software five years ago, replacing all internal tools used by then. The participants are part of the team A (anonymous name) and Team B and work on a project management with Scrum. These teams were previously invited by the organization's director to participate in the study, and all members voluntarily accepted.

Instead of studying participants with individual interviews (as done in the previous work iteration), we decided to conduct a focus group session with the teams and complement this information with data collected using a survey. With these tools, we wanted to understand how Scrum is being implemented in these teams; the challenges they encounter; and possible solutions to address them. We also wanted to study the respondents' gaming habits and how game mechanisms can help addressing challenges in implementing Scrum.

All focus group sessions were held in February 2019 in the participants' organization's office. A semi-structured protocol was followed, meaning we defined the question areas and some questions, but allowing for flexibility and improvisation. The protocol consisted of six parts: interviewee's characterization; level of usage and opinion about Jira Software; experience with Scrum; experience and opinion regarding the time reporting process; experiences with games/gamification; and insights about the proposal. The survey was built using Google Forms [36] and disseminated by e-mail, being open for one month. Three responses were received, all considered valid (response ratio of 0.60). The questionnaire had sections on 1) demographics; 2) Scrum and Jira; and 3) games and gamification. The outputs of both methods are presented and discussed below.

Team A

Team A's (anonymous name) project started in September 2018, five months before the focus group session. Despite being formally focused on service management, the team considers the task performed as more related to product management. This project's client is a government entity, and thus their priorities are often changing, and it is hard to define a work routine.

Team A is composed by five people. According to survey results, three team members are between 26 and 35 years old and have graduate studies. Respondent A1 takes both the Scrum Master and Product Owner roles in the team. She started her career in this organization 10 years ago and completed a PhD while working there. A2 started working in the organization after his graduation, three years and four months ago. A3 and A4 started working in the organization

one and five months ago, respectively, but this was not their first professional experience. All team members hold a degree in IT engineering. All participants agree that the team is very cooperative, and not competitive. All team members are focused on the project's success and work together to achieve their goals.

Team B

Team B's project started on January 2019, but I nmarch resources had to be allocated to another project, and the project slowed down.

This team has three Development Team members. In this focus group session, only the Scrum Master and two Development Team members participated, as one of them (performing Quality Assurance consultancy tasks) was on holidays.

All team members haven been working for more than 10 years in the organization, except for B4, who started three years ago. Except for B2, this was not the first professional experience for any team member: B3 worked in many companies during the 40 years of her professional career, and B4 previously worked in another organization. B2 and B4 hold a degree in IT Engineering, while B3 completed high school and attended an MBA. B2 only spends 35\% of his time in this project and works on functional analysis and solutions architecture in other projects. B3 helps in all projects where she is needed, including this one, performing mainly documentation and management tasks. B4 has 60-70\% of her time allocated to this project and provides maintenance in other projects.

They agree the team is not calm, mainly due to the client's pressure, but they have a good spirit.

Scrum

The organization started using Scrum six years ago, but only in the last six months they have been focusing hard on doing it correctly. Project managers attended a small internal training, but the organization is planning on certifying some people in a near future. A1 and B1 do not believe there is resistance in using Scrum in their teams.

Both teams consider that an issue is done (i.e. their Definition of Done) when automatic tests are completed, and the issue is transitioned to quality assurance tests. After this, the work is delivered to the client, who rarely validate it before going to production. When the client has time, some feedback demos are conducted, but this rarely happens.

Sprints conducted in both teams have a duration of two weeks. Before starting a Sprint Planning, the Scrum Master prepares the sprint and adds all relevant issues. During the event, the team discusses if the allocated work can be done during the sprint, and each member assigns him/herself to the issues (s)he wants to implement. Team A conducts Daily Scrums consistently, where they revise what was done the day before and define what will be done during that day. B1 promotes Daily Scrum with the teams of all projects where she is project manager, comprising around 10 people. The goal is to ensure everyone's work is aligned, but this time is not used to resolve impediments. When some team needs to discuss some topics in more detail, they conduct a "sub-Daily Scrum".

Games and Gamification

All participants have been in contact with games, from board games with their kids (A1 and B1) to regular gamers(A2, B4), casual gamers(B2, B3), and participants that stopped once they started working (A3, A4)

While B1 knows the concept, but has never used it, A1 used gamification during her PhD thesis. She says that while some people were motivated by game elements, others found it childish and did not enjoy the experience. They agree that gamification benefits depend a lot on the person's profile and age. They asked for these teams to participate in this experience because they think they will like it, given their profile. Moreover, they do not consider this will be a distraction from their work nor will have a negative impact on them, as long as they do not link the gamified solution with the evaluation process.

On gamification, only B1 and B3 did not have previous experience with gamification, A3 used the app *Habitica* [37], B2 uses StackOverflow, A2 and B4 experienced a gamified course in college and A4 worked in a organization where the employee's evaluation process was gamified. A1 used gamification during her PhD thesis. She says that while some people were motivated by game elements, others found it childish and did not enjoy the experience.

Summary

No one in this organization is Scrum certified or had training, and only in the last few months the organization started focusing on correctly implementing the framework. Overall, both teams state their sprints last for two weeks and consistently conduct Daily Scrums, which is compliant with Scrum. They state that they select the tests they want to perform, thus teams are self-assigned, and the existence of a QA consultant in each team (despite not participating in the focus group) suggest they are cross-functional.

However, teams differ in some practices. Team A is very motivated to use Scrum, as they like to receive immediate feedback. Yet, oppositely to what Scrum advocates, the same person takes the Scrum Master and Product Owner roles, and the team has difficulties in estimating issues.

Regarding team B, no member is fu*lly allocated to this project, and only one member spends more than 50% of the time working on this project's tasks. They admit the team lacks Scrum knowledge, which does not allow them to improve their practices.

METHOD

After performing the teams and its members task analysis, a baseline study was conducted based on historical data extracted from the teams's Jira Software instance. Then, the teams started using the same gamification solution used in our previous work, and data from this first field study was compared against the results of baseline.

Afterwards, we improved the original solution following the results of our previous work and the task analysis conducted. Some 6D steps were revised, except for steps one (goals for the solution) ,three and five (fun elements). All decisions that are not revised in this section were kept from the original solution.

In the second step, we refined some of the target behaviors and metrics. In the previous study we saw that issues can be assigned to different people through its lifecycle, thus metrics in the practitioner context are not much relevant. Thus, in this work we only studied metrics in sprint context. Other changes in metrics include:

- **Assigned issues:** we now verify this metric is calculated by verifying if the issue has an assignee when its status changes to *In Progress*.
- **Reopened issues:** this metric comprises the number of issues that were open during the sprint, independently on how many times it was open. For example, if one issue is reopened twice in a sprint, this metric is only increased by one.
- **Estimation Precision:** this metrics corresponds to the rate between worklogs recorded for a sprint (i.e. the time reported) and the issue's estimative. This metrics allows to understand the quality of these teams's estimations.

As suggested by the previous work's participants, we automatized the assignment of feature badges in the fourth step. Now, badges are added to the players' featured badges until there are four badges assigned. If players edit this information once, the automatization is disabled. Moreover, we limited the number of notifications to be launched simultaneously to six. If there are more notifications in the pipeline, they are discarded.

In the sixth step, we refined some game elements as to implement some improvements. The score system was simplified to reduce the launched notifications. For example, instead of receiving XP and a badge for an action (which launched at least three notifications), players now only receive XP and one notification.

Instead of receiving the same XPs for resolving issues, different scores were defined based on estimation precision to make the solution more challenging. Players receive increasingly more XP depending on whether they resolved a not estimated issue; an estimated issue with no time reported; or an estimated issue with less than 80% precision; 80% to 89% precision; 90 to 99% precision; or 100% precision. These values were decided with the teams, based on their own success metrics.

Specific game elements were created for Scrum Master and Product Owner roles aiming at varying the rewards offered. As the Scrum Master influences the development team members' performance, the solution now awards the player with this role with XP whenever someone on the team resolves an issue. This value is proportional to the XP received by the other player. The Product Owner receives both XP and rewards whenever a sprint starts with all issues estimated and prioritized.

All notifications were adapted to the previous changes, by creating, deleting, or altering some feedback messages. These changes did not affect the app's visual interface. To improve the analysis of gamification data, we added a logging table to the database that allow us to conduct a deeper data analysis, such as XP evolution over time.

This improved solution was then installed on the teams's Jira Software instance for a period that we will refer to as the second field study. Target metrics were calculated based on the data extracted, and gamification data was analyzed by querying the database with the appropriate management tool. Instead of collecting qualitative feedback in the end of the study like we did in the previous study, we decided to create a communication channel using Hangouts [38], so that participants could provide feedback continuously.

This information collected in the baseline and the two field studies allowed us to evaluate the impact of gamification in this teams's practices (using both the original and improved solution), and to compare the results with those of the teams who participated in the previous work. These results are presented in the next section.

RESULTS

In this section, we first present the results of the baseline, and then those of the two field studies. We end by comparing and discussing the results of all studies. During data analysis, we identified some activity from practitioners that do not belong to the team A. As their activity was residual, we removed them from the analysis. The team members are identified below by the same ID used in Task

Analysis section, being A5 the team member that did not participate in the focus group.

Regarding Team B there was a lack of activity and as such we will not present the evaluation in this abstract.

Statistical analysis of the data was supported by jupyter Notebooks [39], using *pandas* and *SciPy* python packages.

Baseline Study

The baseline study comprises nine sprints between August 2018 and March 2019 (seven months). Mean sprint duration is 23.10 days (sd=20.21). No sprint was completed, and five sprints contain reopened issues (55.6%). All sprints had persistent issues, and in eight sprints some issue was added or removed to the sprint after its start. On average, each sprint has 12.56 issues, where 46.22% are resolved, 75.79% estimated, 70.64% assigned, 8.80% reopened, and 60.79% persistent. Mean velocity is 15.15 days. The most common issue type is Task (avg=6.2;sd=8.4), followed by Document (avg=2.1;sd=1.7) and Design (avg=1.9;sd=1.4). In total, 81 issues were resolved without being assigned to a sprint. No conclusions could be drawn regarding Scrum meetings and roles, as Jira Software does not record that information.

First Field Study

The first field study was conducted between March and May 2019 (two months) and comprised four sprints. Mean sprint duration is 16.17 days (sd=7.14). No sprint was completed, and three sprints contained reopened issues (75.0%). On average, each sprint contains 24.5 issues, where 57.59 % are resolved, 93.19% estimated, 51.68% assigned, 2.97% reopened, and 56.42% persistent. Mean velocity is 22.40 days (sd=3.67). All sprints had persistent issues, and in three sprints some issue was added or removed to the sprint after its start (75.0%). The most common issue type was Task (avg=24.5;sd=5.2), followed by Story (avg=8.0;sd=1.8). In total, 61 issues were resolved without being assigned to a sprint. No Scrum events were created. A1 reported in the Hangouts channel that it was a burden to create and have those extra issues every sprint.

Before analyzing the gamification results, we must explain that this software rewarded another behavior that is out of the scope of this study but might have influenced players performance on gamification. Therefore, we removed the XP attribution and reward unlock events corresponding to that behavior. In this field study, that corresponds to one XP unlock event. Furthermore, an outlier was removed, corresponding to a score attribution of 1172XP to A1 (i.e. she resolved an issue with 293 hours estimate, around seven weeks), for ease of reading.

Figure 2 shows a timeline with the XP earned by each player, where the sprint's complete dates are identified by the vertical bars. These lines are not smooth, as players did not unlock XP every day. The two highest peaks correspond to the score earned by A3 after resolving issues with 40 and 36 hours, respectively (around one week).

Overall, resolving issues was rewarded 44 times and unlocking individual badges was rewarded 13 times. As presented in Table 1, A1 is the player with more XP, followed by A3 and A2. A4 and A5 only have residual XP in this study.

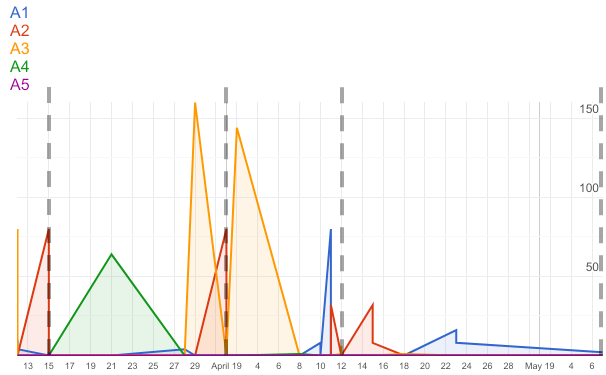

Figure 2. XP attribution over the first field study by player. Sprints complete dates are highlighted with grey vertical bars.

	S1	S2	S3	S4	TOTAL
A1	30	20	1328	46	1424
A2	160	80	32	40	312
A3	90	162	154	1	407
A4	0	64	3	0	67
A5	2	0	0	0	2

Table 1. XP earned by player in each sprint of the first field study.

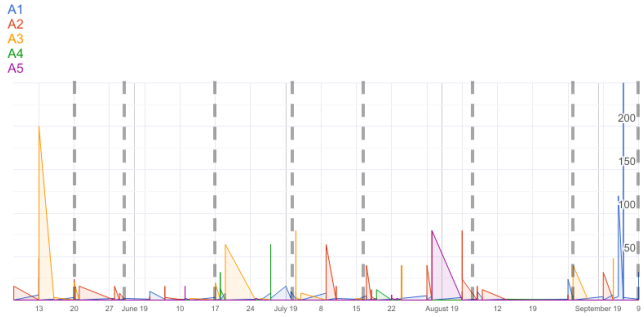

Figure 3. XP attribution over the second field study by player. Sprints complete dates are highlighted with grey vertical bars.Although no gems were unlocked, nine badges were unlocked over the four sprints. A3 unlocked the badge for completing the most estimation units per sprint in all sprints. A1 and A5 were the only ones unlocking the badge for the first time resolving all issues in a sprint, thus other players did not complete any sprint. Consequently, the cooperative badge for all team resolving all issues was not unlocked.

Second Field Study

The second field study was conducted between May and September 2019 (four months) and comprised eight sprints. Mean sprint duration is 15.44 days (st-dev=3.80). No sprint was completed, and all sprints contained reopened and

persistent issues. On average, each sprint has 29.88 issues, where 65.52% are resolved, 76.61% estimated, 54.55% assigned, 9.81% reopened, and 41.39% persistent. Mean velocity is 23.89 days. In all sprints some issue was added or removed to the sprint after its start. The most common issue type was Task (avg=12.8;sd=3.3), followed by Issue (avg=9.0;sd=2.4). In total, 82 issues were resolved without being assigned to a sprint. No Scrum events were created.

Regarding gamification, we removed 65 XP attribution and five reward unlock events corresponding to the behavior outside the scope of this study. Figure 3 shows a timeline with the XP earned by each player, where the sprint's complete dates are identified by the vertical bars. Again, the lines are not smooth and since players did not unlock XP every day. The two highest peaks correspond to the score earned by A3 and A1 after resolving issues with 80 and 250 hours, respectively (around one week).

Overall, 116 events rewarded issue's resolution, including 25 issues resolved with no time was reported), 64 issues resolved with less than 80% precision, four issues resolved with 80-89% precision, eight issues resolved with 90-99% precision, and 15 issues eight issues resolved with 100% precision. The latter included issues estimated with one to eight hours, except for one issue estimated with 80 hours (10 days). A5 did not resolve any of the most precise issues (above 80%). A1 won 231 XP for issues resolved by development team members. As presented in Table 2, A1 is again the player with more XP, followed by A3 and A2. A4 and A5 have more XP in this study.

	S1	S2	S3	S4	S5	S6	S7	S8	TOTAL
A1	66	8	26	65	34	70	65	721	1055
A2	88	48	23	0	80	240	46	32	557
A3	383	0	22	66	95	176	56	48	846
A4	0	0	0	171	0	12	1	0	184
A5	0	0	22	2	0	90	0	0	114

Table 2. XP earned by player in each sprint of the second field study.

Overall, 18 badges were unlocked during the second field study, but no gem was unlocked. The badge for completing the most estimation units per sprint in all sprints was unlocked three times by A3, twice by A2, and once by A1. A1 unlocked the badge for having the lowest velocity negative deviation (i.e. for complying with her issues' estimations). No other badge was unlocked for resolving all issues in a sprint.

After the study started, A1 used the Hangouts channel to share that notifications are now much less intrusive, which improved the solution's experience. A2 also reported some minor bugs, suggesting he does explore the app.

Differences Between Studies

Here we look at the differences between metrics in baseline and field studies. A Shapiro-Wilk normality test was applied to all metrics from all studies. For one metric in the baseline study, three metrics in the first field study, and two metrics in the second field study, we had to reject the null hypothesis that samples follow a normal distribution ($p<0.05$). However, data samples are too small to reach a decision of normality. Following these results, all statistical differences between groups were checked using a non-parametric Mann-Whitney's U test. The statistics calculated are presented in Table 3 for the baseline and first field study, and in Table 4 for the first and second field studies.

The Mann Whitney's U tests applied revealed statistical evidence that most metrics differ between studies (p-value>0.05). However, there is no evidence that the mean total issues between baseline and first field studies and the mean rate of estimated and persistent issues between first and second field studies can indeed be different.

When compared against the baseline, more issues seem to be resolved by the end of the sprint (more than half) and estimated (more than 90%). Additionally, the number of reopened and persistent issues seem to decrease. However, less issues seem to be assigned to some practitioner when their status change to *In Progress*.

Between field studies, the number of issues per sprint seem to have increased, and so did the velocity (around one day). The mean number of resolved issues per sprint also increased. The number of issues assigned slightly increased, but not to the initial values recorded in the baseline study. The number of reopened issues increased to a higher value when compared against the baseline.

	Baseline		1st Field Study		Mann-Whitney U	
	mean	**sd**	**mean**	**sd**	**U**	**p-value**
Total Issues	12,56	7.04	24.5	5.20	2.0	0,016
Estimated Issues [%]	75,79	26.17	93.19	4.71	11.0	0,314
Resolved Issues [%]	46,22	26.53	57.59	6.28	13.0	0,487
Assigned Issues [%]	70.64	20.76	51.68	10.94	29.0	0,105
Reopened Issues [%]	8.80	10.20	2.97	2.11	22.0	0,577
Persistent Issues [%]	60.79	24.37	56.42	11.03	19.5	0,876
Velocity [Days]	15.15	7.72	22.40	3.67	5.0	0,054

Table 3. Statistics for calculated metrics in baseline and first field studies.

	1st Field Study		2nd Field Study		Mann-Whitney U	
	mean	**sd**	**mean**	**sd**	**U**	**p-value**
Total Issues	24.50	5.2	29.88	3.27	8.0	0,200
Estimated Issues [%]	93.19	4.71	76.61	14.91	29.0	0,033

Resolved Issues [%]	57.59	6.28	65.52	13.31	7.0	0,148
Assigned Issues [%]	51.68	10.94	54.55	30.97	12.0	0,552
Reopened Issues [%]	2.97	2.11	9.81	8.48	4.0	0,050
Persistent Issues [%]	56.42	11.03	41.39	7.44	29.0	0,034
Velocity [Days]	22.40	3.67	23.89	8.88	17.0	0,932

Table 4. Statistics for calculated metrics in the field studies.

Overall, most sprints have reopened or persistent issues, showing that not all issues are resolved before the sprint. The number of issues of type Story increased in the first field study but decreased again the second one.

DISCUSSION

Results suggest that some metrics slightly improved after using the solution, particularly after using the improved version. The percentage of issues resolved per sprint increased and the percentage of persistent issues decreased, but most sprints have reopened or persistent issues. This can be related to the team's unstable work routine caused by working for a governmental entity. After introducing the original solution, the number of estimated issues increased, but dropped again after introducing the improved version. Oppositely to what Scrum advocates, the same person takes Scrum Master and Product Owner's tasks in this team. As no Scrum event was registered, we could not confirm if they do conduct Daily Scrums consistently.

All members in this team like to play and have experienced at least one gamification solution. A2 was the participant with more gaming habits and the one providing more feedback, probably because he does spend some time exploring the app. Like in the previous study, the Scrum Master (A1) is the one with higher score in both studies. A3 was the last one to join the team, but currently is the assignee resolving more issues. A4 and A5 are the less active members. Their activity on the app is reduced, and even absent in some sprints. This can be related to the organization projects' allocation, meaning they can be working in other projects' tasks. Overall, gamification results are not yet much balanced between players.

Results show evidence that this team has difficulties in estimating issues. First, because more than half of the issues are resolved with a precision below the organization's standards (above 80%), as seen in the improved version with the layered score system for resolving issues. Moreover, some issues are estimated with periods superior to the sprint's duration, which is unfeasible. We believe that practitioners would benefit if the tool provided some training on how to improve estimations.

Despite enabling a more complete analysis, the new logging system cannot solve the problem of having sparse gamification data. Players do not receive XP or unlock rewards every day, and thus few patterns can be identified in this data. This limitation can be linked with characteristics of the team, like having people allocated to many projects or estimating issues by several weeks. However, it is important to select game elements that promote daily interaction, thus a more continuous usage of the app.

By comparing this work's results against those of our previous work, we believe that the solution impacted positively both teams. In fact, the improvements made based on feedback inputs from the first team seem to be enjoyed by team A. Some challenges could be addressed, like simplifying the complex notifications approach, but the Scrum Master's inflated score remains. None of the teams used the Scrum events functionality. In the future, a simpler mechanism should be studied, for example by creating just one issue for Daily Scrums that resets each day.

While these teams enjoyed the experience, both are composed by young people who like to play games and are willing to use Scrum. In the future, it would be interesting to understand the impact that gamification would have in teams with different profiles, particularly with people more resistant to innovative initiatives.

In the future, motivation theories like SDT could be used to improve the solution, and thus increase motivation. Right now, most of the game elements studied are linked with external motivators, and the only SDT need addressed was relatedness.

Limitations

This study poses some limitations. Although results suggest that gamification positively impacted both teams, the study samples were reduced, thus the tests applied to the differences in metrics might have low statistical power. This means that there is some risk that conclusions drawn can be wrong. Likewise, the statistics presented for the survey analysis are not statistically significant due to the reduced number of answers, and thus cannot be generalized to other settings. Also, this analysis lacks the insights from two team members out of five. As we explained, this software rewarded a behavior not related to Scrum that was not studied but might have influenced players performance on gamification. This impact should be analyzed in the future. Moreover, some practitioners worked on the analyzed project, despite not being officially part of the team. Their activity had an impact (even if little) on the results presented here, and it is important to understand their role on the project. Because little feedback was provided through the Hangouts channel, some results cannot be explained. For example, we cannot be sure if the differences in XP received are due to practitioners' performance or work allocation. A deeper user analysis must be conducted in the future, using methods like interviews or surveys.

CONCLUSION

Addressing practitioners' lack of motivation to adopt agile practices remains a challenge. Many authors have proposed

gamification solutions, but research still lacks empirical validation. We developed a gamification solution as a Jira Software app to increase practitioners' motivation in adopting Scrum practices. The impact of this solution was previously evaluated on a Scrum team, and in this work we evaluated the same solution with a team at a different organization. Based on a task analysis, which allowed us to understand the Scrum practices adopted in this team and its members characterization, and the previous results, we improved and deployed the solution on the same team's Jira Software's instance.

A comparison of data from a baseline and two field studies, extracted based on the metrics discussed, suggest that this team's results slightly improved after using the gamified app, similarly to what was verified in the previous study. Past results allowed us to improve the app and cope with some limitations, like the complex notifications approach, but some challenges remain, such as the Scrum Master's inflated score. The long-term effect of gamification should be addressed in the future.

These results can be of interest not only to researchers in this field, but also to organizations who are looking forward to increasing their workers' motivation in applying Scrum techniques. In the future, these team members' feedback should be collected using methods like interviews or surveys to understand some of the results achieved. Additionally, the app should be tested with more diverse teams to evaluate the impact in real-world Scrum teams.

REFERENCES

1. Gabriel Barata, Sandra Gama, Joaquim Jorge, and Daniel Gonçalves. 2017. Studying student differentiation in gamified education: A long-term study. *Computers in Human Behavior* 71: 550–585. https://doi.org/https://doi.org/10.1016/j.chb.2016.08.049
2. K. Beck, M. Beedle, A. van Bennekum, A. Cockburn, W. Cunningham, M. Fowler, J. Grenning, J. Highsmith, A. Hunt, R. Jeffries, J. Kern, B. Marick, R. C. Martin, S. Mel- lor, K. Schwaber, J. Sutherland, and D. Thomas. 2001. *Manifesto for Agile Software Development*.
3. Flávio Steffens de Castro. 2016. Using Gamification as a Collaboration Motivator for Software Development Teams : A Preliminary Framework. Pontifícia Universidade Católica do Rio Grande do Sul.
4. Bc Martin Češka. 2016. Gamification in the SCRUM Software Development Framework. Masaryk University, Faculty of Informatics. https://doi.org/10.1111/j.1748-5991.2008.00039.x
5. Kieran Conboy, Sharon Coyle, Xiaofeng Wang, and Minna Pikkarainen. 2010. People over process: Key challenges in agile development. *IEEE Software* 28, 4: 48–57. https://doi.org/10.1109/MS.2010.132
6. Mihaly Csikszentmihalyi. 1990. *Flow: the psychology of optimal experience*. New York: Harper and Row.
7. Tom DeMarco and Tim Lister. 2013. *Peopleware: productive projects and teams*. Addison-Wesley.
8. Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. From game design elements to gamefulness. *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments - MindTrek '11*: 9–11. https://doi.org/10.1145/2181037.2181040
9. Sebastian Deterding, Rilla Khaled, Lennart E Nacke, and Dan Dixon. 2011. Gamification: Toward a Definition. In *CHI 2011 Gamification Workshop Proceedings*, 12–15. https://doi.org/10.1007/978-3-642-13959-8_1
10. Alec Dorling and Fergal McCaffery. 2012. The gamification of SPICE. In *Communications in Computer and Information Science*. Springer, 295–301. https://doi.org/10.1007/978-3-642-30439-2_35
11. Daniel J. Dubois and Giordano Tamburrelli. 2013. Understanding gamification mechanisms for software development. In *Proceedings of the 2013 9th Joint Meeting on Foundations of Software Engineering*, 659. https://doi.org/10.1145/2491411.2494589
12. Hassan Hajjdiab, Al Shaima Taleb, and Jauhar Ali. 2012. An industrial case study for Scrum adoption. *Journal of Software* 7, 1: 237–242. https://doi.org/10.4304/jsw.7.1.237-242
13. Eduardo Herranz, Ricardo Colomo-Palacios, Antonio Amescua Seco, and Murat Yilmaz. 2014. Gamification as a Disruptive Factor in Software Process Improvement Initiatives. *Journal of Universal Computer Science* 20, 6: 885–906.
14. Eduardo Herranz, Ricardo Colomo-Palacios, and Antonio de Amescua Seco. 2015. Gamiware: A gamification platform for software process improvement. In *Communications in Computer and Information Science*. Springer, 127–139. https://doi.org/10.1007/978-3-319-24647-5_11
15. Daniel Johnson, Sebastian Deterding, Kerri-Ann Kuhn, Aleksandra Staneva, Stoyan Stoyanov, and Leanne Hides. 2016. Gamification for health and wellbeing: A systematic review of the literature. *Internet Interventions* 6: 89–106.
16. Philipp Lombriser and Roald van der Valk. 2011. Improving the Quality of the Software Development Lifecycle with Gamification.
17. Alexandre Loriggio, Victor Farias, and Pollyana Mustaro. 2013. Aplicações de gamificação e técnicas de motivação à aprendizagem da metodologia ágil scrum. In *VIII International Conference on Engineering and Computer Education*, 326–330. https://doi.org/10.14684/ICECE.8.2013.328-332

18. Anonymous. 2018. Details omitted for double-blind reviewing.

19. Alan McClean. 2015. An Exploration of the Use of Gamification in Agile Software Development. Dublin Institute of Technology.

20. Version One. 2017. *11th annual State of Agile Report*.

21. Erick B. Passos, Danilo B. Medeiros, Pedro A S Neto, and Esteban W G Clua. 2011. Turning real-world software development into a game. In *Brazilian Symposium on Games and Digital Entertainment, SBGAMES*, 260–269. https://doi.org/10.1109/SBGAMES.2011.32

22. Oscar Pedreira, Félix García, Nieves Brisaboa, and Mario Piattini. 2015. Gamification in software engineering - A systematic mapping. *Information and Software Technology* 57, 1: 157–168. https://doi.org/10.1016/j.infsof.2014.08.007

23. Christian R. Prause, Jan Nonnen, and Mark Vinkovits. 2012. A Field Experiment on Gamification of Code Quality in Agile Development. *Psychology of Programming Interest Group (PPIG) Annual Conference* 64: 175–186.

24. Byron Reeves and J Leighton Read. 2009. *Total engagement: How games and virtual worlds are changing the way people work and businesses compete*. Harvard Business Press.

25. Richard M Ryan and Edward L. Deci. 2000. Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist* 55, 1: 68–78. https://doi.org/10.1037/0003-066X.55.1.68

26. Richard M Ryan and Edward L Deci. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology* 25, 1: 54–67. https://doi.org/10.1006/ceps.1999.1020

27. K Schwaber and Jeff Sutherland. 2016. *The Scrum Guide*. https://doi.org/10.1053/j.jrn.2009.08.012

28. Leif Singer and Kurt Schneider. 2012. It was a bit of a race: Gamification of version control. In *2nd International Workshop on Games and Software Engineering: Realizing User Engagement with Game Engineering Techniques*, 5–8. https://doi.org/10.1109/GAS.2012.6225927

29. Jamila Peripolli Souza, André Ricardo Zavan, and Daniela Eloise Flôr. 2016. Scrum Hero: Gamifying the Scrum Framework. In *Brazilian Workshop on Agile Methods*, 131–135.

30. Standish Group. 2015. *CHAOS Report*.

31. Richard Thaler and Cass Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press, New Haven & London.

32. Kevin Werbach and Dan Hunter. 2012. *For the Win: How Game Thinking Can Revolutionize Your Business*. Wharton Digital Press.

33. Murat Yilmaz and Rory V O Connor. 2016. A Scrumban integrated gamification approach to guide software process improvement: a Turkish case study. *Tehnicki vjesnik - Technical Gazette* 23, 1: 237–245. https://doi.org/10.17559/TV-20140922220409

34. Jiraffe. Retrieved June 28, 2016 from https://marketplace.atlassian.com/plugins/com.bugpotion.jiraffe/server/overview

35. Jira Software. Retrieved September 18, 2019 from https://www.atlassian.com/software/jira/features

36. Google Forms. Retrieved September 18, 2019 from https://www.google.com/forms/

37. Habitica. Retrieved September 18, 2019 from https://habitica.com

38. Google Hangouts. Retrieved September 18, 2019 from https://hangouts.google.com

39. jupyter. Retrieved September 18, 2019 from https://jupyter.org/