Improving the Software Homologation Process through Peer Review: An Experience Report on Android Development Environment

Camila Bernardon, Diana Lemos, Mario Garcia, Thiago Souto, Bruno Bonifacio

Abstract—In the current technological market environment, ensuring the quality of new products has become a complex challenge. In this scenario, companies have been investing in solutions that aim to reduce the execution time of software testing and lead to cost efficiency. However, companies that have a complex and specialized testing environment usually face barriers related to costly testing processes, especially in distributed settings. Sidia Institute of Technology works on research and development for the Android platform for mobile devices in Latin America. As we work in a global software development (GSD) scope, we have faced barriers caused by failures detected lately that have caused delays in the homologation release process on Android projects. Thus, we adopt an Internal Review process, using as an alternative to reduce these failures. In this paper it was presented the experience of a homologation team adopting an Internal Review process in order to increase the performance through of improving test efficiency. Using this approach, it was possible to realize a substantial improvement in quality, reliability and timeliness of our deliveries. Through the quantitative analyses, it was possible identify a positive growth in homologation efficiency of 6% after adoption of the process. In addition, we performed a qualitative analysis from the collected data through an online questionnaire. In particular, results show that association between failure reduction and review process adoption provides the most quality that has a positive effect on project milestones. We hope this report can be helpful to other companies and the scientific community to improve their process thereby increasing competitive advantages.

Keywords—Android, GSD, improvement quality process, mobile products.

I. INTRODUCTION

THE process of globalization has become increasingly important for software industry. In view of this, providing quality new products and services represents a major challenge to ensure enhanced competitiveness in new businesses [1]. In this scenario, companies have been investing in solutions that aim to reduce the execution time of software testing and accelerate the time to market for new products, with higher product quality. Even though GSD has been widely disseminated, aiming to decrease costs, with fast production and quality, companies substantially face the barriers related to expensive testing processes which are typically part of a complex and specialized testing environment [2]

Camila Bernardon is with Sidia Instituto de Ciência e Tecnologia, Brazil (e-mail: camila.bernardon@sidia.com).

In order to deliver quality output, the quality aspects need to be maintained during the development software process. Amongst them, one effective way is by including code review as continuous practice during the coding and testing phase. The main goal of reviews is to catch potential market issues, security problems, and bugs before they are introduced to the codebase and prevent them from causing problems in production [3].

Sidia Institute of Technology works on research and development for the Android platform for mobile devices, in partnership with a mobile device manufacturer in Latin America [4]. As we work in GSD settings, we have faced barriers caused by failures detected lately that caused delays in our release process, during homologation phase.

In this paper, we present how we adopt Internal Review testing as part of the homologation phase. By adopting this approach, it was possible to improve quality, reliability and timeliness of our deliveries. In this paper we present our experience, adopting an Internal Review process to increase the efficiency of tests during the homologation of Android software releases. The study was carried out using homologation test requests reports. In addition, we asked developers, through an online questionnaire, regarding their Internal Review considerations and how they balance the benefits and costs Internal Review activity.

Results from quantitative analyses show increased homologation efficiency of 6% afterward the process was adopted. Our qualitative results contribute to identify gaps and improvements related to review adopted by experienced testers. Also, quantitative results were useful to identify difficulties faced by our test team, in improving issues reviews when they adopt ad hoc process. By results, we hope that this report contributes to assist testers and researchers to adopt testing review in development process. In addition, we intend to encourage companies and researchers to adopt review for both, coding and testing process, on development product process.

This paper is structured as follows: Section II provides background related review process improvements. Section III describes our case study, results and important findings. Finally, we present lessons learned and final considerations.

II. BACKGROUND

In the context of GSD, it is even more relevant to reveal the industrial demand for software testing skills, because software testing plays an essential role in development and, moreover,

World Academy of Science, Engineering and Technology
International Journal of Computer and Systems Engineering
Vol:16, No:3, 2022

global software testing has witnessed an increased adoption over the years [13]. Most of the companies in different regions have adopted the GSD to gain maximum benefits by producing quality products with cheap development cost [6]. Moreover, distributed software development affects the quality of the delivered software, caused by missing both code and testing review process.

While code review is process to inspect code that helps identify bugs, code style policy violations, security vulnerabilities, test review is a process to inspect testing results before software released [7]. Test review is often overlooked as an ongoing practice during the development phase; however, several studies show the importance to adopt this procedure since it removes defects at the early stage of development and further reduces the testing and maintenance efforts at the later phase of the development lifecycle [2], [7], [13].

Reference [5] carried out a study to explore impact of test review on unit testing in global software projects particularly in agile practiced teams. The study applied with software developers, quality assurance engineers, project managers and business analysts show that test efficiency, test coverage and test effort had the impact on delivery quality in order of impact. According to the outcome, there is a space to improve using test review, especially automated.

Reference [11] also analyzes process improvements, using testing review. In this context, they evaluate overall projects to identify gaps of communication between developers and testers. In order to deliver quality output, the quality aspects need to be maintained during the development software process. Amongst them, one effective way is by including code review as continuous practice during the coding and testing phase. The main goal of reviews is to catch potential market issues, security problems, and bugs before they are introduced to the codebase and prevent them from causing problems in production [10]. In this sense, several investigations have focused on adoption review process for both, code and testing, however there is still a need to develop effective and efficient methods, techniques and best practices that can lead to enhance the quality of software development [13]. Hence, companies that use approach reviews structured can improve efficient testing or code process to avoid market issues or deadline delays. Thus, this work contributes by making important sharing related how improve testing process, using Internal Review process through peer review, should be handled, and how to achieve massive demand with maximum quality of software product. As Sidia works on research and development (R&D), of one partner mobile device manufacturer, which embeds the Android Operating System (OS), we have a homologation process revised by Google [14]. Due massive demands we faced strong difficulties to meet deadlines caused by issues not detected by testing team during submission to Google Review team. Inspired by this context, we adopt peer Internal Review process. In next subsection we present difficulties faced during homologation release process and how we improve this phase, using Internal Review process.

A. Sidia Homologation Release Process

Sidia is an R&D Institute and has a mobile product area responsible for developing embedded solutions on the Android platform [8]. In this area is a used release process in GSD scope and one important release process phase is homologation release software, conducted by specific SQE team, called Google Approval (GA) team, responsible to validate requirements from Google related security patches, user experience recommended to Open Handset Alliance partners [8].

The aim of GA team is to support homologation process through mandatory tests execution. For this reason, GA team needs to check all patches applied, to verify apps and vulnerabilities to get Google Mobile Services (GMS) certification [10]. GMS is a collection of Google applications and APIs that help support functionality across Android devices [10]. This certificate is the confirmation that a release complies the Google performance requirements and runs correctly Google applications [9].

Sidia Homologation Process starts after release of software binary. As we work on GSD scope, GA testers receive a test request to check if all GMS is applied on software released. We use Jira system to support tests requests of all R&D partners. Afterwards, GA team performs all required tests, as well as the submission of generated results in order to achieve GMS certification, allowing that Android operational system is officially embedded on Android devices. Fig. 1 shows overview process fulfilled to a release get the GMS certification.
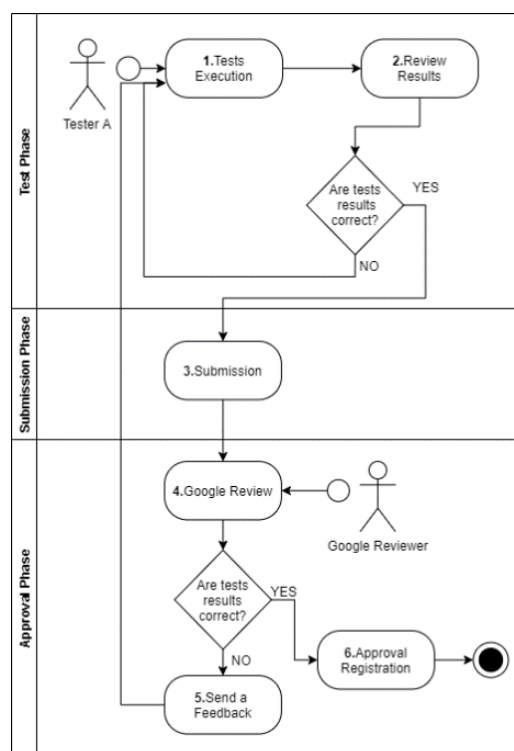

Fig. 1 GA Team process without Internal Review activity

The process starts when GA testers receive a release test

World Academy of Science, Engineering and Technology
International Journal of Computer and Systems Engineering
Vol:16, No:3, 2022

request. According to Fig. 1, the process is composed of following phases:

1) Tests Execution: A homologation request is received by the GA team, which, based on a certain software version, performs the necessary tests.
2) Review Results: From the results obtained, the tester makes a review of the files. If all generated files are correct, the results are ready to be submitted.
3) Submission: The team performs the submission of results obtained in tests required by Google.
4) Google Review: Google receives the files to review and verify if all requirements are being accomplished.
5) Send a Feedback: When Google finds an inconsistency in the obtained results, a feedback is sent to GA team for the correction to be made, generating a new result that will be updated in the submission system.
6) Approval Registration: When all generated results are correct, a GMS certification for the verified software version is sent to the GA team.

Due to massive demands, our GA Team faced strong difficulties to conduct tests and submission to Google review, that occurred increasing amount of received feedbacks. In order to improve quality of homologation test results, we adopt Internal Review during our testing process. In this way, we use peer review of test results. We aimed to avoid issues caused by lack attention or wrong test results. According to Fig. 2, all activities previously existing were maintained and just a new step was added in the process. A review in pairs must be performed by a tester who did not cooperate in the test execution. In case a reviewer finds some inconsistency, it is necessary to inform the main tester (Tester A) so that corrections are made before submission.

In this approach, GA team receives requests to perform tests. Afterwards, GA testers conduct all required tests, if there are no inconsistencies, tester forwards result to another tester and thus carries out peer review. So, if results are correct then results are submitted to Google review. In some cases, when the results have failures, Google reviewer asks tester to fix it. The tester fixes and asks reviewer to check the result again. To analyze the impact of this process improvement, we performed log data analysis and survey with some participants of the GA team. Our study relied on improvements identified by participants combining with tests requests from the data logs at JIRA about Internal Review. The next section describes study design and execution.

III. EVALUATING IMPACT OF INTERNAL REVIEW ADOPTION

As mentioned before, Internal Review process was a solution adopted to get a better timelines and quality of our deliveries. In this scenario, we decide to collect log data from before process adoption to after process adoption. Thus, we plan a study to characterize Internal Review process aiming to provide better understanding related improvements gained. For this reason, we use a Goal Question Metric (GQM) model [14]. GQM is an objective-oriented model aimed to measurement of software products and processes [15]. The goals of this study are shown in Table I.
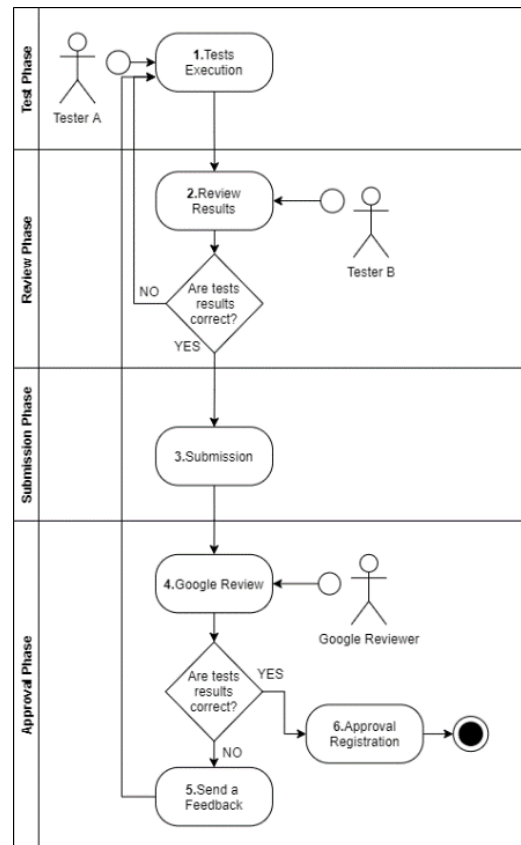

Fig. 2 GA Team process with Internal Review activity

TABLE I
GOAL OF EXPERIMENTAL STUDY USING GQM [15]

Analyze	Internal Review Process
With purpose of	Characterizing
In terms of	Efficiency (number of request with Internal Review / total requests approved without Feedback)
From point of view	GA Testers
In context of	Software Release Process

Following the first goal of the study, we collected data from 2020. We choose this period due increase GA test requests, caused by increasing projects demands. We combined qualitative data with log data collected from Jira, used to log GA test requests during release process [8]. For qualitative analyses, we collected answers from testers about their Internal Review perceptions, which provides us a general perspective how they felt and the activity value. In addition, the case study provides us more details gain of the new activity inserted in the process. It means, the survey and requests report complement each other. In Subsection *A* we described case study planning and Subsection *B* we explained about survey planning.

A. Study Planning

1. Data Collection

Sidia Mobile product area from Sidia uses Jira system as management tool for software release process. During release process, after release done by Software Project Leaders, they request GA tests to check GA requirements [8].

World Academy of Science, Engineering and Technology
International Journal of Computer and Systems Engineering
Vol:16, No:3, 2022

Due to massive demands, we decide to collect data from 2020. In this case, we extract log data from Jira considering the massive demand for the GA team in this period. We use a filter, according to Fig. 3, using the JIRA Query Language (JQL) to obtain Internal Review databases.

```
project = "Team Test Request" AND issuetype = "Scope Type Test" AND status = Closed AND
"GMS Final Status" = (Approved, Cancelled) AND "Test Type" in ("Category Test Request") AND
due >= 2020-1-1 AND due <= 2020-12-31 ORDER BY cf[11227] DESC, due ASC, status DESC,
assignee ASC, createdDate DESC
```

Fig. 3 The JQL query used to collect log data

As results we identify 6055 GA homologation requests along the year. We use to collect data due date between 01/01/2020 and 12/31/2020. From data collection, it was possible to verify that we have increasing demand during last year, as shown in Fig. 4.
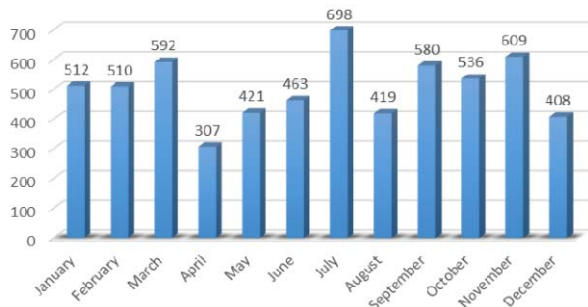

Fig. 4 GA test requests data collected in 2020

Fig. 4 shows that we had high demands due six months, especially on June, caused by releases of OS upgrade systems and other maintain releases. Furthermore, it was possible to see that 270 test requests were approved without review process. In this case, we identified that some requests were tested by experienced testers on emergency demands or due to adaptation process faced with a new activity and nowadays the pandemic scenario. For those reasons, it was not possible to adopt Internal Review to all requests. For this reason, it was not possible to adopt Internal Review. Next subsection presented quantitative results.

2. Quantitative Analysis Results

As indicator to characterize Internal Review process, we use efficiency, defined by reason between number of request with Internal Review and total requests reviewed and approved without Feedback. When analyzing the data, it is possible to see that the efficiency average has a difference positively around 6% between requests that apply Internal Review process, and requests without review. However, such comparison is hard due to the fact that data collected do not show which tests approved detected errors before submission results. Our comparative analyses were around efficiency when Internal Review process was performed or not.

Even though it is an important activity in the homologation process, for some reasons the Internal Review is not performed in 100% of requests. Despite those cases, we realized that GA team had a considerable annual gain, as can be seen in Table II.

TABLE II
GA TEAM PERFORMANCE MONTHLY.

2020	Review	No Review
January	92%	92%
February	93%	87%
March	92%	89%
April	89%	72%
May	90%	81%
June	91%	76%
July	87%	77%
August	88%	89%
September	89%	83%
October	89%	84%
November	93%	91%
December	90%	86%
Average	90%	84%

In Table II, a monthly analysis is shown regarding the performance rate, that it is a relation between approved homologation requests and feedbacks received. The main purpose of this analysis was to understand the relevance of including Internal Review activity in the homologation process.

In the first month of Internal Review, there was no difference between approved requests with Internal Review and without this activity. Only 1 month in the analyzed period showed no improvement when Internal Review was performed. Among the 12 months with Internal Review under analysis, 10 of them showed improvements in team performance.

In general, we can consider that there was an improvement around 6% when homologation requests were reviewed. Thus, the performance values measured in percentages and inserted in the chart were divided in relation to the months of the year studied and then separated in activities that were or were not employed the Internal Review process.

Also, other important results are shown in Table II, it is seen that throughout 2020 the activities that made use of Internal Review have a performance level that varies between 80% and 95%, the same shown in Table I in more detail and is maintained throughout the year studied with an average level of about 90%. However, the activities that did not make use of Internal Review have a greater variation, between 70% and 95%, as shown in Table I with an average level of 84%. Thus, the metric analyzed in relation to the homologation requests that use Internal Review presents a result with greater stability over the months and greater efficiency in relation to the requests without review in the same period. The difference between the annual averages of the performance levels

World Academy of Science, Engineering and Technology
International Journal of Computer and Systems Engineering
Vol:16, No:3, 2022

reinforces that the use of the Internal Review process in a homologation team is efficiency for a better performance of its activities with possible decrease feedbacks received. Furthermore, we decided to collect qualitative data regarding the tester's perception, through survey opinions, to better understand the quantitative results. In next subsection we present qualitative results.

TABLE III
QUESTIONS USED ON SURVEY OPINION

Categories	Questions
Relevant Process Perception	Q1: I consider that Internal Review activity is relevant to my team.
	Q2: I consider that Internal Review activity reduce delay risks.
Utility Perception	Q3: I consider that Internal Review activity is useful to reduce feedback quantity.
Contribution Perception	Q4: I consider that Internal Review activity improve the quality of my submissions

B. Survey

1. Survey Planning

In order to analyze the Internal Review activity, a survey containing four questions was produced using a Likert-scale [12]. Data were collected by creating a questionnaire using the TAM (Technology Acceptance Model) model as a basis. TAM investigates the acceptance of technology by users through the perception of utility and perception of ease of use [15]. This model was used as a basis to create questions about the user's interaction with the technology. The online questionnaire provides us subjective data related to utility, motivation and contribution and engagement perception. We also create two opened questions aimed to collect difficulties and improvement, to help us understand potential added value of the Internal Review process.

2. Participants

Due to massive demands on last year, we included participants that are able to answer the survey according to following criteria: who have been working for last 2 years in Sidia company. In this case, we chose 19 volunteer participants that contributed as a tester on the GA team between 2019 and 2020.

We did not include new members because they started working on the GA team after Internal Review was launched. In this case, they will have strong difficulty to compare differences from before and after Internal Review process released.

3. Qualitative Analysis of Survey

We received responses from all 19 selected participants. Among 19 survey responses, positive answers were provided to most questions, with 2 of them with 100% of positive answers.
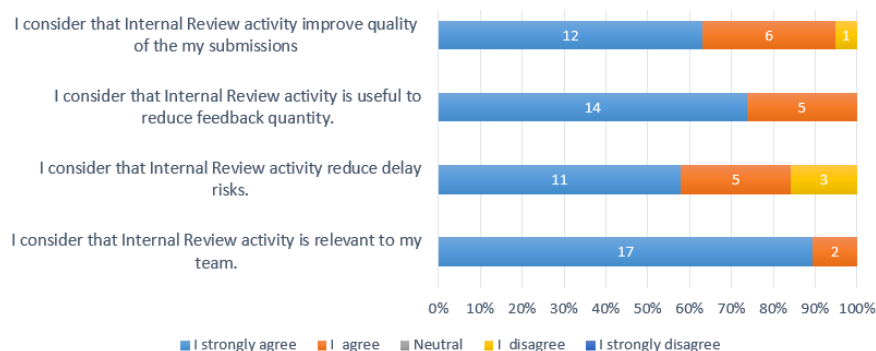

Fig. 5 Testers' perceptions results

Fig. 5 shows results regarding online survey opinion. Considering Relevant Process Perception, 64% strongly agreed that Internal Review activity improve quality of the submissions and other 31% agreed it. Finally, 5% disagreed that Internal Review is relevant. According to opened question, some participants related that in high demand GA team faced strong difficult to maintain Internal Review process, as mentioned by Participant 1, that said "Urgencies, no people to review, Review with a low quality and pressure to submit even without review.". Another important finding was reported by Participant 2, that said "Knowing how and what is important to validate and how to find some information about the doubt." and Participant 4 "Lack of tests results organization standardized.". Participant 9 also confirmed that "Transparent and common review points for everyone. Thereby we prevent or even reduce the possibility of the quality of the review dependent on each person's attention. Thus, each member would follow a standard review instead of each review in their own way.". These findings were important to define a guideline to review results generated by each tester. It will help to newcomers check important points during review testing.

Considering Utility Perception, Fig. 6 shows that of 100% participants consider that Internal Review activity is useful to reduce feedback received by issues. This result was confirmed on opened questions by some participants, as said Participant 7 "A review division into groups, I think can bring benefits to a review. It makes the organization easier and I believe people tend to be more aligned with tests.". However, another answered question captured some difficulties related tutorial outdated and information process. Participant 7 said that "Maintain the wiki updated. Offer workshops and training

World Academy of Science, Engineering and Technology
International Journal of Computer and Systems Engineering
Vol:16, No:3, 2022

with review points for the team from time to time.". Improvement GA team needs to develop a platform for knowledge sharing. As we use Wiki system, we will update some procedure and information to assist newcomers and new team members. Also, as future works we plan new training programs to share knowledge about Internal Review process.

About Contribution Perception, 84% of the participants consider that Internal Review activity reduces delay risks. However, 16% disagreed that activity reduce delay risks. One participant reported that "When there is a high demand it is sometimes hard to review while maintaining the quality due to the limited time to execute this activity". Related to these results, it may be necessary to use another indicator, such as effectiveness, also considering the execution time of each test including peer review.

Even though not being possible to conclude that Internal Review resolves problems related to failures and feedbacks, it is possible to check that we can minimize issues and feedbacks caused by issues on test results, see Fig. 6.
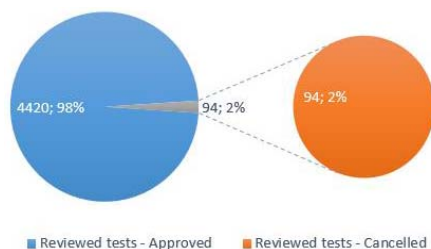

Fig. 6 Reviewed tests - Relation between approved and cancelled tests

As shown in Fig. 6, Internal Review Utility, 2% of all tests reviewed had issues found in the request that caused the request cancel and these issues were found during the internal review activity, reducing the time to find and report some issues and fix them and before the submission activity.

Despite not possible to conclude that Internal Review resolves problems related to feedbacks and failure, there are strong evidences that Internal Review activity brings improvements at homologation process, reducing risk of the delivery inconsistent results. In addition, it was very useful to collected improvement suggestions about the new activity.

Based on these results, there were some lessons learned:

- Lesson Learned #1: The Internal Review process can be considered an "internal body of knowledge" built by a GA team member. In this context, we need to create a guide for team members to follow how they can review each testing result.

With the Internal Review procedure, it was possible to provide greater quality during the homologation process, leading to lower error rates and quicker time-to-market. However, we need to improve the way the GA team needs to conduct peer review results.

- Lesson Learned #2: The Internal Review must be adopted for all homologation test requests, as a mandatory item contributing to the maintenance of process consistency. The historical data can be used to guarantee that the

Internal Review process really improves the homologation process.

The Internal Review has been able to provide standardization on the homologation process. This has led to an alignment between the team, especially during the peer review.

- Lesson Learned #3: We have strong shreds of evidence in the Internal Review process. However, it needs a comparative study considering requests without peer review dropped by failures and requests reviewed dropped by failures.

Continuous improvement is important to evaluate results' reliability. For this reason, it is necessary to extract log data considering failures between two approaches. In addition, another important finding is related to knowledge management by the team. We identified some gaps related to review process that can be minimized with training and workshops which will lead to even more time gain and even faster time-to-testing.

In the next section, we present our conclusion and future works.

IV. CONCLUSION AND FUTURE WORKS

This paper presented an experience report of the efficiency in Internal Review activity in a homologation team in GSD environment. In Sidia, the homologation process is an important and mandatory phase of software development, for this reason the inclusion of Internal Review is a significant improvement in the development process.

In order to validate the proposed activity, we carry out an assessment of efficiency of the Internal Review activity using two approaches. First, we collected data of the JIRA to analyze quantitatively and then we applied a survey in order to analyze and get more detailed answers and perceptions from the members involved, creating an overall perspective and value about the inserted activity in the team. The results from quantitative analysis show an increase in the homologation efficiency of 6% in relation to same period, but when it was not executing Internal Review activity.

For the qualitative results, we collected the testers' perceptions and thus it was possible to identify some difficulties to perform the Internal Review process. According to survey, we need to create a process pattern to review. In addition, we also identified the necessity to implement some tools that to facilitate Internal Review activity.

Finally, the combination between the case study and the qualitative analysis performed and described in Section III demonstrated that the insertion of the Internal Review activity obtained relevant gains of the performance to GA team with more tests approved without feedbacks caused by inconsistent results. Thus, more issues are found before submission phase, reducing time, rework and having approving in shorter time with increased quality.

As future work, we intend to improve the Internal Review activity with the development of auxiliary tools. We also plan to create a Wiki with mandatory review points in order to adopt a pattern process. Another improvement suggestion is

joining some strategies of Internal Review requests distribution to testers. After analyzing collected improvements, we can refine Internal Review activity and compare the results with this work.

REFERENCES

[1] Shrivastav, Supriya, and Maninder Kaur, 2020. "Software Development Testing Using Prioritization." Journal of Critical Reviews 7.17, pp. 2543-2546.

[2] Abdullahi, Shamsu, et al. "Software Testing: Review on Tools, Techniques and Challenges." International Journal of Advanced Research in Technology and Innovation 2.2, pp. 11-18.

[3] Spadini, Davide, et al. "When testing meets code review: Why and how developers review tests." 2018 IEEE/ACM 40th International Conference on Software Engineering (ICSE). IEEE, 2018, pp. 677-687.

[4] Cunha, R. F., Souza, F. B., Lima, F. O., & Bonifacio, B. A. (2020, June). How do newcomers learn work process in global software development (GSD)? a survey study from the perspective of newly project leaders. In Proceedings of the 15th International Conference on Global Software Engineering, pp. 71-75.

[5] Liyanage, Ketanvila Liyanage Kalhan Randula, 2019. Analyzing the Impact of Unit Testing Practices on Delivery Quality in an Agile Environment. Diss. Cardiff Metropolitan University

[6] M. El Bajta et al., 2018, "Software project management approaches for global software development: A systematic mapping study," Tsinghua Sci. Technol., vol. 23, no. 6, pp. 690–714.

[7] Chaudhary, Sarika, 2017. "Latest Software Testing Tools and Techniques: A Review." International Journal 7.5.

[8] Barbosa, Hualter O., et al, 2020. "Developing a release management tool to support global software development: an experience report on Android platform." Proceedings of the 15th International Conference on Global Software Engineering (pp. 117-121).

[9] Android Compatibility Program Overview. https://source.android.com/compatibility/overview (Accessed: 30-Jun-2021).

[10] Mukadam, Murtuza, Christian Bird, and Peter C. Rigby. 2013. "Gerrit software code review data from android." 2013 10th Working Conference on Mining Software Repositories (MSR). IEEE, (pp. 45-48).

[11] Cruzes, Daniela S., Nils B. Moe, and Tore Dybå. "Communication between developers and testers in distributed continuous agile testing." 2016 IEEE 11th International Conference on Global Software Engineering (ICGSE). IEEE, (pp. 59-68).

[12] Joshi, Ankur, et al. "Likert scale: Explored and explained." British Journal of Applied Science & Technology 7.4 (2015): 396.

[13] Hassan, Mujtaba, Muzammil Hussain, and Maham Irfan, 2019 "A Policy Recommendations Framework To Resolve Global Software Development Issues." 2019 International Conference on Innovative Computing (ICIC). IEEE, (pp. 1-10).

[14] SAGP, (2008) Software Acquisition Gold Practice - Goal-Question-Metric (GQM) Approach, https://www.goldpractices.com/practices/gqm/. (Accessed: 12-Jul-2021).

[15] Davis, Fred D. "Perceived usefulness, perceived ease of use, and user acceptance of information technology." MIS quarterly (1989): (pp. 319-340).