Exploring the influence of the emotional state on the efficiency for correcting defects in conceptual models: a live study

Franci Suni-Lopez
Universidad Nacional de San Agustín
de Arequipa, Perú
fsunilo@unsa.edu.pe

Angela Mayhua
Universidad La Salle
Arequipa, Peru
amayhua@ulasalle.edu.pe

Maria Fernanda Granda
Computer Science Department
Universidad de Cuenca, Ecuador
fernanda.granda@ucuenca.edu.ec

Abstract

The purpose of this live study is to investigate the impact of emotional states on the efficiency of correcting defects within a Conceptual Model verification context. We plan to use the CoSTest tool for verifying defects and measure the experimented emotions during the testing tasks through I-PANAS-SF and SUDS instruments. Furthermore, the design, overview and potential threats of this experiment are presented.

1 Introduction

1.1 Research problem

Emotions have an important role in our life and their influence on our work is not an exception. According to Weiss and Cropanzano [WC96], emotions and moods have effects on work performance, affecting the duration of activities and the perceived experience. This situation is not only present in industrial activities; some researchers have been analyzing the influence and importance of emotions in software engineering [Wro13, CSdlB+14, MF15, RFB+19], being the software development process one of the principal centers of interest.

Wrobel [Wro13] analyzed the influence of the emotional state on the performance of programming tasks collecting data through a questionnaire and an interview for each participant, the Job Emotions Scale (JES) was applied to measure human emotions. Romano *et al.* [RFB+19] also investigated in this area, analyzing emotions when novice developers apply the TDD (Test-Driven Development) approach on change tasks; in a similar way, they collected data through self-reported emotions by the participants. In the same way, we propose to analyze the influence of the emotional state on the work productivity based on correcting defects in conceptual models.

In line with this notion, according to Granda *et al.* [GCVP15], defects (*e.g.,* missing, wrong and unnecessary elements) in conceptual models (*e.g.,* UML class diagram) can be located in several ways through Validation & Verification techniques, which use a detection mechanism (based on rules, metrics, and modelling conventions)

In: M. Sabetzadeh, A. Vogelsang, S. Abualhaija, M. Borg, F. Dalpiaz, M. Daneva, N. Condori-Fernández, X. Franch, D. Fucci, V. Gervasi, E. Groen, R. Guizzardi, A. Herrmann, J. Horkoff, L. Mich, A. Perini, A. Susi (eds.): Joint Proceedings of REFSQ-2020 Workshops, Doctoral Symposium, Live Studies Track, and Poster Track, Pisa, Italy, 24-03-2020, published at http://ceur-ws.org

for this purpose. According to the technique nature, this can be statically or dynamically supported by a tool and they can have different scope and limitation depending on its purpose (*i.e.,* detect, prevent and resolve).

1.2 Motivation to conduct the study

In this live study proposal, we start from the hypothesis that the emotional state of workers influences on their productivity when they correct defects during the software development process [WC96]. It could be influenced by the user profile (*e.g.,* background, preferences, experience, personality) or the way how the emotions are experiment for each one. Moreover, we focus on investigating the usefulness and user experience of using CoSTest tool [GCVP17a], which is a tool created to help in the tasks of defect detection and correction.

Therefore, in order to investigate 1) what emotions are experienced by the subjects during the correction of defects using the tool, and 2) how useful the testing tool is perceived for supporting the verification of conceptual models, we present the design and plan of a live study to be conducted with the REFSQ attendees.

2 Scenario

In this section, we present the scenario that illustrates the context of the use of CoSTest tool [GCVP17a], which takes as input a requirements model based on communication analysis to generate test cases and validate a conceptual model (*i.e.,* UML class diagram) evaluating its correctness and completeness.

The software development process includes different tasks, where testing and correction tasks are important to assure quality. Requirements engineers, analysts, and testers have an important role inside these tasks, using a UML modeling editor and the CoSTest tool to detect defects and verify the changes in the UML class diagram. Doing these tasks, they commonly experience a range of emotions (*e.g.,* anger, frustration, stress, anxiety, happiness), that could be generated by the use of the tool; knowing these emotions let us find potential feedback to correct and improve CoSTest tool.

3 Study Design

3.1 Goal and research questions

The live study proposal aims to 1) *analyze* the influence of emotional state on the efficiency for performing defect correction tasks, and 2) *evaluate* the utility and experience of using the CoSTest tool [GCVP17a] as a support in the correction of conceptual model defects.

From this goal, the following research questions are derived:

RQ$_1$: *How does the emotional state influence on efficiency for performing defect correction tasks?*

RQ$_2$: *Is the CoSTest tool perceived as useful to support the verification of conceptual models?*

Type of study: to investigate this research problem, we propose to conduct a *quasi-experiment*, where all participants have the same set of defects to be corrected and will use CoSTest[1] as a support tool.

Variables and metrics: the following variables were identified: **independent variables:** *CoSTest tool* (used to automatically detect defects in conceptual models). Other factors which can impact the results are the selected *Conceptual models (CM)* (see Section 3.3.2 for more details about them) and the *Defects injected* into the CMs. By using the defects classification proposed in [GCVP15], we will use three defect types (*i.e.,* missing, wrong and unnecessary elements) to inject six defects per CM (two for each defect type). As **dependent variables:** *user emotional state* that can be determined by analyzing the self-reported emotions through specific questionnaires (see Section 3.3.1 for more details); and *efficiency for correcting defects* that represents the relationship between the total corrected defects and the total time assigned for correcting all defects.

3.2 Population of interest

3.2.1 Profile of the intended subjects

We plan for 30 participants. Students, senior researchers, and practitioners are very welcome and need to bring their laptops to execute CoSTest tool, which is based on Java (jar file). In order to expedite the installation of CoSTest and avoid compatibility problems with Java versions, we will provide a virtual machine with all the software required in this study. Prior knowledge and experience on modeling UML-based conceptual schemas (*i.e.,* class diagrams) using tools (*e.g.,* UML2Tools editor[2]) is required. However, some tips to remember or

[1] https://costestproject2017.wordpress.com/

[2] https://www.eclipse.org/modeling/mdt/?project=uml2tools

reinforce its use will be included in the training phase. We choose REFSQ'20 to run this experiment thanks to the strong knowledge in testing and knowledge in model-driven development process of REFSQ's attendees.

3.2.2 Benefits to the subjects of participating in the study

The testing tasks are the key leverage point for practitioners (*e.g.,* project managers, analysts, testers) who want to develop software systems with high quality level. In this context, we think that this study might benefit to participants by getting:

- Knowledge and training on a tool to support the verification of conceptual models through test cases, which could be applied on their workplaces or development tasks.
- Experience in running studies to analyze emotions.
- Access to instruments for evaluating and measuring emotions, such as International Positive and Negative Affect Schedule Short-Form (I-PANAS-SF) and Subjective Units of Distress Scale (SUDS), which could be used for the participants in other studies based on human emotions.

3.3 Instrumentation

3.3.1 Questionnaire

We implement a web-based survey using the Qualtrics tool[3], which is composed by the following parts: 1) a demographic and educational background questionnaire (*e.g.,* sex, age, educational degree, domain expertise); 2) CoSTest training, we provide the required material to use CoSTest tool; 3) Relaxing, during this period we ask to participants get relaxed; 4) Testing task, we provide two conceptual models with different defects for correcting, also participants have an option to upload their responses; 5) Emotion state questionnaires; in order to understand the different emotions experimented by the participants during the contest, we use the I-PANAS-SF questionnaire and the SUDS scale. The I-PANAS-SF is a list of 10 adjectives used to describe different emotional states: 5 states of Positive Affect (PA) and 5 states of Negative Affect (NA). The PA scale measures activity and pleasure, while the NA scale relates to fear and stress [KWR11, Tho07]. The Subjective Units of Distress Scale (SUDS) can be used to measure the intensity of anxiety/distress [BOC$^+$10]; this instrument rates on a scale from 0 to 10, so that subject marks a point on the range where he perceives his anxiety/distress state has been located during the different situations presented on the contest. 6) CoSTest evaluation, we propose different questions in order to investigate whether CoSTest tool was useful to test conceptual models, and 7) Experiment feedback, we collect feedback about the study.

3.3.2 CoSTest: testing tasks

CoSTest [GCVP17a] is a tool to verify automatically conceptual models using test cases from an existing approach to communication-oriented requirements engineering. In this live study , we will ask the participants to carry out two testing tasks (*i.e.,* two conceptual models, CM1 and CM2) and run three test cases associated with each one by using the CoSTest tool. Participants will be asking to consider one test case at a time (starting with the first one); if the result of the test case shows the conceptual model has a defect, they have to correct it and rerun the test case to make sure that the defect correction was done well. The participants will be able to perform the next test task when finishing the current one. The total time allocated to perform the two test tasks is 30 minutes, if the participant finish before ending time, he could try to generate new test cases in the tool to find additional defects. The idea is to motivate participants to use the tool so that this will help to measure their perception of the usefulness of CoSTest.

The experimental objects consist of the specifications of two conceptual models: CM1 is a Super Stationery (SS) system, where CM defines the information system of a company that provides stationery and office material to its clients; and CM2 is a Photography Agency (PA) system where a CM defines the information system that manages photographers and their photographic reports for distribution to newspaper publishers. Regardless of the experimental object, we provide the participants with the following experimental material: (i) a brief description for each information system modeled in CM1 and CM2; (ii) a template to create test cases using CoSTest; and (iii) an example test suite (each one with three test cases), developed by the authors to test each CM using CoSTest. We opted for SS and PA as experimental objects because they are often adopted to learn/practice CoSTest and were used in past empirical studies on Mutation Testing [GCVP17c].

[3]https://www.qualtrics.com

3.4 Procedure

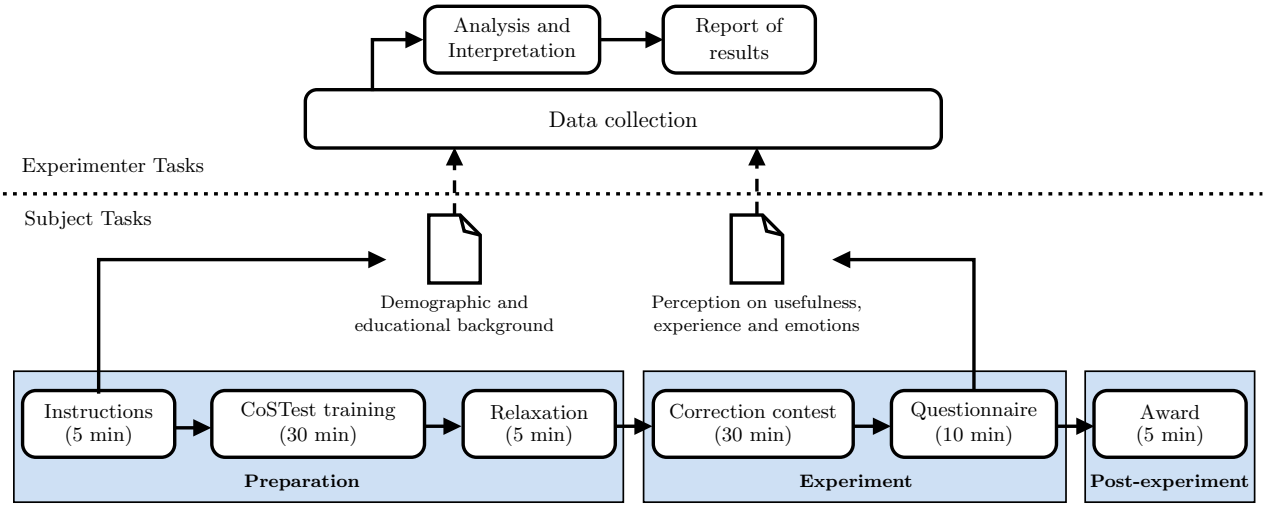

Figure 1: Live study's procedure

The study is composed of three phases, as shown in Figure 1. (i) *Preparation*: first we explain details about the study and request to read and sign the informed consent form, where it is detailed the privacy and confidentiality terms; after we ask to participants with a demographic and educational background questionnaire (five minutes). As the CoSTest could be a new tool for the participants, we will give a training phase about how to use it (30 minutes). As next step, we need to uniform the emotional state of all participants (*e.g.,* someone could came to the experiment already stressed) to avoid the influence of previous emotions in our experiment; for that reason, participants are asked to stay quiet and watch a video during five minutes to get relaxed. (ii) *Experiment*: in this phase, we provide different test cases for using the CoSTest tool (CMs are delivered in random order), participants will have one attempt to send their solution for each task. With the objective to generate competitiveness, during the experiment each time that a participant submit a correct solution of a testing task, we will give a globe (similar to ICPC[4]). After the contest, participants will be asked about their perceived emotions during the experiment, their experience using the CoSTest tool, and their feedback regarding the study (see Section 3.3.1 for more details about these questionnaires). (iii) *Post-experiment*: after to process all submissions of the participants, we will reward three participants who detect and correct more defects in the less time.

4 Threats to validity and ethical issues

Internal validity

The different factors triggered by live study (*e.g.,* place, settings) might affect the observed variables. We mitigate this threat by performing the study in similar conditions for each participant (*e.g.,* material, testing tasks, rules of contest). The tools (i.e. CoSTest and UML2Tools Editor) used in this experiment will be tested to adjust the settings on the virtual machine used for software installation. For this purpose, we will use 5 reference points, taken from other MCs, which are not part of this study. Additionally, our study could be negatively affected if both artifacts (MCs and injected defects) were not properly selected. The MCs have been taken from other experiments that have been carried out with the tool [GCVP17b] and the defects are those generated by the mutation tool [GCVP17c]. Another threat is the emotions of the participants before starting this live study (due to some previous activities or experiences on the day) might affect the perception of their emotional state during this study; to mitigate this threat, we have prepared a relaxing activity to uniform the emotions of all participants before starting this experiment.

[4]https://icpc.baylor.edu/

External validity

This validity is regarding the generalization of our results outside the experiment setting; in this context, a possible threat could be the selection of participants. Nonetheless, it is mitigated by the REFSQ attendees, because they have different personalities, experiences and educational backgrounds, such as master/PhD students, senior researchers, and practitioners. Another possible threat is that the environment where this study will be carried out might not be a typical context for doing software work; to mitigate this threat, we will conduct the study in a quiet environment, trying to limit unnecessary distractions and trying to give comfortable work-spaces for the participants.

Construct validity

As part of this study, our instruments are based on questionnaires with self-reported responses and as consequence, participants could be afraid of giving information about their emotional states or personal information; however, this threat is mitigated through our privacy and confidentiality terms that specify their information and responses are going to be anonymous. Furthermore, the selected instruments are well known and have been used in other works to measure emotions [Epp16, AB18]. Some other possible threat is determining the correctness of a solution for a defect because it could affect the measure of efficiency; this threat is mitigated by the CoSTest tool, which verifies automatically if the defect was solved successfully or not.

5 About the researchers

Franci Suni Lopez is a PhD candidate at National University of San Agustin of Arequipa. His main research focuses on using human emotions to empower the self-adaptation capability of software services. His research interest also includes software engineering for mobile development and self-adaptive software. He executed one empirical studies in the live study track of REFSQ (whose results have been published in [CFML19]) and an experiment at MEGSUS [SLCFMA18].
Angela Mayhua Quispe is researcher at La Salle University (Arequipa, Peru) and PhD candidate at National University of San Agustin of Arequipa. Her main research focuses on data visualization and pattern recognition from images and signals. She has a particular interest in analyzing physiological signals to recognize human emotions, with special emphasis on negative emotions.
Maria Fernanda Granda is a full professor and researcher of the Computer Science Department at University of Cuenca (Ecuador). She obtained her Ph.D in Computing in 2017 at the Universitat Politècnica de València (Spain). Her main research focuses on Requirements Engineering, Software Testing & Quality and Model-driven Development. She has executed empirical studies using mutation testing and the results have been published in [GCVP17b], [GCVP17c] and [GCVP16].

Acknowledgements

This work has been supported by the KUSISQA project (014-2019-FONDECYT-BM-INC.INV) from the National Fund for Scientific and Technological Development (FONDECYT-PERU) and World Bank, and the Dirección de Investigación de la Universidad de Cuenca (DIUC) - Ecuador. Authors also want to acknowledge the anonymous reviewers and Live Study track chairs (Nelly Condori-Fernandez and Luisa Mich) for their valuable feedback.

References

[AB18] A. Alhubaishy and L. Benedicenti. Affect and affective trust in agile requirements engineering. In *2018 1st International Workshop on Affective Computing for Requirements Engineering (AffectRE)*, pages 9–12, Aug 2018.

[BOC+10] Courtney L. Benjamin, Kelly A. O'Neil, Sarah A. Crawley, Rinad S. Beidas, Meredith Coles, and Philip C. Kendall. Patterns and predictors of subjective units of distress in anxious youth. *Behavioural and Cognitive Psychotherapy*, 38(4):497–504, 2010.

[CFML19] Nelly Condori-Fernandez, Denisse Muñante, and Franci Suni Lopez. Exploring users perception on security and satisfaction requirements of context-aware applications: An online survey. In *2019 Joint of International Conference on Requirements Engineering: Foundation for Software Quality*

Workshops, Doctoral Symposium, Live Studies Track, and Poster Track, REFSQ-JP 2019, volume 2376 of *CEUR Workshop Proceedings*. CEUR Workshop Proceedings, 3 2019.

[CSdlB+14] Broderick Crawford, Ricardo Soto, Claudio León de la Barra, Kathleen Crawford, and Eduardo Olguín. The influence of emotions on productivity in software engineering. In Constantine Stephanidis, editor, *HCI International 2014 - Posters' Extended Abstracts*, pages 307–310, Cham, 2014. Springer International Publishing.

[Epp16] Carrie Demmans Epp. English language learner experiences of formal and informal learning environments. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, LAK '16, page 231–235, New York, NY, USA, 2016. Association for Computing Machinery.

[GCVP15] Maria F. Granda, Nelly Condori-Fernández, Tanja E. J. Vos, and Oscar Pastor. What do we know about the defect types detected in conceptual models? In *2015 IEEE 9th IEEE International Conference on Research Challenges in Information Science (RCIS)*, pages 88–99, May. 2015.

[GCVP16] Maria F. Granda, Nelly Condori-Fernández, Tanja E. J. Vos, and Oscar Pastor. Mutation operators for uml class diagrams. In *2016 28th International Conference on Advanced Information Systems Engineering*, volume 9694, pages 325–341, Jun 2016.

[GCVP17a] Maria. F. Granda, Nelly Condori-Fernández, Tanja E. J. Vos, and Oscar Pastor. Costest: A tool for validation of requirements at model level. In *2017 IEEE 25th International Requirements Engineering Conference (RE)*, pages 464–467, Sep. 2017.

[GCVP17b] Maria F. Granda, Nelly Condori-Fernández, Tanja E. J. Vos, and Oscar Pastor. Effectiveness assessment of an early testing technique using model-level mutants. In *2017 21st International Conference on Evaluation and Assessment in Software Engineering (EASE)*, pages 98–107, Jun. 2017.

[GCVP17c] Maria F. Granda, Nelly Condori-Fernández, Tanja E. J. Vos, and Oscar Pastor. A model-level mutation tool to support the assessment of the test case quality. In *2017 Lecture Notes in Information Systems and Organisation*, volume 11, pages 17–37, 2017.

[KWR11] Jahanvash Karim, Robert Weisz, and Shafiq Ur Rehman. International positive and negative affect schedule short-form (i-panas-sf): Testing for factorial invariance across cultures. *Procedia - Social and Behavioral Sciences*, 15:2016 – 2022, 2011. 3rd World Conference on Educational Sciences - 2011.

[MF15] Sebastian C. Müller and Thomas Fritz. Stuck and frustrated or in flow and happy: Sensing developers' emotions and progress. In *2015 IEEE/ACM 37th IEEE International Conference on Software Engineering*, volume 1, pages 688–699, May 2015.

[RFB+19] Simone Romano, Davide Fucci, Maria Teresa Baldassarre, Danilo Caivano, and Giuseppe Scanniello. An empirical assessment on affective reactions of novice developers when applying test-driven development. `https://arxiv.org/abs/1907.12290`, July 2019.

[SLCFMA18] Franci Suni-Lopez, Nelly Condori-Fernandez, and Denisse Muñante Arzapalo. End-user perceptions on social sustainability in context-aware applications: Validation of an experiment design. In *4th International Workshop on Measurement and Metrics for Green and Sustainable Software Systems (MEGSUS)*, pages 30–37, Oulu, Finland, October 2018. CEUR-WS.org.

[Tho07] Edmund R. Thompson. Development and validation of an internationally reliable short-form of the positive and negative affect schedule (panas). *Journal of Cross-Cultural Psychology*, 38(2):227–242, 2007.

[WC96] Howard Weiss and Russell Cropanzano. Affective events theory: A theoretical discussion of the structure, cause and consequences of affective experiences at work. *Research in Organizational Behavior*, 01 1996.

[Wro13] Michal R. Wrobel. Emotions in the software development process. In *2013 6th International Conference on Human System Interactions (HSI)*, pages 518–523, June 2013.