Ooops, I replicated again. Let me tell you why!

Cleyton V. C. de Magalhães
Centro de Informática - UFPE
Av. Jorn. Aníbal Fernandes, s/n –
Cidade Universitária, Recife, Brazil
cvcm@cin.ufpe.br

Teresa Baldassarre
Dipartimento di Informatica – Università degli Studi di Bari, Italy
mariateresa.baldassarre@uniba.it

Fabio Q. B. da Silva
Centro de Informática - UFPE
Av. Jorn. Aníbal Fernandes, s/n –
Cidade Universitária, Recife, Brazil
fabio@cin.ufpe.br

Ronnie E. S. Santos
Centro de Informática - UFPE
Av. Jorn. Aníbal Fernandes, s/n –
Cidade Universitária, Recife, Brazil
ress@cin.ufpe.br

ABSTRACT

What is this study about? The first replications of an empirical software engineering (ESE) study can be traced back to the early 1990s. Since then, the ESE research community has engaged in performing replications and also studies about conceptual issues regarding how to perform it. From these conceptual works, definitions, taxonomies, and guidelines have emerged. In addition, recent and comprehensive mapping studies discussed that the number of replications being published are increasing progressively nowadays. Further, such reviews demonstrated that performing and reporting replicated studies are involving different goals and objectives, leading to an inconsistent use of terminologies, taxonomies and guidelines. In conclusion, researchers are performing replicated studies in SE, but the question is, what are their motivations to replicate?

Why is this study important? There are several studies in the literature addressing issues related to replications, but none of them brings full guidelines with deep information on how to perform replications, the possible changes between them and the original studies or even their purposes. In this context, two recent mapping studies identified and analyzed studies that addressed replications performed between 1994 and 2012, and they stated the following conclusions about replication in SE. In summary, da Silva et al. [1] and Bezerra et al. [2] both arrived at similar conclusions: a) Papers reporting replications did not provide complete information about their original studies, making it difficult to compare results; b) The absolute number of replications (133) and original studies (72) was considered small; c) Most replication sets were small (1 replication per set). Complementary, Magalhães et al. [3] performed another mapping study that analyzed and discussed the contents of a set of studies about replications and identified some issues: a) The number of papers about replication (37 papers were analyzed) could also be considered small; b) A considerable number of studies did not provide a definition replication and those that provided a definition, used different and sometimes non-rigorous and incompatible definitions; c) Types and roles of replications proposed were not uniform. Considering the mentioned concerns, one can observe that when starting a replication in software engineering, researchers could face questions such as "what is the exact role of a replication in SE?" or "what are the motivations to perform a replication and why should a replication be performed in SE?". These uncertainties lead researchers to disagreements about how replications should be undertaken and consequently to an impact on the way they are performed. Therefore, in this paper, our goal is to investigate what motivates software engineering researchers to perform their replications.

How is this study being executed? We performed a cross-sectional survey using 135 replications as the target population. These replications were performed by 295 researchers between 1994 and 2012, identified in two systematic mapping studies. We received 40 complete answers covering 38 replications (28%, 38/135, response rate) in which we identified 5 main motivations to perform replications: a) to confront or confirm the results of a new study in contrast to previous findings; b) to improve the research design and procedures of a previous study; c) to increase the external validity of results obtained in a previous study; d) personal researchers' motivation, in particular regarding the improvement of scientific and research skills; e) to understand the costs associated to the research development and effort reduction in future studies.

Discussions: The main motivation mentioned by the authors was the use of a replication to generalize the results of an original study to a different population. This motivation is consistent to the main goal of replications performed in different research fields and identified in the literature, where the main goal of such studies was also related to generalization. In cases like this, we believe that in some points the authors are consistent when they define their main objective when performing a replication, although, they have mentioned motivations such as "improving skills", which is was not identified in previous studies whatsoever. Therefore, this research reinforces the need for clear cut definition of functions and goals for replications in SE and how to achieve the objectives related to each of them. In conclusion, this paper contributes to a preliminary identification of researchers' motivations and objectives while performing replications, and lead to future research efforts to determine and improve consistency among replications in SE.

Categories and Subject Descriptors

• CCS → Software and its engineering

Keywords

Replication, empirical software engineering, empirical studies.

References

[1] Da Silva, F. Q., Suassuna, M., França, A. C. C., Grubb, A. M., Gouveia, T. B., Monteiro, C. V., & dos Santos, I. E. (2014). Replication of empirical studies in software engineering research: a systematic mapping study. Empirical Software Engineering, 19(3), 501-557.

[2] Bezerra, R. M., da Silva, F. Q., Santana, A. M., Magalhaes, C. V., & Santos, R. E. (2015). Replication of Empirical Studies in Software Engineering: An Update of a Systematic Mapping Study. In Empirical Software Engineering and Measurement (ESEM), 2015 ACM/IEEE International Symposium on (pp. 1-4). IEEE.

[3] de Magalhães, C. V., da Silva, F. Q., Santos, R. E., & Suassuna, M. (2015). Investigations about replication of empirical studies in software engineering: A systematic mapping study. Information and Software Technology, 64, 76-101.