Dealing With Cultural Dispersion: a Novel Theoretical Framework for Software Engineering Research and Practice

Artifact Evaluation Readme

1 Introduction

This is the artifact presentation file of the paper titled **"Dealing With Cultural Dispersion: a Novel Theoretical Framework for Software Engineering Research and Practice"** published at ICSE-SEIS 2024. The main contribution of the published paper is a preliminary and foundational theoretical framework that encapsulates the effect of individual cultural diversity in software development teams, offering strategies to mitigate the challenges faced by practitioners in such contexts. Specifically, the paper delves into investigating three aspects:

1. the impact of *cultural dispersion*—i.e., the degree to which a community is different in terms of its members' cultural attitudes—on the software development lifecycle by means of its *positive* and *negative effects*;
2. the *strategies* practitioners adopt to *mitigate* such negative effects and *support* the emergence of positive ones.
3. the *co-starring factors* influencing the emergence of the effects of cultural dispersion, e.g., other cultural attitudes, working habits, and ambient characteristics.

In order to perform the analysis process, we used *Atlas.ti*, a well-known tool for qualitative data analysis. Such an instrument allowed us to import the transcripts and audio records to manage and organize them better during analysis. It is worth noting that we used it mainly for organizing data, not for analysis (even if this is now possible because of the recent advancements in AI and LLMs); all the analysis processes are entirely independent of the technology. Therefore, we believe it is not crucial to emphasize downloading this tool as a requirement for the replication package. Researchers can use other applications or handle tasks manually with Excel instead. Also, it is worth noting that this platform is not free. Importantly, choosing a different tool should not affect the results that, in any case, come out from a qualitative methodology.

2 Purpose, Objective, and Provenance

In our work, we propose an online appendix supporting the research. Our work consisted of a grounded theory study [1]; the data mainly consisted of (1) the material used for interviews, (2) the codes resulting from the qualitative data analysis (i.e., the codebook), (3) the quotations supporting the reliability of our codebook (inserted in the codebook), and (4) the theory itself. These can be used to conduct other interviews following our same agenda and deductively using our codebook to augment the theory.

In the context of the ICSE 2024 Artifact Evaluation track, we want to apply for receiving both the ACM **available** and **reusable** badge; we relied on the criteria reported on the ACM Artifact Review and Badging site.[1] Moreover, our online appendix [2] contains *non-executable artifacts*. For this reason, we followed the *non-executable* artifacts criteria. The reasons supporting our application are reported in the following paragraphs.

Available Badge. To satisfy these criteria, all the artifacts used in this article have been made publically available in our online appendix [2] on *Figshare* and reachable by using both a DOI[2] and a link.[3]

Reusable Badge. To satisfy these criteria, we have added a readme file (this one) in our appendix that has been formatted following the requirements of the artifact evaluation track. Moreover, we have inserted all the files relevant to our research in the appendix and described how to use them for future research (more details in the following Sections). Furthermore, we followed the ACM Empirical Standards for our research[4] (as also confirmed by one of the paper's reviewer checklists); to assure reliability and reusability, we included all the requested items for qualitative studies both in the appendix and the main research paper.

In order to provide sufficient material to evaluate our data, here we reported a list of useful references:

Paper Pre-Print. It is the pre-print of the published paper; as requested, we inserted it in the main online appendix [2].

Online Appendix [2]. It is the online appendix containing the artifacts produced for our contribution.

3 Data—Artifacts and How to Use

For our study, we adopted semi-structured interviews (for data gathering) and Socio-technical grounded theory for data analysis [1]. Thus, our artifacts consist of the material for conducting the interviews, the codebook, the quotations, memos extracted during data analysis, and the final developed theoretical framework.

[1] ACM Artifact Review and Badging site: `https://www.acm.org/publications/policies/artifact-review-and-badging-current`

[2] Online Appendix DOI: doi:10.6084/m9.figshare.24243928

[3] Online Appendix link: `https://figshare.com/articles/conference_contribution/Dealing_With_Cultural_Dispersion_a_Novel_Theoretical_Framework_for_Software_Engineering_Research_and_Practice_Online_Appendix/24243928`

[4] Available at the following link: `https://github.com/acmsigsoft/EmpiricalStandards`. Given the nature of the study and the currently available standards, we followed the *"General Standard"*, *"Grounded Theory"*, and *"Qualitative Surveys"* guidelines.

There are no technical requirements for using our artifacts since the first can be used mainly for conducting other interviews similar to ours and the second to evolve the theory (the third one) in a deductive approach. Thus, the only required skills are qualitative research knowledge and the grounded theory approach.

Following, we describe each of the primary artifacts included in the online appendix [2].

Interview Protocol and Slides. We formulated an interview protocol comprising of three sections: demographic, introduction, and core. The initial sections were designed to establish rapport and initiate the conversation, while the last section contained questions for data collection. Moreover, we developed a set of 4 slides (file InterviewSlides) that (1) provided a definition of culture and cultural dimensions and (2) showcased examples illustrating how cultural differences can affect technical and social issues as demonstrated in state of the art. Given the semi-structured nature of the interviews, the protocol served as a guide, allowing us to shape the conversation step by step to suit the unique aspects of each case.

Other researchers could use our protocol in the online appendix (file InterviewProtocol) to conduct ulterior interviews and collect new data. Considering that the optimal duration of an interview is around 60 minutes, the interviewer should spend—at most—5 minutes for the demographic part and 5 minutes for the introduction part. During the introduction part, researchers should use and read the first slide to introduce the objective of the research and the second one to provide the interviewees with some foundational information on cultural aspects.

All the remaining time should be used for the core part. Researchers should project the third slide as background during the entire duration of the core section of the interview. In this slide, the list of the interview questions is reported to help the interviewee stay focused and support the interviewer in asking the questions. If the interviewer notices that, at some point, the participant is stuck, the interviewer could use the fourth slide to help return to focus. This last slide contains examples of potential impacts—already documented by state of the art—that culture could have on software development; these could help the interviewee enter the context of the work and provide confirming details from past experiences.

Codebook. Our codebook (file CodeBook) is an excel file that contains the *codes*, *categories*, and *concepts* identified during our qualitative analysis; they have been extracted and used as building blocks of our theoretical frameworks. Moreover, according to the ACM Empirical Standards, we have inserted some of the quotations supporting our coding activity. The details of all the participants, their organizations, and interviews have been kept confidential as per the human ethics guidelines followed in this study. Participants reserve the right to consider sending transcripts to the authors upon request.

The codebook serves dual purposes: firstly, as the fundamental outcomes derived through inductive analysis, showcasing a pivotal contribution and substantiating the essence of our work; secondly, as a tool for deductive exploration in subsequent research endeavors, fostering the evolution of the theoretical framework.

Researchers could take advantage of the *Atlas.ti* feature to import the codebook into their own project ("import codebook" from the top menu after creating a project in the tool) and use it for performing additional qualitative data analysis. Moreover, as reported by the *Atlas.ti* website, the codebook (in an Excel format) could be imported into other qualitative analysis tools (e.g., MAXQDA and NVivo).

Memos. The memos from our qualitative analysis have been compiled into a PDF file accessible in our online appendix (file Memos). These memos encapsulate reflections and ideas that emerged throughout the entire process, steering the discernment of relationships between categories within our framework.

Primarily, these memos are intended to offer fellow researchers valuable insights, enhancing the comprehensibility of our findings and stimulating considerations for future research directions.

Theory of Dealing With Cultural Dispersion. The outcome of our research is the novel theoretical framework encompassing the impact of cultural dispersion on software development. Our framework (1) was developed through the use of memos to link the categories and concepts in the codebook, and (2) provides a cohesive overview of the influence of cultural differences on the software development lifecycle.

As mentioned before, researchers could use our framework—other than clearly as the main output of our research—to augment it and build novel versions in a collaborative effort to advance the state of the art by following a single theoretical structure (file Theoretical Framework).

License. The license under which our artifacts are distributed. This online appendix (TITLE: Dealing With Cultural Dispersion: a Novel Theoretical Framework for Software Engineering Research and Practice—Online Appendix) is made available under the Open Database License: `http://opendatacommons.org/licenses/odbl/1.0/`. Any rights in individual contents of the database are licensed under the Database Contents License: `http://opendatacommons.org/licenses/dbcl/1.0/`.

References

[1] Rashina Hoda. Socio-technical grounded theory for software engineering. *IEEE Transactions on Software Engineering*, 48(10):3808–3832, 2021.

[2] Stefano Lambiase, Gemma Catolino, Bice Della Piana, Filomena Ferrucci, and Fabio Palomba. Dealing with cultural dispersion: a novel theoretical framework for software engineering research and practice—online appendix, 2023. `doi:10.6084/m9.figshare.24243928`.