How do Cultural Characteristics and Software Engineering Practices Interplay?
A Comparative Study Between Indonesia and Sweden
Bachelor of Science Thesis in Software Engineering and Management
AMEERA DARWISH
ANNA HENRYSON

Department of Computer Science and Engineering
UNIVERSITY OF GOTHENBURG
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019
The Author grants to University of Gothenburg and Chalmers University of Technology the non-exclusive right to publish the Work electronically and in a non-commercial purpose make it accessible on the Internet.
The Author warrants that he/she is the author to the Work, and warrants that the Work does not contain text, pictures or other material that violates copyright law.

The Author shall, when transferring the rights of the Work to a third party (for example a publisher or a company), acknowledge the third party about this agreement. If the Author has signed a copyright agreement with a third party regarding the Work, the Author warrants hereby that he/she has obtained any necessary permission from this third party to let University of Gothenburg and Chalmers University of Technology store the Work electronically and make it accessible on the Internet.

© AMEERA DARWISH, June 2019.
© ANNA HENYRSON, June 2019.

Supervisor: MICHEL CHAUDRON & REGINA HEBIG
Examiner: Richard Berntsson Svensson

University of Gothenburg
Chalmers University of Technology
Department of Computer Science and Engineering
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000
How do Cultural Characteristics and Software Engineering Practices Interplay?
A Comparative Study between Indonesia and Sweden

Ameera Darwish
Department of Computer Science and Information Technology
University of Gothenburg
Gothenburg, Sweden
gusdaram@student.gu.se

Anna Henryson
Department of Computer Science and Information Technology
University of Gothenburg
Gothenburg, Sweden
gushenanni@student.gu.se

Abstract— This paper is an empirical study looking at the usage of software engineering practices and culture characteristics through the culture dimensions defined by Geert Hofstede. Using a survey to collect data from software engineering professionals worldwide, in a timespan of three weeks, the data is analysed by taking the responses from Indonesia and Sweden, and mapping them to the selected four culture dimension indices. This study focuses especially on software engineering practices that affect software architecture.

Keywords— software engineering practices, culture, global software development, architecture, design, culture dimensions

I. INTRODUCTION

Global software development (GSD) is a phenomenon in which different national and organizational cultures are involved in the development of software. The benefits include: round-the-clock development, access to a larger labor force, and cost advantages, which is why different companies around the world are employing GSD [1]. Hence, the importance of taking into account the different factors that interplay in software development, especially once development begins to spread globally.

We have decided to focus on the cultural factor in this study. While various definitions of culture exist, we use the Merriam-Webster definition, where culture is viewed as a "set of shared attitudes, values, goals, and practices that characterizes an institution or organization" [2].

Research has been conducted to better understand the relationship between humans and culture, including the roles and impacts of identifying within specific cultures. One framework used to compare cultural differences across nations is the culture dimensions as identified by Geert Hofstede [3]. The six dimensions identified are (i) power distance, (ii) individualism/collectivism, (iii) masculinity/femininity, (iv) long term/short term orientation, (v) uncertainty avoidance, and (vi) indulgence/restraint. These dimensions are a result of a factor analysis that was applied to a survey of 117,000 IBM employees worldwide. Different samples and versions of this study have been conducted following the success of the first study.

Considering the dimensions we have at hand and the rapid globalization of software development, it is interesting to explore whether culture affects people's judgment, perceptions, and use of software practices. Software practices are informal rules that over time have shown to contribute to the development of software with higher quality. Examples of SE practices include, but are not limited to, quality assurance, code reviews, code generation, pair programming, collective ownership, and continuous integration. Since we have an additional interest in software architecture practices, we will be focusing on the practices that correspond to architecture principles in this study.

As a result, we want to focus on the cultural characteristics and to explore whether the use and focus on certain SE practices are a result of cultural factors. Additionally, we will place emphasis on practices pertaining to software architecture and design. These practices will be compared with how the countries score on Hofstede's culture dimension indexes. Of the six dimensions, we have selected the following four dimensions: power distance (PDI), uncertainty avoidance (UAI), long term orientation (LTO), and individualism (IDV). Fig. 1 lists sample countries and their respective scores on the selected culture dimensions.
Long term orientation (LTO) concerns societies that place emphasis on persistence rather than on achieving quick results. A focus on the future is exerted with a willingness to delay short-term success in order to prepare for the future. From this, we imagine that countries with higher scores on LTO more often use practices that ensure the product's maintainability and longevity.

Uncertainty avoidance (UAI) refers to how comfortably a society reacts to unstructured situations by taking into account the society’s tolerance for ambiguity and uncertainty. In our context, we expect that countries with higher uncertainty avoidance to document extensively and make use of practices to ensure minimum risks. Following this line of thought, we expect that countries with low scores on the uncertainty avoidance index follow more agile processes and flexible planning.

Individualism (IDV) refers to a society where the self and immediate connections are prioritized. This defines a loosely-knit society where "I" reflects the self-image, instead of "we". For countries with high individualism, we expect sole responsibilities reflected in practices such as lower shared code ownership and likeliness to use pair programming.

Power distance (PDI) is the measure of the extent to which the lower and poorer ranks in society accept and expect the unequal distribution of power. In our study, we expect to see the effect of countries with higher power distance to be shown in the freedom of decision making when it comes to practices used, as well as the quality of cooperation within practices concerning group work.

The resulting research question is:

How do culture characteristics relate to the use of different software engineering practices?

Samples from different geographical locations could offer an increased variety in culture. For this reason, we tried to compare countries from different regions and with contrasting results on the culture dimensions index. From the resulting survey responses, we have chosen to compare Indonesia and Sweden across all scales.

A. Long Term Orientation

Our theory: Countries with higher LTO display more habits that show consideration of future events by preparing for them in the present, or at least taking them into account.

This theory has then been formulated into a hypothesis:

- \( H_0 \): LTO: Average(Sweden) > Average(Indonesia)
- \( H_1 \): LTO: Average(Sweden) < Average(Indonesia)

B. Uncertainty Avoidance

Our theory: Countries with higher UAI show more tendencies towards documentation, design and planning as to have more control over uncertainties.

This theory has then been formulated into the following hypothesis:

- \( H_0 \): UAI: Average(Sweden) > Average(Indonesia)
- \( H_1 \): UAI: Average(Sweden) < Average(Indonesia)

C. Individualism

Our theory: Countries with lower IDV display habits of sharing responsibility and helping each other.

This theory has then been formulated into a hypothesis:

- \( H_0 \): IDV: Average(Sweden) < Average(Indonesia)
- \( H_1 \): IDV: Average(Sweden) > Average(Indonesia)

D. Power Distance

Our theory: Countries with a higher PDI index exhibit more rule-obeying habits and have more tendencies to agree and obey the hierarchical structures.

This theory has then been formulated into a hypothesis:

- \( H_0 \): PDI: Average(Sweden) > Average(Indonesia)
- \( H_1 \): PDI: Average(Sweden) < Average(Indonesia)

II. MOTIVATION

SE practices are important in software engineering development with real-world implications when used; the reason we have adopted practices is to introduce structure and standards to software quality, as well as increase the shared understanding of the subject under development. The use of these practices influences the development process, making it more efficient. This is especially useful in the world of software engineering where no two products are alike. Thus, the use of SE practices is an important topic to investigate.

The aim of our research is to see if the cultural characteristics influence software development through existing software engineering practices. As little research has specifically looked into SE practices, we aim to uncover whether cultural perspectives impact our use and prioritization of the practices at hand.

Effective and dynamic team collaboration is key to success in any development, and especially within software engineering [4]. An improved understanding of whether culture plays an effect on how people use and prioritize software engineering practices can be applied to take better advantage of the globalizing software engineering industry, particularly within cross-cultural teams.

The scientific motivation is to better understand to what degree the adoption of software engineering practices is dependent upon cultural characteristics. As an example, the
survey responses from China may present a different picture compared to results from a survey done in Sweden [5].

Previous literature, as further explained in related work (see section III), has shown that differences exist when it comes to the company management and organization in the SE industry; particularly for PDI, UAI, LTO, and IDV. For this reason, we are interested to see if culture also plays a role in the more concrete SE practices that are used in development.

Future research could include other cultural dimensions defined by Hofstede and/or other culture researchers. Is it equally interesting to expand the study by investigating more countries.

This research may also have an impact on how software engineering evolves globally, and thus, we can draw from each culture’s strengths in order to minimize weaknesses and misunderstandings.

III. BACKGROUND AND RELATED WORK

Our study uses the culture dimensions defined by Hofstede [3]. We found that most other related work on culture use Hofstede’s culture dimensions as a basis for their research. Other researches have referenced Hall’s [6] or Trompenaars' model of national culture differences [7].

Research into GSD exerts particular interest in culture within software engineering. This includes outsourcing, management, user-interface, communication, culturally different teams, offshoring and offices at multiple sites across different countries. According to Boden, Avram, Bannon & Wulf, having culturally diverse teams will affect communication within the team as members could be used to communicate in different ways and styles [8]. This is primarily seen in countries with large differences in PDI. Additionally, outsourcing and offshoring also affect communication. Offshoring has been shown to differ between Germany and Russia (PDI 93) as the cooperation between the two countries has been difficult. Here it was shown that it was particularly daily stand up meetings that differed due to company culture. In other studies, we see how the mentality to solving problems differ, and how it impacts company meetings. Thus, it is relevant to understand the impacts of the cultural differences between nations.

Culture and its relation to SE practices are not well researched, but it has been shown that the success of agile practices differs depending on the culture, according to Iivari & Iivari [9] and MacGregor, Hsieh & Kruchten [10].

For example, mentioned Relationships between IT department culture and Agile Software T development: An Empirical Investigation by Gupta, Georgeb & Xiaa [11], a country with high PDI, benefits less from adopting agile practices while countries with low PDI do. Culture Differences in Software Engineering by Lavanya [12] brings out differences between countries through observations. One example from this paper suggests that countries with high IDV choose project leaders that are “dynamic, independent and capable of making critical decisions by themselves” [12]. In The Effects of Mentoring by Casado-Lumbreiras, Colomo-Palacios, Soto-Acosta & Misra [13], a study about mentoring newly hired team members and the effects of cultural factors on mentoring within cross-cultural teams, it, for example, showed that in Morocco and Ecuador, the impact of mentoring was much greater than in Argentina and Brazil [13].

To gain a deeper insight into software engineering practices, we have looked into the paper by Kuhnman et. al [14] using the data from the Hybrid dEvElopmEnT Approaches in software systems development (HELENA) survey. It has identified the use of SE practices around the globe. The HELENA survey was sent out to a population of software developers asking participants about their SE practice habits aiming to learn about the use of development processes (agile, traditional, mainstream, or home-grown). The collected data was used to “investigate the use of hybrid development approaches in software systems development - from emerging and innovative sectors to regulated domains.” [14]. The respondents were from a variety of countries, primarily German.

Cultures that behave and think differently need to understand each other's requirements and definitions, as well as company and quality standards to achieve the correct and expected product. Brochers [15] has found that communication within the teams differed depending on the country of origin. Particularly, design, documentation, and project management was influenced by cultural characteristics. Cultures with a higher UAI tended to adopt design more and have an iterative workflow. A high IDV showed a lack of working on tasks not specifically assigned to them, e.g. fixing other’s bugs. Japanese culture (IDV 46), for example, focus more on the group’s success than the individual’s.

IV. METHODOLOGY

The research methodology appropriate for this study is, after considering factors surrounding the scope, a survey. A survey allows us to collect a large amount of empirical data, compared to other data collection methods. It is also easily distributed and can reach a wider range of respondents around the world, which is very beneficial for us since we are comparing different cultures. We created our survey using Google Forms services. This allowed us to distribute the survey to an unlimited amount of respondents, and the service is widely known and reliable. Another advantage was the ability to customize the question format and layout according to our need. We made all questions, except the open-ended question placed at the end, mandatory.

Initially the focus was to collect data from nine countries, seen in Fig. 1. Those countries were selected by looking at countries that stood on opposing scales of each dimension and that were within our possible connections. However, in the limited time set, enough responses were gathered from Indonesia and Sweden. These two countries also lie in very different geographical regions, have
non-identical cultures and for most dimensions, score differently on Hofstede's culture dimensions.

Firstly, a pre-pre-test of our study was conducted which was sent out to the third year students at the Software Engineering and Management BSc, University of Gothenburg. It served as a guide to know which questions yielded useful answers and this later served as a base for the final survey. The final survey was then pre-tested on a group of professional software engineers before being sent out to our sample. We are using cluster-based sampling by collecting data points from a selected number of countries. Within those clusters, random sampling has been applied.

A. Data Collection

The survey was then distributed to both private and professional contacts that met or had access to software engineers that met our criterion. The survey was also released on online software engineering networks and several social networking platforms such as LinkedIn. Our population is professional software engineers around the globe. The criteria we set to the respondent to our survey is that they have professional experience in SE and identify with one national culture. This criterion was placed in the emails, posts and on the landing page of the survey. As an additional level of control we set three questions asking to state nationality, national culture they identify with and the country the respondent has most work experience in.

To attract the largest number of respondents possible, there is no restriction set on the team and/or company size, as well as the industry the respondent works in. We simply ask the respondent to select the checkbox that reflect their current workplace.

After a time span of three weeks we froze the data present and began analysing it. Nonetheless, the survey is still open and regularly receiving new responses that will be used for future research.

B. Data Analysis

The questions used in the survey are constructed by mapping a SE practice with a cultural dimension it could fall under. By looking at other studies and literature such as the HELENA study, we phrased the resulting survey which can be found in the appendix. The questions were designed to give quantitative data, having the respondents answer a 5-point likert scale with a variation of answers ranging from "strongly disagree" to "agree", "never" to "often", "unimportant" to "very important" and "not disciplined" to "very disciplined". We have chosen likert scales since we believe that after post-processing the data, the weights placed on each answer will ease the analysis and comparison of cultural characteristics. In addition, we have two multiple choice questions that fed qualitative data and one open ended question at the end of the survey asking the respondents if they think that "national culture affects the way their team develop software". This amounts to a total of 15 questions.

Once the copy of the frozen data was saved, we removed data deemed invalid such as responses from bi-cultural respondents or incomplete surveys. At the time the data was frozen, we had a total of 132 responses from a wide range of countries covering different continents. However, we decided to include countries with 20 or more responses, to be able to generalise and find more statistically significant results. The resulting valid sample collected is 25 from Indonesia and 43 responses from Sweden.

Questions 9, 11, 12 and 13 were based on likert scales. Since the likert scales have textual labels and those labels differ, we replaced the labels with weights -2 to 2 where -2 translated into a lowdimension index score and 2 a high score. Questions 10 and 14 were multiple choice questions were we counted the amount of recurring answers for each country.

To make sure the data was analysed correctly and evenly across all questions, we had to flip the responses to questions 11A, 11B, 11E, 11F, 11G, 12C, 12H, 12I, and 13. This was done so that an answer reflecting a weight of 2 always reflected a high score on the dimension scale and a -2 being a low score on the dimension scale.

For the descriptive statistics, we look at the combined average response for each question from each country. For the hypothesis testing, we use individual responses per question per dimension in a Mann-Whitney U test. Since Likert-scale is an ordinal variable, it is not normally distributed, which is further supported by the skew present in Fig. 2, 5, 7 and 10, and thus, we do not conduct a normality test. To determine if we can reject the hypotheses we have conducted the Mann-Whitney U test on each question. The Mann-Whitney U test does not rely on the samples being of equal sizes, which our samples are not. The test tests whether a randomly selected value from one sample is greater or less than a randomly selected value from another sample. Our chosen alpha is 0.05, with our confidence level being 95%.

To find differences among cultures we look at the culture dimension index for each country and compared it to answer of each question. We do this analysis for each country.

C. Limitations

Considering the time frame given for this thesis, we did not receive sufficient responses from all nine targeted countries. As a result, the responses analysed in this paper are from software engineers in Indonesia and Sweden. The amount of responses were 43 for Sweden and 25 for Indonesia, limiting the reliability of our conclusions. While we attempt to diversify our data by targeting different countries, our findings are not generalizable because of the small sample size. Nonetheless, considering the factors surrounding this project, this is a good start to see whether there should be further research conducted in this area. In addition, if we take the time frame into consideration, we cannot wait for respondents to take more than the allocated time. Compared to the years 1967 and 1973, when Hofstede conducted his study and sent out the survey, it is considerably more difficult to achieve respondents
engagement in surveys due to the influx of surveys sent out everyday. The result is that people often tend to ignore survey invitations, or lose interest if the survey comes across as demanding.

While designing the survey questions, we were fully aware that our biases might play a role in the formulation of said questions (found in the appendix). We also tried our best to avoid leading or suggestive questions. To decrease this likelihood, we had the survey pretested first. Still, our pretesters are from Europe and thus, we cannot be sure if, for example, Asian or American respondents, would interpret the questions the same way.

Even though we get some significant results, we must be aware of other factors that might be affecting our results, for example, industries might play a role, or even another culture dimension has a strong effect. For this reason, further research investigating the role of industry culture and its effect on software engineering practices is necessary. Also, since we have chosen countries on opposing sides of the culture dimension indices, we could try to choose countries with similar indices in three out of four of the chosen culture dimensions, as a control.

Still, the data collected has a limited amount of responses from each industry, so, it is a threat to the validity of the analysis. 'Finance/Banking/Insurance' and 'Information and Communication' both have sufficient responses (20 & 44 respectively), but the other industries do not. This is something that needs to be taken into consideration.

Another limitation in this study is the amplified focus on software architecture. During our literature review, we found that culture does have an impact on the use of architecture [3]. This means we have modeled our questions to understand more how different cultures utilize and work with software practices aimed primarily at software architecture.

Despite placing prerequisites for participating in the study, some respondents indicated to be bicultural. To generalise the data analysis, we removed their answers. We are including questions to pinpoint which culture the person mostly identifies with, regardless if they live there now, or if they identify with more than one culture. A similar threat is the possible uncertainty respondents faced answering questions related to their team and company size and industry. While we aimed for their average experiences, some respondents might have answered about their current team which they recently moved to after years of developing habits of a team with different dynamics in a different industry.

Finally, the criticism on Hofstede’s culture dimensions cannot be ignored. The main critiques include that culture changes, whereas this analysis and culture dimensions index show the assumption that culture is a "monolithic concept". It also fails to take into consideration the sub-groups that exist within a society making it very simplistic. Finally, different cultures are seen as mutually exclusive. While we do agree that overall these dimensions are presented in a simplistic approach represented on a scale from 1-100, we believe that using these dimensions gives us the first step that allows us to explore ideas further into the future through more thorough research methods. However, since Hofstede’s culture dimensions are well understood and straightforward, we have decided that using them is the best approach [3].

V. Results

In this section the survey is presented using descriptive statistics, graphs, and hypothesis testing to illustrate the results. A high score (1 to 2) in the figures should corresponds to a high culture dimension index, while a low (-2 to -1) score should correspond to a low index.

In addition to the Likert scale questions, questions 10 and 14 were multiple choice questions, asking about the reasons the respondents use documentation and which common practices they use. The result from this is seen in Fig. 18 to Fig. 21. These two questions are more focused on particular practices and thus we can investigate the use of practices in Indonesia and Sweden.

A. Quantitative Data

1) Descriptive Statistics

i) Long Term Orientation

<table>
<thead>
<tr>
<th>TABLE II. DESCRIPTIVE STATISTICS: LTO AVERAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indonesia</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Standard Deviation</td>
</tr>
<tr>
<td>Skewness</td>
</tr>
<tr>
<td>Min</td>
</tr>
<tr>
<td>Max</td>
</tr>
<tr>
<td>Count</td>
</tr>
</tbody>
</table>

Fig. 2: Showing descriptive statistics obtained from comparing the averages of both countries in the long term orientation dimension.
Indonesia and Sweden have similar LTO indices according to Hofstede (53 and 62 respectively). In Fig. 3 we can see that there is a small difference between the two countries.

Fig. 4 shows that there is some spread in answers from both countries. The Sweden sample does, however, have a smaller range than the Indonesia sample.

We can see that respondents from Indonesia and Sweden use practices relating to LTO (Fig. 5) in different amounts. The largest difference here is Risk assessment (Risk) where 86% of the Swedish respondents frequently adopt the practice while only 40% of the Indonesia sample said they do. However, looking at the answers from all questions we can see that they vary. For five of the questions, the Indonesia sample indicate that they adopt the following practice more often than the Sweden sample.

### ii) Uncertainty Avoidance

| TABLE III. DESCRIPTIVE STATISTICS: UAI AVERAGES |
|-------------------------------|-------------------|
|                              | Indonesia | Sweden |
| Mean                          | 0.702      | 0.128   |
| Median                        | 1.080      | 0.460   |
| Standard Deviation            | 0.723      | 0.847   |
| Skewness                      | -0.646     | -0.335  |
| Min                            | -0.230     | -0.920  |
| Max                            | 1.320      | 1.090   |
| Count                         | 5          | 5       |

Fig. 6: Showing descriptive statistics obtained from comparing the averages of both countries in the uncertainty avoidance dimension.
In Fig. 7 we see that Sweden, with the lower UAI (29), consistently scores lower on the average responses than Indonesia (48).

In Fig. 8 we can see that there is a large spread in what respondent's answers for both samples. This is in accord with Fig. 7.

When looking at the answers of questions 10 and 14, as presented in Fig. 9, we can see differences between Indonesia and Sweden. The Sweden sample selected “To make early design decisions” and “Test-first programming” more often than the Indonesia sample. On the other hand, the Indonesian respondents chose “Specify requirements” more often.

### ii) Individualism

<table>
<thead>
<tr>
<th></th>
<th>Indonesia</th>
<th>Sweden</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>-0.673</td>
<td>-0.807</td>
</tr>
<tr>
<td><strong>Median</strong></td>
<td>-0.625</td>
<td>-0.655</td>
</tr>
<tr>
<td><strong>Standard Deviation</strong></td>
<td>0.634</td>
<td>0.641</td>
</tr>
<tr>
<td><strong>Skewness</strong></td>
<td>0.325</td>
<td>-0.429</td>
</tr>
<tr>
<td><strong>Min</strong></td>
<td>-1.400</td>
<td>-1.590</td>
</tr>
<tr>
<td><strong>Max</strong></td>
<td>0.290</td>
<td>-0.150</td>
</tr>
<tr>
<td><strong>Count</strong></td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Fig. 10: Showing descriptive statistics obtained from comparing the averages of both countries in the individualism dimension.
In Fig. 10, the presented data indicates that the Indonesian (14) sample to be more individualistic. The boxplots in Fig. 10 show that Sweden (71) sample has a wider body of opinions towards the questions than Indonesia.

"Pair programming" stands out in Fig. 13, with the Swedish respondents using pair programming twice as often as the Indonesian respondents. The "reason for documentation" among both countries is similar with a tendency of Indonesia focusing more on "Achieving a mutual understanding of the system" while Sweden focuses more on "Help[ing] team-members find the latest information".

iv) Power Distance

TABLE V. DESCRIPTIVE STATISTICS: PDI AVERAGES

<table>
<thead>
<tr>
<th></th>
<th>Indonesia</th>
<th>Sweden</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.181</td>
<td>-0.428</td>
</tr>
<tr>
<td>Median</td>
<td>-0.162</td>
<td>-0.425</td>
</tr>
</tbody>
</table>

Fig. 14: Showing descriptive statistics obtained from comparing the averages of both countries in the power distance dimension.

Fig. 15: Showing the answers from each question relating to the power distance dimension. Note that the range of the y-axis is smaller than LTO, UAI and IDV. This is to show the difference between Sweden and Indonesia more clearly.

In Fig. 15, by comparing the culture dimension visually we see that this graph follows our general theory to a certain degree. Sweden which has a low PDI (31) answered in a way suggesting less hierarchy while the answers from Indonesia sample, which has a high PDI (78), indicated the presence of hierarchical divisions.

Answers of the questions relating to PDI have the smallest spread compared to the other dimensions. This means that respondents from each country agree more with each other and it makes it easier to compare the countries against each other. This is seen in Fig. 16.
we can reject our null hypothesis.

In Fig. 17, we see that the two countries score similarly, but “For negotiation between stakeholders and project managers” we find Sweden display a lower tendency to do so.

2) Hypothesis Testing

i) Long Term Orientation

\( H_0 \): Average(Sweden) > Average(Indonesia)

\( H_1 \): Average(Sweden) < Average(Indonesia)

TABLE VI. MANN-WHITNEY U TEST: LTO

<table>
<thead>
<tr>
<th>Question</th>
<th>Sample Size - Indonesia</th>
<th>Sample Size - Sweden</th>
<th>P-value</th>
<th>Z-value</th>
<th>U-value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>11G</td>
<td>25</td>
<td>43</td>
<td>0.00006</td>
<td>-3.85681</td>
<td>227</td>
<td>( h_0 ) not rejected</td>
</tr>
<tr>
<td>12C</td>
<td>25</td>
<td>43</td>
<td>0.07493</td>
<td>1.43796</td>
<td>382.5</td>
<td>( h_0 ) not rejected</td>
</tr>
<tr>
<td>12D</td>
<td>25</td>
<td>43</td>
<td>0.16354</td>
<td>0.97550</td>
<td>388.0</td>
<td>( h_0 ) not rejected</td>
</tr>
<tr>
<td>12E</td>
<td>25</td>
<td>43</td>
<td>0.04182</td>
<td>1.72633</td>
<td>364.5</td>
<td>( h_0 ) not rejected</td>
</tr>
<tr>
<td>12F</td>
<td>25</td>
<td>43</td>
<td>0.38209</td>
<td>0.29005</td>
<td>469.5</td>
<td>( h_0 ) not rejected</td>
</tr>
<tr>
<td>12G</td>
<td>25</td>
<td>43</td>
<td>0.15866</td>
<td>1.00256</td>
<td>419.5</td>
<td>( h_0 ) not rejected</td>
</tr>
</tbody>
</table>

As seen in Fig. 19, we reject the null hypothesis in questions 9A, 9B, 12A, and 12B. 9C is the only question being not rejected, and the reason for that is because the significance (p-value) is too high.

iii) Individualism

\( H_0 \): Average(Sweden) < Average(Indonesia)

\( H_1 \): Average(Sweden) > Average(Indonesia)

TABLE VII. MANN-WHITNEY U TEST: UAI

<table>
<thead>
<tr>
<th>Question</th>
<th>Sample Size - Indonesia</th>
<th>Sample Size - Sweden</th>
<th>P-value</th>
<th>Z-value</th>
<th>U-value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>9A</td>
<td>25</td>
<td>43</td>
<td>0.01130</td>
<td>-2.28311</td>
<td>357.5</td>
<td>( h_0 ) rejected</td>
</tr>
<tr>
<td>9B</td>
<td>25</td>
<td>43</td>
<td>0.01390</td>
<td>-2.20271</td>
<td>308.5</td>
<td>( h_0 ) rejected</td>
</tr>
<tr>
<td>9C</td>
<td>25</td>
<td>43</td>
<td>0.28434</td>
<td>-0.56625</td>
<td>345.0</td>
<td>( h_0 ) not rejected</td>
</tr>
<tr>
<td>12A</td>
<td>25</td>
<td>43</td>
<td>0.01072</td>
<td>-2.29799</td>
<td>320.0</td>
<td>( h_0 ) rejected</td>
</tr>
<tr>
<td>12B</td>
<td>25</td>
<td>43</td>
<td>0.01743</td>
<td>-2.10731</td>
<td>264.0</td>
<td>( h_0 ) rejected</td>
</tr>
</tbody>
</table>

The questions relating to LTO are 11G, 12C, 12D, 12E, 12F, 12G. As seen in Fig. 18, there are no questions where we can reject our null hypothesis.

ii) Uncertainty Avoidance

The null hypothesis is, in the case of questions relating to IDV, only rejected in one instance: 11A. 12H has a high confidence level, but here the survey results show that Indonesia are more collectivistic than Sweden, which does
not support the culture dimension by Hofstede or our hypothesis.

iv) Power Distance

\( H_0 \) PDI: Average(Sweden) > Average(Indonesia)

\( H_1 \) PDI: Average(Sweden) < Average(Indonesia)

TABLE IX. MANN-WHITNEY U TEST: PDI

<table>
<thead>
<tr>
<th>Question</th>
<th>Sample Size - Indonesia</th>
<th>Sample Size - Sweden</th>
<th>P-value</th>
<th>Z-value</th>
<th>U-value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>11C</td>
<td>25</td>
<td>43</td>
<td>0.06301</td>
<td>-1.52857</td>
<td>359.5</td>
<td>h0 not rejected</td>
</tr>
<tr>
<td>11E</td>
<td>25</td>
<td>43</td>
<td>0.2451</td>
<td>-0.69399</td>
<td>459.5</td>
<td>h0 not rejected</td>
</tr>
<tr>
<td>11F</td>
<td>25</td>
<td>43</td>
<td>0.30153</td>
<td>0.51981</td>
<td>437.5</td>
<td>h0 not rejected</td>
</tr>
<tr>
<td>13</td>
<td>25</td>
<td>43</td>
<td>0.01321</td>
<td>2.22249</td>
<td>328.0</td>
<td>h0 rejected</td>
</tr>
</tbody>
</table>

Fig. 21: Mann Whitney U test on all the questions relating to PDI, comparing the two sample's answers. The result whether the null hypothesis is rejected or not is shown in the result column.

Question 13 reject our null hypothesis. The success of agile practices has been linked to a low PDI according to Livari & Livari [9] and MacGregor et al. [10].

B. Qualitative Data

In addition to the quantitative data, we had one question that were not possible to analyse in the same way. This is a non-mandatory open ended question at the end of the survey asking respondents if their "national culture affect[s] the way that [their] team develop[s] software?" and "if so in what way?".

TABLE X. RESPONSES TO OPEN-ENDED QUESTIONS

<table>
<thead>
<tr>
<th>Indonesia</th>
<th>Sweden</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>38.9%</td>
</tr>
<tr>
<td>Unsure</td>
<td>5.6%</td>
</tr>
<tr>
<td>No</td>
<td>55.6%</td>
</tr>
</tbody>
</table>

Fig. 22: Average response rate from 18 out of 25 respondents from Indonesia and 29 out of 43 respondents from Sweden.

Around 72% respondents from the Indonesia sample and 67% from the Sweden sample answered the open ended question. Fig. 22 shows that the respondents from Sweden think that their national culture affects the way their team developers software more than the Indonesia sample. The words and notions of "personal responsibility", "shared understanding", "consensus", "open discussions" and "non-hierarchical" appear. One response also mentioned the "homogenous educational background" plays a role.

In the Indonesia sample, most answers were, in general, limited to "yes" and "no" answers. One answer mentioned "mutual cooperation", while another response was more focused on the economical side-effects on the production and testing of product.

VI. DISCUSSION

To answer our main research question, *How do culture characteristics relate to the use of different software engineering practices?*, we chose four culture dimensions defined by Hofstede. As such, every dimension has its respective theory and hypotheses based on how both countries, Indonesia and Sweden, score on the index. Furthermore, the qualitative data offers deeper insight into the reasons and practices used. It also helps interpret the data and make connections to our theory with the culture dimensions in mind.

A. Long Term Orientation

Indonesia and Sweden do not have a very large difference in LTO. Sweden's index is 53 and Indonesia's index is 62, and thus, we would expect a tendency of Indonesia being more long term oriented. However, as seen in the result section (see Section V), our data suggests otherwise. We fail to reject the hypotheses for LTO, as the data suggests that Sweden seems to be more long term oriented. This is the opposite of our general theory and does not follow the idea's defined by Hofstede.

As mentioned in Methodology (see section IV), a threat to our data is the fact that there could be reasons other than a specific culture dimension that affects the results. Some culture dimensions may affect others. For example, if we have country A and B, where both have a high LTO index where country A has a high index in UAI while country B has a low UAI, differing factors could lead to a similar LTO scores.

The qualitative data, in Fig. 5, show that Indonesia sample agrees to a greater extent with the theory that developers in Indonesia have habits that show consideration of future events by preparing for them in the present, compared to the Sweden sample. However, we cannot draw direct conclusions from the present data. Even though the differences between Indonesia and Sweden samples are not big, these findings should not be overlooked. We would expect the two countries to answer fairly similarly, which they have. In some questions Sweden sample answers suggesting long term oriented society, while in others, the Indonesia sample did. These differences and the fact that on average Indonesia seem less long term oriented than Sweden could depend on industry and/or the low sample size.

B. Uncertainty Avoidance

Throughout the average responses, Sweden consistently scores lower than Indonesia on all questions that have been classified under the uncertainty avoidance group. Under Hofstede's culture dimension index, Sweden scores 29 while Indonesia scores 48, suggesting that Indonesia is a culture that is more resistant to taking having to adapt in the last minute, and would rather take decisions to minimize risks as early on as possible. By looking at Fig. 7, the trend corroborates our theory (Countries with higher UAI show more tendencies towards documentation, design and
planning as to have more control over uncertainties). In addition, for questions 9A, 9B, 12A and 12B, we reject the null hypothesis, and with significant enough difference we suggest that indeed UAI has a role on the use of certain software practices. There are many reasons that can lead to our results, and an economical factor would lead to more planning and relying on tried and tested tools to ensure somewhat foreseeable results, instead of being open to risk-taking.

We can see in Fig. 8 that the range of answers of questions relating to UAI is wider than other dimensions. While there is a clear difference between the averages of each country, the body of responses overlap. This could be down to individual persons having different opinions.

In Fig. 9, which looks into the reasons for using documentation as well as the use of software engineering practices offers a slightly different perspective. While Indonesia scores higher on "specifying requirements", a practice we attribute with more vigilant planning, we find the Sweden sample scoring higher on "test-first programming" as a practice, and "mak[ing] early design decisions" as a reason for documentation. While we could speculate that the respondents come from industries that have certain practices deemed as essential for the product development, a larger sample or a study focused on different industries is required to come to more conclusive statements.

C. Individualism

Interesting results are to be expected from the individualism index with Indonesia scoring 14 and Sweden scoring 71 on the scale. From our results (Fig. 11), we find fluctuations with both Indonesia and Sweden demonstrating more individualistic responses almost equally. By looking deeper into the questions, we tend to see Sweden is more individualistic revolving questions inquiring the individuals role in the team: (i) if the team members usually specialise in one task, and (ii) if they tend to be concentrated on their own role in the project. This is also reflected in the general inclination that certain people, and not all members of the team, are responsible for the quality of the design. In addition, the Swedish sample was less inclined to ask their colleagues for help. For our hypothesis, we only found statistically significant enough difference to be able to reject the null hypothesis for question 11A, "I find it important to follow certain, predetermined conventions when designing software architecture". The theory we propose, backed up by the data, is that a team with shared conventions will allow the team to communicate more efficiently. Additionally this means that if a team find differences or problems internally, it is easier for them to fix these issues.

From our qualitative data, we find Sweden exhibits more collectivistic use of practices than we expected. Sweden scored higher on using documentation for helping to "achieve mutual understanding of the system" than Indonesia, and scored higher on the practice "pair programming" as well. Nonetheless, the latter could be attributed to the highly agile software development community that can be found in Sweden, a country also scoring quite low on UAI. Independence is also defined differently in more individualistic or collectivistic countries. There is more of a natural tendency to lend a helping hand without being asked to do so in collectivistic countries, versus individualistic countries where people tend to seek help after trying to solve the question at hand on their own first. Following this thinking, we find that Indonesia scores higher by 10% on using documentation "to help team-members find the latest information".

The respondents in both countries have a wide range of opinions, as seen in Fig. 12. A smaller range would have indicated that cultures agree more internally.

From the open-ended question, the respondents from the Sweden sample exert focus on the importance of "personal responsibility" when it comes to their tasks and they take individual responsibility for the quality of the product they develop. This is anticipated coming from an individualistic culture, and thus this qualitative data gives more insight into the nature of working.

D. Power Distance

Indonesia and Sweden have answered in accordance to our general theory in all but one question. However, we only managed to reject the null hypothesis in one of the cases, question 13 - relating to discipline of agile practices. The reason we could not reject the other questions was due to the p-value being higher than our alpha, even though we can see a difference in the graphs.

Sweden’s PDI is 31 and Indonesia’s is 78, and for this reason, we would expect the answers from Sweden to show less of a hierarchical structure than Indonesia. While we find a tendency in which the Indonesia sample's answers are more hierarchical, there is no sufficient data to say with certainty that PDI influence the use of SE practices.

In Fig. 17, we can see that there is no clear difference in the use of coding standards. Sweden has 81% and Indonesia 80%. This is not a difference we can draw any conclusions from. However, while we had expected the Indonesia sample to answer higher on questions relating to PDI, this could be an indication that coding standards has more to do with other cultural dimensions rather than PDI. In the same figure, the reasons behind the use of documentation is fairly similar between the two countries. Sweden seem to be more hierarchical according the Fig. 17, but this could just mean that they use documentation as part of the development process to a higher degree than Indonesia.

Another interesting note is that the respondents from each culture answered similarly to each other, having a smaller spread than any other dimension. This would increase the validity of our findings because there is a general consensus among the respondents, so in the case of question 13 regarding agility, we can with a high probability conclude that PDI impacts the use of agile practices.

The findings are supported by the responses from the Sweden sample to the open-ended question. The importance
of shared understanding, consensus, and discussions are stated multiple times. Team-members are encouraged to question things and openly discuss them. The basis for open discussion is more easily established in a "non-hierarchical" team, where agile practices are also followed more easily. Another reason "attributes [shared understanding of values] to the highly homogenous educational background". This is indeed interesting as young people also learn their culture through their schooling system, and in Sweden, expression of opinions and discussions in addition to teamwork, is highly encouraged. These habits are later carried on to the workforce.

E. Architecture

As stated earlier, in this study we have an amplified interest in practices used for software architecture. From the already presented data, we have composed a graph with all questions relating to architecture (Fig. 23).

![Architecture Graph]

Figure 23: The questions relating to architecture.

There is a significant difference between Indonesia and Sweden when it comes to (i) viewing design documentation as an important part for achieving a high quality system (9A) and (ii) valuing UML models (9B), with Indonesia in favor of it. We also see that Indonesia scores higher on UAI and LTO. If we look at UAI, (i) design documentation shows planning ahead and consideration for future maintainability. The Indonesia sample also answered significantly higher to "create[ing] detailed (UML) models" (12B). We can attribute (ii) the fact that countries with higher UAI prefer set and standardized methods and definitions. This way, there is an official standard the whole company can revert to, instead of each team having formed their own notations. This immediately ties in with "find[ing] it important to follow certain, predetermined conventions when designing software architecture" (11A).

No significant difference between the countries is found in questions 11B, 11C and 12C. 11B and 11C cover the idea of hierarchical domination in the role of a software architect. 12C is more focused on LTO by looking at the preference of "quick and easy fixed" in view of maintainability.

VII. Conclusion

In this paper, our research question looks into the effects of cultural characteristics on the use of software engineering practices. This was conducted through a survey that was distributed to software engineers worldwide. The minimum limit per sample set was 20, and as a result, the data analysed came from Indonesia and Sweden, making this a comparative study.

In conclusion, even though the sample is small, the data indicates interesting parallels worth further investigation. It is interesting to note that the differences in both UAI and PDI culture dimensions followed our theories and hypotheses.

Since both Indonesia and Sweden score similarly on LTO, we did not expect to find any significant differences. The findings support that the two countries are similar in LTO. Long term orientation (LTO) is the dimension that accounts for emphasis on taking actions for future results rather than direct results. The findings from our data analysis also provided results contradicting our theories about IDV, for example, however upon a closer glance, we found that the differences pointed out different mannerisms within individualism. For this reason, subsets of the individualism culture dimension might be necessary to define, since individualism in different contexts yields different responses.

In our current findings, the interplay between cultural characteristics and software engineering practices present new ideas and methods with which to take advantage of the strengths of each culture, especially as the software development field globalises more and more.

VIII. Further Research

In our analysis, we look at differences between countries, but we also noticed that potential future work with the current or with an expanded dataset would be to group all responses according to industry and repeat the analysis method. The industries are i) Finance/Banking/Insurance (20 responses), ii) Governmental (6), iii) Healthcare (6), iv) Information and Communication (44), v) Manufacturing (5), vi) Research and Education (11), vii) Transportation/Logistics (9). We also got responses from Arts/Gaming/Recreation, Defence, Retail/Wholesale, and Travel/Accommodation. These industries did however not get enough replies, averaging around two people, to analyse.

We have included graphs to show differences and similarities between industries. Fig. 24-27 indicate that certain industries differ from each other. For example, we see that the governmental sector stands out in the LTO and PDI culture dimensions. One possible theory is that the software developed for this sector is generally used for a long time, and more planning and attention goes into maintainability and extendability of related products. Another factor is hierarchy since governments are usually run top down, and this is supported by the data shown in Fig. 27.
As seen in Fig. 24, manufacturing scores high on the LTO dimension. Manufacturing encapsulates a wide variety of fields, such as safety or mission-critical software and software for entertainment. We would expect domains where safety or mission-critical software is crucial to score higher on LTO and UAI as these domains need to avoid any possible errors as well as create software that works in the long term. Further research could divide these two domains to better understand safety versus entertainment. If we do find differences within these domains that do not depend on culture, we can adapt and configure practices to fit particular industries. It is also possible that both industry and culture play a big role in the use of software engineering practices, and then future research needs to take both approaches into account when analysing the effects of culture. Certain culture dimensions could affect practices more or less, as we have seen indications of from our findings. For example, the governmental sector in Malaysia, where the PDI is 100, could show a higher indication of a hierarchical system than in China, where the PDI is 80. However, in the ICE sector, China could show indications of acting more hierarchical than in Malaysia. While this is not something we can conclude from our data, it could be interesting to have in mind when doing further research on this topic.

Another interesting observation is that the healthcare industry scores low in the UAI dimension for some questions (see Fig. 25). These are the questions 9A, 9B and 12B which pertain design, and more specifically the use and importance of UML models. This low score is interesting because of the products' direct impact on patients' lives, and thus considerable measurements must be taken. Still, we are not sure from which medical sectors these software developers come. It could be developers developing, for example, a life support system, or developers creating the mobile application for an online doctor’s appointment system, or a mixture of both. These two domains are very different as a life support system needs to work flawlessly while an application for booking appointments can afford bugs, to a certain extent, and thus does not need to use risk assessment as frequently. Similarly as with manufacturing, dividing the healthcare industry up into sub-categories is important when investigating safety-critical systems versus non-safety-critical systems.

In Fig. 26 we can see that ICE scores high on question 11E. Together with healthcare they stand out against the other industries. 11E is “In my team, every team member is comfortable changing any part of the existing code at any time (‘collective code ownership’)”. Why the answer to this question is so different to the other questions should be further investigated.

Overall, we can see that there are many similarities between the industries, but also large differences for some
questions in some dimensions, and a start for future research could be to focus on these differences.

Other research to be conducted is to compare more of the countries on the list in Fig 1. While Indonesia and Sweden have contrasting results, there are other countries that have larger differences in LTO and UAI. Malaysia which has a PDI of 100 is interesting to compare to Denmark, a PDI of 18.

To validate findings more comparisons should be made, not just with countries that have different culture dimension index, but with countries of similar indices, in order to find similarities.

Hofstede’s theory has six dimensions. In addition to LTO, UAI, IDV and PDI there also is masculinity/femininity (MAS) and indulgence/restraint (IND). While comparing our ideas and theories with other literature sources, we found that the four chosen dimensions are the most interesting to look at. Still, to accumulate an understanding of all aspects of cultural differences, studies on MAS and IDN is a possible topic for future research.

We chose to use Hofstede’s culture dimensions as a base for our study, but as mentioned in the introduction, there are other theories on culture that are widely recognised and used. There could be aspects that Hofstede’s theory does not take into account that other theories do, therefore other studies using other culture theories as a base would create a more complete understanding of culture and its effect on SE practices.

References

A. Survey Questions

1. Please state your nationality.
2. Please state the country you have the most years of work experience in.
3. Please state the national culture you most identify with.
4. Please state how many years of professional experience you have in software engineering.
5. Please state which industry you work in.
6. Please enter the number of people in your team.
7. Please enter the number of employees in your organization.
8. Please select what your main role is in your current team.
9. What are your opinions toward the following issues
   (Unimportant - Very Important):
   a. How relevant is design documentation for achieving a high quality system?
   b. How much do you value UML models?
   c. For creating an API, I find it very important to come up with consistent abstractions.
10. In our project developers make documentation (Check all that apply):
    a. As a part of a mandatory process.
    b. For easier re-use of the model in similar software.
    c. To get a quick overview.
    d. For my own understanding.
    e. For negotiation between stakeholders.
    f. To achieve mutual understanding of the system within our team.
    g. To explore complex / risky parts of a design.
    h. To help team-members easily find the latest information.
    i. To keep track of progress for our project manager / customer.
    j. To make early design decisions.
    k. Other.
11. To what extent do you agree with the following statements (Strongly Disagree- Strongly Agree):
    a. I find it important to follow certain, predetermined conventions when designing software architecture.
    b. Every member of the team is responsible for the quality of the design.
    c. In our team, the architect takes the design decisions, regardless if consensus has been reached among all developers.
    d. In my project, people usually specialize in one task (i.e. testing, programming).
    e. In my team, every team member is comfortable changing any part of the existing code at any time (‘collective code ownership’).
    f. At the start of the project, my team is free to decide the software practices we want to follow.
    g. In our team, we prioritize delivering on time over ensuring good maintainability of the code.
12. Please tick the answer that fits best (Never - Always):
    a. When I need to understand a new (part of a) system, I prefer to look at the design first instead of jumping directly into the code.
    b. To what extent do you create detailed(UML) models (abstraction of some aspect of an existing or planned system)?
    c. In our team, we make ‘quick and easy fixes’ that do not conform to architectural design?
    d. How often is version management used for tracking updates to the design of your project?
    e. How often are unit tests conducted as a part of the development process in my team?
    f. Does your team keep track of development risks?
    g. We use retrospectives to learn what we should improve in the upcoming iteration/s?
    h. How likely are you to ask a colleague for help?
    i. How common is it that you fix bugs in your colleague's code?
13. How disciplined is your team in following agile practices (Not Disciplined - Very Disciplined)?
14. Which practices do you follow? (Check all that apply).
    a. Bug tracking
    b. Coding standards
    c. Continuous integration
    d. Documenting design
    e. Pair programming
    f. Refactoring
    g. Risk assessment
    h. Specify requirements
    i. Test-first programming (e.g. Test-driven development)
    j. User stories
    k. Other
15. Does your national culture affect the way that you/your team develop software? If so in what way(s)?