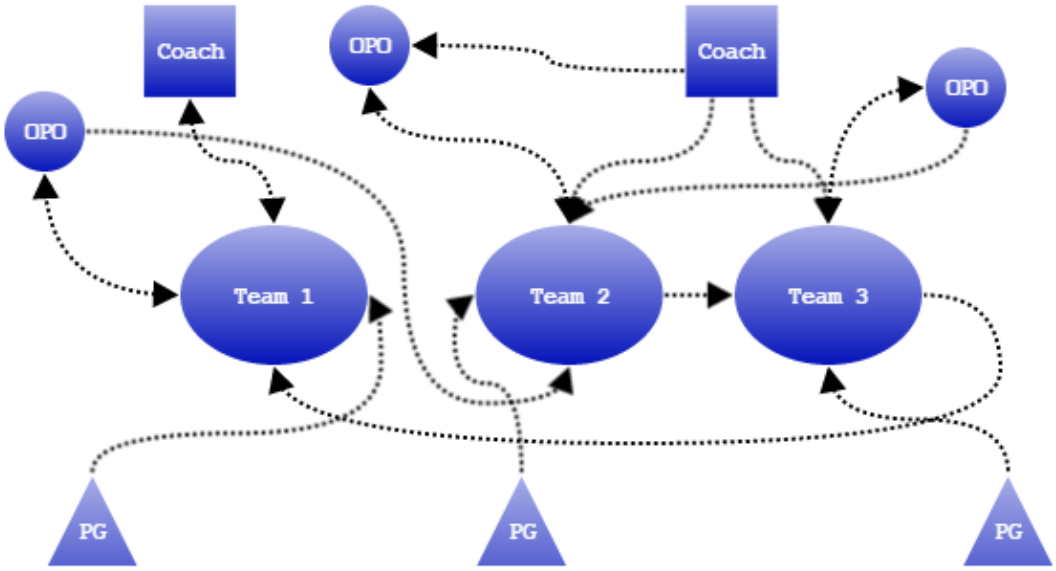
UNIVERSITY OF GOTHENBURG



Improving communication in large-scale agile environments: a quasi-experimental approach

Master of Science Thesis in Software Engineering

JORGE ANTONIO DIAZ-BENITO SORIANO

Department of Computer Science and Engineering
UNIVERSITY OF GOTHENBURG
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden, June 2015

Improving communication in large-scale agile environments: a quasi-experimental approach
Jorge Antonio Diaz-Benito Soriano

Supervisor: Eric Knauss
Examiner: Richard Berntsson Svensson

University of Gothenburg
Chalmers University of Technology
Department of Computer Science and Engineering
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover: Graphical example of the communication overhead at the target environment showing different flows at the lowermost organizational level.

Gothenburg, Sweden 2015

Acknowledgements

A lot of people have been supporting the execution of this research. Specifically I would like to dedicate my special thanks to Eric Knauss for all his help not only during the study itself but also for the help provided for approaching Ericsson. Also I would like to thank Mats Eriksson for being always a nice person ready to answer all of my questions and push me to think differently. In addition, I want to thank everybody I had the luck to work with for their dedication on their interactions with me and I hope that my research has actually contributed to improve the way they work. Finally I would like to thank my parents, whose continuous support has allowed me to enjoy this extraordinary experience.

Jorge Antonio Diaz-Benito Soriano, Gothenburg, June 2015

Abstract

Communication is known to be a problematic-yet-unavoidable concern in software development. Furthermore, in agile contexts, where the distribution of information is intensified and the information itself is less reliable due to the changing nature of agile projects, communication becomes a key factor. In addition, when these characteristics are embraced by the size of a large-scale setup, the aforemencioned concern rapidly becomes chaos. This thesis uses a quasi-experimental approach to attempt to solve some of the communication-related problems in this type of contexts. The results found highlight the importance of the communication between Scrum coaching staff and Scrum teams.

Keywords Agile, communication, large scale.

Contents

1

Introduction

"Industry and technology move too fast, requirements change at rates that swamp traditional methods" (Highsmith, 2000), and "customers have become increasingly unable to definitively state their needs up front while, at the same time, expect more from their software" (Cohen, Lindvall, & Costa, 2004). Among others, these two issues were the cause that made the software development community realize that something had to be done in the field because software projects were failing massively, and going over budget and time too often. Agile methodologies are a reaction to traditional ways of developing software that acknowledge the "need for an alternative to documentation-driven, heavyweight software development processes", and are built on top of the pillars gathered in the well-known Agile Manifesto (Beck et al., 2001).

Nowadays, nobody argues that this attitude towards change has benefited software development. Furthermore, there is empirical evidence of the improvement that Agile methodologies have brought to the field (Arthur, 2013) (Hesser & Tomasini, 2012).

However, and although they have meant one step forward, Agile methodologies have been shown to not to be the definitive silver bullet (Moreira, 2013) due to, among other reasons and despite having been widely approached in research (Kraut & Streeter, 1995) (French & Layzell, 1998) (Hanakawa & Okura, 2004) (Qian & Zhen-hua, 2010) (Jaanu, Paasivaara, & Lassenius, 2012) (MacKellar, 2012), communication. Communication is one of the main sources of chaos in large-scale projects, and it quickly becomes worse when considering the change expectable in agile environments: new requests, procedure modifications, deprecations, etc. are issues that, upon not being properly understood and transmitted across the organisation, will increase the project costs and reduce its value, and may even force its cancellation. So, no matter how well change is perceived, nor how good the preparation is: the fact that such change has to be communicated across the organisation always adds a layer of harm to the one already brought by change itself.

This thesis addresses the subset of problems related to communication that surrounds the environment of Scrum teams in large-scale setups by proposing solutions elaborated upon the results of observation and interviews and evaluating them using a quasi-experiment. First, communication habits at the target environment are analyzed through interviews and informal observation, and the experiment subjects' takes on communication are assessed through a survey that runs daily across one

Scrum Sprint. Afterwards, and based on the knowledge acquired through the interviews and observation and on previous research on the target context (Averianova & Deekens, 2014), guidelines for improved communication are designed and applied as treatment to the experiment subjects, whose performance communication-wise is assessed again over the course of another Scrum Sprint with the same survey than before. Finally, the results of the experiment are presented and analysed (Research Question RQ1) and the lessons learnt from its execution are shown (Research Question RQ2). Therefore, the proposed study has the double goal of improving the efficiency of the communication at the organisation and increasing our understanding of experimental research in real-world large organisations. The thesis develops throughout a four-month-long period in collaboration with a subset of the teams in the PDU LMR PD CAT software organisation at Ericsson AB in Gothenburg, Sweden.

Because of the lack of positive results on research about communication in large-scale agile environments and how important it has proven to be, it is believed that the conclusions extracted from the proposed experiment will help large agile companies to improve in a factor which has a large and diverse impact on software projects.

Furthermore, and independently of the success of the experiment, its results are expected to contribute to advise both companies and future researchers on the way to follow when utilizing experimental approaches on setups with similar characteristics (large-scale agile environments with different cross-functional teams working in parallel), and therefore it is believed that this research could also be of interest for Academia.

The contents of this document are structured in the following way: Chapter 2 reviews previous research on the fields of large-scale software development, Agile methodologies and communication-related fields that are relevant to the thesis. Chapter 3 explains the organizational structure of the target context and the scope addressed. Chapter 4 details the methodology used to design and execute the experiment, including procedures for data collection and analysis. Chapter 5 shows the procedure used to design the experimental treatment and the treatment itself. Chapter 6 lists the results of the experiment and reports some basic interpretations of them. Chapter 7 focuses on both (1) extracting useful conclusions from the results of the experiment and (2) explaining the lessons learnt about performing experimental studies in real-world large-scale Agile environments. Section 7.3 closes the thesis.

2

Literature review

The central phenomenon being approached in this research is the communication occurring in the most proximal context of the development teams in a large-scale agile context. From a literature review point of view, this implies that the research of interest for the proposed study resolves around different topics that derive from either communication or Agile methodologies (both of course in the context of large-scale software development).

Challenges induced by communication

Communication is a broad term. To narrow it down, quoting Kraut and Streeter (1995), "In software development, communication means that different people working on a common project agree to a common definition of what they are building, share information and mesh their activities".

A factor long known to be very influential on whether these agreements are or are not reached is geographical distribution (Herbsleb & Mockus, 2003). Distributed software development is the term coined for software projects carried out by people who are not geographically co-located. This has a variety of issues, like cultural differences, schedule incompatibilities and language barriers (dos Santos, de Farias Junior, de Moura, & Marczak, 2012), which have a harmful impact on communication and, in the end, on the project itself. Furthermore, these consequences happen to occur as well in large-scale contexts due to having a real shared location for huge amounts of developers. There have been some attempts to address these issues, from both the social (Dorairaj & Noble, 2013) and empirical (Korkala & Maurer, 2014) research points of view, although none has gone beyond moderate levels of success. A topic of particular interest with significantly large relevance not only in distributed software development, but also in co-located one, is the usage of mailing lists. In spite of being an old communication method, most relatively big software projects use (at least) one and, although they are mostly used to discuss implementation issues/details, they also happen to be mostly recurred to for discussing usage of the product and even for social interactions not really required for the development of the project (Guzzi, Bacchelli, Lanza, Pinzger, & Deursen, 2013), although surely beneficial for it.

However, although most research in this topic points to communication as a factor commonly involved in problems, it has been shown to be possible that a sufficiently strong positive correlation exists between communication and commitment to the

technical development of the project (Xuan, Gharehyazie, Devanbu, & Filkov, 2012), which means that it could be possible that communication is not a harm int this type of contexts, but a supporing factor instead. Furthermore, in the concrete context of agile practices, there is evidence to acknowledge that iteration reviews, product and sprint backlogs are the practices regarded as having the most positive impact on communication related to feature-requirements dependencies (by "creating new systematic ways to communicate between the development teams and stakeholders, which also helps ensuring that the created product meets the demands of the customer and other stakeholders"), although such impact does not seem to be referable to as completely positive anyway (Pikkarainen, Haikara, Salo, Abrahamsson, & Still, 2008).

As for further research on communication in the concrete context of Agile methodologies, Martini, Pareto, and Bosch (2013) had an initial approach to communication factors creating problems in large-scale agile contexts, but it is believed that the proposed study can, because of the use of a different focus, contribute more specifically to team-centric problems, and Sekitoleko et al. (2014), while not focusing on communication discovered, as a side-effect, that it is ones of the main reasons behind many challenges in large-scale agile software projects. Averianova and Deekens (2014) show a very clear identification of some of the problems related to communication in the same context that the proposed experiment intends to be developed on and identifies a few suggestions to address them, but they are not given with enough detail, are too costly resource-wise and are not tested in this scale, which is expectable to bring even more obstacles into the way. Because of their bigger scope, large-scale projects are not ideal for Agile methodologies because, among other factors, it is difficult, costly and time-consuming to adapt them to change quickly enough, it is hard to find the right amount of documentation needed and it is usually not possible to collaborate with the customer as much as Agile guidelines prescribe (Beck et al., 2001).

This variety of causes is reflected on the related research: some put their focus on technical challenges (Søvik & Forfang, 2010), others into dependencies (Sekitoleko et al., 2014), and others believe that the most important and hard to overcome hurdle is finding the right people (Moore & Spens, 2008).

Fortunately, wherever there is a problem in research there are researchers trying to solve it: research exists providing guidelines on how (Paasivaara, Väättänen, Hallikainen, & Lassenius, 2014) and when (Power, 2014) to adopt Agile in large-scale organizations, although they do not seem to be widely established yet (Rohunen, Rodriguez, Kuvaja, Krzanik, & Markkula, 2010).

There are also more holistic, generic proposals to address the problem, like the framework suggested by Soundararajan and Arthur (2009) based on an Agile philosophy of requirements gathering and a flexible development process, that can be adapted to any size.

Socio-technical congruence

Socio-technical congruence is a relatively young approach that refers to the ideal measurement between dependencies that people hold on each other derived from their interdependent tasks and the coordination activities that they carry out in order to approach them (Cataldo, Wagstrom, Herbsleb, & Carley, 2006). It was developed to address the limitations of modularisation in regards to dependencies among software development tasks, as traditional software modularisation techniques seem to rely on using only a subset of the technical dependencies of a software system (Garcia et al., 2007).

As Cataldo, Herbsleb, and Carley (2008) mention, the socio-technical congruence framework allows for systematic and deep examination and location of both direct and indirect dependencies among people working in a software project, this meaning that it is able to expose who should interact with whom based solely on the dependencies among their tasks (which takes their arguments to a very reliable state). Complementing this information with the improvements on how to perform these interactions in an optimal manner built upon the actual results of the real-world experimentation that this research addresses could lead to highly efficient dependency management.

Software Process Improvement

Software Process Improvement (SPI) refers to the set of activities that take place in a software development environment with the goal of enhancing their software development habits at an organizational level. Several formal approaches to its assessment and development exist, like the ISO/IEC 15504 "SPICE" standard or CMMI (Capability Maturity Model Integration), but, although the benefits that following their guidelines is assumed to create are widely acknowledged as convenient, it is very costly, complex and resource-demanding and, although there are examples of adaptations of formal SPI standards to small companies (Tuffley, Grove, & McNair, 2004), it is not usually the case. Instead, it is not uncommon to find lightweight alternatives for reduced benefits at reduced costs (Espinosa-Curiel, Rodríguez-Jacobo, & Fernández-Zepeda, 2013), (Ñaupac, Arisaca, & Dávila, 2012), (Buchalcevova, 2011).

In addition, SPI procedures face another limiting factor: they are usually strongly bound to the context which, along the aforementioned size dependency, is the major reason for the wide spectrum of different proposals available. Unfortunately, there is not too much literature about standardised or widely accepted concrete methodology to develop SPI procedures when required; only maybe Kuhrmann (2015) can be highlighted and it is still a very recent work that, as the author acknowledges, needs reviews and replications in order to become polished and solid enough to gain mainstream acceptance.

As for what concerns to Agile practices in SPI, a significant part of the little existing research focuses on large scale, giving concrete instructions on steps to follow and reporting results unusually positive in Agile contexts this size (Auvinen, Back,

Heidenberg, Hirkman, & Milovanov, 2006), (Kettunen & Laanti, 2008). It is also observable how there seems to be a slow trend of proposing Agile SPI methodologies for general adoption (Abdel-Hamid & Abdel-Kader, 2011) (Salo & Abrahamsson, 2007), of which unfortunately none seem to have had an outstanding impact to date.

3

Case description

3.1 Organizational tree

After the partial adoption of agile practices, Ericsson (the target company) presents an organizational hierarchy consisting of several levels that expands horizontally as navigated deeper (Averianova & Deekens, 2014). As for what concerns to the concrete case, only the PDU LMR CAT (Product Development Unit for Long Term Evolution and Multistandard Radio Base Stations Common Architecture Technology) is described, since it is the environment the study is restricted to.

The PDU LMR CAT is located under the Business and Development units in the global organizational tree. All teams in the PDU LMR CAT, including the two teams whose individuals constituted the set of experiment subjects, utilize a flexible adaptation of Scrum and, organizationally speaking, share a similar structure:

- **Developers**. From five to nine constitute the core of the team. They are the ones who have the strongest and most direct impact in the success of the company as a whole since they are the only ones who get to materially contribute to the development and maintenance of products that can actually be sold by Ericsson and therefore generate direct profit.

- **Scrum Master**. Each team of developers has its dedicated Scrum Master, who features the roles of a generic Scrum Master. This is, they support their team of Developers to ensure that they apply the methodology as it is expected to benefit from it and removes obstacles from them so they can stay focused on producing profitable assets for the company. Also, it shall be mentioned that it is not uncommon to have the Scrum Master executing the role of a Developer as well.

- **Operative Product Owner**. Because of evident size-related issues, it is not feasible to have only one Product Owner role per product in large scale companies. To approach this issue in Ericsson, a three-level hierarchical organization exists, so that the topmost level is closest to the client while the bottommost, the Operative Product Owner, is closest to the developers. Each Operative Product Owner is usually associated to two or three teams, and is responsible for user story prioritization and similar tasks, but always on the scope of the team backlog only.

- **Team Coach**. Since the adoption of agile methodologies is still rather fresh

and could even be considered incomplete, Team Coaches, who feature deep knowledge in the area, are needed to support the resolution of problems related to the organizational change and continue enhancing its adoption when possible.

- **Product Guardian**. In an interesting approach to agile quality assurance, the Product Guardian is the role that takes responsibility for ensuring that the level of deliveries meets acceptable standards. Generally, the Product Guardians are technical experts in the domain of the products they handle (usually from two to four per Product Guardian) and therefore are entitled to support the developers in decision-making activities.

Other roles, both derived and not derived from Agile guidelines, can be found in a regular basis according to expectable patterns across the organisation and some of them in the PD LMR CAT unit in concrete, but they are not situated close enough to the teams and therefore it was decided to exclude them from the study.

3.2 Organizational scope of the study

As described above, the scope of this study is narrowed down to the Developers of the teams considered and the very concrete set of roles that are situated the closest to them (namely Scrum Master, Operative Product Owner, Team Coach and Product Guardian). The reasons motivating this decision are shown below:

- These roles picture the complete scene of a professional Cross-Functional Team.

- Different roles, as the combined presence of each one of them, introduce numerous communication patterns and ways to interact, both among themselves and with the roles considered. While it is undeniable that it would be very interesting to include more, if not all of the roles that affect communication to increase the generalisability of the treatment applied, the number of different types of communication-related interactions would become so high that it would become impossible to handle.

4
Research methodology

4.1 Research purpose

The purpose of this study unfolds as a double intent composed of both (1) enhancing the procedures for distribution of information within the smallest organisational units in large-scale industrial agile contexts and (2) contributing to the body of knowledge through the execution of a experimental study in a real-world large-scale agile setup.

4.2 Research questions

The research questions stated below are aimed at addressing communication challenges in large-scale agile environments and exploring the field of experiments in real and large-scale agile setups to contribute to the body of knowledge.

RQ0 What improvements can be proposed to address communication concerns in the target environment?

In response to the communication problems identified in the target environment a proposal will be elaborated on how they can be addressed, setting the base for the rest of the study, i.e. Research Questions RQ1 and RQ2.

RQ1 What was the actual result of implementing the proposed improvements?

Based on the metrics defined in Section 4.3.1 and on the experiment subjects' satisfaction with the experimental procedure it will be possible to report an empirical evaluation on its success. The answer to this question will constitute empirical evidence of whether the experimented changes worked, not only providing the company with a solid basis for making a decision on this regard, but also by indirectly contributing to understand how communication works in this type of contexts.

RQ2 What was learnt about the execution procedure of quasi-experimentation studies in large-scale agile environments?

This question will support upcoming research by providing (1) example guidelines to be followed on future experimental research in similar contexts and (2) a reference for improving in-house quasi-experimental procedures in the

large-scale agile software development industry (always having in mind the validity concerns of the study described in Section 4.6).

4.3 Experiment design

4.3.1 Experiment variables

As of Thompson (2008), variables in quantitative research are defined as references "to characteristics or attributes of an individual or organisation that can be measured or observed and that varies among the people or organisation being studied, this meaning that scores in a given situation fall into at least two mutually exclusive categories". Depending on how they influence the experiment, variables can be categorized as *independent*, *dependent*, *mediating*, *moderating*, *control* or *confounding* variables (Creswell, 2008).

Independent variables

"Independent variables are those that (probably) cause, affect or influence outcomes" (Creswell, 2008). They are the ones manipulated by the experimenters and, in this study, the next one was identified:

I1 **Procedure for distribution of information to and among the teams**. This was the central focus of Research Question RQ1 and was considered as the treatment variable in a binary basis (this is, it uniquely reflected which ones of the traditional and the treatment guidelines were being used).

Dependent variables

"Dependent variables are those that depend on the independent variables; they are the outcomes or results of the influence of independent variables" (Creswell, 2008). In this study, the next dependent variables were identified:

D1 **Information relevance**: During previous research on the same environment (Averianova & Deekens, 2014), evidence was found indicating that one of the main causes of the communication-derived problems was people getting information that was not really important to them. Therefore, measuring how the relative relevance of information varies between the traditional and proposed communication procedures was unarguably important.

D2 **Information completeness**: The initial interviews revealed how, given that a source had information of actual relevance to give, there were no complaints about too large amounts of relevant information being received (and Dependent Variable D1 exists to evaluate if the information being received is actually useful). However, certain sources which were expected to provide information were not actually doing so. Investigating the concrete details of these communication shortages was considered extremely relevant not only to enhance the current situation but also to prevent a negative trend that could cause communication issues to go out of control.

Mediating variables

"Mediating variables stand between the independent and dependent variables, and they mediate the effects of the independent variable on the dependent variable" (Creswell, 2008). In this study, the next mediating variable was identified:

M1 **Amount of information transferred to and among the teams as for what concerns to the relevant roles (Developers, Product Guardians, Scrum Masters, Operative Product Owners and Team Coaches) in their closest context**. Because previous research in the same environment (Averianova & Deekens, 2014) found evidence to support that most of the problems related to communication were caused either by inadequately large amounts of information being distributed or by inaccurate choices of recipients, it was reasonable to expect that, when smaller amounts of information were handled, the experiment subjects would unwittingly perceive the procedure they are using at the moment as more adequate because the aforementioned causes of problems were minimized, even if it was not because of the procedure used being better but rather because the amount of information handled being lower.

Confounding variables

Confounding variables are not "actually measured or observed in a study". They exist, but their "influence cannot be directly detected" (Creswell, 2008). The following confounding variable was identified in this study:

C1 **Adjacent communication sources' behaviour**. Mainly because of the large-scale character of the target organisation, it was not feasible to involve all people at the organisation in the experiment. However, because individuals had to be left out of it, a situation was created in which subjects that were utilising the experimental guidelines would need to either send information to or receive information from people that would be using the traditional ones instead, as Figure 4.1 shows.
This means that, because communication is a bidirectional act, the effects of the treatment, either positive or negative, on the experiment subjects, could be reduced, to a certain extent, by the fact that they have to communicate to individuals that are not part of the experiment.

No remarkable moderating nor control variables were identified for this study.

4.3.2 Experiment procedure

This study was constructed utilizing a temporarily linear design laid out in three stages of data collection (Figure 4.2) (namely Contextual information gathering, Baseline and Treatment measurements and Satisfaction assessment, see below) plus a fourth one for data analysis and discussion.

1. **Contextual information gathering**. During this stage the experiment subjects were interviewed individually to obtain concrete information about the

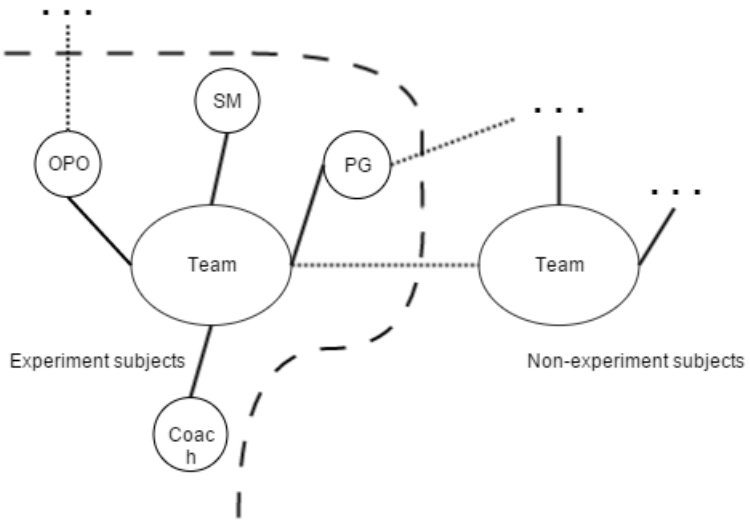


Figure 4.1: Example of communication on the setup boundaries.

communication-related problems that they face in order to build an appropriate treatment.

2. **Baseline measurements**. In this stage the measurements considered relevant for the study were recorded for later comparison against the ones taken when the treatment was applied. At the same time, the treatment was designed with the information gathered during the previous stage. This stage lasted for a full Sprint, which means a total of three weeks, so that a complete basic cycle of Scrum (the methodology utilized at the target environment) is studied. In fact, it would haven been interesting to study as many complete Sprints as possibles to strengthen the reliability of the research and increase its generalizability, but after consideration it was deemed too risky to use more than one Sprint in the concrete case as the same amount of time is required for the next stage for obvious reasons and choosing to use two Sprints for each stage would have tightened the schedule too much (not to mention the increased complexity of the study due to concerns like learning interactions).

3. **Treatment measurements**. With a length similar to that of previous stage, this one was utilized to take the same aforementioned relevant measurements but with the treatment applied to the subjects.

4. **Satisfaction assessment**. Because of how relevant to both the company and the subjects themselves satisfaction with the treatment is (even if the reliability of this type of assessment is moderate), it was considered essential to perform a satisfaction survey as a final data collection stage.

5. **Data analysis and discussion**. A statement on the comparison of the measurements taken during the two previous stages (Research Question RQ1) was elaborated. Also a discussion on the particularities of experimentation on industrial contexts (Research Question RQ2) was reported.

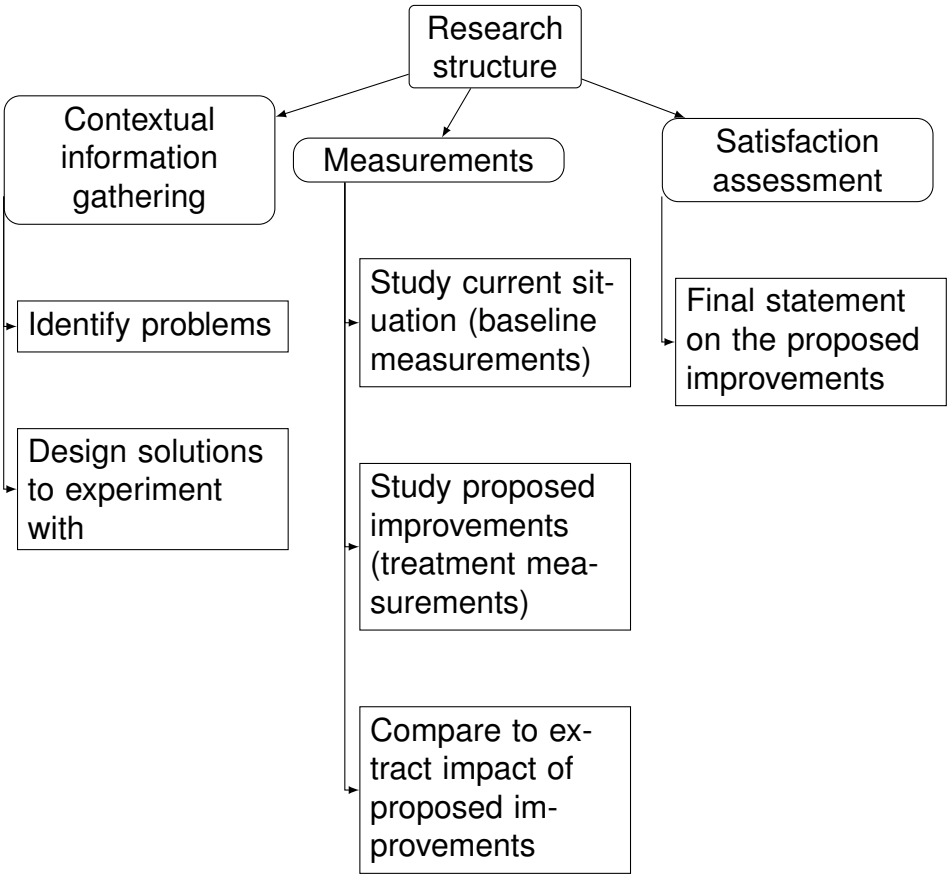


Figure 4.2: Research structure.

More details on the execution of each stage can be found in Chapters 6 and 7.

The design of the research was based on a quasi-experimental setup. This is because both the context was impossible to control as it would be required for a true experiment and the sampling method was convenience due to both (1) sampling being performed by Ericsson representatives according to their internal guidelines and (2) not all potential experiment subjects having the time and/or being willing to participate in the experiment. Every subject was observed for baseline measurement, then learnt about the adoption of the treatment (through a workshop for each of the Scrum Masters plus an additional one for each of the complete teams) and then he/she was observed again, as Figure 4.3 shows (Juristo & Moreno, 2010).

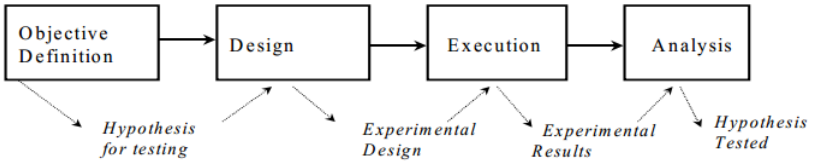


Figure 4.3: Process of experimentation in Software Engineering.

Therefore, the setup was a Single-Group Interrupted Time-Series design (Creswell, 2008), with the design specified in Table 4.1.

Group A	O___X___O
Group B	O___X___O
	⋮
Group N	O___X___O

Table 4.1: Per-group experimental setup, sample size N.

Finally, it shall be noted that, because of the complexity that the large-scale character of the target organisation introduced in the context of this study, it was not intended to develop as a strict, scientific true-experiment (nor it should be understood as such), as the desirable situation of having a controlled setup would have been totally unfeasible. More details on this context-induced particularity can be found in Section 7.2.

4.4 Data collection

The sample from which the data is collected throughout the execution of the experiment consists of the developers of two of the teams in the PD LMR CAT unit, plus a given set of their closest roles, Scrum Master, Operative Product Owner, Team Coach and Product Guardian. More information on the organizational scope of the study can be found in Section 3.2.

As for the procedure, two different ones are considered: semi-structured interviews (namely "initial interviews"), which are used at the beginning of the study for deeper understanding of the problems that have driven Ericsson to request this study in order to attempt to maximise the efficiency of the treatment designed, and surveys. Two surveys can be distinguished in this study:

1. The measurement survey. This survey is used for collecting empirical measurements on both the baseline and the treatment guidelines. Such measurements are the relevance (Dependent Variable D1) and correctness (Dependent Variable D2) of information transferred, and they are gathered daily through a role-independent online form (using Typeform[1]) during the course of one Sprint. It shall be understood that in a setup like the proposed one, in which a need for communication can arise at any given moment, having one researcher observing each test subject, which would be needed in order for the survey observations to be fully performed by the researchers, is not feasible at all. This survey is longitudinal: it is filled on a daily basis during a preset period of time (one Sprint), and then, after the treatment is applied, it runs for another period of time (which, in the concrete case, is enforced to be of the same length). In addition, it was tested with a small set of MSc-level students before its real usage to ensure that there were no understandability problems even with subjects that are not fully knowledgeable of the situation at the target organization.

[1] `http://www.typeform.com/`

2. The satisfaction survey is a regular cross-sectional *Pen and Paper* inquiry. It is performed at the end of the experiment in two different per-team sessions.

Finally, observation is utilised during the whole execution of the experiment with the intents of (1) clarifying doubts, (2) increasing the quality of the study through deeper knowledge of the workspace, (3) extracting information useful for addressing Research Question RQ2 and (4) improving the interactions between the researcher and the experiment subjects by learning their schedules and habits.

4.5 Data analysis

Firstly it shall be mentioned that, because one of the most important hurdles to communication at the target organisation seemed to be channel heterogeneity (Averianova & Deekens, 2014), it would have been counterproductive to have deep adaptations of the experimental procedure to every role in the groups the experiment is run on, so the same procedures for analysis of the data were utilised independently of the role being studied. However, because of their diverse communication needs, it was likely that the different roles would assess the experimental procedure differently and according to different criteria. This is taken into account when judging the success of the experimental treatment.

Below is shown how the information obtained through each of the aforementioned data collection methods is analysed.

4.5.1 Initial interviews

Brown and Stockman (2013) reasonably used thematic analysis to analyze the data gathered during a study on habits of HCI focused on communication among family members that, despite sharing a close context, feature incompatible schedules and different roles in their *organisational structure* that force them to depend on technology to communicate. Because of the strong similarity with the context of the study, as far as to what is relevant to this research (communication) is concerned, it has been decided to use thematic analysis as well, supported with a software tool for processing of qualitative data called NVivo[2].

As Braun and Clarke (2006) detail, thematic analysis is, in most cases, a procedure consisting of six phases. The adaptation of each of the phases to the concrete study is detailed below.

1. **Familiarization with the data**. All notes collected during the interviews were read twice in a sequential manner; this is, first all of them were read once, and then once again in the same order. A special focus was put on identifying reiterated and/or strong references to same topics.

[2] `http://www.qsrinternational.com/products_nvivo.aspx`

2. **Generation of initial codes**. Codes were extracted from the data gathered without establishing any limitations on the sources and in an ambitious pattern (this is, selecting as many codes as possible) for later filtering, based on the repetitions patterns mentioned before. Seven different codes, like Information lacking practical value and Communication with the Team Coach, were identified in this study.

3. **Search of themes**. All seven codes found were grouped under two themes according to the scope of communication they refer to.

4. **Review of themes**. The review was performed on a code-level as there was no data available to discard whole themes without deeper reasoning. All codes were reviewed attending to quantitative data about the amount (and characterization, if applicable) of the corresponding references and only the two that clearly stood out as truly troublesome were kept.

5. **Definition and naming of themes**. This step was used as a consistency review of the themes previously generated. No actual changes occurred.

6. **Production of the report**. The final thematic map was reported and discussed in a narrative argument that is able to support both (1) the understanding of the communication problems at Ericsson Lindholmen and (2) the appropriate design of the treatment, which effectively were the goals of this set of interviews. This narrative is reported in Section 6.1.

These interviews were performed during the early days of the study with the main objective of gathering information on what the concrete problem with communication among the relevant roles (Developers, Product Guardians, Operative Product Owners, Scrum Masters and Team Coaches) is.

According to Hove and Anda (2005), certain information should be described in studies reporting interviews to assess the quality of the results obtained. In concrete:

- Nineteen subjects across two teams were theoretically available for this round of interviews. The teams were selected for their predisposition to try innovative methodologies for work and communication, and the subjects were filtered out if they were not one of the following relevant roles: Developer, Scrum Master, Team Coach, Product Guardian or Operative Product Owner (more details on the reasoning behind this role-based filtering rule can be found in Section 3.2).

- It was found that, in a very marginal amount of cases, the subjects were acting both as Developers and as another, different role, like Scrum Master for example. For these cases, the subjects were studied and accounted for only as the least common role as (1) the amount of occurences of this situation was extremely low and (2) it was understood that their Developer views were different from those of pure Developers because of the implicit influence that having also another role could have on them.

- Regarding interview scheduling, the subjects were approached individually through e-mail as it was regarded as the established way for appointment arrangement in the target environment. Upon the initial invitation being rejected, a new one was sent after review of the subject's calendar, but not a third one (instead it accounted for mortality because it was considered, based on the others' subjects average confirmation time of under an hour, that the interview was not of enough priority to gain a slot in their schedule). Lack of confirmation of the meeting after one week of the initial approach was accounted for mortality as the corresponding subject(s) were part of the study until then.

- The interviews were performed in a meeting room in the same floor where the teams usually work (in other words, in their comfort environment) and were always kept under a maximum duration of twenty minutes to minimize the disturbance caused on the subjects' regular schedules. There was only one interview with each subject as it was considered that repeating it would not have yielded an amount of new information sufficiently large to compensate the costs it would have taken.

- There was only one interviewer, the main author of this study.

- It was chosen to not to record the interviews to avoid both altering the data gathered because of the interviewees feeling intimidated and increasing opt-outs. Since the goal of these interviews was to gather information that would be utilized only once and with a creative purpose (designing the experimental treatment) rather than a descriptive one, having an exact transcription of every interview was not as relevant as having a large amount of sincere information that would allow the researcher to develop an experimental treatment that could help solving (some of) the issues with communication in the target context.

- The interview guide can be found in Appendix A. It was reviewed by both the company and the university supervisors before execution.

Most of the guidelines given by the same authors to address interaction issues (facing subjects that barely talk or deviate from the interview questions) were followed when performing the interviews, except having two interviewers instead of one as the majority of the arguments given on its favour are rather weak and/or context-dependent, as the own authors acknowledge (Hove & Anda, 2005).

4.5.2 Measurement surveys

The measurements provided by these surveys are shown both independetly and as a comparison. Per-role graphics are utilized for the assessment of both Dependent Variable D1 and Dependent Variable D2.

4.5.3 Satisfaction survey

The analysis of the data extracted from the satisfaction survey is performed both on a role-independent and role-dependent basis because, and as described in Chapter 3, although the communication requirements of every role were different and therefore it was reasonable to expect their satisfaction with the treatment to be different as well, not all roles have the same level of problems with communication. For this reason, the relevance of each role's satisfaction with the treatment is different as well. It is reported in a tabular manner with the intent of providing a firm statement to assess the feasibility of the proposed treatment.

4.5.4 Observation

The analysis of the data gathered through observations responds to an informal rationale procedure for conclusion elaboration. Information built upon this data is mentioned as deemed influential for both the design of the treatment and Research Question RQ2.

4.6 Threats to validity

The following threats to validity have been identified, following Wohlin, Runeson, Höst, Ohlsson, and Regnell (2012)'s guidelines.

Conclusion validity

"Threats to the conclusion validity are concerned with issues that affect the ability to draw the correct conclusion about relations between the treatment and the outcome of an experiment" (Wohlin et al., 2012).

The most important threat to the conclusion validity ot this study is the reliability of the measurements, because (1) human judge is strongly involved (in the filling of the surveys) and (2) due to non-completely controllable character of the quasi-experimental setup there are some details of the treatment that cannot be fully/realistically implemented. To account for this, each of the subjects was always enforced to attempt to maintain a constant criteria during the measurement stages, even if this criteria was different across the different individuals.

Reliability of treatment implementation and random irrelevancies in the experimental setting are considered inherent to quasi-experiments as they are consequences of not being able to control every variable on a setup, which is also the cause of random heterogeneity of subjects in the concrete case. Reliability of treatment implementation was ensured by using similar procedures to apply the treatment on all subjects, but random irrelevancies were intentionally unaccounted for, as they are an important factor of the day-to-day of the target environment and, if removed, the results of the study would, indeed, gain in reliability, but become useless in practice.

Low statistical power and violated assumptions of statistical tests are not relevant to this study as there is no intention to claim statistical significance, nor so are fishing and the error ratio as the data gathered during each stage is treated as a whole set, not as an analysis repeated several times.

Construct validity

"Construct validity concerns generalizing the result of the experiment to the concept or theory behind the experiment. Some threats relate to the design of the experiment, others to social factors" (Wohlin et al., 2012).

A threat to the construct validity of this study is interaction of testing and treatment: because the subjects know that the treatment that they are put through is oriented to improving their communication habits and it involves judging the actual value in their communications, they may unconsciously try to increase (or decrease) their quality. Accounting for this was impossible as the subjects had to modify their behaviour actively with the treatment, and therefore there was no way to have them treated without them knowing (this is, performing a blind test).

Evaluation apprehension is not considered to apply to this study as the subjects have been explicitly chosen because of their availability and interest on participating in this type of studies. Neither is hypothesis guessing, as the subjects were knowledgeable from the beginning about the goal of the experiment.

In addition, it is worth mentioning that, during the execution of one of the surveys one of the relevant roles, the Product Guardian, changed person for one of the teams, which had implications for the study as this new Product Guardian had not been involved in the research until then. Because this brings a relatively large set of additional confounding factors into play without adding a significant amount of relevant information (especially when taking into account that the communication between the Team Coach and the Product Guardian is expected to be very little when compared to that between the Team Coach and any other role in this study), it was decided to discard the data gathered from the former Product Guardian when the change happened and not to gather any from the newer one to protect the reliability of the rest of the information.

Finally, the author acknowledges the possible existence of experimenter expentancies, this is, a bias created by an inherent desire of obtaining successful results that may increase the interest of the study for both researchers and practitioners and lead to future work and further development of the treatment guidelines. To address this type of threat usually a technique called double-blind experiment is generally used. This forbids the researcher from distinguishing the control and treatment groups while the treatment is being applied, suppressing any kind of bias oriented towards favouring any of the groups on such application. However, and as it can be seen on the experiment design (Table 4.1), every group receives the treatment, so using a double-blind technique is not possible. As an alternative, unbiased reviews will be

requested before turning into converting the study into a publication that can be used as a reference by future research, as schedule issues make it impossible to have them requested earlier.

Internal validity

"Threats to internal validity are influences that can affect the independent variable with respect to causality, without the researcher's knowledge. Thus they threat the conclusion about a possible causal relationship between treatment and outcome" (Wohlin et al., 2012).

One of the most relevant threats to the internal validity of this study are confounding variables. As discussed earlier, confounding variables are those whose effect, despite influencing the outcome of the experiment, cannot be determined (Juristo & Moreno, 2010). Confounding variables are very common, but not always extremely relevant if the interactions between them and the studied variables are small when compared to the ones between the studied variables themselves. This type of verifications are usually performed at the end of the study to validate the results and increase their precision, as it would occur with Confounding Variable C1 in the particular case, if this was not one of the particularities of experiments in large industrial contexts as discussed in Section 7.2.

Other threats that can be considered to limit the internal validity of this study are mainly selection bias and regression toward the mean. Because the amount of information usually transferred in the workplace that is target of the study is not in the scope of this study (and descriptions usually given are 'a lot', 'too much' and 'chaotically large', which are terms that carry little to no scientific meaning at all), it is impossible to ensure that the subjects selected for the experiment constitute a sample communication-wise representative of the general population, nor that it does not deviate from the mean over the course of the experiment (which could be considered a maturation threat if the length of the experiment was not more than just a few weeks).

Nevertheless, to account for this a reasonable measurement of the amount of information transmitted in the target context before commencing the study would have been needed in order to deliberately prepare a representative sample, and then the problem of selection bias would have appeared instead.

External validity

"Threats to external validity are conditions that limit our ability to generalize the results of our experiment to industrial practice" (Wohlin et al., 2012).

The major external threat affecting this study is its generalizability. It shall be clearly stated that this study is not generalizable by itself on a reliable basis. This study is a quasi-experiment and therefore it cannot be guaranteed that the treat-

ment was the reason for the differences between the values of the measurements and, even if it actually was, the study was performed in a single, very concrete environment, which is also the same in which previous research relevant to it (Averianova & Deekens, 2014) was executed, so it is risky to attempt to generalize the results of this research out of this concrete context before any replications are performed.

History-treatment effect could be an issue because the design of the procedure was slightly influenced by the results of Averianova and Deekens (2014) in what concerns to problems found, which means that it moderately depends on data gathered some time ago. Fortunately, this influence was relatively light and the amount of time that had past was not too long so the fact that a few months have passed by should not have had relevant, if any, influence on the communication habits of the targeted environment.

Finally, setting-treatment and selection-treatment interactions are not considered as concerning because the proposed study is not intended to be generalized before replications are performed.

Other

Other threats related to the social aspect of the research environment build upon the possible friendship relationships that can appear between the researcher and experiment subjects, and the 'friendly' behaviour that the latter may decide to approach the experiment with in order to benefit the former. However, because the teams would be adopting the experimental procedure should the study show positive results, this issue is discarded: it is reasonably unlikely that the subjects would be willing to adopt communication procedures for their day-to-day job that they have found to be worse just to benefit the researcher.

5
Treatment design

As it can be read with more detail in Section 6.1, the results from analyzing the data gathered in the interviews list seven possible sources of communication problems in the context studied, and therefore of issues prone to be experimented on. Below can be found discussions on which of these are issues that (1) are considered irrelevant or not relevant enough and therefore are excluded from the study and that (2) are actually considered relevant and therefore selected to be addressed during the experiment by being addressed through the treatment.

5.1 Discarded codes

The following codes were discarded. Along each code can be found the reason(s) why they were discarded.

- *Communication with the Scrum Master and Developers.* The results of the initial interviews (Section 6.1) clearly showed that the candidate issue behind this code was the best regarded one, and the margin for improvement was so tiny that trying to address it would have created a situation with very little room for improvement when compared to that for declination. Therefore it was deemed that attempting to enhance the issues pointed by this code was not compensated by the risk it implied.

- *Communication with the Product Guardian.* Because, as discussed in the results of the initial interviews (Section 6.1), Team B could be considered an outlier in the target environment and Team A was experimenting on its own with new ways of communicating with the Product Guardian, the data gathered about this code was rather inconclusive. Therefore it was considered that sufficiently strong reasons for justifying experimentation with changes that address the issues grouped under this code could not be found.

- *Communication with the Operative Product Owner.* The data associated to this code reports results very similar to those of *Communication with the Scrum Master and Developers* (Section 6.1), and therefore the same rationale is executed: it is considered that the risk that trying to enhance communication with the Operative Product Owner does not compensate the extremely small improvement that can be achieved.

- *Communication channel.* It is well known, and supported by research (Lipinski-Harten & Tafarodi, 2013), that face-to-face communication is overall more effi-

cient and less time-consuming than computer-mediated communication (CMC). Fortunately, the results of the initial interviews (Section 6.1) showed that face-to-face communication seemed to be the most recurred method for communication. The question of which among usage of phone and CMC should be most prioritised was considered a minor detail that would probably require major focus on each subject's current task and the concrete targets each they communicate with, reasons why it was understood out of the scope of this study. Therefore, once again no strong reasoning for proposing improvements related to this code was found. Because the choice of communication channel was up to each individual, disrupting their comfort habits without a clear justifying statement would have likely turned counterproductive and also harmed the rest of the experiment.

- *Awareness of alien work.* According to the results of the initial interviews (Section 6.1), only a relatively small amount of the subjects interviewed wanted a change on the communications that they got related to other teams' work situation. Moreover, it was not possible to find a consensus on the interviewee's opinions about which alien information they are interested in nor how it should be given to them (multi-team meetings, using some of the already existing roles to route it, filter relaxation, regular presentations, etc.). This combination of uncertainty and lack of interest caused this code to be discarded.

5.2 Selected codes

The following codes were selected. Along each code can be found the reason(s) why they were selected.

- *Communication with Team Coach.* Clearly regarded as the code with the largest room for improvements (Section 6.1), communicating with a Team Coach was impossible in practice as it was a role that apparently only existed in paper. However, it is not straightforward to know if (1) the lack of this communication has supposed a problem or harm to the teams' and/or their work or (2) it did not simply because the Team Coach as a role was never really needed. It was acknowledged that selecting this code would allow to confirm either of those.

- *Information relevance.* As Section 6.1 shows, a large percentage of the interviewees were annoyed by or not interested in regular, global product-related communications that are sent on a regular basis to a wide spectrum of employees "just in case". Preventing the subjects from having this type of communications (or, at least, minimizing them as much as possible) and studying their reactions was considered interesting as it would allow to start building up a solid, argument-backed empirical reasoning towards reducing the amount of irrelevant information transferred in large-scale contexts similar to the one under study.

5.3 RQ0. Treatment

The design of the treatment responded to elaborating proposals on how to enhance the issues derived from the two selected codes listed above. With this in mind, this section is structured in two subsections that break the treatment down according to which of the aforementioned codes they address.

5.3.1 Enhancing the communication with the Team Coach

As described above, it was straightforward that the Developers lacked communication with the Team Coaches, but it is unclear if this communication was really necessary, as they already had a Scrum Master whom they communicated with very fluidly and that had knowledge of Scrum at least equal to that of the Team Coaches, since these were just Line Managers who had gone through an adaptation workshop but lacked the time to keep on learning on their own. This limited the extent to which they could actually coach the teams in practice.

As a candidate solution it was proposed to defer the coaching responsibilities from the Line Manager role and empower the Scrum Master one, which would become a full-time role free to also support Backlog work through direct contributions if time allows for it but that is not committed to do so as there may be impediments or similar issues that he/she must address with higher priority. This would bring two main benefits:

1. **The Line Manager role would become a full-time one** again, which seemed to be a necessity as the attempt made by the company to combine it with coaching seemed to be totally unsuccessful.

2. **The communications between the coaching entities and their teams would become extremely fluid**, which would also cause the teams to improve at their usage of Scrum practices.

In order to implement this the next actions were executed:

- Instruct the Scrum Masters about their temporary new role and let they know that contributing to pulling items from the backlog should become their lowest prioritized task during the Sprint in which the treatment is applied. This instruction was performed through a workshop whose material can be found in Appendix C.

- Let the Developers know that the Team Coach role was embedded into the Scrum Master one during the duration of the Sprint in which the treatment was applied. As an approach to simulate the experience that a coaching role is usually characterized by and which would be unfeasible to teach within the scope of the study, the Scrum Masters were allowed to consult and discuss with the researcher about Scrum practices as he had more theoretical knowledge and time to research about them.

- Let the actual Team Coaches know about the fact that the teams experimented with would not appeal nor report to them as Team Coaches (interactions with them as Line Managers would not suffer any changes) during the whole duration of the Sprint in which the treatment was applied. Even if the Team Coaches dropped off of the study at the very beginning, they did not know that the treatment would stop all communications coming (supposing there were some) from the two teams experimented with. For obvious reasons, they should know about this and were therefore notified accordingly.

5.3.2 Increasing information relevance

It was observed that all the information described as annoying, uninteresting or lacking a practical value happened to be distributed through e-mail. Therefore, probably the most rational approach to suppressing these unnecessary communications would be to have the uninterested addressees let the senders know about it. However, in large-scale contexts, like the one studied, actions apparently simple may become surprisingly complex and/or slow, and to look for all sources of value-lacking product-related information, to explain them the study and to have them remove the concrete set of recipients only for the specified period of time would have probably been a very good example. Instead, it was decided to simulate such request using automated e-mail filters that would silently place global and poorly-valued product-related information in a separate folder allowing the individuals to skip any undesired interactions with them but keeping them stored in case they were needed later. For confidentiality reasons the details of these filters can not be revealed, but it was ensured that they were the same for all subjects across the experiment during the whole duration of the exposure to the treatment.

6
Results

6.1 Initial interviews

Nineteen subjects were initially considered. One of them (a Developer) was on parental leave and accounts for mortality of the study. One additional Developer on each team, one Operative Product Owner and both Team Coaches did not respond to the aforementioned approach procedure and account for mortality of the study. This sums up for a total of thirteen subjects interviewed and kept in the study, distributed as Table 6.1 shows. As it can be seen, Team Coaches were conspicuous by their absence. Further in the analysis it is shown the reason of this and why it is not a critical matter of concern.

Table 6.2 shows the thematic map created as a result of performing the corresponding analysis (Braun & Clarke, 2006) explained in Section 4.5.1.
Detailed information on the information collected using this map is reported below.

Communication with the Team Coach

This code includes all references, either implicit or explicit, that try to describe, quantify or qualify the communication (or lack thereof) between the Team Coach and other roles in a team.

Ten out of thirteen interviewees referenced this code a total of eleven times. The average coverage of this code among the aforementioned ten interviewees (including the interviewer's dialog) is 6.0%, with extremes of 1.52% and 12.98%.

Figures 6.1 and 6.2 show the distributions of the references based on how they qualify the communication with the Team Coach globally first, and then filtered by team and by role.

As it can be observed, the current position of the Team Coach role from the point of view of communication is considered mostly negative. Having also in mind that, despite the fact that every Team Coach is assigned to several teams, both teams in this experiment happened to have their Team Coaches workplace located along theirs, it is reliable enough to assume that the interviewees who did not address this topic were probably so used to this lack of communication that no longer worried about it. Summarising, these results clearly show a demand for increasing the communication with the Team Coach. This data shows that this code is a potentially

Role	Subjects interviewed	Subject participation ratio
Operative Product Owner	1	0.5
Developer	8	0.72
Product Guardian	2	1
Team Coach	0	0
Scrum Master	2	1
Total	13	0.65

Table 6.1: Demographic distribution of the subjects interviewed.

Global theme	Organising themes	Codes
Communication in the context of agile teams in a large scale environment	Within-teams communication	Communication with Team Coach Communication with Scrum Master, Developers Communication with Product Guardian Communication with Operative Product Owner Communication channel
	Between-teams communication	Information lacking practical value Awareness of alien work

Table 6.2: Thematic map of interview data.

good candidate to be improved, with the most likely cause of this (as deemed by the informal observations performed) being the fact that the Team Coach role usually co-exists with another one in the same person. This is the Line Manager role, a well-established figure that has existed in the target organisation since the days before the agile transformation. During the interviews, most subjects mentioned that the Line Manager is a role known to have too many duties and, therefore, very little time. Because the line management duties are seemingly essential, this causes the Team Coach role to be too close to nonexistent in practice. Fortunately, the other roles interviewed were fully conscious of this situation and gave their points of view and wishes for it to change ("Very low interactions with the team coach as a role [...]", "And (with the) coach they don't talk [...]", "Wants some more communication with the coach, delimit his role a little bit better. More of an active personality [...]"), so the total lack of responses from the Team Coach as a role was not an issue as critical as it could seem at the beginning.

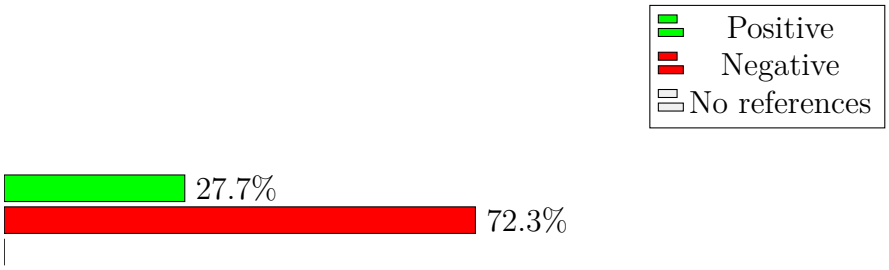
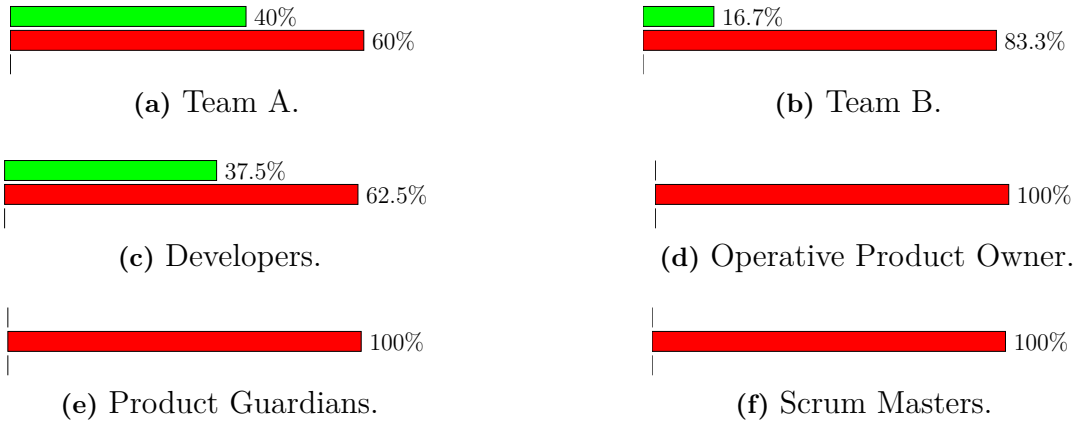


Figure 6.1: Global characterization of the references to *Communication with Team Coach*.



(a) Team A. **(b)** Team B. **(c)** Developers. **(d)** Operative Product Owner. **(e)** Product Guardians. **(f)** Scrum Masters.

Figure 6.2: Detailed breakdown of the characterizations of the references to *Communication with Team Coach*.

Communication with the Scrum Master and Developers

This code includes all references, either implicit or explicit, that try to describe, quantify or qualify the developers and Scrum Master's communication (or lack thereof) with other, less central roles.

Nine out of thirteen interviewees referenced this code a total of ten times. The average coverage of this code among the aforementioned nine interviewees (including the interviewer's dialog) is 5.75%, with extremes of 1.80% and 19.73%.

Figures 6.3 and 6.4 show the distributions of the references based on how they qualify the communication with the Scrum Master and the developers first globally, and then filtered by team and by role (excluding the Team Coach since its response ratio is 0 as explained above).
As it was expected, communication in the context of the Developers and the Scrum Master is regarded mostly positive. The lack of references from both Scrum Masters and the Operative Product Owner is probably caused by the fact that they share the same workplace with the teams in the experiment, which makes face-to-face communication trivial and therefore easy to take for granted when describing communication habits.

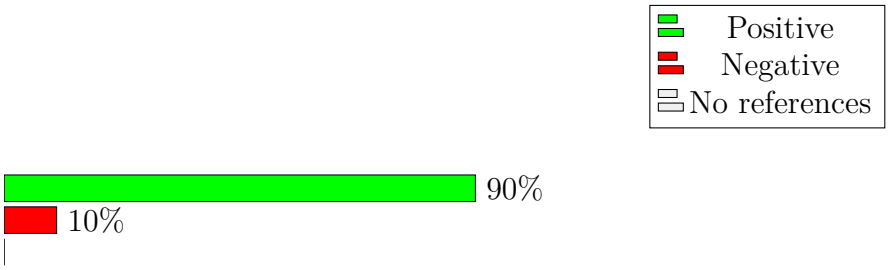
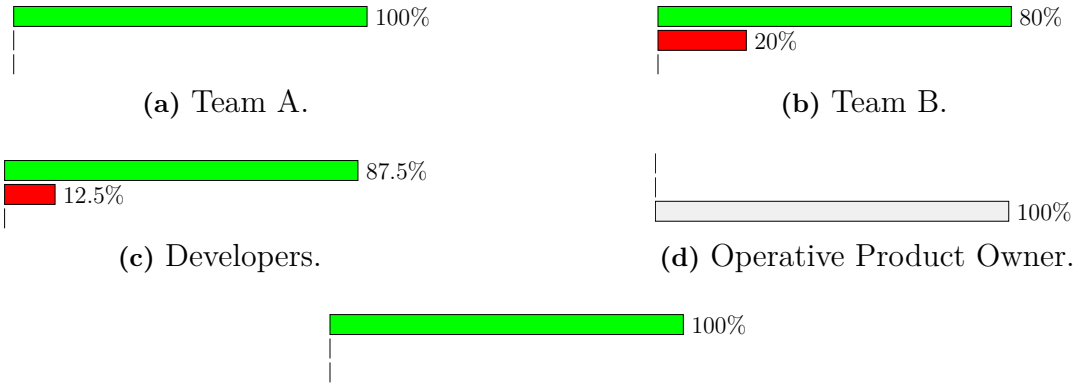


Figure 6.3: Global characterization of the references to *Communication with Scrum Master and Developers.*



(a) Team A.

(b) Team B.

(c) Developers.

(d) Operative Product Owner.

(e) Product Guardians.

Figure 6.4: Detailed breakdown of the characterizations of the references to *Communication with Scrum Master and Developers.*

Communication with the Product Guardian

This code includes all references, either implicit or explicit, that try to describe, quantify or qualify the communication (or lack thereof) between the Product Guardian and other roles in the close context of a team (the rest of roles on this study).

Nine out of thirteen interviewees referenced this code a total of ten times. The average coverage of this code among the aforementioned nine interviewees (including the interviewer's dialog) is 5.58%, with extremes of 1.27% and 13.83%.

Figures 6.5 and 6.6 show the distributions of the references based on how they qualify the communication with the Product Guardian first globally, and then filtered by team and by role (excluding the Team Coach since its response ratio is 0 as explained above).
It can be seen that overall the results are inconclusive, as the only filter that provides a basis for clear, straightforward judgement is the Scrum Master role, and since only one Scrum Master referenced this code the result of the characterization would unavoidably be 100% or 0% either way.

However, some knowledge of the context of the teams studied can contribute to

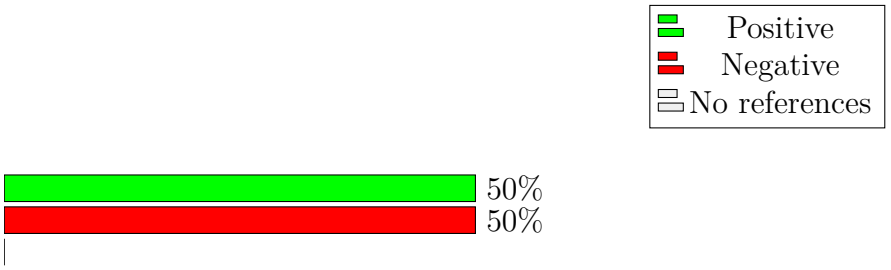
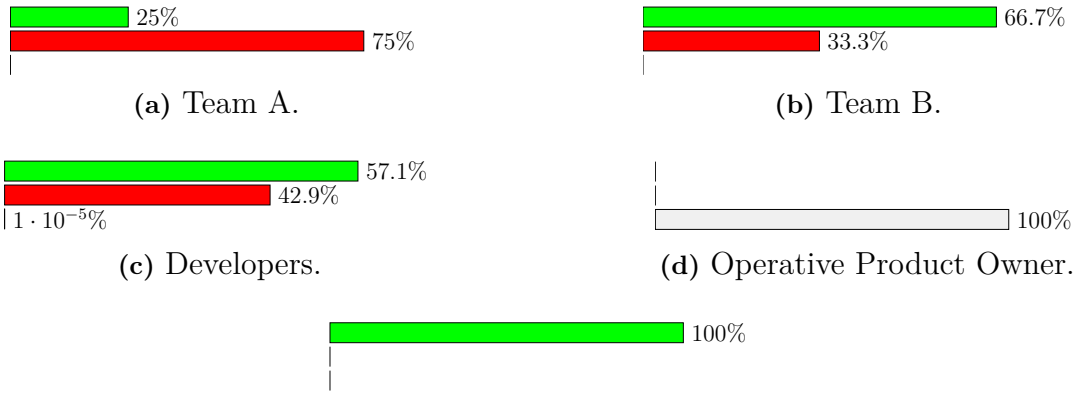


Figure 6.5: Global characterization of the references to *Communication with Product Guardian*.



(a) Team A.

(b) Team B.

(c) Developers.

(d) Operative Product Owner.

(e) Scrum Masters.

Figure 6.6: Detailed breakdown (per Team A, Team B, Developers and Scrum Masters) of the characterizations of the references to *Communication with Product Guardian*.

enlighten this situation of uncertainty quite a lot by clarifying what the gathered data means. In concrete, the helpful parts of this knowledge are listed below:

- Team A has explained to the researcher that, for them, communication with the Team Coach was becoming so chaotic just prior to the beginning of this study that they decided to act on their own and establish an experimental procedure for it which, although seems to be better liked than the previous situation, is still far from being optimal.

- The situation of Team B with regards to the Product Guardian is different from normal as in this this role sits in their workplace, and therefore happens to have an unusually large amount of face-to-face communication with them and is able to join their meetings on a frequent basis.

In summary, the team that has decided to address the problem by itself has shown very tiny improvement, and the team that features an ideal situation shows acceptable results that would probably be good to spread to other teams before trying to enhance them, as it is hard to assess the feasibility of this having only one team in such situation.

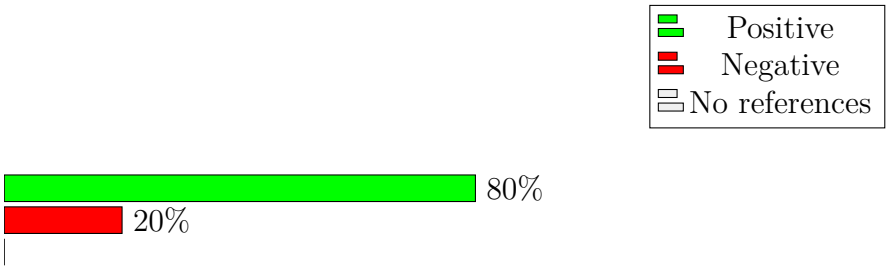
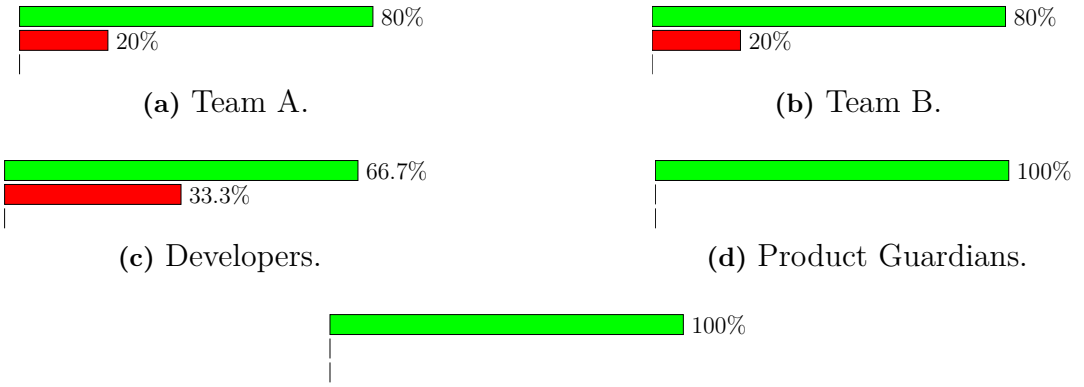


Figure 6.7: Global characterization of the references to *Communication with Operative Product Owner*.



(a) Team A. **(b)** Team B. **(c)** Developers. **(d)** Product Guardians. **(e)** Scrum Masters.

Figure 6.8: Detailed breakdown (per Team A, Team B, Developers, Product Guardians and Scrum Masters) of the characterizations of the references to *Communication with Operative Product Owner*.

Communication with the Operative Product Owner

This code includes all references, either implicit or explicit, that try to describe, quantify or qualify the communication (or lack thereof) between the Product Guardian and other roles in the close context of a team (the rest of roles on this study).

Nine out of thirteen interviewees referenced this code a total of ten times. The average coverage of this code among the aforementioned nine interviewees (including the interviewer's dialog) is 5.76%, with extremes of 1.80% and 13.83%.

Figures 6.7 and 6.8 show the distributions of the references based on how they qualify the communication with the Product Guardian first globally, and then filtered by team and by role (excluding the Team Coach since its response ratio is 0 as explained above).

As opposed the data of the previous code, this time the results are much clearer and straightforward to interpret. The perception of the communication with the Operative Product Owner is really close to completely positive. Furthermore it should be recalled that, as explained above, the team that reported the largest

amount of negative references, Team A, happened to be experimenting with their own proposal to enhance communication with the Product Guardian by using the Operative Product Owner as a router, which means that this role featured less time than usual for its regular tasks.

Communication channel

This code includes all explicit references that mention the different procedures for communication used at the target environment. Communication channel priority is not a dichotomous variable and therefore determining implicit references is impossible in practice, so they are left out.

All thirteen interviewees referenced this code a total of seventeen times. The average coverage of this code among all the interviewees (including the interviewer's dialog) is 7.14%, with extremes of 0.99% and 17.83%.

Figure 6.9 shows the distributions of the priorities that each communication channel is assigned by the studied individuals.
From the global figure it is understood that the majority of the communication is carried out face-to-face whenever possible, falling back to e-mail, then IM (Instant Messaging services) and phone.

The per-role breakdown graphs show similar patterns and are therefore omitted to minimize clutter.

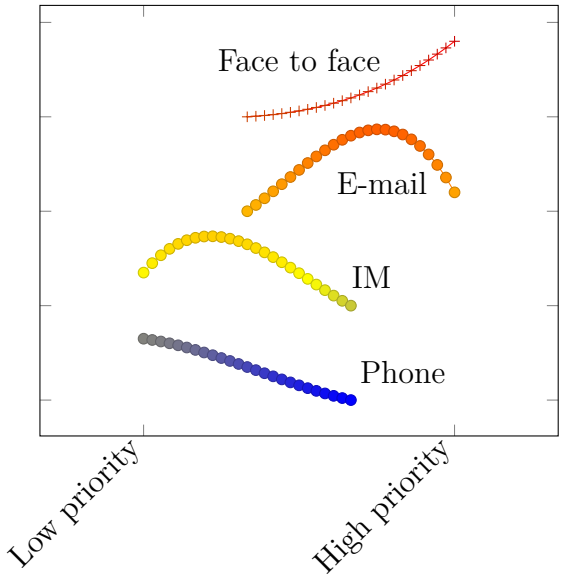


Figure 6.9: Distribution of channel priorities among the interviewees according to the code *Communication channel*.

Information lacking practical value

This code includes all explicit references that mention one or more sources of information lacking practical value for its receiver(s) (from each receiver's point of view). Again, the source referenced is not a dichotomous variable and therefore determining the source from an implicit reference is impossible in practice, so they are left out.

Eleven out of thirteen interviewees referenced this code a total of eleven times. The average coverage of this code among all the interviewees (including the interviewer's dialog) is 7.15%, with extremes of 1.21% and 15.03%.

Figure 6.10 shows the sources considered to deliver the largest amount of information without practical value overall.
In the concrete case no further breakdown is provided as it is considered unnecessary. Only one of the subjects, a Developer from Team B, does not find any of the information received unecessary or non-valuable, and just another one of the remainders disagrees on which is the problematic source.

Nevertheless, it is safe to conclude that product-related information is the least-valuable overall as products in large-scale companies are huge assets handled by considerable amounts of people who usually work in distributed setups and in very different parts of it, to the extent that they know almost nothing, nor barely need to, beyond their team and its closest environment.

Awareness of alien work

This code includes all explicit references tat mention desire and/or demands for knowing the duties of other teams and their statuses.

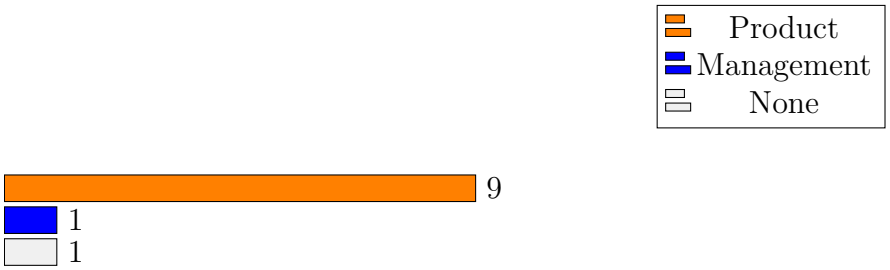


Figure 6.10: Global identification of sources of poorly-valued information according to the code *Information lacking practical value*.

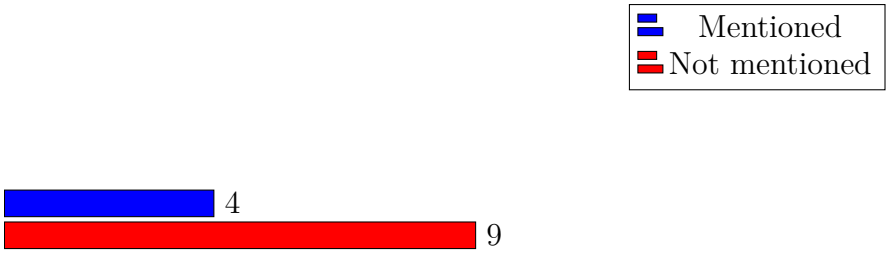


Figure 6.11: Global description of the mentions to interest in alien work according to the code *Awareness of alien work*.

Only four of the interviewees referenced this code, and they did a total of four times (therefore one each). The average coverage of this code among all the interviewees (including the interviewer's dialog) is 11.33%, with extremes of 8.00% and 19.75%.

Figure 6.11 show a comparison between the amounts of individuals who mention this code and those who do not. Because this code is referenced relatively seldom, the figures below include not only the references but also the lack of them, as it could happen that the latter overwhelm the former, therefore minimizing the relevance of this code as a source of problems.

Because the reference ratio of this code is so low (below half of the interviewees did not mention it), it is hard to characterize it as actually relevant. The distribution of the results of this code is listed in a tabular manner in Table 6.3 but no explicit graphic shape is reported.

	Team A	Team B
SM	1	1
Developer	1	1

Table 6.3: Demographic distribution of the interviewees for the code *Awareness of alien work*.

6.2 Baseline measurements

To continue with the study a daily survey should be filled daily by the experiment subjects.

In accordance with what Kelley, Clark, Brown, and Sitzia (2003) state, "when reporting survey research, it is essential that a number of key points are covered". Most of these points are addressed below, except for the ones oriented to conclusion withdrawal, as this survey was intended to obtain data for later comparison with the treatment measurements, but not to be able to build up reasoning for any conclusions at all. Concrete information on how this comparison was performed can be found in Section 4.5.2.

1. The intent of this survey was to perform measurements on the way information is distributed at the target context.

2. This survey was necessary to obtain baseline measurements to compare the treatment against so that its efficiency and overall success could be effectively and impartially evaluated.

3. The potential subjects for the survey were identified by the representatives of the research at Ericsson based on their predisposition to the study. The subjects were initially approached by the researcher on their regular workplace and then took the survey daily on the Internet and on their own.

4. A document-type version of the survey utilized (including extended explanations) can be found in Appendix B. It was reviewed by both the company and the university supervisors before execution.

5. The response ratio is shown in Table 6.4 (the Team Coach role is not included as both of them dropped off the research during an earlier stage). During this survey one of the Product Guardians dropped off as the role moved to another person until then not included in the research. More details on this decision can be found in Section 4.6.

Role	Subjects involved	Response ratio
Operative Product Owner	1	0.26666
Developer	8	0.55
Product Guardian	1	0.2
Scrum Master	2	0.56
Total	12	0.49033

Table 6.4: Baseline measurement survey, response ratio.

6.2.1 Information gathered

Below are shown the results of the measurements gathered every day during the base period in what concerns to the two problems addressed. It should be mentioned that in order to process the qualitative character of the questions related to the problem *Communication with the Team Coach*, they are converted into the equivalent value in a scale of 1 to 4 (where 1 corresponds to "Strongly disagree" and 4 corresponds to "Strongly agree"). "Neither agree or disagree" responses are discarded as they do not constitute mid-range choices, but rather neutral ones. In addition, it is worth noting that not all subjects were able to approach the survey daily (consequence of the real-time environment in which this study is carried).

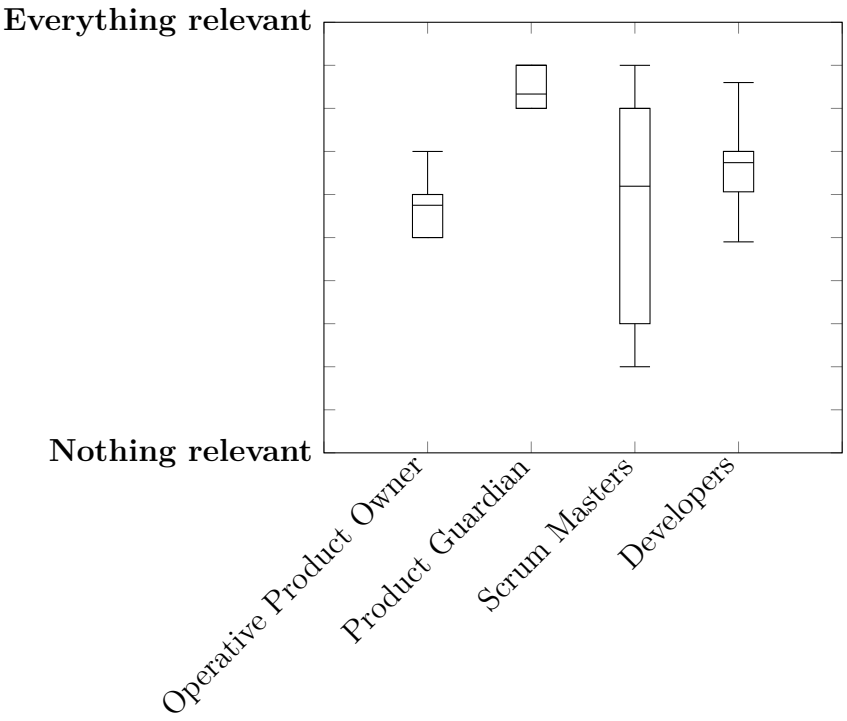


Figure 6.12: Per-role breakdown of the perceived percentage of relevant information received during the measured Sprint.

Figure 6.12 shows the data gathered for the four roles remaining in the study about the first question in the survey, which was intended to cover the problem of the amount of undesired information received.

The data related to the Operative Product Owner and the Product Guardian is little and therefore, although it seems to be stable, it would be mistaken to take its meaning as relevant on its own.

Fortunately the case with the data gathered from the Scrum Masters and Developers is totally the opposite. The Scrum Masters seemed to experience a rather unstable situation which, if studied from the more holistic point of view of the mean relevance of the total information received, shows that almost a third of it (which is a rather high amount) could be deemed as useless. As for what concerns to the Developers, the results seem to behave more in a more stable manner, which could be due to the fact that this role requires considerably larger amounts of information than any other. Nevertheless, the average percentage of irrelevant information received is still too high if we take into account that all this information is originated in controlled environment with clearly-defined boundaries.

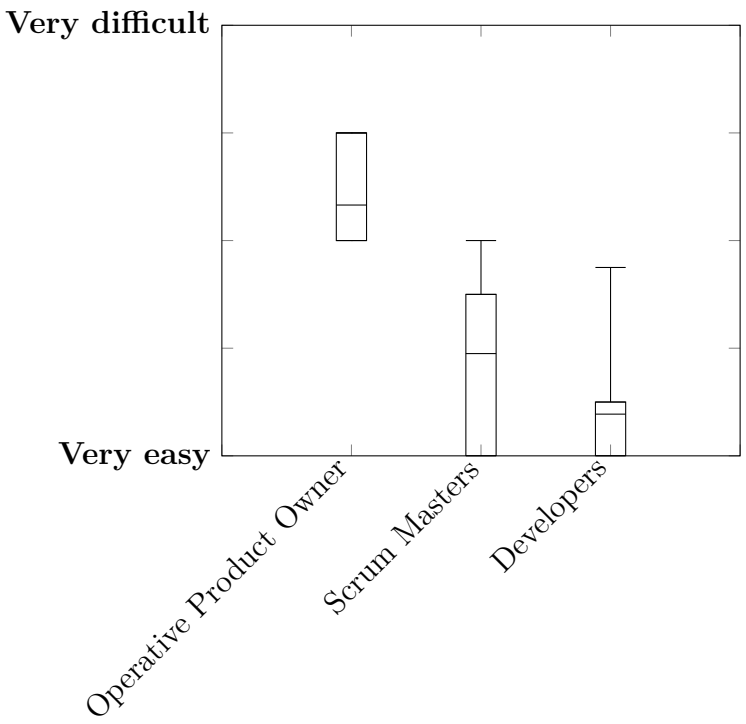


Figure 6.13: Per-role breakdown of the perceived feeling of impossibility to communicate with the Team Coach in situations in which such communication is needed during the measured Sprint.

Figure 6.13 shows the data gathered for the four roles remaining in the study (except the Product Guardian, as all its responses to this question corresponded to the neutral choice) about the second question in the survey, dedicated to research on the availability of the Team Coach.

To begin with, it is interesting to highlight how the Operative Product Owner scores rather high, which means that it missed being able to communicate with the Team Coach. Unfortunately, in the particular case only one Operative Product Owner filled the survey and very few days and, therefore, the results can only be interpreted as an indication.

As for what concerns to the Scrum Masters, they seem to feature a relatively stable behavior that shows how they generally miss being able to communicate with the Team Coach, but also includes 'spikes' showing more critical communication demands, probably related to concrete issues about the Scrum methodology that they are not able to figure out how to address on their own. And for the Developers, they apparently miss communications with the Team Coach just slightly, which is rather straightforward as they are in close contact with the Scrum Master.

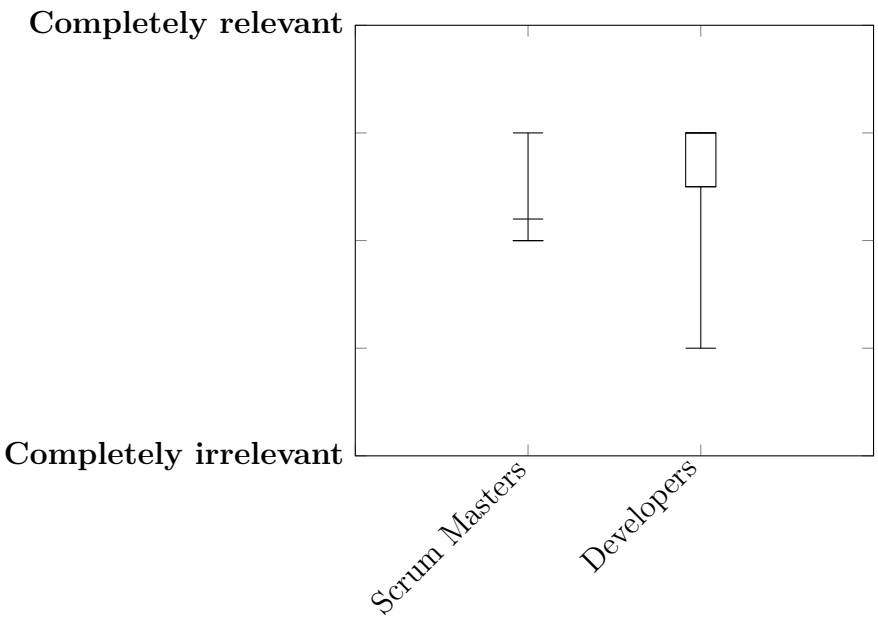


Figure 6.14: Per-role breakdown of the perceived utility of the communications with the Team Coach during the measured Sprint.

Finally, Figure 6.14 shows the data gathered for the four roles remaining in the study (expect Operative Product Owner and Product Guardian due to very small available data) about the third question in the survey, focused on studying the utility of the Team Coach.

Although it is important to mention that the amount of responses for this question is noticeably lower (due to the fact that there is a high amount of neutral ones), their score is very important. It can be observed how almost all responses state that communicating with the Team Coach is useful to some extent, which means that the knowledge that this role brings into the company is indeed appreciated, independently of whether it is or not properly placed in the organizational structure.

6.3 Treatment measurements

The survey used for these measurements was similar as that for the ones before. Therefore, its description applies also here but it is not replicated for obvious reasons.

As for the form of execution, this survey was performed in a daily basis similar to the aforementioned one, but during a period of time in which the subjects had the treatment applied to them, this is, the next Sprint after the one in which the baseline was measured.

The same points (Kelley et al., 2003) detailed for the baseline survey are addressed

Role	Subjects involved	Response ratio
Developer	8	0.5375
Scrum Master	2	0.45
Total	10	0.52

Table 6.5: Treatment measurement survey, response ratio.

below for this one:

1. The intent of this survey was to perform measurements on the treatment.

2. This survey was necessary to obtain treatment measurements to compare the treatment against the baseline so that the efficiency of the treatment and its overall success could be effectively and empirically evaluated.

3. The potential subjects for the survey were identified by the representatives of the research at Ericsson based on their predisposition to the study. The subjects were initially approached in a workshop for each team (in addition to one only for the Scrum Masters) in their regular workplace during which the researcher explained what they should change during this period (in scientific terms, the treatment). Then they took the survey daily through the Internet on their own. The response ratio is shown in Table 6.5 (the Team Coach role is not included as both of them dropped off the research during an earlier stage). It shall be noted that the remaining Operative Product Owner and Product Guardian dropped off during this stage.

6.3.1 Information gathered

Below are shown the results of the measurements gathered every day during the treatment period in what concerns to the two problems addressed. The procedure utilized is identical to the one described in Section 6.2.1.

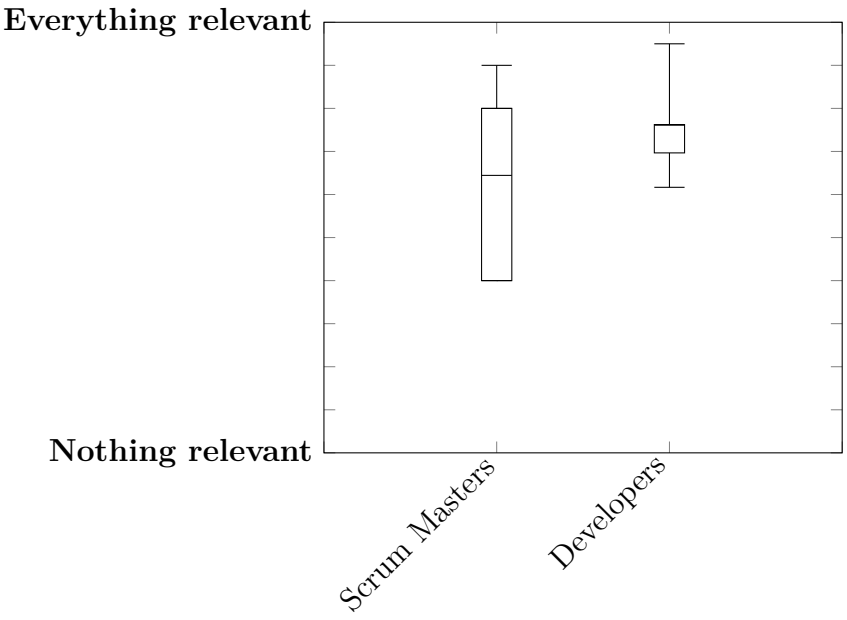


Figure 6.15: Per-role breakdown of the perceived percentage of relevant information received during the measured Sprint.

Figure 6.15 shows the data gathered for the two roles remaining in the study about the first question in the survey, which was intended to cover the problem of the amount of undesired information received.
Both Scrum Masters see slight reliability and number increases on the final results, although the differences are not large enough to withdraw any unarguably strong conclusions.

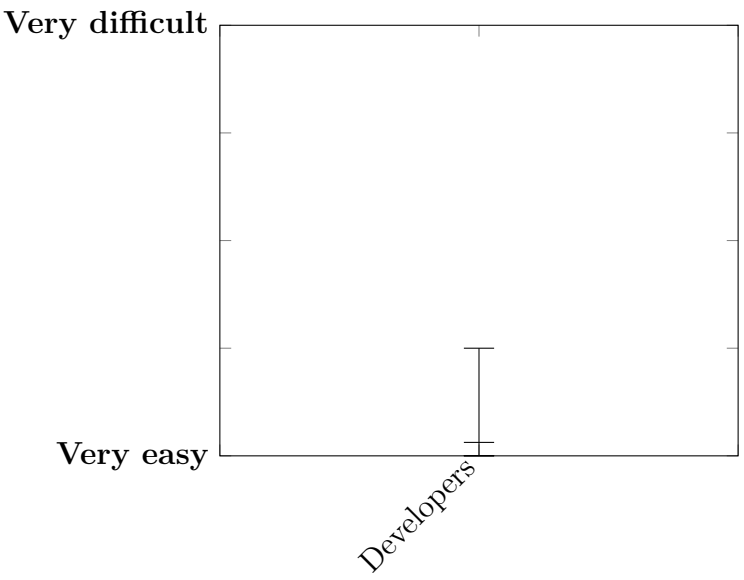


Figure 6.16: Per-role breakdown of the perceived feeling of impossibility to communicate with the Team Coach in situations in which such communication is needed during the measured Sprint.

Figure 6.16 shows data about how the Developers, the only remaining role implied with the difficulties to communicate with the Team Coach for this Sprint (as the Scrum Masters were doubling as Team Coaches for its duration), answered the second question in the measurement survey, aimed at evaluating the difficulty of communicating with the Team Coach. The results are quite good and, in fact, the only datapoints that slightly harm them correspond to the first two days of the Sprint, when the Scrum Master's agenda (which, because of the treatment, is also the Team Coach's) is filled with routine meetings.

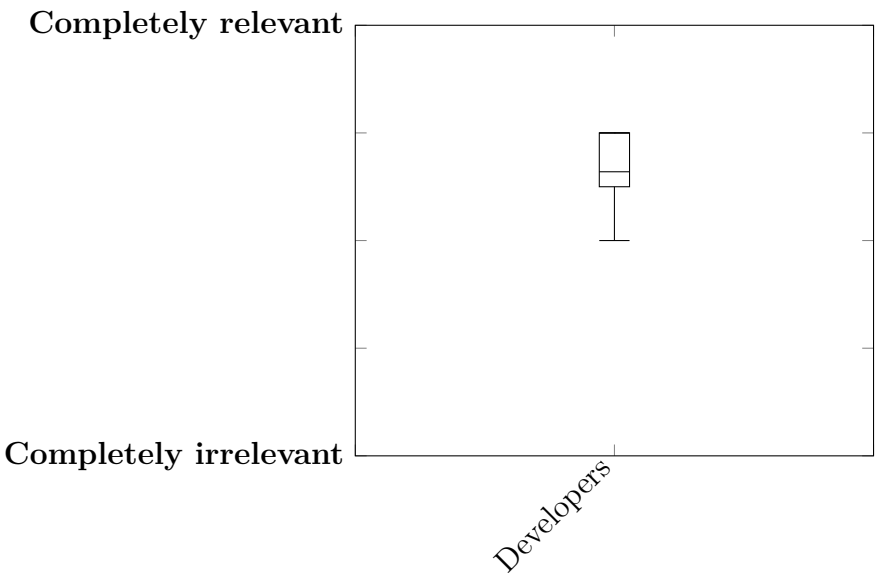


Figure 6.17: Per-role breakdown of the perceived utility of the communications with the Team Coach during the measured Sprint.

Finally, Figure 6.17 shows data about how the Developers (as once again it would be a non-sense to include the Scrum Masters), answered the last question in the measurement survey, aimed at evaluating the utility of the communications with the Team Coach. From the figure it can be seen how the treatment has increased the average utility of these communications and, furthermore, across a set of data points that is 50% larger than that of the baseline less variability is shown, which is an important factor to consider when evaluating the robustness of any conclusions withdrawn.

6.4 Satisfaction measurements

The survey used for these measurements was performed at the end of the experiment. It was performed only once and on the same subjects as the others.

Again in accordance with Kelley et al. (2003), the next concrete details are reported about the survey:

1. The intent of this survey was to measure the subjects' satisfaction with the treatment.

2. This survey was necessary to gather data that could support the decision of adopting the treatment as regular practise or discarding it. Because of the criteria at Ericsson, satisfaction was the most influencing factor on this decision and therefore this survey was of extreme importance for them.

3. The potential subjects for the survey were identified by the representatives of the research at Ericsson based on their predisposition to the study. The

Role	Subjects involved	Response ratio
Developer	8	1
Scrum Master	2	1
Total	10	1

Table 6.6: Satisfaction survey, response ratio.

subjects were initially approached on a per-team workshop during which they took the survey. The response ratio is reported on Table 6.6.

4. Detailed information on the methods used for analysis of the data collected with this survey can be read on Section 4.5.3.

5. The concrete procedure for data analysis can be found in Section 4.5.3.

6.4.1 Information gathered

Tables 6.7 and 6.8 show the normalized results of the satisfaction survey. As the example form in Appendix D shows, all questions rate on a standard 5-point Likert scale.
While the satisfaction level of the part of the treatment that relates to the relevance of the communication handled is acceptably good, the second part of it cannot be deemed as so satisfactory, specially as far as it concerns to the Developers, which is a curious result having in mind that the effort that they had to deploy for this treatment, as opposed to that put by the Scrum Masters, was void.

Question	Normalized score (Scrum Masters)	Normalized score (Developers)
1	3	3.84
2	4	4
3	5	4
4	4	4.3
Overall	4	4.035

Table 6.7: Results about the *Communication relevance* section of the satisfaction survey.

Question	Normalized score (Scrum Masters)	Normalized score (Developers)
1	4	3.5
2	4.5	3.5
3	4	2.66
4	3.5	3.34
5	4	2.83
6	4	2.33
Overall	4	2.86

Table 6.8: Results about the *Communication with the Team Coach* section of the satisfaction survey.

7

Discussion and conclusion

7.1 RQ1. Result of the improvements

The proposed approach to address the first problem, *Communication relevance*, consisted of setting some automatic filters with the purpose of reducing the overhead of information transferred through e-mail in the organization. The comparison presented in Figure 7.1 shows how the mean agreement percentage over time (obtained as the mean of the mean agreement percentages every day) is slightly better for both Developers and Scrum Masters. While this can definitively be considered good, the differences are small and so was the sample size which, along the low response ratio characteristic of online surveys, converts this result into a mere indication requiring repetitions and confirmatory studies to acquire relevance and become reliable.

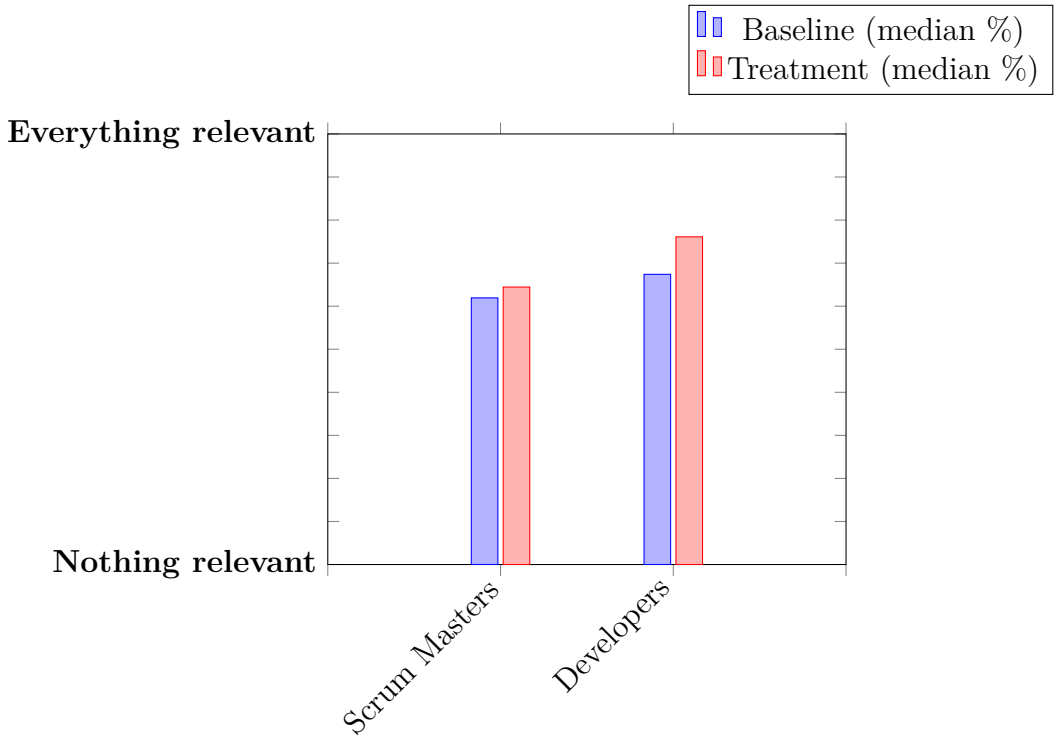


Figure 7.1: Comparison of the percentage of relevant information received between the baseline and the treatment methodologies.

The corresponding satisfaction results (Table 6.7) are overall good for both roles. This *could* be understood as an indication that the subjects considered the filters good *because* their interest in the content that they expected to be filtered was low.

As for the second one, *Communication with the Team Coach*, a modification of the organizational structure was suggested in which the Team Coach role was merged into the Scrum Masters, deferring some load from the busy schedules of Line Managers (who, in the current setup, also play the Team Coach role) while at the same putting the role in a figure that is closer to the team. The comparisons (Figures 7.2 and 7.3) show, opposite to previously, acceptably larger differences. In addition, and although it should still be kept in mind that the sample was small and the period of time measured rather short, there is a pattern that strengthens the idea that this approach was right - as Figure 6.16 shows, the experimental approach eliminated all hurdles in (what refers to communication attempts with the Team Coach) except for the very beginning of this Sprint, which makes sense as the Scrum Masters' availability during this time was probably lower than usual.

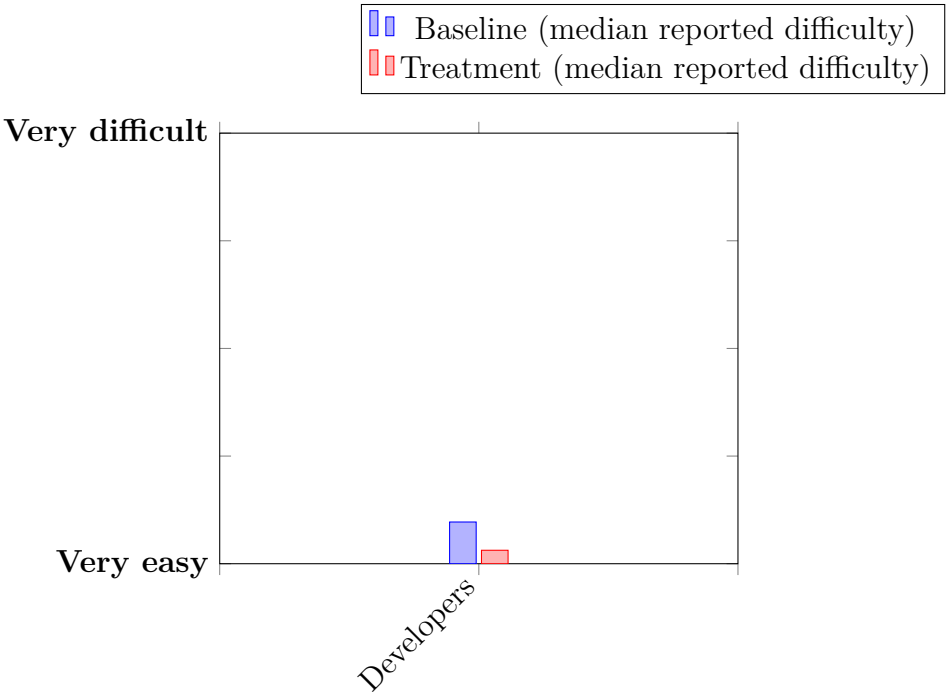


Figure 7.2: Comparison of the difficulty to communicate with the Team Coach between the baseline and the treatment methodologies.

Finally, and although the satisfaction results for this approach (Table 6.8) are quite good for the Scrum Masters, the same cannot be said from the Developers, which is important to take into account as they constitute an arguably more important set of individuals in any organization (due to size, basically). It is hard to understand this contradiction with the results of the treatment, but it should be taken into account that there is no "baseline satisfaction" data to compare against so it is possible that

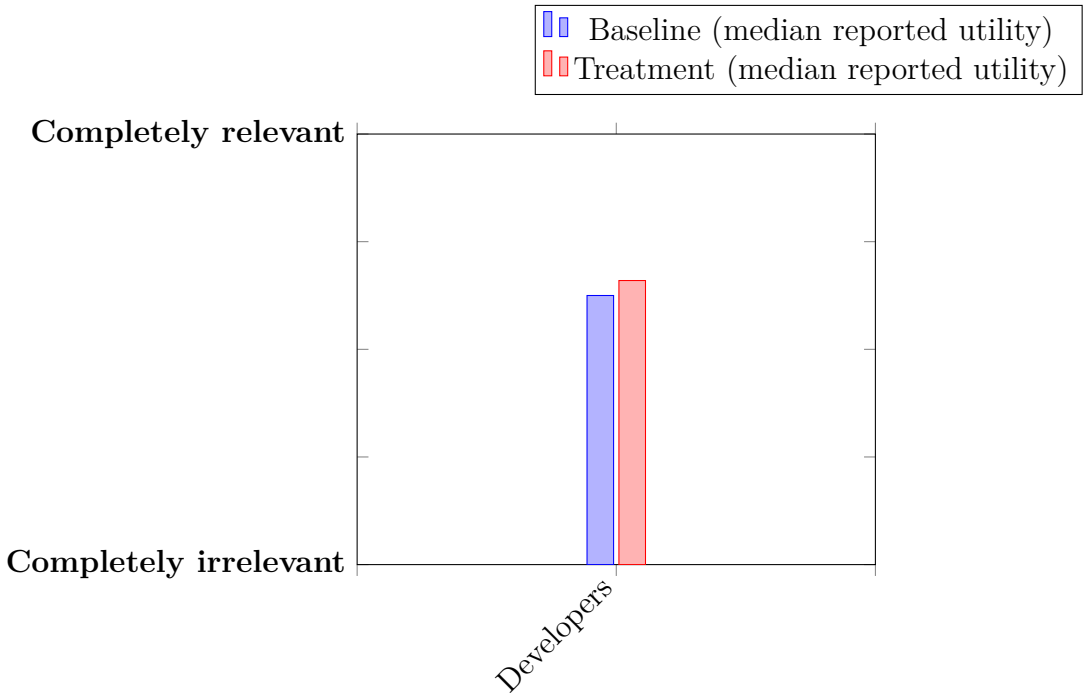


Figure 7.3: Comparison of the utility of the communications with the Team Coach between the baseline and the treatment methodologies.

it is still possible that it is successful satisfaction-wise.

7.2 RQ2. Lessons learnt

During the execution of the study a few observations were made about the particularities of experimental research in large industrial software development environments. The reliability of all of these is of course not high in different setups but, because they are topic-independent, they should be applicable to future experimental research in this same type of contexts, even if it focuses in studying completely different variables.

- **Setups are impossible to control as much as it would be desirable in a true experiment, which is the reason why most large-scale industrial experimentation studies are quasi-experiments**. Opposite to what happens with case studies, where it is particularly interesting to be able to perform raw observations on the targeted environment, for experiments it is preferred to be able to narrow the scope to an easily controllable size and to account in a very detailed way for, if not remove completely, any other external factors that could affect the execution. Unfortunately, in large scale environments this does not seem to be feasibly possible: for example, in the concrete case, and since communication is a bidirectional action, it would be desirable to isolate the experiment subjects from interacting with information that is not distributed as according to the treatment guidelines during the treatment application (at least), but this is a condition understandably impossible to impose as it would unavoidably isolate them from subjects they need to communicate with, just because such subjects are not taking part in the experiment.

  A quick approach to try to address this problem would be increasing the scope of the experiment so that all subjects in the targeted environment (this is, the concrete office where the experiment is executed at) happen to participate. This way, the factors that could influence the experiment without being explicitly addressed are limited to influence from other places, like communications sent from other offices, in the concrete case, which seems to constitute an insignificantly small percentage of the entirety of factors, at least in communication-centric experiments. However, the viability of this is strongly limited because it both (1) carries a huge increase in the amount of resources needed to perform the experiment and (2) creates a high-risk-high-reward situation for the company which could turn either extremely well and extremely bad, a situation that currently Ericsson does not need to risk.

- **Time is a matter of very high concern**. Independently of the experimental subjects' interest on contributing to the study, they have their occupations and responsibilities, which are specially numerous in setups this big. Since they acknowledge this and having into account that side-performed studies are not high priority when compared to working in something the company can get more straightforward and tangible profit, it is hard to take on their time, not because they do not want to collaborate, but because they lack the time to do so. Therefore availability of the subjects is limited to very short time lapses

and activities like asking for in-advance preparation for interviews or alike are barely feasible.

- **Subjects who volunteer for experiments are surprisingly interested not only in how the results affect them, but also in learning why they show as they do**. During this study in particular, several subjects repeatedly expressed their interest in seeing what it would reveal, in addition to what improvements it could bring to their regular work habits. Interestingly enough this could be a hint indicating the subjects' desire for developing their self-improvement skills, to possibly become able to enhance their own practises in innovative ways and also more decoupled from external agents.

- **Undocumented social networking can be a very influent factor on the subjects' perception of the environment.** Specially in stages in time in which deep, complex and slow organisational changes are taking place, it is not unusual to meet situations in which information needs to be distributed to people who are not available, doubts about certain fields need to be solved but it is not clear who should solve them, and the like. With this in mind, having spent relevant amounts of time in the organisation seems to be very helpful due to the fact that it provides knowledge on who, exactly, should be contacted for every situation and, more importantly, how this person can actually be reached in practice. Because this can have a strong impact on studies in this type of environments, this should be deemed as a threat to validity important to account for.

- **Large-scale contexts have a distorting effect on meta-information, which creates a necessity for performing confirmatory research on the target environment before commencing with the experiment**. Involvement of lots of diverse roles with incompatible schedules, different tasks in an organisational structure that is hard to keep perfectly stable during long periods of time requires special dedication to stay up to date with the details of such structure and with to what extent it is actually implemented. Unfortunately, the attention required from every individual to address this issue is far too resource-demanding, and it does not seem to be cost-efficient to dedicate a role to it (and there is no documentation about attempts being made).

  As a consequence, it is not unfeasible to find situations in which the organisational structure and/or guidelines that should be in place are not really being utilised due to nimble reasons like small communication misunderstandings between the roles driving organisational changes and the roles adopting them.

- **Borders between variables are fuzzy**. Because, as mentioned before, these environments feature a high level of complexity due to their size, tremendously large amounts of interactions are continuously happening between the individuals that are in the experiment and the ones that are not. Obviously it is unfeasible to try to isolate a subset of these interactions (which would formally be known as "variable") and modify it as desired. A possible approach to work around this issue is boundary-located simulation, this is, setting up

simulated actions on the boundaries between the experimental subjects and the individuals that are out of the context, so that the latter do not have to change how they interact with the former, while these receive altered or mocked-up information. However it should be taken into account that, in addition to being a slight risk to reliability of the study (as experiments are meant to focus on study the real world), it also poses a high risk for the company because interactions between subjects are manipulated and only one of the participant sides knows about it (the experiment subjects), while the other one is not necessarily aware of this (in practice, it would be impossible to inform all of the subjects out of the experiment of this situation).

7.3 Conclusion

In this study a set of initial interviews revealed two major problems with communication in the target environment. Then, alternative approaches were designed and experimentally evaluated for both of them through a comparison between measurements performed on the base situation and on the treatment. In addition, the satisfaction of the study subjects with the alternative approaches was assessed.

The treatment addressed two concerns: the distribution of information without practical value, through the simulation of an appropriate criteria for addressee selection, and the lack of communication with coaching staff, through the reallocation of coaching duties into the Scrum Master role. Both parts showed mostly positive results and, although there are some caveats that are possibly worth being further investigated in the future, the results suggest that the treatment is probably interesting to, as minimum, try out.

Finally, interesting findings were observed during the execution of the study, such as subject interest for the experiment and a possible candidate for threat to validity attached to large-scale contexts.

7.4 Implications for Academia

The information gathered about the first problem (*Communication relevance*) and the proposed way to address it are not considered to carry a strong meaning for Academia mostly because, as it was discussed in Section 7.1, the data gathered shows only small differences and falls short in strength to make any kind of statement convincingly. Fortunately, this part of the study would be very easy to replicate and the results of these replications would surely help to withdraw more reliable conclusions.

However, the data gathered about the second problem (*Communication with the Team Coach*), the proposed way to address it and the results of the subjects' satisfaction with it pose a very interesting question: if, as the comparison of the measurements performed on both the traditional and treatment setups clearly shows, the Developers find much easier to talk with the Team Coach when this is their own

Scrum Master and, furthermore, their discussions are more useful, what keeps them from feeling more satisfied with it? Nevertheless, note that this does not strictly mean that they are less satisfied with the proposed structure than with the traditional one as, as mentioned in Section 7.1, no satisfaction assessment was performed for it and therefore there is no base data to compare against. Because of this, it is believed that it would be very interesting to replicate this part of the study with a larger set of subjects, assessing the satisfaction with the traditional structure and, should the results be similar, perform some exploratory research to find the reason(s) behind.

7.5 Implications for Practitioners

Because of the result-oriented character of this sector, it is believed that two conclusions are of main interest to Practitioners:

- **Massive communications are useless in practice**. Because they are not topic-driven, they are irrelevant to most of their targets, too frequent, and often just ignored, plus they contribute to create clutter that obstructs people's access to the information that is actually important.
- **The Team Coach needs to be close to the team just as the Scrum Master does**. The results obtained in regards to this are straightforward. And they actually makes sense, as it is understable that having the Team Coach closer to the core of the team eases communication, availability and trouble resolution.

Bibliography

Abdel-Hamid, A. N., & Abdel-Kader, M. A. (2011). Process increments: An agile approach to software process improvement. In *2011 agile conference, AGILE 2011, salt lake city, utah, usa, august 7-13, 2011* (pp. 195–200). Retrieved from http://dx.doi.org/10.1109/AGILE.2011.26 doi: 10.1109/AGILE.2011.26

Arthur, T. (2013). Agile adoption: Measuring its worth. SAS Institute. Retrieved from http://support.sas.com/rnd/papers/2013/AgileAdoptionPaper.pdf

Auvinen, J., Back, R., Heidenberg, J., Hirkman, P., & Milovanov, L. (2006). Software process improvement with agile practices in a large telecom company. In J. Münch & M. Vierimaa (Eds.), *Product-focused software process improvement* (Vol. 4034, p. 79-93). Springer Berlin Heidelberg. Retrieved from http://dx.doi.org/10.1007/11767718_10 doi: 10.1007/11767718_10

Averianova, A., & Deekens, T. (2014). *Information and communication and their impact on productivity determinants of xfts in a large-scale agile environment: a case study.* (Unpublished)

Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., ... Thomas, D. (2001). *Manifesto for agile software development.* Retrieved from http://www.agilemanifesto.org/

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. Retrieved from http://www.tandfonline.com/doi/abs/10.1191/1478088706qp063oa doi: 10.1191/1478088706qp063oa

Brown, N., & Stockman, T. (2013). Examining the use of thematic analysis as a tool for informing design of new family communication technologies. In *Proceedings of the 27th international bcs human computer interaction conference* (pp. 21:1–21:6). Swinton, UK, UK: British Computer Society. Retrieved from http://dl.acm.org/citation.cfm?id=2578048.2578078

Buchalcevova, A. (2011). Software process improvement in small companies as a path to enterprise architecture. In *Information systems development, reflections, challenges and new directions [proceedings of ISD 2011, heriot-watt university, edinburgh, scotland, uk, august 24 - 26, 2011]* (pp. 243–253). Retrieved from http://dx.doi.org/10.1007/978-1-4614-4951-5_20 doi: 10.1007/978-1-4614-4951-5_20

Cataldo, M., Herbsleb, J. D., & Carley, K. M. (2008). Socio-technical congruence: a framework for assessing the impact of technical and work dependencies on software development productivity. In *Proceedings of the sec-*

ond international symposium on empirical software engineering and measurement, ESEM 2008, october 9-10, 2008, kaiserslautern, germany (pp. 2–11). Retrieved from http://doi.acm.org/10.1145/1414004.1414008 doi: 10.1145/1414004.1414008

Cataldo, M., Wagstrom, P., Herbsleb, J. D., & Carley, K. M. (2006). Identification of coordination requirements: implications for the design of collaboration and awareness tools. In *Proceedings of the 2006 ACM conference on computer supported cooperative work, CSCW 2006, banff, alberta, canada, november 4-8, 2006* (pp. 353–362). Retrieved from http://doi.acm.org/10.1145/1180875.1180929 doi: 10.1145/1180875.1180929

Cohen, D., Lindvall, M., & Costa, P. (2004). An introduction to agile methods. In *Advances in computers* (Vol. 62, p. 1 - 66). Elsevier. Retrieved from http://www.sciencedirect.com/science/article/pii/S0065245803620012 doi: 10.1016/S0065-2458(03)62001-2

Creswell, J. W. (2008). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Sage Publications Ltd.

Dorairaj, S., & Noble, J. (2013, Aug). Agile software development with distributed teams: Agility, distribution and trust. In *Agile conference (agile), 2013* (p. 1-10). doi: 10.1109/AGILE.2013.7

dos Santos, A., de Farias Junior, I., de Moura, H., & Marczak, S. (2012, Aug). A systematic tertiary study of communication in distributed software development projects. In *Global software engineering (icgse), 2012 ieee seventh international conference on* (p. 182-182). doi: 10.1109/ICGSE.2012.42

Espinosa-Curiel, I. E., Rodríguez-Jacobo, J., & Fernández-Zepeda, J. A. (2013). A framework for evaluation and control of the factors that influence the software process improvement in small organizations. *Journal of Software: Evolution and Process*, *25*(4), 393–406. Retrieved from http://dx.doi.org/10.1002/smr.569 doi: 10.1002/smr.569

French, A., & Layzell, P. (1998, Nov). A study of communication and cooperation in distributed software project teams. In *Software maintenance, 1998. proceedings., international conference on* (p. 146-154). doi: 10.1109/ICSM.1998.738503

Garcia, A., Greenwood, P., Heineman, G., Walker, R. J., Cai, Y., Yang, H. Y., ... Zhao, J. (2007). Assessment of contemporary modularization techniques - acom'07: workshop report. *ACM SIGSOFT Software Engineering Notes*, *32*(5), 31–37. Retrieved from http://doi.acm.org/10.1145/1290993.1291005 doi: 10.1145/1290993.1291005

Guzzi, A., Bacchelli, A., Lanza, M., Pinzger, M., & Deursen, A. v. (2013). Communication in open source software development mailing lists. In *Proceedings of the 10th working conference on mining software repositories* (pp. 277–286). Piscataway, NJ, USA: IEEE Press. Retrieved from http://dl.acm.org/citation.cfm?id=2487085.2487139

Hanakawa, N., & Okura, K. (2004, Nov). A project management support tool using communication for agile software development. In *Software engineering conference, 2004. 11th asia-pacific* (p. 316-323). doi: 10.1109/APSEC.2004.8

Herbsleb, J., & Mockus, A. (2003, June). An empirical study of speed and commu-

nication in globally distributed software development. *Software Engineering, IEEE Transactions on*, *29*(6), 481-494. doi: 10.1109/TSE.2003.1205177

Hesser, H., & Tomasini, A. (2012). Experiences of the ericsson mobile core agile transformation - tales of agile change.. Retrieved from http://www.agilealliance.org/files/session_pdfs/Ericsson_Success_Story_agile2012.pdf

Highsmith, J. (2000). E-business application delivery. In C. Consortium (Ed.), (p. 4 - 17).

Hove, S., & Anda, B. (2005, Sept). Experiences from conducting semi-structured interviews in empirical software engineering research. In *Software metrics, 2005. 11th ieee international symposium* (p. 10 pp.-23). doi: 10.1109/METRICS.2005.24

Jaanu, T., Paasivaara, M., & Lassenius, C. (2012, Sept). Effects of four distances on communication processes in global software projects. In *Empirical software engineering and measurement (esem), 2012 acm-ieee international symposium on* (p. 231-234). doi: 10.1145/2372251.2372293

Juristo, N., & Moreno, A. M. (2010). *Basics of software engineering experimentation* (1st ed.). Springer Publishing Company, Incorporated.

Kelley, K., Clark, B., Brown, V., & Sitzia, J. (2003). Good practice in the conduct and reporting of survey research. *International Journal for Quality in Health Care*, *15*(3), 261–266. doi: 10.1093/intqhc/mzg031

Kettunen, P., & Laanti, M. (2008). Combining agile software projects and large-scale organizational agility. *Software Process: Improvement and Practice*, *13*(2), 183–193. Retrieved from http://dx.doi.org/10.1002/spip.354 doi: 10.1002/spip.354

Korkala, M., & Maurer, F. (2014). Waste identification as the means for improving communication in globally distributed agile software development. *Journal of Systems and Software*, *95*(0), 122 - 140. Retrieved from http://www.sciencedirect.com/science/article/pii/S0164121214001009 doi: http://dx.doi.org/10.1016/j.jss.2014.03.080

Kraut, R. E., & Streeter, L. A. (1995, March). Coordination in software development. *Commun. ACM*, *38*(3), 69–81. Retrieved from http://doi.acm.org/10.1145/203330.203345 doi: 10.1145/203330.203345

Kuhrmann, M. (2015). Crafting a software process improvement approach—a retrospective systematization. *Journal of Software: Evolution and Process*, *27*(2), 114–145. Retrieved from http://dx.doi.org/10.1002/smr.1703 doi: 10.1002/smr.1703

Lipinski-Harten, M., & Tafarodi, R. W. (2013). Attitude moderation: A comparison of online chat and face-to-face conversation. *Computers in Human Behavior*, *29*(6), 2490 - 2493. Retrieved from http://www.sciencedirect.com/science/article/pii/S0747563213001969 doi: http://dx.doi.org/10.1016/j.chb.2013.06.004

MacKellar, B. (2012, April). A case study of group communication patterns in a large project software engineering course. In *Software engineering education and training (csee t), 2012 ieee 25th conference on* (p. 134-138). doi: 10.1109/CSEET.2012.21

Martini, A., Pareto, L., & Bosch, J. (2013). Communication factors for speed and reuse in large-scale agile software development. In *Proceedings of the 17th international software product line conference* (pp. 42–51). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/2491627.2491642 doi: 10.1145/2491627.2491642

Moore, E., & Spens, J. (2008, Aug). Scaling agile: Finding your agile tribe. In *Agile, 2008. agile '08. conference* (p. 121-124). doi: 10.1109/Agile.2008.43

Moreira, M. E. (2013). Three case studies in adopting agile being agile. In *Being agile: Your roadmap to successful adoption of agile* (1st ed., chap. 24). Berkely, CA, USA: Apress.

Ñaupac, V., Arisaca, R., & Dávila, A. (2012). Software process improvement and certification of a small company using the NTP 291 100 (moprosoft). In *Product-focused software process improvement - 13th international conference, PROFES 2012, madrid, spain, june 13-15, 2012 proceedings* (pp. 32–43). Retrieved from http://dx.doi.org/10.1007/978-3-642-31063-8_4 doi: 10.1007/978-3-642-31063-8_4

Paasivaara, M., Väättänen, O., Hallikainen, M., & Lassenius, C. (2014). Supporting a large-scale lean and agile transformation by defining common values. In T. Dingsøyr, N. Moe, R. Tonelli, S. Counsell, C. Gencel, & K. Petersen (Eds.), *Agile methods. large-scale development, refactoring, testing, and estimation* (Vol. 199, p. 73-82). Springer International Publishing. Retrieved from http://dx.doi.org/10.1007/978-3-319-14358-3_7 doi: 10.1007/978-3-319-14358-3_7

Pikkarainen, M., Haikara, J., Salo, O., Abrahamsson, P., & Still, J. (2008). The impact of agile practices on communication in software development. *Empirical Software Engineering*, *13*(3), 303-337. Retrieved from http://dx.doi.org/10.1007/s10664-008-9065-9 doi: 10.1007/s10664-008-9065-9

Power, K. (2014). A model for understanding when scaling agile is appropriate in large organizations. In T. Dingsøyr, N. Moe, R. Tonelli, S. Counsell, C. Gencel, & K. Petersen (Eds.), *Agile methods. large-scale development, refactoring, testing, and estimation* (Vol. 199, p. 83-92). Springer International Publishing. Retrieved from http://dx.doi.org/10.1007/978-3-319-14358-3_8 doi: 10.1007/978-3-319-14358-3_8

Qian, W., & Zhen-hua, S. (2010, July). Research on multi-perspective communication management of software development project based on theory of project management. In *Signal processing systems (icsps), 2010 2nd international conference on* (Vol. 3, p. V3-192-V3-195). doi: 10.1109/ICSPS.2010.5555844

Rohunen, A., Rodriguez, P., Kuvaja, P., Krzanik, L., & Markkula, J. (2010). Approaches to agile adoption in large settings: A comparison of the results from a literature analysis and an industrial inventory. In M. Ali Babar, M. Vierimaa, & M. Oivo (Eds.), *Product-focused software process improvement* (Vol. 6156, p. 77-91). Springer Berlin Heidelberg. Retrieved from http://dx.doi.org/10.1007/978-3-642-13792-1_8 doi: 10.1007/978-3-642-13792-1_8

Salo, O., & Abrahamsson, P. (2007). An iterative improvement process for agile software development. *Software Process: Improvement and Practice*, *12*(1), 81–100. Retrieved from http://dx.doi.org/10.1002/spip.305 doi: 10

.1002/spip.305

Sekitoleko, N., Evbota, F., Knauss, E., Sandberg, A., Chaudron, M., & Olsson, H. H. (2014). Technical dependency challenges in large-scale agile software development. In *Agile processes in software engineering and extreme programming - 15th international conference, XP 2014, rome, italy, may 26-30, 2014. proceedings* (pp. 46–61). Retrieved from http://dx.doi.org/10.1007/978-3-319-06862-6_4 doi: 10.1007/978-3-319-06862-6_4

Soundararajan, S., & Arthur, J. (2009, April). A soft-structured agile framework for larger scale systems development. In *Engineering of computer based systems, 2009. ecbs 2009. 16th annual ieee international conference and workshop on the* (p. 187-195). doi: 10.1109/ECBS.2009.21

Søvik, H., & Forfang, M. (2010). Tech challenges in a large-scale agile project. In A. Sillitti, A. Martin, X. Wang, & E. Whitworth (Eds.), *Agile processes in software engineering and extreme programming* (Vol. 48, p. 353-361). Springer Berlin Heidelberg. Retrieved from http://dx.doi.org/10.1007/978-3-642-13054-0_38 doi: 10.1007/978-3-642-13054-0_38

Thompson, B. (2008). *Foundations of behavioral statistics: An insight-based approach*. Guilford Press. Retrieved from http://books.google.se/books?id=yLbNKLCSIwQC

Tuffley, A., Grove, B., & McNair, G. (2004). Spice for small organisations. *Software Process: Improvement and Practice*, *9*(1), 23–31. Retrieved from http://dx.doi.org/10.1002/spip.191 doi: 10.1002/spip.191

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., & Regnell, B. (2012). *Experimentation in software engineering*. Springer. Retrieved from http://dx.doi.org/10.1007/978-3-642-29044-2 doi: 10.1007/978-3-642-29044-2

Xuan, Q., Gharehyazie, M., Devanbu, P. T., & Filkov, V. (2012). Measuring the effect of social communications on individual working rhythms: A case study of open source software. In *Proceedings of the 2012 international conference on social informatics* (pp. 78–85). Washington, DC, USA: IEEE Computer Society. Retrieved from http://dx.doi.org/10.1109/SocialInformatics.2012.17 doi: 10.1109/SocialInformatics.2012.17

A

Initial interview guide

[THIS IS READ TO EVERY INTERVIEWEE UPON INTERVIEW START.]

The intent of this interview is to gain some insight on:

- how the information that flows around the within the team' closest context is currently distributed,
- what is the reasoning behind such distribution procedure and
- how you, the people who use such information, believe that the procedure for its distribution could be improved.

Note that this interview focuses only on information on the domain of the PDU CAT, and only on the point of view of the role currently interviewed (therefore your current role).

The data gathered through this interview, concretely the notes that you will see me taking, will be available only to me and associated to your current role in your team and the current task of your team, but never to any other information, like your name. If you do not agree this, you can opt out of the interview at any time. Should this happen, the notes gathered would be discarded and this interview would account towards mortality rate.

[IF THE INTERVIEWEE DECIDES TO OPT OUT, FINISH.]

Questions

1. Very briefly, what is Team X set to do during the current Sprint?
2. How would you, as for what concerns your role in the current Sprint, improve communication to and among teams, Operative Product Owners, Scrum Masters, Product Guardians and Team Coaches?
    - For example, how would you alter this communication? Would you approach the fact that quite often you receive information that is irrelevant to you? Maybe even replicated? Or what?
3. As of your role in the current Sprint, when you want to distribute information, how do you choose how to do it and, most importantly, who to send it to?
4. Between product-related, project-level, line-related (management and administration issues) and other sources of information, which would you say is/are the one(s) disturbing you the most with information that lacks a practical value for you?
    - Could you elaborate a little bit on why that one rather than the others?
5. Enumerate as many ways of distributing information to Product Guardians, Operative Product Owners, Scrum Masters, Team Coaches and other teams as you can think of that it is reasonable to expect that you are using now or will use in the short future to either send or receive information within these roles.
6. Is there anything else you would like to mention?
7. What is your role?
    - Answer should be one of *Developer*, *Scrum Master*, *Operative Product Owner*, *Product Guardian* and *Team Coach*.

Farewell

[THIS IS READ TO EVERY INTERVIEWEE UPON QUESTIONS FINISHED.]

Thanks for participating. Please do not talk about the contents of the interview with your colleagues unless they are done with it as well.

The information provided will help improving how information is distributed at your workplace.

[FINISH.]

B

Measurement survey

The intent of this survey is to perform empirical measurements on the way information is being distributed within teams and their closest context. By comparing the measurements taken on the traditional and on the experimental information distribution procedures, it will be possible to elaborate a statement on the effectiveness of the experimental improvements suggested, which could drive to evaluate their adoption.

Note that this survey focuses on just the closest context to the teams, this meaning the Developers, the Team Coaches, the Scrum Masters, the Product Guardians and the Operative Product Owners. Note that this survey focuses only on the point of view of the role currently being surveyed (therefore your current role) within this context. If you happen to have the Developer role plus another one (like Scrum Master, for example), please fill this survey as if you were just **the least common role** of these (just Scrum Master in this example). If you happen to be assigned to several teams, address this survey only with either Team A or Team B in mind (the one you count among yours obviously) as much as possible.

Also please remember to **respond these questions attending only to your needs and not to what you understand and/or accept**; for example, if somebody is too busy to talk to you and you believe you need this communication, please respond attending to this need, even if you perfectly understand that the other person's role implies being rarely available to you.

The data gathered through this survey will be available only to the researcher and associated to your current role in your team, but never to any other private information, like your name nor your team's. If you do not agree this, you can choose to not to submit the survey at all, none of the days. Should this happen, you as an individual would be dropped off the rest of the study and account as a semi-anonymous unit (as your role would be kept still) towards the mortality rate of the study (a measurement of the subjects who leave the study throughout its execution).

Questions

1. What is your measurement survey id?
2. Please estimate the percentage of information among that which you got today that had an actual practical value for you.
   _____ % was relevant.
3. Please indicate to what extent you agree/disagree with the following statements:
   - "I missed being able to communicate with my Team Coach today".

     Strongly disagree. . . . . ☐
     Somewhat disagree. . . . . ☐
     Neither agree or disagree. . . . . ☐
     Somewhat agree. . . . . ☐
     Strongly agree . . . . . ☐
   - "My communications with the Team Coach today were useful".

     Strongly disagree. . . . . ☐
     Somewhat disagree. . . . . ☐
     Neither agree or disagree (for example you did not talk to your Team Coach today). . . . . ☐
     Somewhat agree. . . . . ☐
     Strongly agree . . . . . ☐

Farewell

Thanks for participating.

The information provided will help improving how information is distributed at your workplace.

C

Scrum Master empowering workshop

ERICSSON

TREATMENT GUIDELINES: SCRUM MASTERS

Enhancement of communications in the context of Scrum teams

GOAL

› Communication with coaching staff is deficient because coaching staff is non-existent in practice.

› Therefore coaching information must be provided from somewhere else.

HOW TO ACHIEVE IT? (I)

› Because Team Coaches are non-existent in practice, SMs become the figures most experienced in Scrum and, therefore, they should be the ones providing coaching.

› This means that you will dedicate as much as possible of your time not only to your obligations as Scrum Master but also to researching the details of Scrum to become an expert worth learning from.

HOW TO ACHIEVE IT? (II)

› Therefore, contributing to pulled Backlog items / **pulling them yourself must be the lowest priority** task for you to do, no matter what.

› If you pull something from the Backlog and during a Daily Scrum somebody in your team believes that you are putting aside your SM/TC role for it they are empowered to force you to put it back to the backlog and go back to your SM/TC tasks.

SCRUM SUBTLETIES (I)

Sprint Planning:

› Features a really high ROI, so do not fear making it as long as necessary to address everything you need to.

› Book a room and have it seated.

› A good length is that which is acceptable with at most two breaks. If you would need more, probably you need to rethink what you want to address.

SCRUM SUBTLETIES (II)

Daily Scrum:

› The reason why it is said to be had while standing is to keep it short. If it is not working (the duration is going over 15 min. too often), try other alternatives.

› The "Done/Problems/ToDo" strategy is **a guideline, not a rule**. As a Scrum Master you should ensure that everybody talks, not give everybody turns to respond the three questions. Just be the lubricating oil in the wheel that is your team: the perfect Daily Scrum does not need a Scrum Master.

SCRUM SUBTLETIES (III)

Importance of estimations (I):

› Scrum is a formal, rather inflexible iterative methodology.
  – Thus it assumes that you are able to quantify your tasks so that you can commit to quantities, evaluate if you were able to accomplish such commitments and improve by acting accordingly in the future.

› So, **if you do not have something like velocity or points, do not go Scrum**.
  – The data you will gather during those "Sprints" will be useless.
  – The only thing you can find in commitment to non-estimated items is trouble.

SCRUM SUBTLETIES (IV)

Importance of estimations (II):

› Summarizing, do not use a knife in a soup. They are a tool-problem pair, but do not belong to each other.

› In situations in which estimates cannot be given consider a temporary switch to more flexible methodologies: you can always go back to Scrum and discard the data, which is the same thing that would have happened if you have used "point-less Scrum", with the difference that you will have worked with procedures that are better-suited for these types of situations.

SCRUM SUBTLETIES (V)

Sprint Retrospectives:

› The Sprint Retrospective is an informal meeting that should feel as a **reward to the team** for the completed Sprint. This reward is a privilege that allows them to retrospectively learn from experiences.

› Unless you face the extreme case of the meeting going over two hours, you do not want to limit its duration. Therefore, do not have it standing. The team must feel comfortable to spend whichever reasonable amount of time they want discussing how they did things.

PLUS WHAT WE'LL LEARN

› I will be around either of your team places researching on Scrum to keep you up-to-date so that you can carry on allowing your teams to improve continuously as a real Team Coach would do.

› Whatever problem you find, topic you want me to investigate, or issue you want to discuss, you are very welcome to reach me!

ERICSSON

D

Satisfaction survey

The intent of this survey is to perform an overall evaluation of the satisfaction of the suggestions experimented experimented.

The data gathered through this survey will be available only to the researcher and associated to your current role in your team and the current task of your team, but never to any other information, like your name. If you do not agree with this, you can choose to not to submit the survey. Should this happen this survey would still be associated to your current role and the task of your team, but it would account towards mortality rate.

This survey is based on Bargas-Avila et al.' *s Intranet satisfaction questionnaire: Development and validation of a questionnaire to measure user satisfaction with the Intranet* (a model for Intranet satisfaction questionnaires), since the information flows happening at the organisation are similar to those happening in an Intranet, in the sense of being narrowed down to a concrete and controlled environment.

Questions

Answer these questions from your point of view, and try to not to include experiences with information outside of the scope of the Lindholmen office.

Communication relevance

1. With the communication filters I feel more likely to miss important information (when compared to not using them).

Strongly disagree				Strongly agree
☐	☐	☐	☐	☐
1	2	3	4	5

2. The communication filters allow me to work fast and efficiently.

Strongly disagree				Strongly agree
☐	☐	☐	☐	☐
1	2	3	4	5

3. I encounter the information separated by the communication filters to be presented in a format that I can easily handle.

Strongly disagree				Strongly agree
☐	☐	☐	☐	☐
1	2	3	4	5

4. Overall I am satisfied with the communication filters.

Strongly disagree				Strongly agree
☐	☐	☐	☐	☐
1	2	3	4	5

Communication with the Team Coach

1. The empowered Scrum Master structure is easy to understand.

Strongly disagree				Strongly agree
☐	☐	☐	☐	☐
1	2	3	4	5

2. The empowered Scrum Master structure is easy to execute and, complexity-wise, it would be acceptable to adopt it.

Strongly disagree				Strongly agree
☐	☐	☐	☐	☐
1	2	3	4	5

3. The empowered Scrum Master structure facilitates coaching-related communications (when compared to the traditional method).

Strongly disagree				Strongly agree
☐	☐	☐	☐	☐
1	2	3	4	5

4. The empowered Scrum Master structure hinders coaching-related communications (when compared to the traditional method).

Strongly disagree				Strongly agree
☐	☐	☐	☐	☐
1	2	3	4	5

5. The empowered Scrum Master structure allows me to work fast and efficiently.

Strongly disagree				Strongly agree
☐	☐	☐	☐	☐
1	2	3	4	5

6. Overall I am satisfied with the empowered Scrum Master structure.

Strongly disagree				Strongly agree
☐	☐	☐	☐	☐
1	2	3	4	5

Farewell

Thanks for participating.

The information provided will contribute to optimize the way information is distributed at your workplace. I hope to work again with you soon! :)