



UNIVERSITY OF GOTHENBURG

# **The role of Distributed Version Control Systems in team communication and learning**

**Corina Diana Deaconu**

**Master of Communication Thesis**

**Report no. 2014:095**

**ISSN: 1651-4799**

# Table of Contents

Abstract.....	3
1. Introduction.....	4
1.1. Research question.....	5
1.2. Motivation for the study.....	5
1.3. Literature review.....	6
2. The Affordances of the Distributed Version Control System.....	8
3. Theoretical approaches to the study of distributed version control systems.....	13
3.1. Media Richness Theory.....	13
3.2. Media Synchronicity Theory.....	15
3.3. Andragogy - a theory of adult learning.....	18
3.4. Preliminary conclusion.....	21
4. Methodology.....	23
4.1. Content analysis.....	23
4.1.1. Balanced Payments Ltd.....	25
4.2. Survey research.....	25
4.2.1. Survey design.....	26
4.2.2. Survey participants.....	26
4.2.3. Survey distribution.....	27
4.3. In-depth interviews.....	27
4.4. Ethical considerations.....	28
5. Results and data analysis.....	29
5.1. Content analysis.....	29
5.2. Questionnaire results.....	35
5.3. Analysis of questionnaire data.....	44
5.4 Interviews.....	49
6. Discussion.....	53
6.1. General findings.....	54
7. Limitations and future research.....	56
References.....	57
Glossary of IT terms.....	60

Appendix 1: Content analysis data.....	61
Appendix 2: Online questionnaire.....	61
Appendix 3: Interview transcripts.....	61

## **Abstract**

The present thesis is a descriptive study on the usage patterns and perceived action possibilities of distributed version-control systems. The project offers an overview of the technology with a focus on its role in communication, information sharing and learning in IT teams or organizations. As such, the thesis fills a research gap in the field of computer-mediated interaction, by analysing distributed version-control systems in professional contexts, rather than in academic or educational ones.

The thesis bases its claims on established theories on communication technology and adult learning. The data collection and analysis in the project consists of a combination of qualitative methods, namely content analysis and semi-structured interviews, and quantitative methods – the questionnaire.

Distributed version-control systems play an important role in information sharing and developing understanding of voluminous or complex data. In the context of IT professionals working in teams, this technology can improve co-operation and increase the efficiency of interpersonal communication.

**Keywords:** distributed version-control, source control, communication, IT, learning, media richness, synchronicity, computer-mediated interaction

# 1. Introduction

Interpersonal communication is one of the requirements for the existence any organization, ultimately playing a decisive role in the efficiency and overall well-functioning of the organization, as a group of employees working towards a common goal. Similarly, the processes of continuous development, learning, skill-acquiring and sharing can be claimed to be other key elements upon which the success of the organization might rest.

In nowadays' technologized western society, a majority of organizations are aware of the importance of learning, knowledge sharing and communication and, as a result, different types of information and communication technology (ICT) systems are being employed to support these processes. Despite the modern workplace being dominated by digital artifacts, there is however often little understanding at a theoretical level concerning the ways in which these systems affect employee communication and the manner in which users actually work with the system's features. This fact is made apparent by the increasing number of empirical studies emerging in the field of team communications and digital, or computer-mediated learning and communications.

One type of organizations in which the study of technology usage and digital artifact communication is of particular relevance are those organizations having IT as their main area of business. IT organizations can be defined as “the department within a company that is charged with establishing, monitoring and maintaining information technology systems and services.” (Rawson, 2013). Consisting mainly of IT professionals, such as programmers, software engineers or testers, these organizations have technology as their work object, not simply as a tool supporting their work tasks. In this context, it is proposed that the present research should be conducted on IT teams, with focus on *team communication, information creation, sharing and learning*.

However, analysing a complex of information and communication systems is a task which might not be attainable in a research study of the present time span and magnitude. In order to avoid having diffuse results and a general, rather than concise, analysis of communication, the study will instead focus on one single type of computer system, which has not been analyzed thoroughly to date, namely the *distributed version control system*.

The concept of *version control* system is defined, according to the GIT Manual (Scott, 2009), as “*a system that records changes to a file or set of files over time so that you can recall specific versions later*”. The benefits of using a version control system in an IT company are obvious, as this would allow employees to “revert files back to a previous state, revert the entire project back to a previous state, review changes made over time, see who last modified something that might be causing a problem, who introduced an issue and when, and more”. Thus, version control ensures that no information inside the system is lost and that several employees may work simultaneously on the same piece of code without interfering with one another.

Several models of version, or revision control systems evolved throughout the years, the main ones being the local, centralized and distributed systems. While in the case of local version control systems all data is stored either on a single computer or on a central server, distributed models give all users, or peers local access to the entire project they are working on. Moreover, unlike their counterparts, distributed models allow tracking of all of a user's history, from small changes, to commands typed and identity, thus making them more suitable for research. As a consequence, as well as due to the fact that the distributed peer-to-peer version control model is the most widespread at present, with an adoption of over 36% among IT organizations (Cochez, 2013), this has been chosen as the technology for analysis in the current research project.

### **1.1. Research question**

As version control systems represent a crucial element in organizations whose main area of business is IT and programming, an interesting topic to follow would therefore be that of the perceived usage of such systems and of the communication patterns afforded by the interaction with this technology. Thus, the present research project aims to provide answers to the following two questions:

- 1. What is the role of distributed version control systems in organizational communication, information processing and learning?*
- 2. What strategies might users of the system employ in order to ensure optimal usage of its capabilities, with a maximized learning and communication experience?*

It is believed that by answering these two questions the project will provide a thorough understanding of the organizational and individual learning processes and communicative practices afforded by this technology. Furthermore, the fast-growing adoption of distributed version-control systems will be analysed and explained as a result of the study.

### **1.2. Motivation for the study**

Despite the fact that version control systems are not regarded as communication systems *per se*, it is the author's conviction that this technology in fact mediates and promotes learning, sharing and communicating. The project is meant to be a contribution to the relatively limited amount of research on the role of version control systems (and distributed version control systems in particular, henceforth referred to as DVCS) in learning, cooperation and communication. As stated in Cochez et. al. (Cochez, 2011), it is widely believed in the Computer Science academia that DVCSs contribute positively to learning and cooperation. The articles presented in the literature review section stand to prove this assumption correct. However, no study to date, as uncovered by the researcher, has analysed the process of DVCS-mediated-learning in professional settings inside an organization. Unlike users in educational settings, such as pupils or students, who have a limited experience of using DVCSs, IT professionals interact with the DVCS on a daily basis. The project can thus be claimed to fill a research

gap, by bringing forward a new perspective on the study of DVCS.

### **1.3. Literature review**

There is a limited amount of research focusing on learning and communication in the context of version control systems. As a result, the project will be based not only on literature on learning and communication mediated by DVCS, but also on a combination of literature on (adult) learning on the one hand and literature from the IT field regarding the DVCS on the other.

One of the first articles considered is *Learning by Doing: Introducing Version Control as a Way to Manage Student Assignments* by Reid & Wilson (Reid, 2006), which discusses learning in an academic context, analysing the effects of the introduction of a DVCS in the freshman courses syllabus. As the authors point out, the DVCS appeared to have improved students' cooperation while simultaneously enabling them to better explain their code in writing commit messages. Another relevant result is the fact that student teams were observed checking their work with other teams, in order to learn new implementations and solutions to a common problem. Thus, although the article does not feature experienced professionals, these results can be used to support the project's assumption that DVCSs have a positive effect on learning and communication.

Another article focusing on learning as mediated by the DVCS is *Version Control in Project-based Learning* (Milentijevic, 2008). Similarly to the article presented previously, this study also focuses on learning in an academic setting, rather than inside an organization. The authors follow the use of a DVCS in a group of students, with the purpose of identifying how cooperation and learning improves through project-based learning, as well as due to the medium's affordances. The perspective on learning adopted by the authors is a constructivist one, as it is assumed that learning should be based on experience. While a large part of the article discusses a proposed implementation of the DVCS in project-based curricula, several relevant points for the present project are mentioned. First off, the authors claim that by using a DVCS, students were able to observe and learn from each other's code design and architecture, as well as make use of components designed by other students (Milentijevic, 2008). Another interesting conclusion that the article presents is the fact that the DVCS allowed mentors or supervisors to have access to students' entire development process, rather than only the final product.

Cochez et al. (Cochez, 2011) provide a thorough analysis of the usage patterns of the DVCS in several Computer Science academic courses. In an attempt to observe students' committing patterns, working style, group leadership and system understanding, the authors resort to complex quantitative methods. Of particular interest for the present project is Cochez et al.'s analysis of the commit messages. The authors devise a commit taxonomy, by dividing messages into useful, trivial and nonsensical (Cochez, 2011). Their analysis revealed that some groups of students provided lengthy and detailed commit messages, indicating that the DVCS was being used as a group-communication tool.

Media Richness Theory (MRT) is one of the first theories to look at the role of media and communication technology in an organizational setting. It started out in 1984, as a theory of information richness (Daft & Lengel, 1984) and was later adapted to include newer types of media. The main premise of the theory is that success in organizations is directly connected to managers' ability to process and cope with information richness, uncertainty and equivocality, as will be more thoroughly discussed in a later chapter.

A different theory, Media Synchronicity Theory (MST), proposed by Dennis, Fuller and Valacich, focuses on the fit between the ability of different media to support what is claimed to be the main communicative processes of any group task, conveyance and convergence. The theory is based on MRT and it attempts to improve upon it, in order to better stand up to empirical evidence in the field of computer-mediated interaction.

The theory of Andragogy, introduced by Knowles is used in addressing learning from an adult perspective. The name of andragogy is a compound of the latin *androgia* and the word *pedagogy*, and it literally translates as adult pedagogy. The term has been purposely coined in order to “differentiate [andragogy] from youth learning” (or pedagogy), according to Knowles (Knowles & Shepherd, 2005). The theory, in its refined version, takes into account previous research from the fields of psychology, sociology, education and human resources in an attempt to set forth “a set of core learning principles applicable to all learning situations” (Knowles & Shepherd, 2005). Particular attention should be paid to the perspective on learning adopted by the theory. Admittedly a vast and multifaceted concept, **learning** is defined by Knowles as an “act or process by which behavioral change, knowledge, skills, and attitudes are acquired” (Knowles & Shepherd, 2005). An important distinction is made between learning and education, as the latter focuses on an agent/educator transmitting knowledge to a disciple, rather than on the process of acquiring information. Although not an all-encompassing and exclusive definition, the one above appears to capture the nature of learning. Moreover, this definition is reminiscent of MST's process of conveyance, thus suggesting that learning is an encompassing part of many human interactions.

As the literature review has revealed, the identified studies combining the topics of communication, learning and DVCSs have been conducted in an academic setting, in which students are either learning to make use of the DVCS's capabilities, or have a rather limited experience of using this technology. As the focus of the current project is communication and learning in a professional setting rather than an academic one, additional literature and theories on both learning and technology-mediated communication will be employed in order to support and respond to the chosen research question.



## 2. The Affordances of the Distributed Version Control System

Version control systems are generally viewed as playing a central role in IT/developer teams' interaction with their code, as expressed in McChesney (2004), Reid (2006), Milentijevic (2008) and others. This section of the paper will attempt to explain the reasons which make the DVCS a technology widespread among developers, by identifying and providing an in-depth characterization of the communicative affordances of the technology, as they are supposedly readily perceived by IT professionals in an organizational setting.

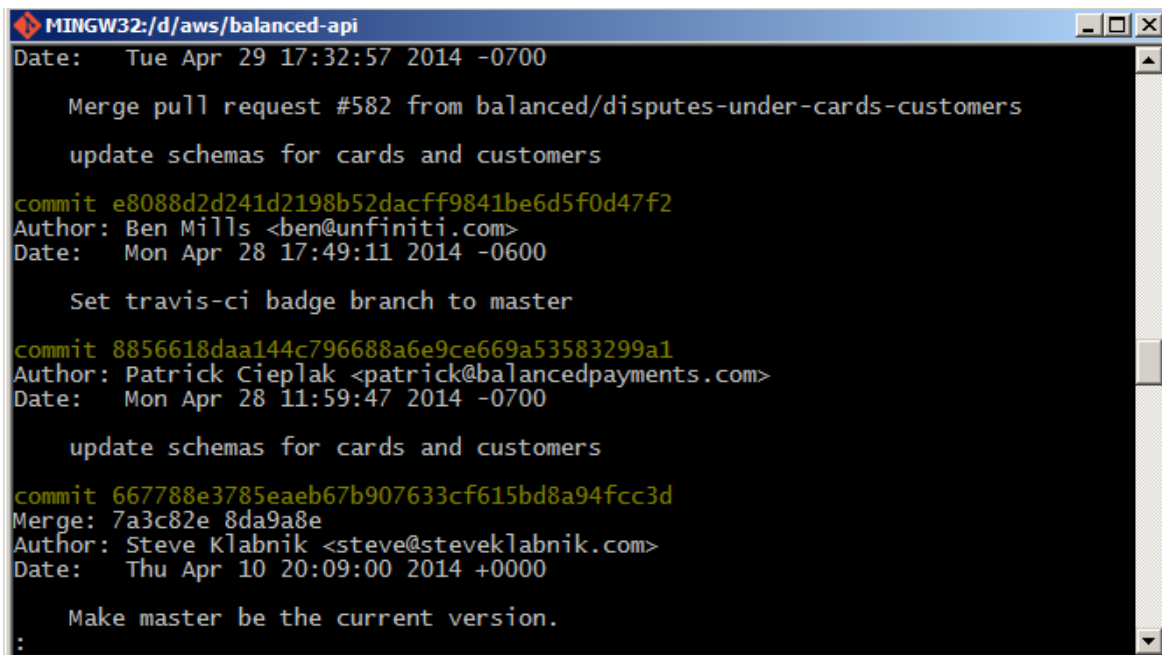
The current research project adopts the mechanical **communication model** developed by Shannon and Weaver (Shannon & Weaver, 1949). The model defines the act of communicating as the transmission of a message from a sender to a receiver, which might be obstructed by noise (Shannon & Weaver, 1949). Although not without its criticism, the model appears to be suitable for the purposes of the present research project as it can be applied to communication mediated by technology.

To begin with, the concept of **affordance** will be introduced. One possible definition, provided by Norman (1988), states that affordances refer to the action possibilities of an artifact or technology, as they are perceived by the user in the course of an interaction. Following this definition, it can be said that a ball has the affordance of throwing, while a button affords pressing. Furthermore, this definition is in accordance with the theoretical standpoint adopted in the paper, which assumes the role of technology in society can be identified through a balance between technological and social determinism. According to Oliver (2011), technological determinism "is the belief that technology shapes society in some way – which includes social practices such as learning". Social or cultural determinism is identified as focusing on "the social shaping of technology or political economies of technology" (Oliver, 2011). Another, later and perhaps more refined definition of affordances is that these are the specific characteristics of an artifact, which are stable with regard to the needs of the user and which differentiate the artifact from similar ones (Hutchby, 2009). By combining these two subtly different perspectives, affordances will be defined as the inherent properties of an artifact, which exist independent of the user, but which can be perceived only depending on the context.

Having established a working definition of the concept of affordance, several related terms need to be mentioned. To begin with, as expressed in Whittaker (2003), the literature on digital communication technologies focuses primarily on two main types of affordances, namely on modalities and interactivity. Thus, a majority of the influential theories in the field of communication technology employ these categories in order to make predictions and draw general conclusions on the effects of particular types of technology on society or the individual (Whittaker, 2003). The term modality refers to the types of cues a particular technology supports, such as, for example: visual, linguistic, verbal and non-verbal. An example of a technology having affording visual cues is a video-calling software. On the other hand, interactivity is concerned with the nature of the interaction a technology promotes. When considering the degrees of interactivity afforded by a technology, the focus is twofold: once on

whether communication mediated by the technology in question is synchronous or asynchronous and secondly on whether the communicative act is co-located or at a distance. Synchronicity is a concept which refers to a communicative artifact's ability to allow users to communicate concurrently, in real time (Alan, 2008) and similarly, asynchronicity characterizes technologies in which there is a delay between the time a message is sent and the time it is received by its intended interlocutor. A typical example for illustrating these concepts would be the telephone as affording synchronous communication and traditional mail as asynchronous technology.

Turning back to the DVCS as a communication technology, the affordances of the system are identified in what follows. First of all, taking into consideration that all interaction with the system, as well as with other users is in written form, either as command-line instructions or commit messages, it is clear that the system affords the transmission of linguistic information. The basic unit of information in the DVCS is, therefore, the commit message. No other types of cues except for linguistic ones can be transmitted using this technology and as a result it could be claimed that the DVCS is a modality-lean medium. In this respect, the DVCS is similar to another popular communication technology—the e-mail. Consider the example images below representing screen-shots of a DVCS (in this case Git).

A screenshot of a terminal window titled 'MINGW32:/d/aws/balanced-api'. The terminal displays a Git log for a merge pull request. The log shows three commits. The first commit is by Ben Mills, dated Mon Apr 28 17:49:11 2014, with the message 'Set travis-ci badge branch to master'. The second commit is by Patrick Cieplak, dated Mon Apr 28 11:59:47 2014, with the message 'update schemas for cards and customers'. The third commit is by Steve Klabnik, dated Thu Apr 10 20:09:00 2014, with the message 'Make master be the current version.' The terminal also shows the merge pull request number #582 and the branch name 'balanced/disputes-under-cards-customers'.

```
MINGW32:/d/aws/balanced-api
Date: Tue Apr 29 17:32:57 2014 -0700

Merge pull request #582 from balanced/disputes-under-cards-customers
update schemas for cards and customers

commit e8088d2d241d2198b52dacff9841be6d5f0d47f2
Author: Ben Mills <ben@unfiniti.com>
Date: Mon Apr 28 17:49:11 2014 -0600

Set travis-ci badge branch to master

commit 8856618daa144c796688a6e9ce669a53583299a1
Author: Patrick Cieplak <patrick@balancedpayments.com>
Date: Mon Apr 28 11:59:47 2014 -0700

update schemas for cards and customers

commit 667788e3785eae67b907633cf615bd8a94fcc3d
Merge: 7a3c82e 8da9a8e
Author: Steve Klabnik <steve@steveklabnik.com>
Date: Thu Apr 10 20:09:00 2014 +0000

Make master be the current version.
:
```

*Illustration 1: Git DVCS log using the command line interface*

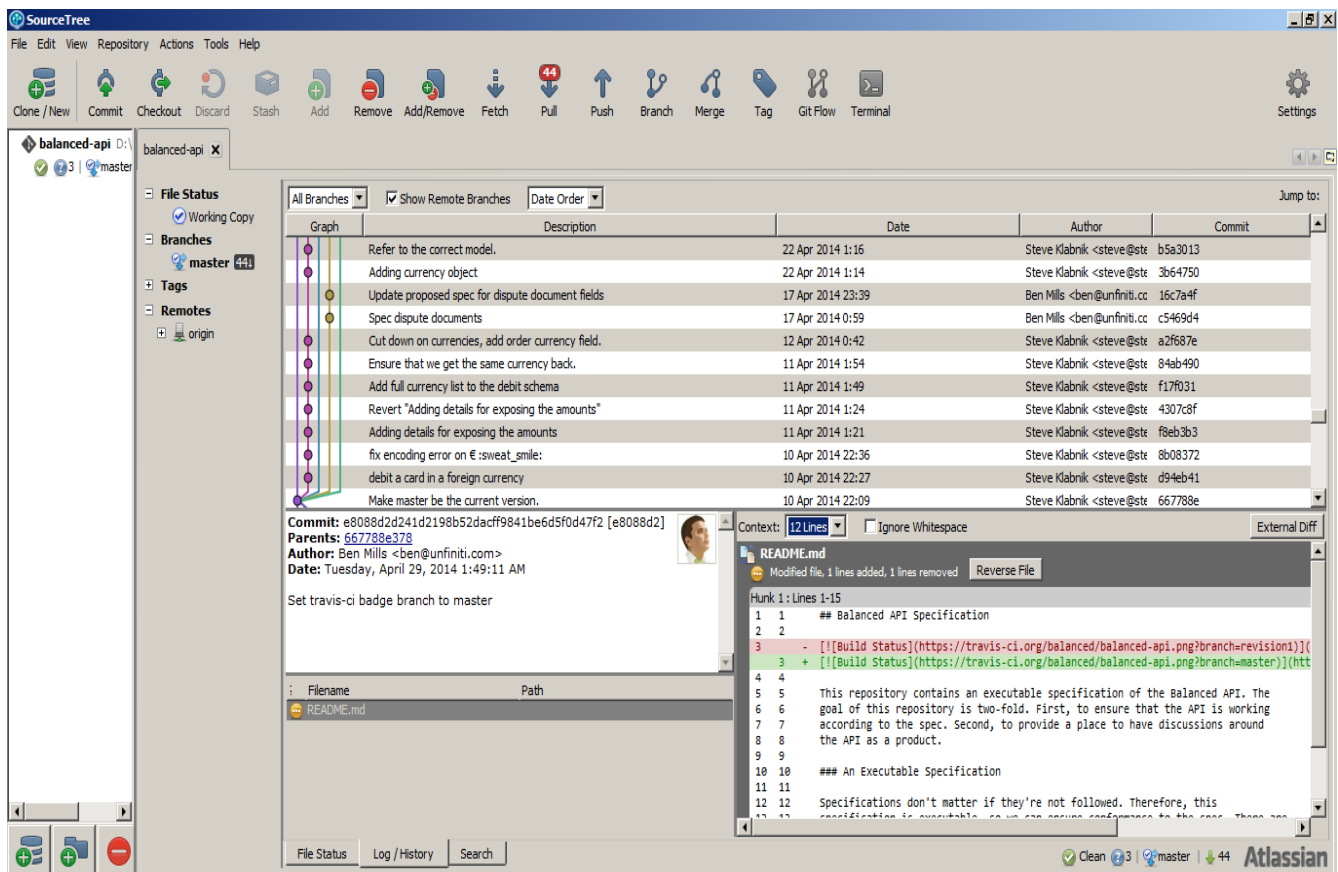
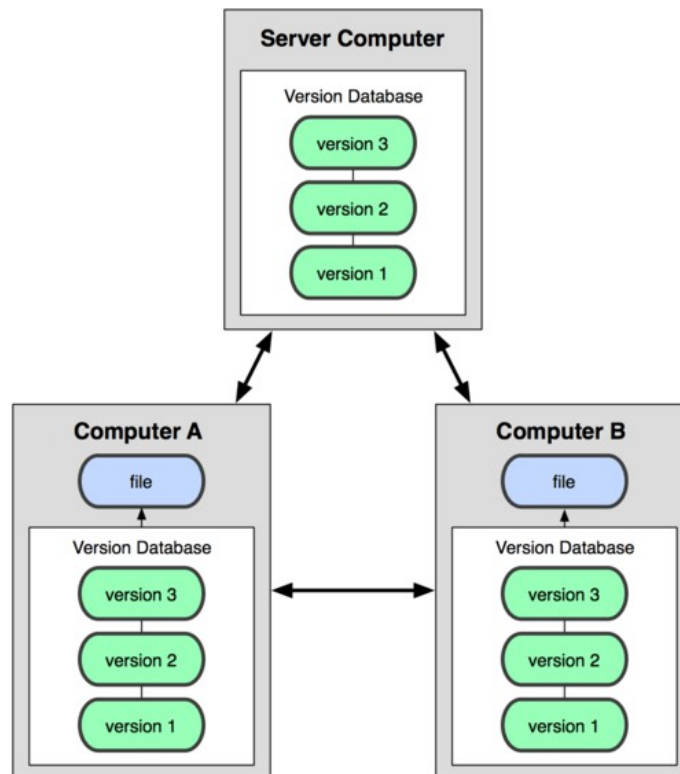


Illustration 2: Git DVCS log using a graphical user interface

With respect to the degree of interactivity the DVCS allows, it can be argued that it affords synchronicity in some respects, although it can also be viewed as an asynchronous technology. From a synchronous perspective, users are permitted to see changes made by others in real time, by requesting the latest version of a particular file. Moreover, the modification of files and sending of commit messages visible to all other contributors to the project is almost instantaneous, unaffected by any delay. Compared with the process of editing a file or piece of code without the support of the DVCS, it is assumed that the affordance of synchronicity is one of the factors which might have led to this technology's wide adoption among IT professionals. However, the DVCS does not offer any possibility for users to engage in conversation and as a result it falls into the asynchronous category. Moreover, on account of the distributed nature of the technology, all users keep on their devices a full copy of the projects and files they are working on and no external update is performed unless an explicit request for the latest version is made. This affordance might imply that several users are able to work simultaneously on the same file, at their own pace and without interference from one another. The concept of distributed version-control is represented in the diagram below, as taken from Chacon (2009).



Finally, another affordance related to interactivity is the fact that the DVCS promotes interactions at a distance. More precisely, users may modify files and write commit messages which are made available to others regardless of the physical distance between the interactants. This affordance, combined with asynchronicity, suggests that this technology can be used successfully in geographically distributed teams, across different time-zones.

Although the general affordances identified above allow for predictions of the manner in which the DVCS might be used and as such influence communication, these do not explain what makes this technology different from any others. It can be mistakenly assumed that the DVCS is quite similar to e-mail or perhaps wikis, as all three technologies can be described as affording the (a)synchronous transmission of linguistic information across geographical and temporal barriers. It is therefore necessary to identify the affordances or characteristics which distinguish the DVCS from other technologies, in an attempt to justify its widespread choice and usage in the IT and programming community.

As a first, one property specific to the DVCS is the fact that information is kept in the system indefinitely (alternatively, for a very long period of time). As the name of the technology suggests, the DVCS's main functionality is to allow users to keep track of different versions of a particular file or project, as well as to view the entire history of a file, with all its intermediary stages and changes. The fact that information, and communication, in the system is not ephemeral brings forth other related

possibilities of action. Thus, due to the fact that users may read past commit messages, the DVCS can be used as a learning technology. By having access to the entire history of a file or program, users may observe how their peers develop code, identify past problematic situations and observe how they were amended or learn alternative styles of coding.

The access to a conversation and version history points to yet another affordance of the DVCS, namely seamless access to all changes made in a project by all contributors. In a community of programmers where code is constantly changed and updated, the opportunity to follow all changes made by contributors in real time can be both time-saving and beneficial for the well-functioning of the development process. As an example, the DVCS may prevent a situation in which users constantly inquire about the files others have been working on, in an attempt not to perform the same modification twice.

Similarly, as the complete history of a file is stored in the system, it can be claimed that the DVCS allows users to experiment and learn by trying new coding styles, algorithms etc. More precisely, users can at any time revert back to previous versions of a file for example in case a mistake was made or a particular addition to the file is no longer required. This flexibility in restoring previous versions and the security of having permanent access to the history of a file can be claimed to contribute to the DVCS as a learning and self-reflecting tool.

All in all, this section has served to identify several of the main affordances of the DVCS. It has been proposed that this technology allows the transmission of written linguistic information, both synchronously and asynchronously as well as at a distance. Moreover, the affordance of information being permanently stored in the system has been identified central to the technology, with other affordances stemming from it, such as the ability to review and learn from past actions/communicative acts, the ability to follow peers' work and work process and the freedom to experiment in a loss-free environment. On account of these properties several assumptions can be made regarding the role this technology plays in IT/programming teams.

### 3. Theoretical approaches to the study of distributed version control systems

The previous section identified and explained the affordances of the DVCS and this section will shift focus towards several theories from the field of communication and technology studies, in an attempt to establish several basic premises of the current study, as well as in order to provide an account of the possible reasons behind the adoption of this type of technology in IT organizations. Based on the affordances of the technology, several theories are to be considered in order to successfully predict the adoption of the technology, its usage patterns and its potential effects on communication.

The section begins by introducing a theory which focuses on analysing media usage and technology choice based on the range of modalities and cues it provides users with (Daft & Lengel, 1986). This theory, known as Media Richness theory (MRT), has been widely influential in the communication field and as such, it is the starting point of several other theories. Once the premises of MRT are presented, another closely related theory is introduced, namely Media Synchronicity theory (MST). Rather than considering the range of modalities a particular technology affords, MST turns to the context in which technology is used and the situational factors in play (Dennis & Valacich, 1999). Due to the attention paid to the types of affordances which might render technologies more suitable in certain situations than others, MST will be part of the theoretical framework of the study. Finally, a third theory will be discussed, this too becoming an important part of the theoretical framework of the study. This last theory focuses on adult learning and can be applied to organizational development, providing a background for the hypothesis that the DVCS can be used to improve learning. Thus, the theory of andragogy, proposed by Knowles (1968) will round up the current subchapter.

#### 3.1. Media Richness Theory

The term media richness is defined as “the potential information-carrying capacity of data” (Daft & Lengel, 1984). With reference to communication media, “Media can be characterized as high or low in "richness" based on their capacity to facilitate shared meaning” (Daft & Lengel, 1987). Thus, the more types of communicatively relevant information a technology provides, such as non-verbal cues, the richer it can be considered. In order to better illustrate the concept, the authors provide a 5-item continuum, listing out several types of media, from rich to lean, as follows. Face to face communication is considered the richest type of information-conveying medium, followed by telephone conversations, written, descriptive documents and finally numerical documents.

The richness of a technology/medium can be assessed based on several criteria, as proposed in Daft & Lengel (1984). First, the **feedback capability** of a medium should be considered. Media which afford immediate feedback are described as richer, due to the fact that unclarities can be resolved and corrections can be made swiftly. This criterion is correlated to the affordance of synchronicity,

presented in the previous section. A second criterion for judging richness is the variety of **communicative channels**, or cues it can convey. In this case, it is claimed that media which affords multiple cues, such as visual, auditory, gestures, voice inflection etc. is richer and therefore better suited to carry complex information. Furthermore, the variety of symbols and **language** allowed by the medium is another indicator of richness—natural language is deemed to be richer than numerical expressions, which are instead suitable for communicating clear, quantifiable data. Finally, the **source**, or the personal/impersonal nature of the communication afforded by a medium represents an important factor in establishing its richness.

The main reason behind the need to ascertain the richness of a medium is, according to the authors (Daft & Lengel, 1984), due to the direct correlation between the degree of complexity of the group or managerial phenomena which need to be communicated or discussed and the richness required in order to achieve success. Thus, it is claimed that the choice of medium at managerial levels is influenced by whether the information task is simple, such as a routine check, applying a rule in a specific situation, or more complex, which requires interpretation, negotiation or clarification. The theory builds upon the idea that rich technology/media is best suited for interpreting information in the organizational environment, coordinating complex tasks, reducing uncertainty and equivocality and establishing a shared view of events (Daft & Lengel, 1984). Richer media is therefore predicted to be used in situations where the information task is uncertain and the organization is complex, while lean media should be encountered in less complex, more straightforward situations.

Moreover, in order to support this hypothesis, the authors propose a model for analysing the complexity of an organizational situation. The model focuses on two situational characteristics deemed vital for classifying information tasks, namely uncertainty and equivocality. Although related, the two concepts are subtly, but fundamentally different. Uncertainty is said to arise due to “the absence of information” (Daft & Lengel, 1984). A lack of information can be solved “through objective analysis” (Daft & Lengel, 1984), meaning that uncertainty can be solved swiftly by providing the additional information. Equivocality, on the other hand, is a concept which characterizes ambiguous situations, in which multiple valid interpretations may be identified. As the authors state, equivocality is reduced through negotiations and discussions. Considering these concepts, it is clear that MRT predicts rich media to be used in equivocal situations, while leaner media to be employed in situations defined by uncertainty.

Furthermore, another assumption is that information and media distribution in an organization propagate on two dimensions, following a vertical and a horizontal path, respectively. At a vertical level, the authors suggest that the higher up in an organizational hierarchy one is placed, the more equivocality is encountered. As a result, rich media should be used at managerial level, while less rich media should be employed as lower levels of the organization are reached. From this perspective, the pattern of usage of DVCS can be predicted and somewhat supported. Thus, by being a technology typically associated with developers in the IT sector, who can be claimed to be lower in the hierarchy than managers, the fact that the DVCS is lean in communicative cues is on par with the hypothesis that

less rich media is suitable at this lower level. However, it should be noted that this is unlikely to be the only reason behind the adoption of this technology inside the community of practice under analysis.

At a horizontal level, the authors state that the more interdependent particular divisions, or teams, in the organization are, and the more “divergent frames of reference” they have (Daft & Lengel, 1984), the richer a medium is required in order to achieve coordination. By transmitting rich information differences may be overcome. Considering the DVCS as an example, it appears to support this hypothesis—due to its affordances, it can be considered a medium suitable for usage across IT teams. However, due to the fact that it is suitable for keeping track of mostly written documents, such as programming code, it can be predicted that this medium will not be used across departments which have no relation to programming.

Although MRT is an influential theory in the area of computer-mediated communication, its description of the information task and the relation between equivocality and media choice has not been fully supported by subsequent empirical studies. In an article by Markus and El-Shinnawi (1997), for example, MRT is under critique, as it appears to wrongly predict users’ choice of video conversations above written messaging. Moreover, the theory is questioned with regard to its applicability and consideration of new media technologies. Similarly, MRT does not appear to give consistent results in the case of more traditional media either—it appears that under certain circumstances media which differ in richness lead to equally effective results and task completion times (Suh, 1998). Due to the perceived inconsistency of MRT, the current research project does not make use of criteria for establishing the richness of a medium to any large extent. These criteria serve merely as guidelines and have not been the main focus of the analysis. The main hypothesis drawn from the theory is the one that media low in richness is suitable for communication at horizontal levels in an organization. Taking into account the criticism brought to MRT, it is necessary to consider another theory, which might be better suited to explain the role and usage of the DVCS in interactions in the IT field.

### **3.2. Media Synchronicity Theory**

As an alternative and a complement to the previously presented theory, a theory concerned with the effect using media has on communicative performance is introduced in what follows. Media Synchronicity Theory (Dennis, Fuller & Valacich, 2008) focuses, more precisely, on “the ability of media to support synchronicity, a shared pattern of coordinated behaviour” among team members (Dennis, Fuller & Valacich, 2008).

While MRT attempts to account for media choice, claiming that performance is codependent with the well-matched richness of a medium with the information task needs, MST dismisses this idea, arguing that “not one medium [is] better than another” (Dennis, Fuller & Valacich, 2008). As a reaction to MRT’s weak empirical findings, MST turns its focus to new media, starting from the hypothesis that the appropriation and use of new media, and eventually even mediated communicative performance, depends on a successful fit between media capabilities and the specific needs of different



communicative processes. Both theories analyze technology and media from a communicative task perspective. However, MRT defines a task as a “set of communication processes needed to generate shared understanding” (Dennis, Fuller & Valacich, 2008), rather than a single process, for which only one media would be suitable.

Furthermore, the theory stresses the importance of identifying media capabilities which support, or conversely, discourage what are identified as the two main communicative process types. According to the authors, these main types of communicative processes are convergence and conveyance, respectively. By identifying the dominant process, as well as the capabilities of the media at one’s disposal, it is claimed that communicative performance can be improved. **Conveyance** can be defined as a process of transmission of information, making sense, enabling the interlocutor to “create and revise a mental model of the situation” (Dennis, Fuller & Valacich, 2008). As an example, a senior programmer telling a newcomer about the coding standard in the company would be an interaction dominated by conveyance. In this case, the DVCS might be used in order to convey information due to its affordance of storing and allowing access to large quantities of written data. Due to the fact that conveyance typically implies individual and in-depth processing, it is proposed that the transmission of information may be slow without affecting the process negatively. The term of **convergence** refers to discussions, debates on the meaning of a previously interpreted situation, with the purpose of reaching an agreement, or a common mental model. Convergence is typical in situations in which team or organizational members need to choose one item out of a list of possible options. Examples of situations in which the convergence process is dominant are when a team of designers need to decide on which layout to have for a new website or when a group of programmers need to choose which new feature should be added to their software project. That which characterises convergence is a rapid succession of opinions and arguments, which leads to the assumption that there is a need for fast information processing.

A key concept which the authors propose is directly linked to identifying a good fit between media and the communicative processes of conveyance and convergence is synchronicity. Although synchronous communication, as presented in the previous subchapter, is related to synchronicity, the two terms are not synonymous. Thus, while media may be used synchronously or asynchronously, depending on its capabilities or on the needs of the situation, as in the case of the DVCS, **synchronicity** is a “state in which actions move at the same rate and exactly together” (Dennis, Fuller & Valacich, 2008). More precisely, synchronicity implies a common focus between conversation interactants, as well as carefully coordinated behaviour. In the context of media and technology, media synchronicity is described as the extent to which the capabilities of a particular medium afford synchronicity in human interactions. Based on this definitions, a correlation between the level of synchronicity and the processes of conveyance and convergence can easily be established. High synchronicity implies engaged interaction, the fast transmission and evaluation of messages, as well as nearly instant feedback. Additionally, a reduced effort to decode or encode messages can also be attributed to high synchronicity, thus supporting the hypothesis that high synchronicity is typical of convergence processes. On the other hand, low synchronicity appears to be better suited for conveyance processes,

as it presupposes a longer time for sending/receiving messages, non-immediate response and an overall decreased level of interaction. As conveyance typically implies the processing of complex, lengthy or diverse information, it is obvious that by employing media low in synchronicity more time is afforded to process, analyze or develop the information. Moreover, conveying information through media low in synchronicity has the added advantage of allowing the sender to compose the message carefully, by taking into account contextual factors and possible misunderstandings.

Having established that high synchronicity is beneficial for convergence processes, as well as low synchronicity for conveyance, the authors devise a set of properties on the basis of which a medium's degree of synchronicity, or "capability to support information transmission and processing" (Dennis, Fuller & Valacich, 2008) can be established. These properties are derived from the classical model of communication by Shannon and Weaver, where a sender sends a message through a channel to a receiver. The first capability to be considered is the **transmission velocity** of the medium. This physical characteristic of media refers to the speed with which a message can be transmitted and reach the receiver. As an example, written mail has a low velocity, while the telephone has a high velocity, as the message reaches the intended receiver almost instantly. Transmission velocity is a component of synchronicity, as it directly influences the level of interactivity, speed of feedback and the conversation-like nature of an exchange. It is claimed that high velocity improves synchronicity. Considering that the DVCS was identified in the previous subchapter as affording the immediate transmission of information, it will be described as having high transmission velocity.

A second property which is said to determine media synchronicity is **parallelism**, or the number of simultaneous interactions a medium allows senders to engage in. Also known as the width of a medium, parallelism implies the sending and receiving of messages from multiple interactants at the same time and, according to MST, with no need to manage turns or sequences. Due to the multidirectional nature of communication when using media rich in parallelism, synchronicity is reduced, as a common focus is not easily achievable. However, wide media appears to be useful in conveying large amounts of information quickly. The DVCS as a technology affords parallelism highly, by allowing users to receive, read and analyze information and commit messages from multiple authors simultaneously.

Another capability which is said to influence a medium's afforded degree of synchronicity is the set of symbols it provides. **Symbol sets** refer to the number of ways in which information can be encoded by using a particular medium, similarly to the *number of cues* and *language variety* features in MRT. It is postulated that media with a wide range of symbol sets is more suitable for convergence, and implicitly affords synchronicity, while media lean in symbol sets promotes a reduction of social presence. As established previously, the DVCS only allows information to be sent in written form and therefore it can be viewed as a medium suitable for conveying information rather than debating on meaning.

The last two capabilities which can be used to determine a medium's level of synchronicity pertain to

the individual use of the medium rather than to its physical capabilities. Thus, **rehearsability** is defined as the extent to which a sender is allowed to compose, rehearse, edit or fine grain a message before sending it. While face-to-face or telephone communication prompts for immediate replies and feedback, more asynchronous technologies, among which the DVCS, allow senders to compose the message in their own time. Although positive in situations in which complex information needs to be transmitted, rehearsability leads to delays and as such deters synchronicity. Finally, **reprocessability** is concerned with whether or not the receiver is allowed to re-read, examine or process the message during or after it has been sent. A medium which affords reprocessability is not expected to promote synchronicity, and thus would not be best used in situations where convergence is desired. However, by allowing interactants to revisit messages, information processing and decoding can be done more thoroughly. Moreover, new conversation participants or system users can gain access and understanding of previous activities. Reprocessability is afforded by the DVCS, as identified previously—users of the system have constant access to all previous information and no message may be completely removed from the system. Moreover, access to previous data is encouraged through the existence of a dedicated command.

Based on the previously presented characteristics, it can be safely claimed that the DVCS is a technology which has capabilities better suited for interactions requiring low levels of synchronicity, more precisely, conveyance processes. The delay in feedback typical of DVCSs has been shown to promote a more thorough and deep understanding of the information exchanged through this channel, which is beneficial for conveyance. Moreover, as a wide medium, the DVCS gives users access to information coming from multiple sources simultaneously, thus being theoretically suitable for managing large volumes of data. Due to the rehearsability of the messages transmitted using this medium, information is expected to be well-structured and fine-tuned, leading to improved communicative performance in conveyance processes. Finally, the high level of reprocessability afforded by the DVCS points not only to its appropriateness for usage when large volumes of information, or difficult to process data, needs to be transmitted, but also to its potential as a learning tool. By keeping a history of all messages and information exchanged through the system, this technology facilitates the understanding of previous conversations and of the development process. In order to fully develop the concept of learning and on the learning affordances of the DVCS, a new theory is required. The following subsection presents an educational approach to the technology under analysis, as a complement to the hypotheses raised by the previous two theories.

### **3.3. Andragogy—a theory of adult learning**

As the previously introduced theories illustrate, different media and technologies may influence communication and information exchange. However, these theories of media pay little attention to the process of learning and to the manner in which technologies mediate or affect knowledge and skill gaining. The topic of learning, is instead prominent in all the articles identified as dealing with version control from a communication and mediated interaction perspective, thus supporting the need to address this topic in the current research paper as well. MST briefly mentions learning in arguing for

the advantages of reprocessability. It could, however, be argued that the process which MST calls *conveyance*, building a mental model based on new information, is in fact similar to a type of learning, as developed further in the subsection. Therefore, a theory of learning is required in order to account for this perceived affordance of the DVCS. Due to the fact that the DVCS is a technology used preponderantly by adults, a theory of adult learning is preferred, and Knowles' theory of andragogy (Knowles & Shepherd, 2005) has been chosen for that purpose.

The main assumptions, or defining characteristics to be taken into account when designing successful learning situations, are introduced in what follows. To begin with, it is proposed that both young and adults learners' motivation is influenced by six main factors and that by making the correct assumptions in connection with the needs of the learner and the situation at hand, the success of learning can be predicted. By translating this hypothesis to the field of computer-mediated learning, and in the current case, the DVCS, it can be suggested that a technology possessing affordances which satisfy the motivational needs of group of learners will be more suitable for learning experiences in that group. These factors are as follows: the learner's **need to know**, the **self-concept**, or degree of self-direction of the learner, **prior experience**, the **readiness to learn**, as well as the **orientation to learning** and finally the type of **motivation** (Knowles & Shepherd, 2005). As adult learners are believed to have interests and abilities which differ from those of young pupils, the assumptions regarding child learning differ from those regarding adult learning, as presented in what follows. An adult is defined by the theory as an individual whose psychological self-concept is self-directed, responsible and independent.

In the case of young learners, or when dealing with pedagogy, the assumption is that learners do not have a strong need to know why they are learning a certain skill, but rather they follow the instructions provided by an authoritative figure. Moreover, the self-concept of young learners is dependant and personal experience does not play an important role in the learning process, due to its limited amount and low quality. In pedagogy, learners' readiness appears to be determined by their desire to obtain good marks or pass a course and the manner of acquiring knowledge and skills is usually systematic, divided into clear subjects based on logic. Finally, the motivation of young learners is claimed to be mostly extrinsic, meaning that external factors such as parents' opinions, grades or teachers' attitudes dominate the learning process. By applying these assumptions to mediated learning, it might be concluded that technology which guides the user and closely monitors user actions and provides feedback would be suitable for young learners. Additionally, media affording a clear top-down transmission of knowledge or data (from teacher to student) would also appear to fit the needs of pedagogy.

On the other hand, the needs and assumptions regarding adult learners differ largely from the pedagogical model. First of all, adults are claimed to have an acute need to know the reasons behind undertaking a learning activity. Knowles claims that adults carefully weigh the advantages and disadvantages, or the benefits and costs of learning a new skill or piece of information (Knowles &

Shepherd, 2005). Furthermore, by being responsible for their own lives, adults are said to require a large degree of self-direction in learning, as impositions and restrictions are perceived as negative. Moreover, as adults have both more and qualitatively better experience than youths, efficient learning implies the use of personal experience, through processes such as problem solving, case studies and peer-tutoring. Related to the readiness to learn of adults, andragogy assumes that only knowledge which is deemed necessary for accomplishing or coping with everyday situations is readily learned. An important proposal is that “exposure to models of superior performance” (Knowles & Shepherd, 2005) may induce a readiness to learn. While youths acquire knowledge best in a structured form, adults are claimed to learn best from real-life situations, such as a work problem. The motivation behind undergoing a learning process differs also in adults from young learners. Although extrinsic incentives such as a better salary or work position can to some extent motivate learning, Knowles suggests that adults’ main learning drive is intrinsic, ranging from job satisfaction, an increase in self esteem to any other type of personal gain.

Although the assumptions mentioned above may not hold under all circumstances, depending on, among others, individual differences, situational factors or the goals and purposes for learning (Knowles & Shepherd, 2005), it is safe to assume that they accurately describe the appropriate conditions and manner of adult learning in general. Taking these assumptions into account, some hypotheses can easily be constructed with regard to adult learning as mediated by technology. Thus, a first hypothesis would be that if a technology possesses affordances which reflect and recreate the conditions which are assumed to characterize successful adult learning experiences, then that technology will also afford learning. More specifically, the claim is that media which meets the principles of andragogy is a suitable learning tool for adults. As the previous chapter has anticipated that the DVCS would afford learning, the principles of andragogy can now be used to verify this affordance.

Starting by turning to the need-to-know-assumption, it is clear that the DVCS does not directly offer any suggestions or guidance regarding what information or skill the user should learn. On the other hand, by providing comparison tools and a timeline of data and commentary additions and modifications, the technology might help users decide and set their own learning goals. From this perspective, the DVCS appears to comply with the proposal that adults need to know the reason for learning in order to be motivated. Moreover, due to the diverse features and commands of the DVCS, it can be argued that users are fully responsible for the type of information they have access to in the system. By not imposing any restrictions on the type or amount of information available to particular users, this medium promotes self-directed actions, and thus also adult learning. As a concrete example, users can choose to print out a list of commit messages either with the default formatting or in a custom manner. Options range from specifying a desired time-span for the messages, printing only messages older or newer than a certain date or relative time (e.g. older than two weeks), sorting by author, e-mail address etc. to only returning messages that follow a desired pattern, or commits from a particular branch of the file system.

Another grounding hypothesis of andragogy is that adults learn more effectively when their experience is acknowledged and put to use. From this perspective, the DVCS can be labeled as a powerful learning tool. As mentioned previously, the system keeps a record of all information and thus users can access their old contributions at any time. This feature may be regarded as an opportunity for users to review their coding styles, problem-solving techniques and as a result, learn from their past successes and failures. Additionally, the access to the complete history of a project may support not only individual learning, but also peer-assisted one. As a distributed system, all information is available to all users, which suggests that learning from the experience of others is possible. Moreover, through its distributed and open nature, the DVCS might simplify peer-helping, as it makes the discovery of errors, failures or mistakes by peers working on a common project more accessible. Another consequence of using this medium might be the “exposure to models of superior performance” (Knowles & Shepherd, 2005), which the theory of andragogy claims to induce a readiness to learn in adults. More specifically, it is expected that if users observe that their peers solve tasks more effectively, in a different manner, or that their writing/coding style is more robust or attractive, then they will be more willing to learn and adopt the model perceived as superior. Finally, as the main purpose of the DVCS is to give access to and store data for either personal or work-related projects, it can be claimed to pair up well with the adults’ task- or problem-centered orientation to learning.

All in all, based on the principles of andragogy, it has been shown that the DVCS can be described as a suitable tool for adult learning. The medium has been argued to afford learning through its features, such as the diverse list of commands, the ability to share data with multiple users simultaneously, the ability to review data historically and without restrictions or the ability to view and compare peer data and problem-solving techniques, which create fruitful conditions for learning.

### **3.4. Preliminary conclusion**

The current chapter has discussed several theoretical stances on communication, learning and technology in relation to the focus of the present research paper. The theories presented were chosen with regard both to their relevance and general adoption in their respective field and to their particular relevance to the topic of the paper at hand. Thus, Media Richness Theory is part of the theoretical framework of the project due to its pioneering advances in technology-mediated communication and organizational theory. By applying the richness measurement criteria proposed by MRT to the DVCS, the richness of the communication technology can be established. As pointed out in the previous section, the DVCS can be considered synchronous, thus ranking high on the feedback capability feature. However, due to its limited cue range, this medium is far leaner than face-to-face communication, the telephone or other media which afford more than written verbal communication. With regard to the variety of symbols the medium allows, the DVCS can be described as rich, as it is not limited to abstract, or numerical language. Similarly, due to the fact that any message in the system is attributed to a particular individual, the technology may also be said to be rich in the category of personal interactions. Overall, on the continuum proposed by the authors of MRT, the DVCS could be placed between the telephone, a rather rich medium, and written documents, which are leaner. Media

Richness Theory has been selected due to its closeness to the previous theory and with regard to its alternate approach to media capabilities. By combining these two theoretical frameworks, a more thorough and complete analysis of the DVCS as a communication technology is predicted. The third theoretical approach of the current project, the theory of adult learning—*andragogy*, was selected in order to make predictions and pertinent observations related to the learning aspect of the medium under analysis. As the focus of the project is both mediated communication and learning among IT professionals, this theory was deemed the most appropriate, as it is a dominating theory in the domain of adult learning. Based on the theoretical framework built in the current chapter, empirical data has been gathered and analysed, ultimately leading to a formulation of an answer to the research questions of the project. The methodology, data and analysis make the subject of the coming chapters.

## 4. Methodology

An important part of any scientific research paper is the presentation and argumentation of the chosen methodology for the study. Thus, it is aim of the current chapter to introduce the type of research which was conducted, the methods employed and the reasoning behind each of them. The chapter starts off with the introduction of the concepts of quantitative and qualitative research, accompanied by additional general methodological aspects which concern the project in its entirety. Succeeding this section, separate sections are dedicated to each of the three methods of data collection and analysis adopted by the project, namely message content analysis, the questionnaire and the semi-structured interview.

To begin with, the type of research conducted in the present project can be classified as descriptive. As defined in Bhattacharjee, descriptive research “is directed at making careful observations and detailed documentation of a phenomenon of interest.” (Bhattacharjee, 2012:15). As the project aims to provide a close observation and in-depth analysis of the role of the DVCS in communication and learning in IT organizations, this type of research is deemed most appropriate. Having established the nature of the study, consideration is given to the methodological approach towards data collection and analysis. Two distinctive methodologies of data collection are identified in the scientific community, namely the quantitative approach and the qualitative one. A simplified definition of the terms may be that “Qualitative analysis is the analysis of qualitative data such as text data from interview transcripts.”, while “quantitative analysis, which is statistics driven and largely independent of the researcher” (Bhattacharjee, 2012:113). A similar distinction is proposed by Dey: “Whereas quantitative data deals with numbers, qualitative data deals with meanings.” (Dey, 1993:11). However, although “qualitative researchers claim that their aim is to provide rich description so as to achieve understanding” and “quantitative scientists aim for prediction” (Sechrest, 1995), the present project does not intend to be limited by adopting a single methodological approach. As suggested by Sechrest, “good science is characterized by methodological pluralism, choosing methods to suit the questions and circumstances” (Sechrest, 1995). A similar position is adopted by Denzin, who introduces the term of **triangulation** to refer to the “combination of methodologies in the study of the same phenomena” (Denzin and Norman, 1978). According to Dezin, a qualitative approach is necessary in order to reach a clarity of meaning and to uncover the themes and direction for research. In the case of the current project, qualitative data can be used in order to point out the relevant properties of the DVCS which influence team communication and learning. In addition, quantitative data can be used in order to reduce the bias and thus provide the project with increased validity. Taking into account all the factors presented above, it has been decided that a combination of quantitative and qualitative methods would be employed, as further elaborated in the following sections.

### 4.1. Content analysis

The method of content analysis is a qualitative method, defined as “the systematic analysis of the content of a text” (Bhattacharjee, 2012:115). In conducting this type of analysis the focus lays on



textual meaning and the design of communicative messages (Downe and Wamboldt, 1992). According to Bhattacharjee, the first step in conducting content analysis is to sample “a selected set of texts from the population of texts for analysis.” (Bhattacharjee, 2012:115). Thus, the type of data collected for analysis would come in textual form and would consist of the complete history of commit messages which follow a particular project in an IT organization. The choice of analysing messages from a single organization was taken with regard to the span of the current research paper – it was considered that in order to perform an analysis of messages from more than one organization would a longer period of time would be required, and in addition it might lead to diffuse results. It could, therefore be argued that the content analysis in the present research project resembles a case study of an IT organization.

In choosing the project whose messages to analyze several factors have been considered, namely: ease of access to the contents of the IT project, complexity of the commit messages, diversity of the messages and finally the number of contributors, or authors of commit messages. Thus, after considering several alternatives, it was decided that an open source project would best fit the requirements of the project. In order to decide on a particular project, the researcher consulted GitHub, a website which hosts open source projects along with their publicly available version controlled history. In addition to the fact that the website ensured that the IT projects hosted were all using a DVCS, another advantage was the opportunity to browse through a multitude of different projects and to observe their commit messages, in order to ultimately choose a project whose messages would come close to the perceived standard in the industry.

The list of projects to choose from was narrowed down to the 14 projects featured in the “Open source organizations”-showcase on GitHub, as the focus of the research paper is on communication inside organizations. Out of those 14, only projects which could be considered as active, namely those receiving commits on a regular basis, rather than sporadically (a few times per year) were deemed appropriate for the purposes of the research study. Moreover, repositories which featured both simple, short and more complex and lengthy commit messages were sought after, as these would result in a more rounded analysis. Similarly, projects featuring messages which could be categorised into multiple topics were preferred to projects dominated by mainly one or two types of commit messages, as the former would allow for a more complex coding scheme (Bhattacharjee, 2012:115). As the purpose of the analysis is to gain insight into the communication patterns of a team of IT professionals, another key factor in choosing a repository was the number of contributors, or committers to it. The concept of team is interpreted to imply more than two individuals and as a result, only projects consisting of three or more contributors were taken into account. Thus, based on all the factors presented above, a choice had to be made between the following projects: CFPB, Adobe central hub for open source, GitTip.com and Balanced Payments. The project Balanced Payments, introduced in more detail in the following subsection, was the one chosen for the purposes of content analysis as it featured the largest number of contributors within the organization, as opposed to independent, external ones.

Following the selection of a relevant body of messages for analysis, the process of unitizing may begin. As proposed in Berg, textual units “vary according to the nature of the research and the particularities

of the data”(Berg, 2008) and they can be chosen at the level of “words, phrases, sentences, paragraphs, sections, chapters, books, writers, ideological stance, subject topic or similar elements relevant to the context” (Berg, 2008). Considering the fact that the commit messages have a rather unitary structure in themselves and that they are already ordered based on the data and time of creation, the unit for analysis in this section of the project will be the commit message.

#### **4.1.1. Balanced Payments Ltd.**

As previously stated, the commit messages from a single open-source project make up the data which is subject to content analysis in the research paper. The project chosen is a banking application called Balanced Payments (henceforward referred to as Balanced). By being open source, all the content of the application, as well as their version controlled history and commit messages are available to the public. However, unlike many such projects, Balanced is the product of the established company with the same name which has a permanent IT development team consisting of 12 IT professionals (<https://www.balancedpayments.com/about>). According to information on the website, a number of five independent coders have contributed to the project up present.

As a self-declared *open company*, the main values of Balanced are openness and transparency. The company also stands for driving innovation and purpose, building passion in the community and being committed to drive global commerce, according to their website. The main vision behind the project is “through payments—improve the global economy” (Matin Tamizi, CEO). Inspired by software companies which release part of their assets and code freely, Balanced embraces this philosophy completely. While not without its advantages, going open source appears to have also brought about several challenges. Firstly, the developers admit that they need to develop faster and have an increased feeling of accountability for the quality of their work. Moreover, the internal processes in the company are more formalized and decisions are often reasoned and argued for. The advantages Balanced mention are, among others, the ability to receive feedback from the public even before deciding to implement or modify a feature or piece of code. In addition, new features can be evaluated more easily and moreover, outside professionals may contribute to the project.

The DVCS as a technology plays a central role in Balanced’ open-company-strategy. To begin with, the company uses a DVCS internally, to keep track of changes in code and documentation. Furthermore, the complete version history of the project, along with all the information stored in the DVCS is also made publicly available on the freely accessible hosting website GitHub (<https://github.com/balanced>). Thus, Balanced appears to be using the DVCS as a tool and a means to fulfill their goals and values, by sharing information openly both internally and with the outside world.

## **4.2. Survey research**

The content analysis data is complemented by the results of a survey based on the technique of the standardized online questionnaire. A questionnaire can be defined as “a research instrument consisting of a set of questions (items) intended to capture responses from respondents in a standardized manner”

(Bhattacharjee, 2012:74). The questionnaire is a method generally associated with the quantitative methodology, as it evaluates patterns of behavior in a standardized manner, as opposed to other more qualitative research methods. Moreover, this method is deemed appropriate for collecting “data about people and their preferences, thoughts, and behaviors in a systematic manner.” (Bhattacharjee, 2012:73). In the context of the present project, the questionnaire is used to gather information relating to respondents’ habits and expressed preferences in connection to the DVCS.

#### **4.2.1. Survey design**

The standardized questionnaire was divided into four sections, each focusing on particular themes and consisting of similar types of questionnaire items. By dividing the questionnaire it is expected that the respondents would follow the information more easily, thus improving the response rate of the method. Moreover, with the same purpose, as suggested in Bhattacharjee, the questionnaire items are disposed in a logical manner, starting with easy-to-answer questions and following with more complex items which might require reflection from the part of the participants. The first section consists of single-answer, fixed-choice, screening questions which are meant to establish whether the questionnaire participants are eligible for the study. These questions refer to the professional experience of the respondents. Similarly, the second section of the questionnaire is made up of fixed questions which gather data on the respondents’ familiarity with the DVCS and partly their habits of using the technology. The main security question “Are you currently using a distributed version control system for work-related projects?” is also placed in the section. A third section contains a list of Likert-type items, meant to measure respondents’ attitudes and opinions with regard to the affordances of the DVCS, as well as the connection, if any, between the DVCS and communication and learning inside an organization. Likert items are “simply-worded statements to which respondents can indicate their extent of agreement or disagreement on a five or seven-point scale ranging from “strongly disagree” to “strongly agree”” (Bhattacharjee, 2012:47). A five-point scale was chosen for the purposes of the project, as it was assumed to offer the desired level of detail with regard to respondents’ attitudes. The points on the scale were the following: strongly agree, agree, neutral, disagree and strongly disagree. The fourth and final section of the questionnaire consisted of two questions with demographic character, which could provide additional insight into the variety and distribution of the sample of respondents.

#### **4.2.2. Survey participants**

As the purpose of the paper is to observe and account for the role and usage of the DVCS as a communication and learning technology among IT professionals, it can safely be claimed that the target population for the online questionnaire is IT professionals using a DVCS for work-related purposes. However, as the target population is dispersed geographically around the globe and can not realistically be contacted to take part in the study, a smaller, accessible sampling frame has been selected. The sampling frame has been built on the basis of several techniques, as detailed in what follows. First, several IT companies in the Stockholm area were contacted and asked whether they were using a DVCS and whether they would be interested in participating in the study. The willing companies provided the researcher with a list of e-mails to be used for questionnaire distribution. The

sampling frame was further populated with members of social-media groups aimed at either IT professionals or DVCS users. The social-media sources employed during this phase were Facebook, Google Groups and LinkedIn. In addition to the previously named sources, potential questionnaire respondents were also selected based on membership in the technology-communities on the web-platform Meetup. The survey was not sent to the authors of the commit messages making up the content analysis data as the current research paper aims at obtaining a broader perspective on DVCS usage, rather than presenting the opinions and usage patterns of a single organization. It is believed that a diverse sampling frame may reduce the risk of obtaining biased results. With regard to the sampling technique, simple random sampling was chosen, thus ensuring each member of the sampling frame had an equal chance of receiving the questionnaire.

#### **4.2.3. Survey distribution**

The questionnaire was distributed online, through two different channels. On the one hand, the questionnaire was uploaded on social-media groups which were IT- or DVCS-themed, on the social networks of Facebook, Google Groups and LinkedIn. In addition, the e-mail addresses obtained from participating IT companies and the platform Meetup were used to send the questionnaire to potential respondents' personal e-mail addresses. The distribution method for the questionnaire was chosen with regard to both convenience and the assumed characteristics of the target group. More precisely, as IT professionals have the personal computer as a main tool for work, it was assumed that a significant portion of the population would have access to the Internet and to e-mail.

### **4.3. In-depth interviews**

The final method which the current project employed for gathering data was the qualitative interview. According to DiCicco and Bloom "The purpose of the qualitative research interview is to contribute to a body of knowledge that is conceptual and theoretical and is based on the meanings that life experiences hold for the interviewees." (DiCicco and Bloom, 2006). For the purposes of the project, the interview would provide additional insight into the perceived affordances of the DVCS. Moreover, the role of the technology in communication or learning might also be uncovered during the interview phase. Following the recommendations of DiCicco, a semi-structured interview was designed, consisting of several open-ended questions based on the main topics of learning, communication and team interaction through the DVCS.

Starting from the sampling frame composed for the purposes of the online questionnaire, potential interview participants from the Stockholm area were contacted by e-mail. Only two e-mail recipients agreed to participate in the study and thus, two qualitative interviews were conducted. The participants were both IT professionals, employed in Stockholm, at two different companies whose main area of business is IT/software development. They were chosen based on convenience and with regard to the same factors as in the case of the online questionnaire, namely connection to the IT field, experience in IT and familiarity with the DVCS. The interviews were conducted after the results of the questionnaire had been analysed, at locations convenient to each of the voluntary participants and lasted between 15-

20 minutes each.

Turning to the structure of the interview, it was designed as an inverted funnel, starting with easy-to-answer, single answer questions and gradually introducing more open-ended, reflective ones (Treadwell, 2014:133). Thus, the first questions offered interviewees the opportunity to introduce themselves and their experience with the IT domain and the technology under analysis. Following these questions were questions closely related to the research study, which served as a reinforcement for the questionnaire results, by providing a more nuanced perspective on the topic. Although the questionnaire consisted of predetermined items, a follow-up question was used in order to allow one of the interviewees to clarify and expand on their expressed opinion.

#### **4.4. Ethical considerations**

Each part of the data collection process can be associated with particular research ethics. To begin with, particular research ethics should be taken into consideration with regard to the fact that the content analysed commit messages in the project are based on one particular IT organization (Balanced Payments). As an open source project, the researcher had unrestricted access to the organization's main IT project, including source code and commit messages. It is, however, the duty of the researcher to guarantee that no names or information which might lead to the identification of the contributors or users of the Balanced project will be published in the final project report. The same anonymity was guaranteed for questionnaire respondents, as no personal questions which might be traced back to any particular individual were asked. Moreover, contact details were provided in the introductory part of the questionnaire, in order to receive eventual complaints or reflections with connection to the study being undertaken. Similarly, participants in the interviews were assured of their confidentiality, more specifically by being presented with a confidentiality agreement which stated that their answers will not be used in any other purpose than for the present project. Moreover, as requested, no mention of the interview participants identity was made public in any document pertaining to the project.

## 5. Results and data analysis

Having presented the research methodology and the logic behind it in the previous chapter, the current chapter aims at presenting the results of the research process. The chapter is divided into several sections, each of which is dedicated to the results obtained through a particular method. First, the data gathered for the purposes of content analysis will be introduced. This data consists, as mentioned in the previous chapter, of commit messages from a particular project. Due to the large volume of data, as well as due to the nature of the information contained in the messages, the section will focus primarily on categories of messages. A second section is concerned with the results of the online questionnaire. Based on the responses gathered, relevant statistics will be provided, along with a thorough explanation of their possible implications. Finally, the third section is a summary of two in-depth interviews on the topic of communication, learning and the DVCS. This section will be concerned with several main themes and motives that have been identified during the interviews. The chapter ends with a partial conclusion and some observations on the data.

### 5.1. Content analysis

The commit messages which will be analysed in what follows span over a period of 532 days, or approximatively one and a half years. The first commit message dates back to August 2012, representing the initial submission into the DVCS, while the last message to be considered is dated April 2014. The number of messages on which the analysis is based is 742. There are in total 17 unique commit message authors and contributors to the system, which will only be referred to by means of their initials.

To begin with, several general observations will be made with regard to the structure and form of the commit messages in Balanced's application. First off, each commit to the DVCS consists of the file or piece of code to be uploaded, followed by the commit message, as well as other information (such as the author's name, date of the commit, file path etc.). The message itself is made up of a title, which (usually) appears to sum up or meaningfully describe in short the nature of the file to be uploaded in the system, and a message body, in which a lengthier description, additional details or other communicative data is transmitted. While the message title is mandatory, accompanying all of the modifications registered in the DVCS, the body is often left empty. The short nature of the messages could denote a lack of richness of the medium, but it might also be the case that this is an organizational choice, taken in order to make browsing the repository/code faster and finding particular items in the system more efficient. With regard to the writing style of the messages, a semi-formal register is prevalent throughout the project—as supported by the presence of contracted forms (“didn’t”, “can’t” etc.) and informal vocabulary (“stuff”) intertwined with full forms (“does not”).

A convention adopted by Balanced is that if a message contains information in the body, then the title of the message ends with a full stop (“.”). As the body is not usually visible from the main system log, this convention can be claimed to have beneficial effects, by indicating to users that more information is available. The same convention holds even for commits in which a text file (such as the technical

documentation) is added, modified or updated. Another observation is that files committed to the DVCS generally have a modular or unitary structure, meaning that they often consist of a single feature, one fixed error/bug or one module. As a consequence, the content of the commit message usually covers a single topic. This might be another strategy users employ in order to make browsing the history or a repository more accessible. The presence of clusters of commits made in a short span of time suggests that users willingly break their work into individual units or modules, as illustrated below:

Time and date	Author's initials	Commit message (title)
2013-12-17, 19:49:52	SK	Implement PUT in the client.
2013-12-17, 19:50:03	SK	New feature: update card

Having concluded with the general observations, a closer analysis of the commit messages follows. Upon a close reading of the commit history from the period previously mentioned, it was observed that a majority of the data consists of modular messages, whose main topic can easily be identified. Moreover, the overt content of the commit messages revealed the existence of recurring themes. In order to present the data in a succinct form as well as in an attempt to identify the communicative capabilities of the DVCS as illustrated by the commit messages, the analysis begins with a classification into the main identified themes. A first category consists of **addition messages**, which usually accompany new features in the project, new functionalities, modules or any other addition to the version controlled files. Another clearly delimited category is that of the **fixing messages**. This comprises all bug fixes and error corrections, as well as other editing messages, such as rewriting documentation or fixing typos. Furthermore, a considerable number of messages can be labeled as **progression reports**. More specifically, these messages are part of commits which, taken individually, do not make up a unit, but are rather only part of a feature, complex function, bug fix or document. Finally, the last comprising category is that of the **merging messages**. Merging is the process of “join[ing] two or more development histories together” (Git documentation). In addition to the messages which fit in the above described categories, there are a number of odd messages which can be grouped under the label **miscellaneous**. Examples of such commit messages are one-word or abbreviated, unintelligible ones, non-specific messages or messages written in a conversation-like style.

Based on the categories outlined above a number of representative messages were chosen for a closer analysis. The messages included in the analysis were selected based on the presence of keywords considered representative for each category, such as: “new”, “add”, “fix”, “typo”, “change”, “progress”, “update” and “merge”. In addition to messages containing the a fore mentioned keywords, several messages were selected through a close reading of the data. Messages not selected for analysis were those which could not be understood or properly interpreted by the researcher, such as messages consisting of a single string of abbreviated content, or messages in which the communicative content could not be properly identified.

The content of the selected messages was interpreted from a communicative perspective, focusing (whenever possible) on elements which might suggest, display or indicate interactivity, interpersonal exchanges or any other relevant aspect of communication. To begin with, addition messages are considered below.

*(2014-01-02, 22:45:18) R.: New features: customers and orders*

This is a typical addition message in which the author informs his/her team-members of the creation of a new feature for the application. Despite the standard and short form, this type of message is believed to play an important role in the development process, as it allows users to be constantly up to date with the latest modifications. Moreover, by using the DVCS and logging all of the addition messages from a particular period, a clear picture of the structure and capabilities of the application emerges. Not all addition messages are brief, however. Consider the example below:

*(2013-12-31, 03:31:13) SK.: New feature: refunds go to original funding source.*

*I made the last scenario mis-understanding the way that the YAML scenarios worked.*

*Often, they have duplicate stuff with new names just to make sure that things work.*

*That means I don't need to translate them 1-1, I have to figure out which*

*thing is actually important and just do that.*

While the title of the commit message keeps the format pointed out in the previous example, this message also consists of a body. The message body, in this case, is not directly related to the accompanying commit file. Instead, the author comments on his misunderstanding of an organizational convention (“the way that the YAML scenarios worked”). This admission can be interpreted as an apology to any team-members that might have observed the mistake. In addition, the author also provides an explanation and clarification of the nature of the misunderstanding. It is clear that this part of the message is addressed to team-members who might have the same unclarities, thus teaching or reminding them of the proper procedure for writing scenarios.

Whenever an addition is made which is not a new feature to the application, the commit message resembles to the following:

*(2014-02-07, 00:17:45) MFL: adding requests json schemas*

The following category, the fixing messages, can be divided into two subcategories, namely code and bug fixes on the one hand and spelling or typos on the other. A majority of the fixing messages which refer to code modifications are structured such as the examples below:



(2013-09-11, 02:01:21) MFL: fixing makefile so it can be used to easily parallel results  
or

(2014-04-02, 21:09:46) ST: Fix up insufficient funds test

Not only was the number wrong, but I wasn't testing for the right things.

Almost all messages in this subcategory mention in their title the item which has been fixed. It is, however, also a recurring practice to argue/explain the reason behind the fix, or to point out which part of the project was malfunctioning and in what respect. In addition to their obvious role in informing team-members of the correction made in the code, these messages may also prompt readers to recheck the fixed items by comparing them with the old versions. Thus, users may choose to analyze the nature of the error, as well as the identified solution closely. This interpretation is especially relevant in cases in which an error is fixed by someone other than its original author.

A particularity of several fixing messages is the fact that the author expresses his/her mood or feelings, either directly or indirectly, recurring in some cases to the use of emoticons. The following are some examples of this phenomenon:

(2014-04-10, 22:36:06) SK: fix encoding error on € :sweat\_smile:

or

(2013-01-17, 4:55:02) MFL: I did it, its done, no more failing

or

(2013-09-06, 8:45:46) MT: yay! dry run passes

A possible explanation for the presence of emotional expressions in the commit message may be linked to the fact that errors and bugs are undesired, unpredictable events, which might raise strong negative feelings, such as anger and frustration, in those tasked with remediating them. As the three examples presented above illustrate, once the error has been corrected, users may express their relief and share their accomplishment with the peers in the form of a commit message. The emoticon-like “:sweat\_smile:” in the first message is directly associated with relief. Similarly, the repetition “I did it, its done” marks the ending of a possibly tedious or lengthy bug fix. Finally, the third message contains the interjection “yay”, generally interpreted as an expression of happiness, achievement or relief.

Turning to the spelling correction messages, they tend to be rather short, generally consisting of a single word, namely “typo”. It might be the case that these modifications to the version-controlled data are not of major relevance to the team, thus explaining their brief nature.

The following major message category, namely progress reports, are similar in structure with the addition messages. In resemblance to the latter, progression messages typically accompany additions to the code or document. However, while addition messages indicate that a complete new feature or

independent piece of code has been uploaded into the DVCS, progress reports are connected to partly developed code, unfinished functions or any other pieces of information which the author deems as unfinished. More specifically, these commit messages describe code which can not be interpreted as unitary or stand-alone. As a consequence, such messages usually appear in clusters of two or more. This differentiates progress reports from any other type of commit, as the others can generally be interpreted independently. Below are several examples illustrating clusters of progress reports from the data gathered:

(2014-01-17, 2:27:04) MFL: the end is nearing  
(2014-01-17, 3:40:04) MFL: one more to go  
(2013-01-17, 4:55:02) MFL: I did it, its done, no more failing  
or  
(2014-01-07, 1:35:03) SK: Moving forward on 'create an order'  
There's still a failure, unsure which part is right or wrong.  
(2014-01-07, 2:22:03) SK: progress on checking escrow on orders.  
Same delivery address issue  
(2014-01-07, 20:06:09) SK: Making some progress on 'checking escrow of order'  
Erroring for the same reason the other Orders are.  
(2014-01-07, 20:14:55) SK: Making progress on 'orders can't be credited more...  
... than escrow balance.  
(2014-01-07, 20:22:27) SK: Making progress on 'Create a refund'  
(2014-01-07, 20:22:27) SK: Update interpolation syntax  
I missed this one because the scenario fails before it hit this part.  
But it's pretty obvious it needs fixed.  
(2014-01-08, 12:45:18) SK: New feature: underwrite a customer.  
Closes #468

As the examples above demonstrate, progression report clusters may finish with either a message informing of the correction of an error or with a message stating that a new feature/function/etc. has been added to the repository. The role and importance of progression messages can be interpreted from more than one perspective. To begin with, these reports usually follow progress over a span of several hours up to a day or more. It may therefore be inferred that the choice of breaking a lengthy, complex task into several smaller ones ensures that the progress is safe in the DVCS, rather than being stored on a single computer. Furthermore, these messages also have the typical role of informing other team members of the changes made to the repository/code. In addition to the direct information conveyed by the content of the message itself, progress messages may play a secondary role, namely in indicating to other team members the fact that the committer is engaged in an activity, rather than being idle. As the DVCS makes information available to the entire team, it is believed that progress messages may be a suitable solution for showing oneself as occupied, as well as for preventing other team members to start working on the same piece of code, effectively performing the same task twice.

Having concluded the analysis of the progression messages, the focus is now on merging, or automatic messages. As defined earlier in the section, merging implies the joining of a file from a branch, or repository, into another branch. This procedure, made possible by the DVCS, is widely used in the IT community. It allows users to work on, for example, unstable, experimental branches and to later combine their work into a main, generally stable master branch. As merging does not imply the addition or removal of any data from the system, but rather the movement of one file/piece of code, commit messages in this case are typically generated automatically by the DVCS. Consider the examples below:

*(2013-07-17, 1:48:21) MFL: Merge remote-tracking branch 'matin/revision1' into rev1*

or

*(2012-10-31, 2:40:10) A: Merge remote-tracking branch 'matin/revision1' into rev1*

Finally, the last category of commit messages, the miscellaneous one, is briefly discussed in what follows. This category consists, as stated previously, of messages which do not easily fit into any of the previous categories. While such commit messages include abbreviated or difficult to understand messages (in-team jargon), other, more standard messages may also have their place here. A particular type of message from this category is the “thank-you” message, illustrated in the examples below.

*(2014-05-02, 22:25:23) SK: Let's just stick with the current month.*

*We're already bumping the year, and this calculation gets complex.*

*Thanks @jtdowney!*

or

*(2014-04-11, 1:54:02) SK: Ensure that we get the same currency back.*

*Thanks @matthewfl*

or

*(2013-12-08, 22:00:52) SK: refactor to use the client*

*Thanks @matthewfl*

It is proposed that thank-you messages in the DVCS have a social function, rather than an informational character. In such instances the commit message enables users to transmit greetings to each other.

To sum up, the data gathered for the purposes of content analysis was divided into several main categories, namely into addition messages, fixing messages, progress reports, automatic and miscellaneous messages. This classification has served to strengthen the hypothesis regarding the existing affordances of the DVCS, as well as reveal patterns of communication in the Balanced team. Thus, all types of messages were observed to play a role in the conveyance of information. Furthermore, addition messages were also interpreted to verify the presence of reprocessability, as this type of messages can be used to gain an overview of the features and additions in a project inside the

DVCS. The category of fixing messages has been claimed to reveal, among others, the learning capabilities of the system, as these types of commit messages may contain explanations and guidance from the authors to their team members. Progress reports were observed to add a social dimension to the technology, thus appearing to support the hypothesis made by MRT that media low in richness is suitable for communication on a horizontal level in the organization, in this case among members of a team. Each category of commit message has been analysed closely, on the basis of relevant examples from the data. Focus has been laid on both the general structure of the messages and on the particularities of the content. Moreover, the communicative role and importance of each message type has been taken into account and presented. Remarks have been made on the possible implications each commit message might have with regards to informing team-members, such providing explanations or apologies or expressing gratitude.

## 5.2. Questionnaire results

The content analysis of the commit messages performed in the previous section has revealed some of the potential roles that the DVCS plays in team communication. However, the a fore mentioned analysis only takes into account the perceptions and attitudes members of a single organization have towards the DVCS as a learning and communication tool. As the current research study aims at providing a broad overview of the technology, it was deemed necessary to gather data which would reflect the attitudes and usage patterns of a larger population, rather than being confined to one organization. The current section presents the results of an online questionnaire conducted in an attempt to strengthen and refine the findings of the previous section.

The online questionnaire ran between March 13, 2014 and May 14, 2014, and was sent to 133 potential respondents. A total of 64 answers were gathered during this period. With regard to the source of the responses, a third of them are estimated to come from IT professionals employed in interested companies in Stockholm, while the remaining responses were gathered from respondents belonging to social-media groups or the social-hub Meetup. Out of the total number of respondents, 4 (6.5%) replied negatively to the check-question, by stating that they were not actively using a DVCS for work-related purposes. As a consequence, the main relevant number of questionnaire answers was reduced to 60. Taking into account the sample size, with a desired confidence level of 95%, it was calculated that the confidence interval for the questionnaire data is of 12.8 percent (Treadwell, 2011). The confidence interval reached is deemed appropriate for a project of the present scale and resources. Additional details on the method for data gathering have been provided in the previous chapter.

The questionnaire was designed based on several themes, chosen with regard to both the theoretical framework of the project and to the research question. The main themes are **behavior and habits of using the DVCS**, the DVCS as a **team communication technology**, DVCS's **influence in organizations** and finally the **medium as a learning tool**. In addition to the theme-related questions, two questions regarding demographic data were placed at the end of the questionnaire.

The first section of the questionnaire consists of a series of questions pertaining to the respondents' professional background. This type of data has been considered relevant as it might reveal a correlation between professional background and history and respondents' attitudes and declared usage of the DVCS. In an attempt to structure the questionnaire logically into sections, the main check-question was only made available at the end of the first set of questions. Thus, questions 1 to 4 received answers from all 62 respondents. To begin with, the first question, Q1 - "Are you currently employed in the IT or Computer Science field?" received 58 positive responses and 3 negative ones. In case of a positive answer to Q1, the follow-up question Q2 - "How long have you been employed in the IT field?" was provided. The available answers to Q2 were: less than 1 year, 1-4 years, 5-10 years and over 10 years. Out of the 64 respondents, 37.5% (24 respondents) stated they had between 1 and 4 years of experience in IT, 28.1% (18 respondents) chose the option 5-10 years, while 25% (16 respondents) replied they had over 10 years of experience in IT or Computer Science. A minority of 9.4% (6 respondents) stated they had less than a year of work experience in the field. A visual representation of the responses for Q2 is provided in Figure 1.

The next question, Q3 - "How long have you been employed in your current organization?", is a logical follow-up to Q2. The options to Q3 coincided with those provided in Q2 and the responses were as follows. A majority of the respondents (39 or 60.9%) replied they had been part of their current organization for less than one year. The next option with the highest amount of responses was 1 to 4 years, with 18 respondents or 28.1%. A minority of the respondents indicated they had been employed for more than 5 years in their current organization, with 7.8% (5 respondents) choosing the 5-10 years option and only 3.1% (2 respondents) choosing the remaining option. Figure 2 illustrates the data obtained from Q3.

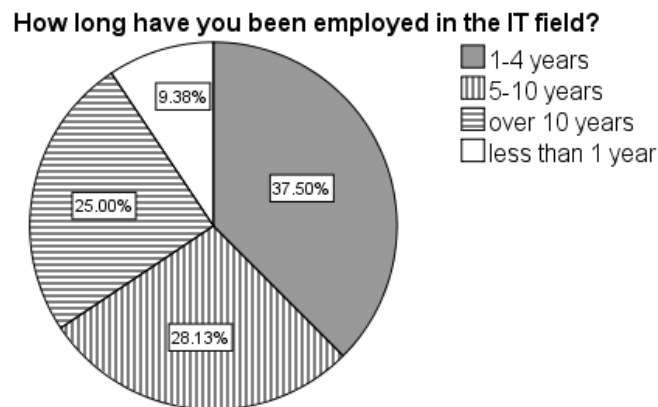


Fig. 1

**How long have you been employed in your current organization?**

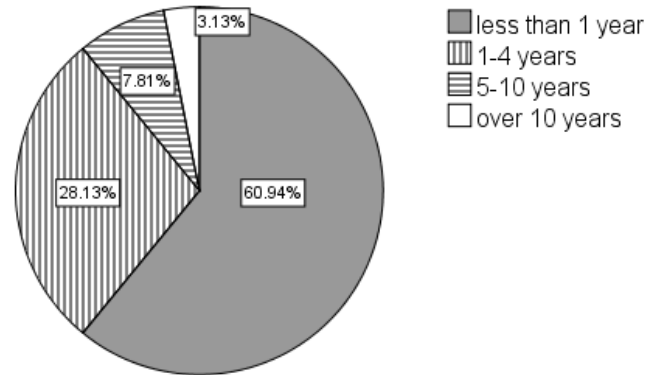


Fig. 2

Finally, the last question pertaining to the first section of the questionnaire referred to the position the respondents occupy inside their organization: Q4 - “What title describes your current position inside the organization most accurately?”. The question provided the following answer options: application analyst, computer scientist, database administrator or developer, programmer/developer, software analyst, web developer. In addition to the fixed answers, an open “Other”-option was provided, allowing respondents to input an own response. As presented in Figure 3 below, the majority of the respondents (44 out of 64, or 68.7%) revealed their position as programmer or developer. Additionally 10.9% (7 respondents) identified themselves as web developers, while 18.8% or the respondents provided their own answer, thus considering that none of the provided options was accurately describing their current position. One respondent, representing 1.6% of the total, chose the computer scientist option, while no respondent chose the options of application analyst, database administrator or developer or software analyst.

**What title describes your current position inside the organization most accurately?**

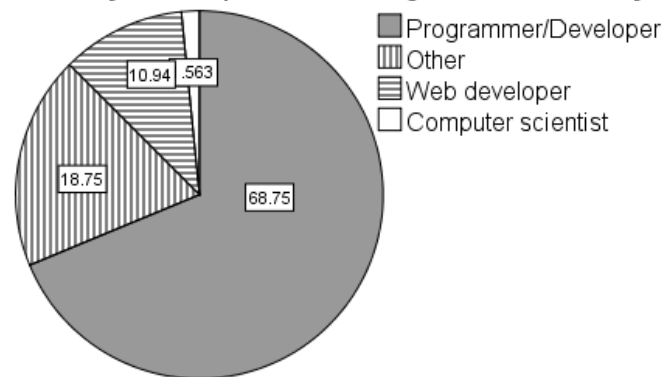


Fig. 3

The second section of the questionnaire comprised the three questions, Q5 to Q7, and focused on the frequency and use of the DVCS among IT professionals. The section begins with the check-question Q5 - “Are you currently using a distributed version control system for work-related projects? (ex. Git, BitKeeper, Mercurial)”. As mentioned previously, 93.5%, or 58 respondents replied positively, while 6.5% (4 respondents) chose the negative option. The remainder of the questionnaire was only made

available to the 58 respondents who replied positively to Q5.

Q6 - “How long have you been using a DVCS?” received the following responses: 38 respondents (63.3%) indicated they had been using a DVCS for a period of between 1 and 4 years. A percentage of 25 (15 respondents) out of the 60 respondents replied they had less than a year’s experience of using the technology under survey. An additional 8.3% (5 respondents) chose the option of between 5 and 10 years of experience and finally 3.3% of those questioned (2 respondents) stated they had over 10 years of experience in using a DVCS. The percentages are illustrated in Figure 4.

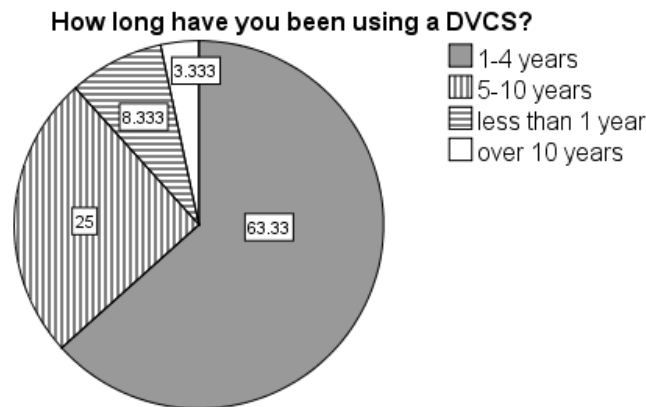


Fig. 4

The following question, Q7 - “How often do you commit changes to your work projects?” provided the following possible answers: several times a day, once a day and several times a week. The first option was chosen by 50 of the respondents (83.3%), while 5 respondents (8.3%) indicated they were committing changes into the DVCS at least once a day. Moreover, an equal number of respondents (8.3%) replied that they were performing this action several times a week.

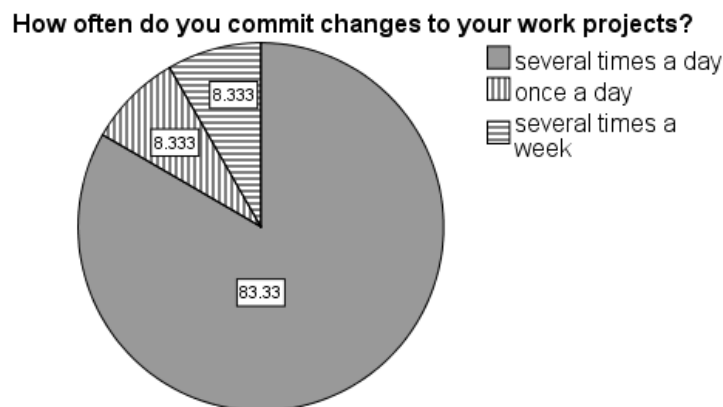


Fig. 5

The following section of the questionnaire consisted of a series of 9 Likert-type items, with a five-degree scale with the following options: Strongly agree, Agree, Neutral, Disagree and Strongly disagree. The section gathered data on organizational and team communication, as well as on learning mediated by the DVCS. The results of this section have been illustrated through bar charts, due to the

ordinal nature of the data. The first item, I1 - “I write my commit messages according to a team/organizational standard.” received the following responses: 9 respondents strongly agreed with the statement, representing 15% of the total. An equal number of respondents, namely 21 (35% each), indicated agreement and neutrality towards the statement, respectively. 8 respondents (13.3%) chose to disagree, while one respondent (1.7%) strongly disagreed with the item.

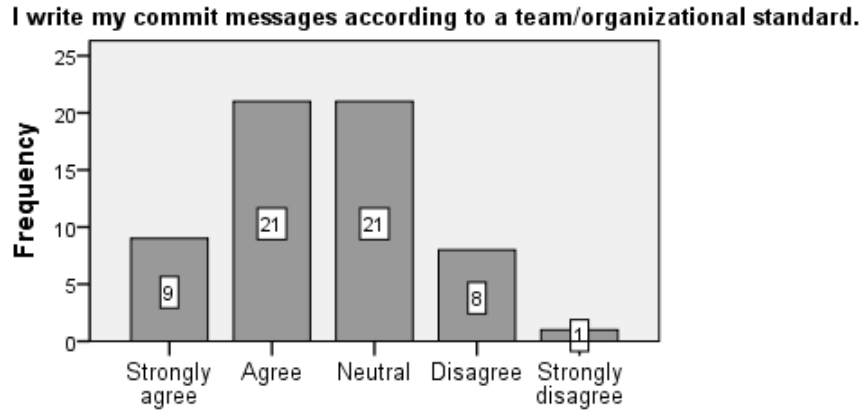


Fig. 6

The second Likert-type item, I2 - “Summarizing my changes in commit messages helps me reflect on the work I have done.” gathered 18 strongly agree-responses (30%). 33 of the respondents (55%) indicated they agreed with the statement, while 7 (11.7%) were neutral. Only 2 respondents (3.3%) disagreed, with none of them strongly disagreeing.

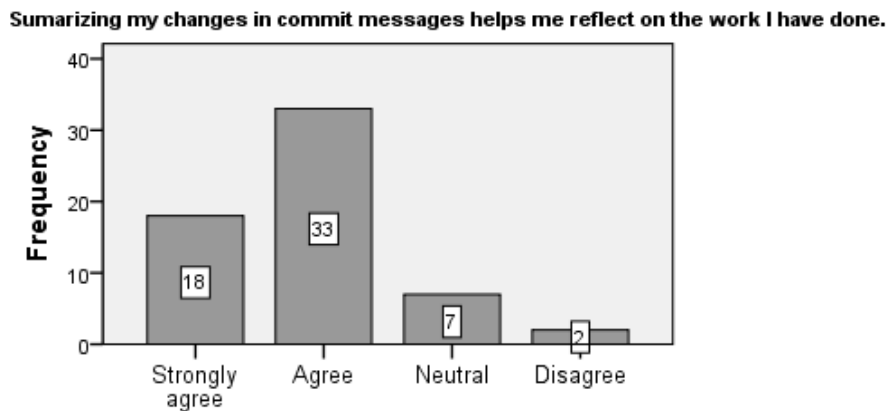


Fig. 7

The responses obtained for the third item, I3 - “I use the DVCS to familiarize myself with the coding standards at the workplace.” were as follows. Out of the 60 respondents, 26 (43.3%) agreed with the statement. An additional 18 (30%) indicated neither agreement or disagreement. Moreover, 14 respondents (23%) disagreed with the statement. The options “Strongly agree” and “Strongly disagreed” were chosen by one respondent (1.7%) each.



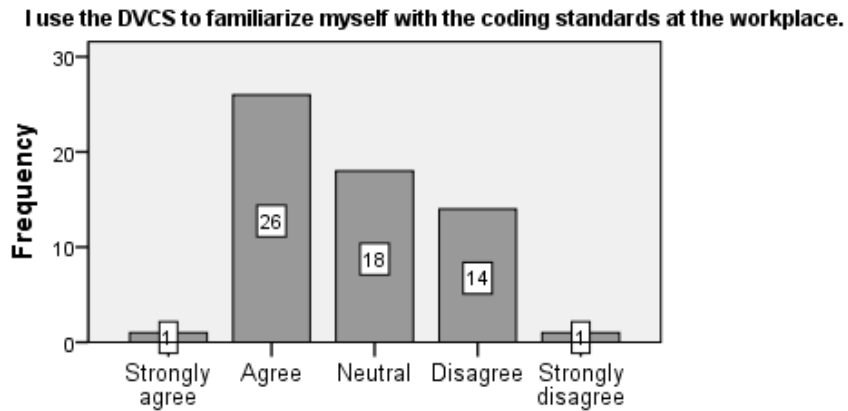


Fig. 8

The following statement, I4 - “I write commit messages in order to remind myself or the team what needs to be done next.” received only a single “Strongly agree” answer (1.7%) and 12 agreement replies (20%). 13 respondents (21.7%) chose the “Neutral” option, while 25 (41.7%) disagreed with the statement. An additional 9 respondents (15%) expressed strong disagreement with this item.



Fig. 9

With regard to item 5, I5 - “Using a DVCS improves my co-operation with the other team members working on the same files.”, 31 of the 60 respondents (51.7%) strongly agreed and an additional 20 (33.3%) expressed agreement. A minority of 6 respondents (10%) were neutral towards the statement and 3 responses (3%) of disagreement were recorded. No respondent disagreed strongly with the statement.

**Using a DVCS improves my co-operation with the other team members working on the same files.**

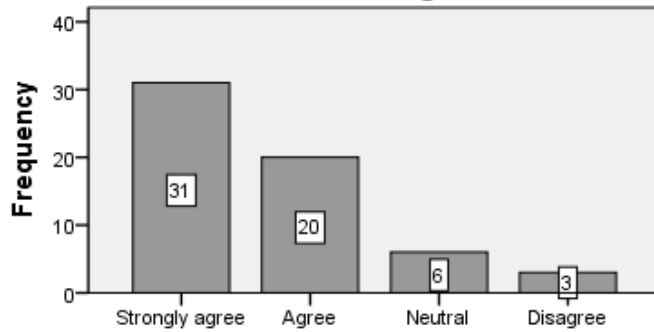


Fig. 10

The attitudes expressed concerning item 6, I6 - “I learn new implementations, techniques or designs by checking my teammates' code through the DVCS.” are presented in Figure 11. As illustrated, 6 respondents (10%) indicated strong agreement towards the statement. A total of 29 respondents (48.3%) stated they agreed with the item. There were 20 (33.3) “Neutral” answers and 4 (6.7%) “Disagree” ones. One respondent (1.7%) disagreed strongly with the idea expressed in the statement.

**I learn new implementations, techniques or designs by checking my teammates' code through the DVCS.**

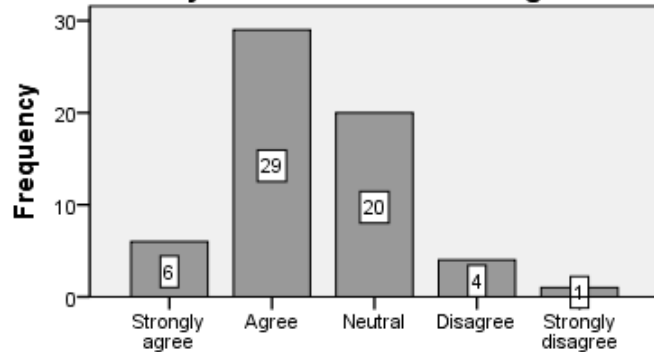


Fig. 11

Item 7, I7 - “The workflow in my organization is dictated by the DVCS.” received 7 (11.7%) responses of “Strongly agree” and 18 (30%) of “Agree”. 15 of the respondents (25%) were neutral to the statement, while 16 (26.7%) expressed disagreement. Furthermore, 4 respondents (6.7%) strongly disagreed.

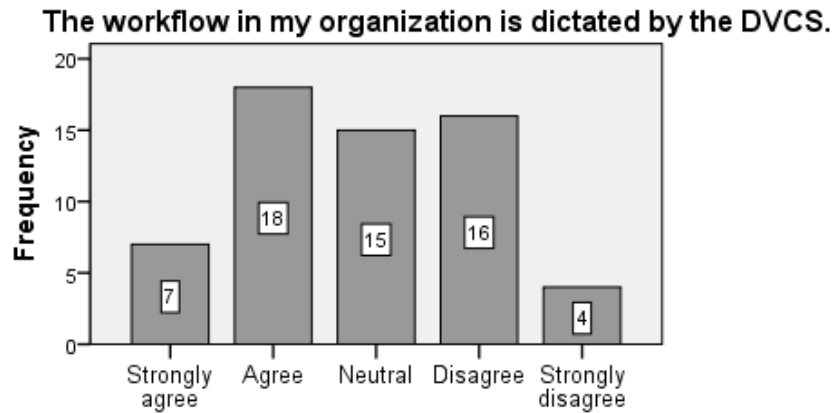


Fig. 12

The following item, I8 - “My coding style in work-related projects is different than in private (non-public) projects.” was rated by 5 respondents (8.3%) as “Strongly agree”. Moreover, 22 respondents (36.7%) expressed agreement and 8 (13.3%) chose the neutral option. Furthermore, 18 (30%) replies of “Disagree” were obtained, as well as 7 (11.7%) of “Strongly disagree”. Figure 13 illustrates the results of I8.

The last Likert-type item from this section of the questionnaire was I9 - “I can easily track down the changes I am looking for in the code by reading the commit messages.” The following responses were obtained: 5 respondents (8.3%) expressed strong agreement, 35 (58.3%) agreed with the statement, 8 (13.3%) were neutral and finally, 12 respondents (20%) disagreed. No strong disagreement was expressed.

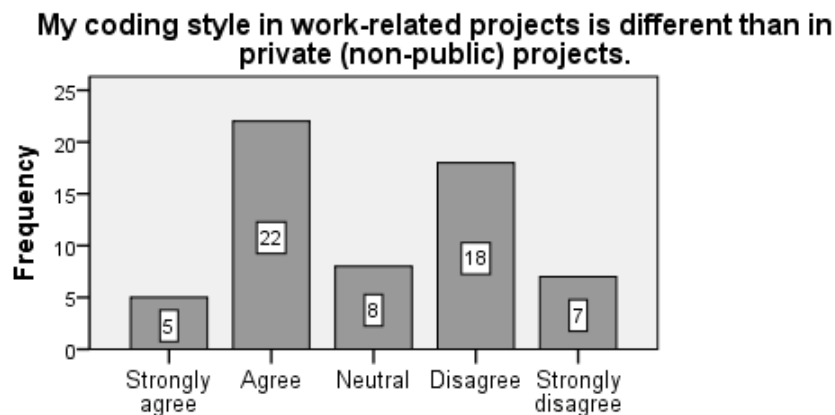


Fig. 13

**I can easily track down the changes I am looking for in the code by reading the commit messages.**

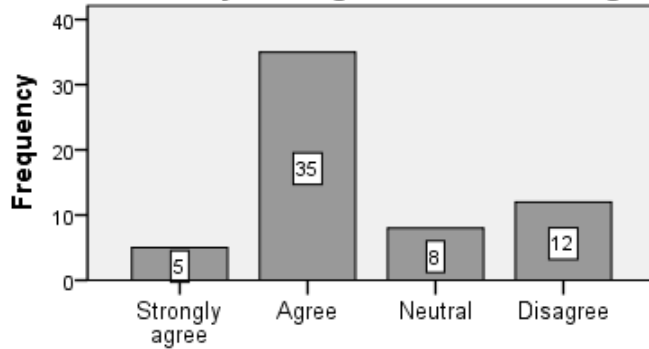


Fig. 14

The final section of the questionnaire consisted of two questions meant at gathering demographic data. In referring to these questions, the numbering from the first section is continued. Thus, question 8, Q8 - “What age-group do you belong to?” has the following distribution of responses: 63.3%, or 38 respondents, indicated they belong to the 26-39 age group, 21.6% (13 respondents) are aged between 18 and 25, 10% (6 respondents) are between 40 and 49 years of age and finally, 5% (3 respondents) are over 50.

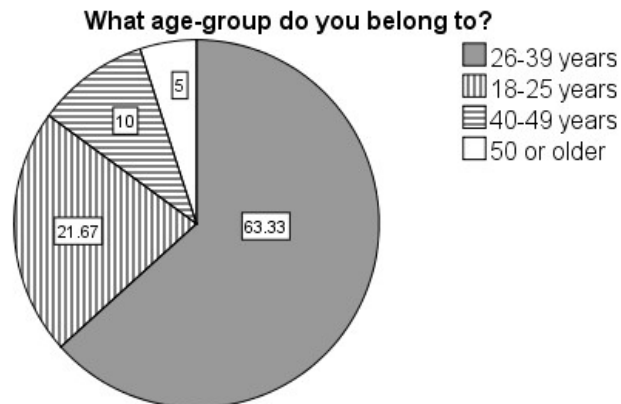
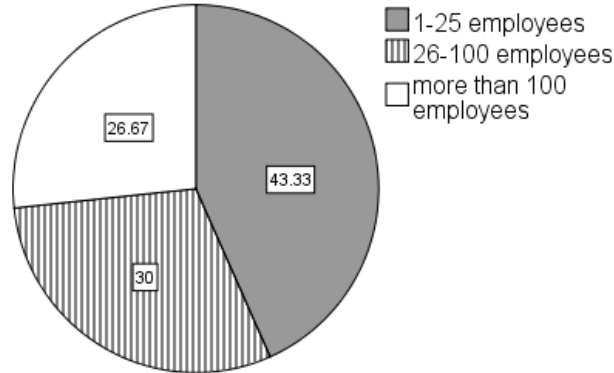


Fig. 15

The questionnaire ended by requiring respondents to approximate the size of the organization they were employed in: Q9 - “What is the size of the organization you are working in?”. The options provided for this question were: “1-25 employees”, “26-100 employees” and “more than 100 employees”. Out of the 60 respondents, 43.3% (26 respondents) revealed they were employed by a small organization, with less than 25 members. Furthermore, 30% (18 respondents) of the replies consisted of the “26-100 employees” option and the remaining 26.6% (16 respondents) of “more than 100 employees”.

What is the size of the organization you are working in?



### 5.3. Analysis of questionnaire data

The current section will provide an in-depth analysis focusing primarily on establishing the connection between the data and the theoretical framework of the project. The section begins with a descriptive analysis of each of the four sections of the questionnaire, focusing on each questionnaire item, in order of appearance in the original layout. The reasoning behind each item will be presented, followed by additional remarks on possible implications of the results with regard to the three main theories which make up the basis of the current research project. The individual descriptive analysis is, whenever relevant, accompanied by a correlation analysis, featuring all of the relevant bivariate correlations which were identified based on the questionnaire data.

To begin with, the first section of the questionnaire inquired with regard to respondents' professional background and experience in the IT field. The four questions which were part of the section had two primary purposes. On the one hand, the questions were intended to **measure respondents' familiarity with the IT field** and their experience of working in IT teams. It was expected that correlations might be identified between the professional experience of the respondents and their attitudes towards and perceived affordances of the DVCS. A secondary purpose of the questions was to **assess the reach of the questionnaire** and whether or not the results obtained might be relevant only to particular subgroups of experienced professionals. The first question Q1 was a check-question, referring to whether the respondent was an IT professional. The following question, Q2 inquired about the working experience respondents had in the IT field. The responses received showed a rather equal distribution of experience among respondents, with the least represented group being that of professionals with under one year of experience. The third question referred to respondents' experience in their current organization. In addition to measuring familiarity with the field, the responses received were meant to check respondent's hypothetical readiness to learn. It was believed that new members of an organization will need to learn new social and organizational rules and as a result might respond more positively to questionnaire items connected to learning and socializing. The obtained responses showed that over 60% of those who responded had been employed for less than a year in their current organization. In itself, the response can be interpreted twofold: on the one hand, it might point towards the fact that the questionnaire data is biased, by one group of respondents being overrepresented.

However, on the other hand, it might simply indicate a natural tendency of IT professionals to change their workplace often. Based on the evenly distributed results obtained from Q2, the latter alternative appears to be more accurate. Finally, Q4 required respondents to identify with a professional role from a given list. The DVCS is a tool which the researcher deems appropriate for usage in teams of professionals whose role involves the creation of textual data, such as documents or programming code. Based on this assumption, it was expected that out of the group of respondents who use a DVCS, a majority of them would hold a position typically associated with team interactions and coding/programming, such as the position of programmer/developer. The results obtained verify the assumption, with close to 80% of the respondents indicating their current position as programmer/developer or web developer.

The second section of the questionnaire was designed to introduce the topic of the DVCS and to gather data on the usage of the technology among the respondents. The primary check question Q5 was meant to filter out all respondents who were not actively using a DVCS for professional purposes. The following question Q6 gathered information on respondents' experience in using the technology under survey. As such, it was expected that the responses obtained would reveal whether a majority of the questionnaire participants were well familiar with the technology. The hypothesis was that users of less than one year might not perceive all of the affordances of the DVCS and as a consequence might provide biased replies. The obtained results showed that a clear majority of the respondents had over one year of experience with the DVCS, with 63.3% having between 1-4 years and a total of 33.3% having over 5 years of experience. A minority of 3.3% indicating less than a year of DVCS usage is interpreted as a reassurance of the validity of the questionnaire results.

The final question of the section, Q7, referred to how often respondents used the technology at the workplace, by requiring them to estimate how often they committed changes to the system. The question evaluated the role of the DVCS in participants' daily work routine. With a detached majority of 50 respondents (83.3%) stating they commit changes several times a day and an additional 5 (8.3%) admitting to regular usage at least once every day, it can be claimed that the DVCS is indeed a highly employed technology among IT professionals.

The following section of the questionnaire consisted of 9 Likert-type items which were designed to measure participants' attitudes towards the DVCS. More specifically, these statements attempted to verify the extent to which the theoretical claims and assumptions identified in the theoretical framework of the project were accurate. Moreover, the items in the section were also intended to evaluate which affordances are generally perceived by users and which action possibilities play a latent role in user interaction. In formulating the statements particular focus lay on the perceived affordances, rather than the objective physical affordances of the DVCS. It was assumed, following Dennis and Fuller (Dennis, Fuller & Valacich, 2008), that physical affordances will be inherent to the technology and thus all users will readily identify them. The characteristics which MRT identifies as relevant for determining the richness of a medium are therefore not evaluated by the questionnaire. The feedback

capability, channel and language variety, as well as the nature of the source can all be labeled as physical properties of the DVCS. On the other hand, the main claim of the theory, namely that richer media is more appropriate for usage in complex organizational tasks, is covered by several items. Similarly, in the case of MST a majority of the capabilities which the theory claims influence communication processes—symbol sets, parallelism and transmission velocity—will not be tested by any particular item in the questionnaire, as they have already been analysed with regard to the DVCS in a previous chapter. However, the theoretical claims of MST are covered by several questionnaire items. Finally, with regard to the theory of adult learning (andragogy), it has already been established that the DVCS is theoretically a suitable technology for learning. The questionnaire items connected to andragogy aim at identifying the perceived potential of the DVCS as a learning medium.

The first Likert-type item in the section, I1 - “I write my commit messages according to a team/organizational standard.“, is connected to Media Richness Theory. The DVCS has been identified as a moderately-rich medium, not as rich as either video conferencing or face-to-face conversations, but above the richness of written letters or numerical information. In such a case, MRT would predict that the technology can be used either for simple, routine tasks or for more complex, creative ones. By inquiring whether respondents have an established standard for commit messages, either on a micro-scale, at team-level, or in the organization as a whole, the routine nature of the tasks in which the technology is used could be inferred. The results were divided, with 50% of the participants expressing agreement and strong agreement and the remaining half being either neutral or in disagreement with the statement. The distribution of responses can be interpreted as a confirmation of the level of richness of the technology, thus supporting the predictions made by MRT. There appear to be organizations or users which perceive the DVCS as leaner, implementing as a result a standard structure for commit messages, while in organizations in which the DVCS is interpreted as a rich medium, more freedom of expression is allowed.

The responses gathered from the following item of the section, I2 - “Summarizing my changes in commit messages helps me reflect on the work I have done.” can be interpreted from two different theoretical stances. Analysed from the MST perspective, the statement evaluates the perceived affordances of rehearsability and reprocessability. The act of summarizing changes is connected to creating, or rehearsing, the message, while by reflecting on previous changes the information is effectively being reprocessed. The statement received a total of 51 positive replies (85%), with a majority of 33 (55%) of the questionnaire participants agreeing and 18 (30%) of the remaining ones strongly agreeing. The fact that only 9 respondents of the total (15%) expressed either neutral or disagreeable attitudes can be said to prove that users perceive the two affordances mentioned previously. The results can be further analysed from an andragogical perspective, as the statement can also refer to learning, more specifically to users’ need to know and their ability to learn from own experience. Thus, by reflecting on an accomplished task, learners can be assumed to reevaluate the benefits and the outcome of their work, while simultaneously reviewing a past experience. As the DVCS has been deemed to have the necessary affordances for promoting adult learning, it was expected that the statement I2 provided positive results. In light of the expressed attitudes towards this

item, the data predictions appear to be in line with the theoretical assumptions made by the theory of andragogy.

The third item of the section is concerned primarily with learning in teams or organizations: I3 - “I use the DVCS to familiarize myself with the coding standards at the workplace.”. The phrase “familiarize myself with the coding standards” refers to the learning of organizational rules which typically occurs to a higher degree in the first phase of becoming an organizational member. It was predicted that a higher rate of agreeability with the statement would indicate respondents’ perceiving the DVCS as a technology for learning. Additionally, by agreeing to the statement, respondents would also indicate the extent to which they perceived the affordance of reprocessability in connection with the DVCS. According to MST, media which possesses this affordance might be used by new participants in an activity in order to understand past activities, norms and rules of the group. Although a majority of the questionnaire participants responded that they agreed with the statement, a considerable number of participants also expressed either disagreement or neutrality. With 26 responses being in agreement (45%), 18 neutral (30%) and a remaining 15 in disagreement (24.7%), no relevant conclusion can be drawn.

The following Likert-type item in the questionnaire used a reverse technique in order to test the proposal, based on MST, that the DVCS is better suited for conveyance purposes rather than for convergence. Thus, I4 - “I write commit messages in order to remind myself or the team what needs to be done next.”, places the DVCS in a context in which the dominant task type would be that of convergence. With a majority of the respondents (25, or 41.7%) indicating disagreement with the statement, it might be suggested that the results follow the expected development, thus supporting the prediction made based on MRT.

Item 5 of the questionnaire refers to both MRT and MST simultaneously. The statement I5 - “Using a DVCS improves my co-operation with the other team members working on the same files.” would be a reverse test for MRT’s prediction that less rich media are not successfully used in complex tasks. On the other hand, I5 may also be claimed to test MST’s proposal that conveyance is one of the primary processes of group co-operation<sup>1</sup>. By analysing the received responses to the item, it is observed that a striking majority of the participants indicate strong agreement (31 respondents, or 51.7%), followed by 20 respondents stating they agree (33.3%). Thus, with a total of 51 (85%) of those questioned expressing positive attitudes towards the claim that the DVCS is beneficial in co-operation processes, the hypothesis of MST appears to be accurate. Furthermore, as respondents declared that the DVCS improves their team-co-operation, it may be concluded that MRT’s predictions were not met. Despite a limited number of communicative cues, the medium is chosen for complex tasks involving co-operation.

Another statement which is intended to measure MST’s predictions is item 6: I6 - “I learn new

---

<sup>1</sup> I6 checks whether the DVCS is suitable for conveyance



implementations, techniques or designs by checking my teammates' code through the DVCS.”. As observed from the content of the statement itself, adult learning is also a parameter analysed through this item. With regard to MST, the item, similarly to I3, checks the perceived affordance of reprocessability. The act of checking others' code may also be an indicator of DVCS's appropriateness for conveyance processes, as it would function as a hub for receiving information. Turning to andragogy, the statement directly inquires the extent to which the respondents perceive the technology as suitable for learning in an organizational setting. The theory predicts that learners' readiness to learn would increase when exposed to superior models, as well as in cases in which the learning outcome is connected to everyday, working situations. Moreover, learning should also be associated with exposure to peers' experience and alternative mental models, all of which are conveyed in I6. Thus, a positive reaction to the statement was expected. The actual results obtained for the item appear to support the theories, as 58.3% of the responses expressed positive attitudes (10% strongly agree and 48.3% agree). It should be mentioned that a relevant number of respondents, namely 20 (33.3%), chose the neutral option for the statement. This does not, however, necessarily indicate that the DVCS is not a learning medium for these respondents; it might be a consequence of the fact that informal, situational learning<sup>2</sup> does not always occur consciously (Eraut, 2000).

Considering the following item, I7 - “The workflow in my organization is dictated by the DVCS.”, the impact of the DVCS was being measured. While the statement does not directly reflect any of the theories considered in the project, it was designed with MRT in mind. The workflow of a team of IT professionals is supposedly established at a higher managerial level, rather than at a lower employee level. Thus, a high level of agreement with the statement would contradict MRT which claims that less rich media is not generally considered at higher levels in the organization. The responses gathered showed a slight tendency towards agreement, with 18 (30%) of the respondents indicating agreement and 7 (11.7%) - strong agreement. However, a comparably large number of responses were either neutral or negative: 15 (25%) neutral, 16 (26.7%) and 4 (6.7%) disagreeing and strongly disagreeing, respectively. These mixed responses can neither support nor dismiss the theoretical claims of MRT.

The eighth item in the current section of the questionnaire focused on the social aspect of mediated-interaction. Statement I8 - “My coding style in work-related projects is different than in private (non-public) projects.” attempted to check whether the fact that the DVCS makes one's code available to an entire team has any effect on the style of coding displayed. Although the results were inconclusive, with 27 (45%) positive responses (agree and strongly disagree), 8 (13.3%) neutral and 25 (41.7%) negative responses (disagree and strongly disagree), running a bi-variate correlation test revealed a .288 correlation with the work experience in the IT field. The correlation can be interpreted to mean that the social aspect of the DVCS is perceived gradually, as users become familiar with affordances of the system. This might be relevant as it appears to indicate that more experienced users may discover new functionalities within the technology.

---

<sup>2</sup> Informal learning is defined as “Learning which takes place in the work context, relates to an individual's performance of their job and/or their employability, and which is not formally organized into a programme or curriculum by the employer.” (Dale & Bell, 1999)

The last Likert-type item on the questionnaire is I9 - “I can easily track down the changes I am looking for in the code by reading the commit messages.”. The statement directly refers to MST and the assumption that the DVCS is a technology best used in conveyance-dominated processes. The act of searching, or tracking down, changes can be interpreted as an act of information seeking. The statement implies that information/data in the system is purposely shared by authors with their peers. As expected, the item received predominantly positive responses, having 66.6% of all questionnaire participants express either agreement or strong agreement. As only a minority of 20% of the respondents indicated disagreement and none of the total disagreed strongly, it can safely be assumed that the DVCS is generally perceived as an information seeking and sharing technology.

Finally, the fourth and final section of the questionnaire gathered demographic data by means of two questions. The purpose of the section was to assess the extent to which the questionnaire responses were affected by situational factors. The two factors which were deemed relevant to consider were the age of the respondents and the size of the organization they were employed in. The age distribution among questionnaire respondents has not been even. Thus, a majority of the participants are aged between 26 and 39 (63.3%), while the remaining age groups appear to be underrepresented - 21.6% aged 40-49, 10% between 18-25 and 5% older than 50. An uneven age distribution does not necessarily imply that the questionnaire obtained biased results. As the target group was IT professionals who used a DVCS, it could be said that the results accurately reflect reality. While the age group 18-25 might be underrepresented due to a higher number of (future) professionals conducting their academic education, the older age-groups might consist of professionals which have advanced in their careers, possibly to managerial positions, thus no longer having the need to use a DVCS. In analysing the responses received for the last question, namely “What is the size of the organization you are working in?”, an even distribution can be observed.

## **5.4 Interviews**

As the questionnaire results presented previously were interpreted in a quantitative manner, an additional data source was chosen in order to obtain a deeper view of the role that the DVCS plays in team communication and learning. Thus, the current section introduces the results of two interviews conducted with IT professionals, with a focus on the DVCS. The interviews were structured based on the same themes as the questionnaire, namely: perceived affordances of the DVCS, team communication as mediated by the DVCS and finally, learning and decision making with the technology. Moreover, the section also interprets the data and draws logical conclusions between the interview participant’s expressed opinions and the theoretical stance adopted by the project. The analysis resorted to the narrative technique, as it was deemed most suited for this type of qualitative data. Furthermore, interview responses are interpreted in order of their appearance in the transcript. Whenever necessary, parallels are drawn between the findings of the interview and those uncovered during the analysis of previous data, namely the commit messages and the questionnaire.

The interview began with a series of screening questions and a security question. The purpose of the screening questions was to establish whether the participants were representative for the target group relevant for the project. The following questions were designed: Q1 - “Could you tell me what work experience you have in the IT field?”, Q2 - “What is your position in the company?” and Q3 - “Are you using a DVCS? How would you evaluate your proficiency in using it?”. Both interview participants provided satisfactory responses to the first question, one of them by indicating an employment history of over one year and the other of approximately 4 years in IT. Considering that these experience groups were also highly represented in the questionnaire, it can safely be claimed that the respondents fit the target group from this perspective. Furthermore, when scrutinized in relation to the position they were holding, the interviewees both mentioned highly creative positions (software developer and software engineer), which mirror the findings of the online questionnaire.

The first question of the main section of the interview was Q4 - “How often would you say you are using the DVCS?”. Both respondents provided similar answers, by stating they were using the DVCS several times a day. The following question was Q5 - “What are the main tasks you use the DVCS for?”. The reasoning behind posing the question was in an attempt to allow the interviewees to reveal the main affordances and capabilities of the technology under analysis. One of the two responses received can be interpreted to mean that the interviewee perceives the DVCS as an information storing and sharing technology: Interviewee 1 indicates this by admitting one of the main tasks he/she performs with the medium is “to add code to the existing code base”. A common code in both cases was the identification of the history and document storing affordance: “I use it for tracking the latest changes that my colleagues do” (Interviewee 1) and “...to keep track of my code, and my co-workers’ (Interviewee 2). By mentioning the information tracking capability of the system, the interviewees’ responses appear to confirm previous claims and findings that the DVCS is suitable for conveyance purposes. In addition to tracking, Interviewee 2 also brings up the affordance of reprocessability, in the form of accessing previously stored information and restoring it: “if I make some change that just doesn’t work I know I can just find an old version that did, and restore it”. The interview finding strengthens the obtained questionnaire results, which pointed towards the same perceived affordance of reprocessability.

A second question related to the perceived affordances of the DVCS (Q6) required interviewees to reflect on three main reasons, or advantages of using the technology at the workplace. Thus, additional perceived affordances were sought after. Although the interview participants each responded to the question differently, several theoretically assumed affordances and properties of the system were mentioned. The first interviewee mentions the ability to reasonably manage code and information in a team of professionals as a core advantage of the DVCS. The response can be interpreted to mean that the technology is appropriate for conveyance processes, thus reinforcing the results of the questionnaire. Another positive remark towards the technology was the fact that it is reliable, which in the context can be viewed as an allusion to the fact that information is permanently stored in the system and can be reprocessed at any time: “it’s reliable, [pause] way better than the alternatives [pause] and it’s a civilized tool for managing source changes in an IT company of more than 1 programmer.” (Interviewee 1). Furthermore, the phrase “way better than the alternatives” which Interviewee 1 uses may be interpreted as an admission of the unique affordances and capabilities of the DVCS. Turning to the second interview participant, the first identified property of the system is the high transmission velocity: “It’s fast, [pause] you can share files without any hassle... and the repository always comes

with me, if I want to change something when I'm at home, let's say". As both MRT and MST refer to velocity as a main factor influencing the adoption and usage of a medium, the response appears to fall in line with their predictions. Starting on a similar theme as the previous interviewee, Interviewee 2 comments on the ease of sharing files with the team, thus discussing the conveyance of information towards cooperation and the affordance of reprocessability. Finally, the second interview participant mentioned the ability to constantly have access to work data, independent of the location. This references an affordance which has been identified in Chapter 3, thus confirming its existence.

Closely related to the previous question, Q7 - "Can you think of something that could be improved with the DVCS?" required interviewees to suggest how the DVCS could be improved. The aim of the question was to identify affordances which the participants perceive as missing or lacking in quality. The first interviewee could only come up with one item, namely the learning curve. According to Interviewee 1, the DVCS has a steep learning curve, which might refer to the multitude of commands available, as well as to the manner of structuring data adopted in version controlled repositories. The response provided by the second interviewee, namely that the DVCS is "as good as its wielder—it's as good as the people who use it", may refer to the fact that not all affordances of the system are readily perceivable to new users. Moreover, the interviewee also stated that "I think there's more room for improvement in teams' usage patterns", which is interpreted to mean that the multitude of affordances and the flexibility of the system, otherwise said, its richness, might not be beneficial in all contexts.

The eighth question was a reflective one, requiring the interviewees to reflect on alternative technologies which might serve a similar function as the DVCS. The question was formulated as follows: Q8 - "If you didn't have access to a DVCS, what alternate technologies would you use to get a similar functionality?" and it was also connected to MRT and user perceived affordances. Thus, by requesting that the interviewees think of alternative technologies which might be used in the absence of a DVCS, the researcher attempted to observe where the technology would be placed on a richness scale. The responses received from both interviewees were similar: "[smile] I guess I'd resort to e-mails and archives, I mean zip-files—it would be tedious..." (Interviewee 1) and "Much of what we do with version control could be done with some manual mailing list, but that would be difficult to follow, [pause] at least for new-comers..."(Interviewee 2), namely that the e-mail could be used as a version control system. Analysed from an MRT perspective, the e-mail is also a rather lean medium, sharing, as explained in Chapter 3, several affordances with the DVCS. The interview responses can therefore be said to confirm the assumed richness level of the technology under survey. Moreover, as both participants expressed concern, or mistrust, with regard to the success of employing e-mail as a replacement for the DVCS, it can be claimed that a perception of the additional affordances of the latter technology exists.

Turning aside from medium affordances, the theme of learning mediated by the DVCS was addressed by the following question of the interview, Q9 - "Can you think of a situation when you learned something through the DVCS? Describe!". Thus, the interviewees were required to recall a situation in which they learned with the help of the DVCS. Neither of the respondents had any difficulty in assessing whether the medium was beneficial for learning or not. Moreover, in support of the hypothesis that the medium is suitable for learning, both interview participants responded positively to the question. Interviewee 1 provided a more vague response, which however does appear to check the predictions made by andragogy: "Plenty! When I'm tracking the evolution of a feature, how it's added in the code base. I can't give you one example, I do this every time I'm tracking a feature". The

assumption that adults are task-centered in their orientation to learning is illustrated by the interviewee stating he/she learns every time a feature needs to be tracked in the work context. Moreover, the interviewee's statement can also be an argument in favor of the assumption that learning readiness is increased when it is necessary to cope with every-day situations (Knowles & Shepherd, 2005). With regard to the second interview response, a more specific learning outcome is identified. The respondent directly references peer-learning, or learning from experience: "I'm learning Angular whenever I have to review some of my team members' code". Furthermore, the statement may also be interpreted to mean that learning can be triggered by being exposed to different conceptual models. Considering the two responses, it is proposed that the theoretical assumptions of andragogy are further verified by the interview.

Nearing the end of the interview, a direct question related to team communication was addressed to the participants. In answering the question "What would you say is the role of the DVCS in team communication?" Interviewee 1 expressed an opinion which appears to reinforce MRT's claims that in situations in which interactants have highly similar frames of reference and are highly interdependent, less rich media should be used to reduce equivocality: "We don't have to waste time describing what and when we did.. whatever changes or features got added, they get recorded into the project's "history" and they can be viewed by anyone. [pause] You don't have to communicate changes on a person-basis...". As the DVCS is considered a less rich medium, it was predicted to be suitable for use at the horizontal level in an organization. In developing the response to the question, Interviewee 1 also addressed several issues pertaining to MST: "You need to communicate a lot of changes to a lot of people, you may become exhausted and they may start forgetting." - the phrase can be interpreted to reference the concepts of transmission velocity, parallelism and reprocessability by stating that the medium permits the communication of multiple pieces of information to several users simultaneously. Similarly to the first interviewee, Interviewee 2 mentions the importance of information sharing and transmission in team communication: "Everyone in the team gets informed when a change is made so they will all know what is the status of project". Moreover, the second participant appears to refer to the conveyance process as a precursor of convergence, by stating that "when you want to discuss a change with your team, you know everybody can participate". The role of conveyance, as presented by MST, is thus strongly reinforced by the two interviewees.

The final question of the interview required participants to reflect on the role, if any, of the DVCS in decision-making in their team or organization. Although the topic of the question was partly covered by previous responses, the aim of the question was to further inquire with regard to co-operation and the predictions made based on the theoretical framework. The two interviewees consented that decision-making is affected by the DVCS, although their interpretations differed. The first interview participant mentioned being able to learn from previous mistakes as a factor which might have an influence on decision-making: if you look in the history and see that a similar feature was added in the past and it didn't work out well, you can avoid making the same mistake twice". Based on its content, the response can be connected to andragogy, as it appears to emphasize the role of learning from experience. Moreover, a reference to the ability to review old information, a vital component of the DVCS, supported by MST, can also be observed in analysing the response. On the other hand, the other participant mentioned access to information and sharing code as an opportunity to build a common ground for decision-making: "when you get together to make a decision it's good if everybody saw the code and knows what's going on [pause] then you take decisions easier".<sup>3</sup>

3 The interview concluded with the participants being offered the opportunity to ask their own questions, as a means to alleviate any possible misunderstandings or concerns. Neither of the two participants chose to enjoy this opportunity.

## 6. Discussion

The current chapter represents the conclusion of the research paper and during its course the interpretations of all the three different data types discussed in previous chapters is combined in a common brief analysis. The main purpose of the current research project has been to observe and analyze usage patterns and role of a chosen technology, namely the distributed version control system. The project has been driven by two research questions, as repeated below:

- 1. What is the role of distributed version control systems in organizational communication, information processing and learning?*
- 2. What strategies might users of the system employ in order to ensure optimal usage of its capabilities, with a maximized learning and communication experience?*

The previous chapters have each represented particular steps in the research process, starting with the exploratory phase during which the research question was formulated, along with the motivation and general introduction of the topic. Moreover, this phase also introduced previous literature and research conducted on the topic of distributed version control systems. The following stage in conducting the study has been the identification of the affordances, or action possibilities of the medium. While also having an exploratory component to it, this chapter served mainly a descriptive purpose, by offering insight into the technology to be analysed, its potential usage, its features and possibilities. Furthermore, this chapter familiarizes the reader with the terminology relating to the DVCS. Having presented the technology in brief, the theoretical framework of the project is outlined in a new chapter. The three theories which make up the theoretical framework are each chosen with consideration to the research questions. Thus, Media Richness Theory and Media Synchronicity Theory are similar, but complementary theories which make predictions on the role a technology may play in communication and information processing in a group or team. The first theory in particular makes reference to the managerial and employee levels in organizations, thus appearing to be directly connected to the chosen research questions. The third theory in the chapter, the theory of andragogy, offers a theoretical basis for observing and analysing the effect the DVCS has on adult learning in an IT team. The chapter following the theoretical framework is concerned instead with the methodological aspect of the project. The chapter introduces methodological concepts relevant for the study and describes in detail the process of data collection and analysis. A preference for triangulation, or the combination of several methodologies, is made apparent in the introductory section of the chapter. Further on, the two types of qualitative methods used in the project - content analysis and semi-structured interview - along with the quantitative method of online questionnaire are presented. The remaining two steps in the research process, namely data presentation and data analysis, or interpretation, are covered in a fifth chapter which precedes the current one. Based on the data gathered through the methods outlined previously, interpretations are made with regard to each data set individually. The analysis is driven by the research questions and is based on the theoretical framework presented previously.

Although the gathered data is thoroughly analysed in the previous chapter, the analysis is not holistic, as it focuses on one type of data at a time. Thus, despite comparing the results of the in-depth interview with those of the online questionnaire, for example, the analysis has not yet fully answered the research questions. The present chapter intends to formulate the answers to the two questions, as a final stage of the research process.

## 6.1. General findings

To begin with, the first question will be answered based on the data gathered and its interpretation:

*1. What is the role of distributed version control systems in organizational communication, information processing and learning?*

One of the first findings of the project was that the DVCS is primarily employed in the process of **conveying, sharing and tracking information** in IT teams. The dominant type of commit message identified in the system was the addition message, having the role of informing team members of a change or introduction of new information in the system. Moreover, both questionnaire responses and interview analyses revealed that reprocessability, or the ability to review previously sent information, was one of the most generally perceived affordances of the technology. The same pattern of information sharing and communication was predicted by MRT and verified by the collected data—as a rather lean medium, the DVCS is useful in contexts in which users have similar frames of reference—in this case in IT teams. Thus, with regard to information sharing, the main role of the DVCS would be to **reduce uncertainty and afford conveyance**.

A second finding of the project was connected to communication mediated by the DVCS. It was observed that the commit message in the DVCS can have several functions, depending on the context. In general, users can write messages with the purpose of **informing** team-members of changes or **recording their progress** and thus showing proof of activity. In addition, users are also observed to **communicate feelings** through the medium, by expressing joy, frustration or gratification. In addition, the DVCS was identified to have a role in team communication by **perceptibly improving co-operation** in work-related tasks. More precisely, users of the DVCS have constant access to their co-workers' changes and work, as well as the entire repository history. As a result, as predicted by the theoretical framework and supported by the data gathering results, co-operation emerges from a combination of the ability to send and receive high volumes of data simultaneously, the affordance of rehearsing a message before sending it to the team and the ability to review and process information at one's own pace, thus developing a deep understanding of it. Furthermore, as the system promotes transparency and openness, it was also discovered to have an influence on decision making, during the information gathering phase.

On the topic of learning, the DVCS can be argued to be one of the main facilitators for learning in the IT team. The data has verified that the technology allows users to learn from their own and others'

experience. Moreover, a readiness to learn is induced when browsing the DVCS and coming upon a mental model which is perceived as better. The results of the questionnaire confirmed that users of the system are to some extent aware of the learning process which can take place whenever they are using the DVCS. Furthermore, the role of the DVCS as a learning tool is readily recognized by the interview participants. Taking into account the data analysis results, it can be concluded that the main role of the DVCS in learning is to **act as a hub for knowledge, facilitating the transmission** and adoption of new techniques, styles or models between members of a team.

Having provided an answer to the first research question, a collection of general indications or recommendations will follow, in order to accurately answer the second research question. These guidelines are meant to serve new users of the DVCS, which might not yet be familiar with its affordances.

*2. What strategies might users of the system employ in order to ensure optimal usage of its capabilities, with a maximized learning and communication experience?*

A first strategy which emerged from the research could be formulated as follows: communication through the DVCS is generally **effective on a horizontal level in an organization**, rather than on a vertical one. As suggested by MRT and further confirmed by the research data, the technology under analysis is most appropriate for usage in an IT team in which all members have similar information needs and common frames of reference. Thus, users of the technology should be aware that trying to communicate through commit messages with interlocutors having highly divergent frames of reference would most likely prove to be ineffective. With regard to communication upward or downward in an organization, it is suggested that a richer medium would be more suitable.

Another guideline which users should consider in order to effectively use the capabilities of the DVCS is to **take into account the readers** when uploading files followed by commit messages to the system. The analysis of the commit messages has identified a recurring practice among several members of the IT team, who were **providing explanations** for the changes made and additional details whenever appropriate. Such a practice could on the one hand facilitate understanding for other team members and on the other, improve the information tracking speed of the users. From a managerial perspective, the introduction of a commit messaging standard in the IT team might have the effects exposed above. Closely connected to the previous guideline is the proposal to **commit changes which are unitary**, accompanied by easy-to-follow messages which accurately and briefly describe the change. As one of the main perceived affordances of the DVCS was tracking the history of data, following this recommendation is believed to simplify the process, thus resulting in faster access to the desired piece of information.



## **7. Limitations and future research**

As can be expected, the present research project is not without its limitations. One of the major limitations arises from the high volume of qualitative data making up the body of the project. Although such data can offer a more nuanced and complete overview of the role the DVCS plays in communication and learning, its findings should be backed up by quantitative, precise data. As the single quantitative method employed in the project, the online questionnaire may not be sufficient to support the analysis. Moreover, the relatively low number questionnaire respondents (64 out of 134 possible respondents) might not afford a high enough certainty coefficient and as a consequence, a replicated study might not come to the same results as the current one. Furthermore, considering the fact that a large amount of the data is qualitative, and thus requires the researcher's analysis rather than the use of objective mathematical systems, a possible limitation of the project might be caused by the author's unintentional subjectivity when drawing conclusions.

Another possible limitation which has been identified is connected to the population in the sampling frame for the online questionnaire. As mentioned previously, only a third of the respondents were estimated to come from IT organizations in the Stockholm area, with the remaining responses coming from IT professionals engaged in social media groups. A consequence of this choice might be that only a particular type of respondents have been taken into consideration, namely those who are active on the social media sites taken into consideration by the researcher.

Having stated some of the possible limitations of the current research study, the opportunities for future research may now be presented. To begin with, the exploratory and descriptive nature of the research performed in the project may be used as a basis for further analytical studies, which might reach a deeper understanding of the role of the DVCS in team communication. Moreover, further research is needed on the effect using a DVCS in a team of professionals has on individual learning, as the current research project did not delve into the topic deeply.

Another suggestion of a possible follow-up study would be an analysis and comparison of commit messages from several different organizations and projects. Such a study might uncover communication differences between IT teams, as well as general patterns of communication in using the DVCS as a mediating technology. Finally, future studies might use the DVCS and the patterns of usage created by the medium in order to test different theories from the field of computer-mediated communication and human-computer interaction.

## References

- Berg, Bruce L. *Qualitative Research Methods for the Social Sciences: International Edition*. London: Pearson, 2008.
- Bhattacharjee, Anol. *Social Science Research: Principles, Methods, and Practices*. Vol. 3. Florida: Global Text Project, 2012. Textbooks Collection.
- "Bug", "Console", "GUI." *Gartner IT Glossary*. Gartner Inc., n.d. Web. 08 June 2014.
- Chacon, Scott. *Pro Git*. N.p.: n.p., 2009.
- Cochez, Michael, Ville Isomöttönen, and Jonne Itkonen. "How Do Computer Science Students Use Distributed Version Control Systems?" *9th International Conference on ICT in Education, Research, and Industrial Applications*. Proc. of ICTERI 2013, Ukraine, Kherson. (2013): 210-28.
- Christensson, Per. "Application.", "Console.", "Log." *PC.net*. N.p., n.d. Web. 08 June 2014.
- Daft, Robert L., and Richard H. Lengel. "Information Richness: A New Approach to Managerial Behavior and Organizational Design." *Research in Organizational Behaviour* 6 (1984): 191-233.
- Daft, Robert L., Richard H. Lengel, and Linda K. Trevino. "Message Equivocality, Media Selection, and Manager Performance: Implications for Information Systems." *MIS Quarterly* 11.3 (1987): 355-66.
- Dale, Margaret, and John Bell. *Informal Learning in the Workplace*. London: Department for Education and Employment, 1999. DfEE Research Report 134.
- Dennis, Alan R., Robert M. Fuller, and Joseph S. Valacich. "Media, Tasks, and Communication Processes: A Theory of Media Synchronicity." *MIS Quarterly* 32.3 (2008): 575-600.
- Denzin, Norman K. *The Research Act*. New York: McGraw-Hill, 1978.
- Dey, Ian. *Qualitative Data Analysis: A User-friendly Guide for Social Scientists*. Oxon: Routledge, 1993.
- DiCicco-Bloom, Barbara, and Benjamin F. Crabtree. "The Qualitative Research Interview." *Medical Education* 40.4 (2006): 314-21.
- Downe-Wamboldt, Barbara. "Content Analysis: Method, Applications, and Issues." *Health Care for*

*Women International* 13.3 (1992): 313-21.

El-Shinnawy, Maha, and Lynne M. Markus. "The Poverty of Media Richness Theory: Explaining People's Choice of Electronic Mail vs. Voice Mail." *International Journal of Human-Computer Studies* 46.4 (1997): 443-67.

Eraut, Michael. "Non-formal Learning, Implicit Learning and Tacit Knowledge in Professional Work." *British Journal of Educational Psychology* 70 (2000): 113-36.

Glassy, Louis. "Using Version Control to Observe Student Software Development Processes." *Journal of Computing Sciences in Colleges* 21.3 (2006): 99-106.

Hutchby, Ian. *Conversation Technology: From the Telephone to the Internet*. USA: Blackwell Publishers, 2001.

Knowles, Malcolm S., Elwood F. Holton, and Richard A. Swanson. *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development*. 6th ed. Oxford: Elsevier, 2005.

Loeliger, Jon. *Version Control with Git*. 1st ed. CA: O'Reilly Media, 2009.

Milentijevic, Ivan, Vladimir Ciric, and Oliver Vojinovic. "Version Control in Project-based Learning." *Computers & Education* 50.4 (2008): 1331-338.

Norman, Donald A. *The Design of Everyday Things*. New York: Basic Books, 1988.

Oliver, Martin. "Technological Determinism in Educational Technology Research: Some Alternative Ways of Thinking about the Relationship between Learning and Technology." *Journal of Computer Assisted Learning* 25.5 (2011): 373-84.

Reid, Karen L., and Gregory Wilson V. "Learning by Doing: Introducing Version Control as a Way to Manage Student Assignments." *Proceedings of the 36th SIGCSE Technical Symposium on Computer Science Education*. SIGCSE 2005, New York. N.p.: n.p. (2005): 272-76.

Rodríguez-Bustos, Christian, and Jairo Aponte. "How Distributed Version Control Systems Impact Open Source Software Projects." *Proceedings of the 9th Working Conference on Mining Software Repositories*, Switzerland, Zurich. (2012): 36-39.

Rouse, Margaret. "Workflow", "Code." *Tech Target*. N.p., Apr. 2005.

Satish, Laika. "Identifying the Dissimilarities Based on Working of Programs among Versions in DVCS (Distributed Version Control Systems)." *International Journal of Computer Applications* 36.6

(2011): 25-29.

Sechrest, Lee, and Souraya Sidani. "Quantitative and Qualitative Methods: Is There an Alternative?" *Evaluation and Program Planning* 18.1 (1995): 77-87.

Shaikh, Maha, and Tony Cornford. "Version Control Tools: A Collaborative Vehicle for Learning in F/OS." *Proceedings of the 4th Workshop on Open Source Software Engineering*. Collaboration, Conflict and Control, Scotland, Edinburgh. (2004): 87-91.

Suh, Kil Soo. "Impact of Communication Medium on Task Performance and Satisfaction: An Examination of Media-richness Theory." *Information & Management* 35.5 (1999): 295-312.

Treadwell, Donald. *Introducing Communication Research: Paths of Inquiry*. 2nd ed. California: SAGE Publications, 2014.

Whittaker, Steve. *Theories and Methods in Mediated Communication*. New Jersey: AT&T Labs-Research, 2003.

## Glossary of IT terms

**application (software):** software program that runs on a computer or similar computing device (mobile phone, tablet etc.). “The word "application" is used because each program has a specific application for the user.” (PC.net Glossary)

**bug:** “unexpected problem with software or hardware. Typical problems are often the result of external interference with the program’s performance that was not anticipated by the developer.” (Gartner IT Glossary)

**code (noun):** “term used for both the statements written in a particular programming language - the source code, and a term for the source code after it has been processed by a compiler and made ready to run in the computer [...]. To code (verb) is to write programming statements - that is, to write the source code for a program.” (Rouse, 2005)

**coding style:** “a set of rules that a programmer uses for choosing an expressive form to use in a given situation. Usually these rules are aesthetic, but sometimes there are efficiency issues involved” (MIT gnu.org)

**console:** “The part of a computer used for communicating between the user and the system.” (Gartner IT Glossary); also, a text-based interface used to send commands to the system (PC.net Glossary)

**development (software):** “the act or process of creating something [a piece of software] over a period of time” (Mirriam-Webster dictionary)

**Git:** a type of distributed version control system

**graphical user interface:** “graphics-based operating system interface that uses icons, menus and a mouse (to click on the icon or pull down the menus) to manage interaction with the system. ” (Gartner IT Glossary)

**log (verb):** to keep track of certain events (PC.net Glossary)

**workflow:** “term used to describe the tasks, procedural steps, organizations or people involved, required input and output information, and tools needed for each step in a business process.” (Rouse, 2005)

## Appendix 1: Commit message data

The nature of the data does not permit it to be appended to the thesis in its entirety. Below are two samples of the data, for illustrative purposes. In order to access the full commit message repository, visit the following link:

<https://github.com/balanced/balanced-api>

2014-05-05 18:57:31 +0000	SK	Remove requirement for events to exist.	Sometimes, there will be zero events.
2014-05-05 10:07:29 -0700	SK	Merge pull request #590 from EPSILON/patch-1	Change endpoing with endpoint
2014-05-02 16:49:12 -0500	A	Change endpoing with endpoint	Is this a typo?
2014-04-30 21:16:31 +0000	SK	Remove extraneous @focus.	Good job, past me. :frowning:
2014-04-30 21:05:14 +0000	SK	Regression spec for #572	
2014-04-30 13:50:21 -0700	PC	Display request id in test message	
2014-04-29 17:32:57 -0700	PC	Merge pull request #582 from balanced/disputes-under-cards-customers	update schemas for cards and customers
2014-04-28 17:49:11 -0600	BM	Set travis-ci badge branch to master	
2014-04-28 11:59:47 -0700	PC	update schemas for cards and customers	
2014-04-10 20:09:00 +0000	SK	Make master be the current version.	
2014-04-10 20:07:57 +0000	SK	Nuke old version.	
2014-04-07 09:36:23 -0700	SK	Merge pull request #533 from matthewfl/add-marketplaces-schema	adding marketplaces schema validation
2014-04-03 14:10:03 -0700	MFL	Merge remote-tracking branch 'origin/revision1' into add-marketplaces-schema	
2014-04-03 14:09:00 -0700	MFL	Merge pull request #556 from balanced/clarify_appears_on_statement_as	Fixes balanced/balanced-docs#335
2014-04-03 14:05:22 -0700	MFL	make it wait up to 3 minutes for the dispute to come through	
2014-04-02 23:29:55 -0700	MFL	mark card scenario as failing, due to issue #552	
2014-04-01 18:04:42 -0700	MFL	fixing json issues	
2014-03-24 15:40:59 -0700	MFL	spaces	
2014-03-18 09:02:38 -0700	MJ	Merge branch 'revision1' into creation	Conflicts: features/checkout_flow.feature, Features/credits.feature
2014-03-14 17:15:44 -0700	MFL	adding marketplaces schema validation	
2014-03-14 13:45:08 -0700	MFL	fixing appears_on_statement_as for credits	
2014-03-14 11:13:18 -0700	SK	Merge pull request #532 from matthewfl/update-appears-on-statement-as	add full description to appears on statement as
2014-03-13 12:38:29 -0700	MFL	add full description to appears on statement as love long strings in json ;)	
2014-03-11 17:00:07 -0700	MFL	some minor changes	
2014-03-11 16:32:03 -0700	MFL	Merge branch 'fix-docs-forms2' into revision1	
2014-03-11 16:09:48 -0700	MFL	adding descriptions to schemas for docs, and changing scenarios to demonstrate more r	
2014-03-11 12:27:27 -0700	MFL	Merge pull request #524 from matthewfl/fix-docs-forms	fixing forms description for the docs and external account fi

## Appendix 2: Online Questionnaire

### Section 1

Q1: Are you currently employed in the IT or Computer Science field?

Q2: How long have you been employed in the IT field?

Q3: How long have you been employed in your current organization?

Q4: What title describes your current position inside the organization most accurately?

### Section 2

Q5: Are you currently using a distributed version control system for work-related projects? (ex. Git, BitKeeper, Mercurial)

Q6: How long have you been using a DVCS?

Q7: How often do you commit changes to your work projects?

### Section 3

I1: I write my commit messages according to a team/organizational standard.

I2: Summarizing my changes in commit messages helps me reflect on the work I have done.

I3: I use the DVCS to familiarize myself with the coding standards at the workplace.

I4: I write commit messages in order to remind myself or the team what needs to be done next.

I5: Using a DVCS improves my co-operation with the other team members working on the same files.

I6: I learn new implementations, techniques or designs by checking my teammates' code through the DVCS.

I7: The workflow in my organization is dictated by the DVCS.

I8: My coding style in work-related projects is different than in private (non-public) projects.

I9: I can easily track down the changes I am looking for in the code by reading the commit messages.

### Section 4

Q8: What age-group do you belong to?

Q9: What is the size of the organization you are working in?

## Appendix 3: Interview transcripts

### Interview 1

Interview Participants	Message Transcription
Researcher	Hello! Could you tell me what work experience you have in the IT field?
Interviewee	I've been working in IT for a year and a bit
Researcher	What is your position in the company?
Interviewee	I'm a software developer
Researcher	Are you using a DVCS? How would you evaluate your proficiency in using it?
Interviewee	Yes, I'm using Git. I'm fairly acquainted with it
Researcher	How often would you say you are using the DVCS?
Interviewee	Several times a day
Researcher	What are the main tasks you use the DVCS for?
Interviewee	I use it for tracking the latest changes that my colleagues do [pause] and ultimately using it as a means to add code to the existing code base.
Researcher	What are the best things about using a DVCS at work?
Interviewee	Best thing about using a DVCS?! [hmm] it's reliable, [pause] way better than the alternatives [pause] and it's a civilized tool for managing source changes in an IT company of more than 1 programmer.
Researcher	Can you think of something that could be improved with the DVCS?
Interviewee	[hesitation] A leaner learning curve
Researcher	If you didn't have access to a DVCS, what alternate technologies would you use to get a similar functionality?
Interviewee	[smile] I guess I'd resort to e-mails and archives, I mean zip-files—it would be tedious...
Researcher	Can you think of a situation when you learned something through the DVCS? Describe!
Interviewee	Plenty! When I'm tracking the evolution of a feature, how it's added in the code base. I can't give you one example, I do this every time I'm tracking a feature
Researcher	What would you say is the role of the DVCS in team communication?



Interviewee	[hmm] We don't have to waste time describing what and when we did.. whatever changes or features got added, they get recorded into the project's "history" and they can be viewed by anyone. [pause] You don't have to communicate changes on a person-basis...
Researcher	Can you elaborate on that?
Interviewee	You need to communicate a lot of changes to a lot of people, you may become exhausted and they may start forgetting. Verbal communication—think storytelling—for this sort of thing is not ideal.
Researcher	Does the DVCS play any role in decision making in your team?
Interviewee	With decision making... [pause] having a history of changes—if you look in the history and see that a similar feature was added in the past and it didn't work out well, you can avoid making the same mistake twice
Researcher	Thank you for your time! I don't have any other questions, if you would like to ask me something, feel free to do so now.

## Interview 2

<b>Interview Participants</b>	<b>Message Transcription</b>
Researcher	Hello! Could you tell me what work experience you have in the IT field?
Interviewee	Yes, I'm in IT since [hmm] almost 4 years now, I think...
Researcher	What is your position in the company?
Interviewee	Software engineer
Researcher	Are you using a DVCS? How would you evaluate your proficiency in using it?
Interviewee	Yes, sure, I'm a normal user, I know enough to get things done...
Researcher	How often would you say you are using the DVCS?
Interviewee	I commit changes every day, [pause] maybe on average 4-5 changes a day
Researcher	What are the main tasks you use the DVCS for?
Interviewee	...to keep track of my code, and my co-workers', [pause] if I make some change that just doesn't work I know I can just find an old version that did, and restore it
Researcher	What are the best things about using a DVCS at work?

Interviewee	It's fast, [pause] you can share files without any hassle... and the repository always comes with me, if I want to change something when I'm at home, let's say
Researcher	Can you think of something that could be improved with the DVCS?
Interviewee	As any other tool, it's as good as its wielder—it's as good as the people who use it... I think there's more room for improvement in teams' usage patterns
Researcher	If you didn't have access to a DVCS, what alternate technologies would you use to get a similar functionality?
Interviewee	Much of what we do with version control could be done with some manual mailing list, but that would be difficult to follow, [pause] at least for newcomers...
Researcher	Can you think of a situation when you learned something through the DVCS? Describe!
Interviewee	There are many things I've learned through version control... what comes to mind now is how I'm learning Angular whenever I have to review some of my team members' code
Researcher	What would you say is the role of the DVCS in team communication?
Interviewee	Everyone in the team gets informed when a change is made so they will all know what is the status of the project... Then when you want to discuss a change with your team, you know everybody can participate and that's good for communication
Researcher	Does the DVCS play any role in decision making in your team?
Interviewee	Well, as I said, when you get together to make a decision it's good if everybody saw the code and knows what's going on [pause] then you take decisions easier
Researcher	Thank you for your time! I don't have any other questions, if you would like to ask me something, feel free to do so now.