



NTNU – Trondheim
Norwegian University of
Science and Technology

Effective e-procurement implementation in the public sector

*A framework covering critical success factors
and project phases*

Andrè Bernhard Bakland
Eyvind Sæhl Kilvik

Industrial Economics and Technology Management

Submission date: June 2015

Supervisor: Luitzen de Boer, IØT

Norwegian University of Science and Technology
Department of Industrial Economics and Technology Management

Problem formulation

In this master's thesis we will employ the framework developed in our pre-diploma thesis as the point of departure. The aim is to validate the framework through empirical evidence, and assess its relevance for end-users with purchasing rights. The empirical evidence will be gathered using a survey deployed to the case company, St. Olav's Hospital. Therefore, this thesis will be of a quantitative nature and employ a research design appropriate for this research strategy.

Preface

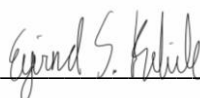
This master's thesis is the result of a combination between TIØ4911 Strategic Purchasing and Supply Management and TIØ4950 Strategic Change Management at the Department of Industrial Economics and Technology Management at NTNU. The thesis was written during the Spring of 2015.

The purpose of this thesis is to develop a survey in order to assess the relevance of a framework covering critical success factors for the different phases of an electronic procurement implementation project. The framework was developed in the pre-diploma thesis, and forms the theoretical basis for this thesis. The survey is deployed to 803 end-users at the case company, St. Olav's Hospital.

First, we wish to thank our supervisor Luitzen de Boer at the Department of Industrial Economics and Technology Management at NTNU for support and guidance throughout the work with this thesis. In addition, we wish to thank Kyrre Svarva, a senior advisor at the Faculty of Social Sciences and Technology Management at NTNU, for guiding us through the statistical tests and results. We would also like to express our gratitude to Bjørn R. Albrigtsen at Helse Midt-Norge IT for introducing us to the case company and assisting us in defining the scope of the thesis. Moreover, we would like to thank Baard Ervik, a change manager at St. Olav's, for introducing us to the sample of end-users and helping us understand the context of the implementation.

Trondheim, June 10, 2015

Authors,



André Bernhard Bakland and Eyvind Sæhl Kilvik

Abstract

Currently there is a trend towards increased focus on the importance of the purchasing function and costs of maintenance, repair, and operation (MRO) goods. This has led to companies investing in electronic procurement tools in order to reduce those costs and increase efficiency. However, studies have shown that only 20-27% of these investments are successful. To ensure the success of such investments, it is beneficial to consider critical success factors (CSFs) covering important aspects of the implementation.

The aim of this master's thesis is to validate and assess the relevance of a framework covering eleven CSFs for three distinctive project phases of an electronic procurement implementation. The main research question is "*Are different CSFs more or less relevant for different types of end-users in an e-procurement implementation context?*", but is limited to some parts of the framework considering its comprehensiveness. The purpose of the framework is to aid managers in decision-making by presenting a framework with clear managerial implications. In addition, the framework is also intended to assist the ongoing implementation of an electronic ordering system at the case company, St Olav's Hospital, to which both this master's and our pre-diploma thesis were written in collaboration with. The framework in focus was developed by us in our pre-diploma thesis, but although looking promising, it needed further validation. This master's thesis will continue the work by employing a case study survey research design and quantitatively analyzing the results of a survey deployed to 803 end-users at the case company, capturing the end-user perspective of the implementation.

The findings of the thesis show that some of the investigated CSFs are indeed important to the sample of end-users, and that there are differences, especially regarding sex and age, in how important end-users consider the CSFs to be. For example, training is found to be more important for both females and for older respondents, while communication is found to be important for all end-users. Moreover, an exploratory factor analysis suggests that there is room for improving the framework, by finding that the perceptions towards electronic systems are important to consider. Furthermore, these findings imply that companies should not only consider the CSFs in the framework, but also the differences that may exist between end-users.

Sammendrag

Det er for tiden en trend mot økt fokus på viktigheten av innkjøpsfunksjonen og kostnader for varer tilknyttet vedlikehold, reparasjon og drift (MRO varer). Dette har ledet selskaper til å investere i elektroniske innkjøpsverktøy for å redusere disse kostnadene og for å øke effektiviteten av kjøp. Likevel viser studier at bare 20-27% av disse investeringene er vellykkede. For å sikre at slike investeringer blir vellykkede, er det fordelaktig å vurdere såkalte kritiske suksessfaktorer (KSFer) som dekker viktige aspekter av gjennomføringen.

Målet med denne masteroppgaven er å validere og vurdere relevansen av et rammeverk som dekker elleve KSFer for tre karakteristiske prosjektfaser i implementeringen av et elektronisk innkjøpssystem. Forskningsspørsmålet er som følger: "*Er forskjellige KSFer mer eller mindre relevante for ulike typer sluttbrukere i implementeringen av et elektronisk innkjøpssystem?*", males er begrenset til enkelte deler av rammeverket siden det er omfattende. Hensikten med rammeverket er å hjelpe ledere i beslutningsprosesser ved å presentere et rammeverk med klare implikasjoner. I tillegg er rammeverket også ment å hjelpe den pågående implementeringen av et elektronisk bestillingssystem hos casebedriften, St. Olavs Hospital, som både denne master- og prosjektoppgaven er skrevet i samarbeid med. Rammeverket ble utviklet av oss i vår prosjektoppgave, men selv om det så lovende ut, var det behov for ytterligere validering. Denne masteroppgaven vil fortsette arbeidet ved å anvende et forskningsdesign kalt "case study survey research", og kvantitativt analysere dataene fra en spørreundersøkelse distribuert til 803 sluttbrukere i casebedriften for å fange sluttbrukernes perspektiv av gjennomføringen.

Funnene i denne oppgaven viser at noen av de undersøkte KSFene faktisk er viktige for utvalget av sluttbrukere, og at det er forskjeller, spesielt på tvers av kjønn og alder, for hvor viktige sluttbrukerne vurderer KSFene til å være. For eksempel viste sluttbrukertrening seg å være viktigere for både kvinner og for de eldre respondentene, mens kommunikasjon ble funnet å være viktig for alle sluttbrukere. En eksplorativ faktoranalyse viste at rammeverket har et forbedringspotensial da oppfatningene av elektroniske systemer er viktig å vurdere. Videre impliserer disse funnene at selskap ikke bare bør vurdere KSFene fra rammeverket, men også forskjellene som kan eksistere mellom sluttbrukere.

Innhold

1.	Introduction.....	1
1.1	Case company - St. Olav's Hospital.....	2
1.2	Continuation of pre-diploma thesis.....	5
1.3	Research questions.....	6
1.4	Limitations of the thesis.....	8
1.5	Structure of the thesis.....	9
2.	Theory, framework, and hypotheses.....	10
2.1	Methodology underlying the framework.....	10
2.2	Purchasing and E-procurement research.....	12
2.3	Presentation of the framework.....	14
2.4	Change management and project phases.....	15
2.5	Critical success factors in the framework.....	18
2.5.1	End-user uptake and training.....	21
2.5.2	Supplier adoption.....	23
2.5.3	Business case and project management.....	23
2.5.4	System integration.....	24
2.5.5	Security and authentication.....	24
2.5.6	Re-engineering the process.....	24
2.5.7	Performance measurement.....	25
2.5.8	Top management support.....	25
2.5.9	Change management.....	25
2.5.10	E-procurement implementation strategy.....	26
2.5.11	Technology standards.....	26
2.6	Hypotheses.....	27
2.7	Summary.....	30
3.	Methodology.....	31
3.1	Research design and research questions.....	31
3.2	Research Method.....	33
3.2.1	Survey development.....	33
3.2.2	Response bias.....	40
3.2.3	Conducting the survey.....	40
3.3	Quality criteria.....	46

3.3.1	Reliability	46
3.3.2	Validity	47
3.3.3	Replication	49
3.4	Statistical analysis.....	50
3.4.1	Data analysis	50
3.4.2	Alternative data analysis	52
3.5	Summary.....	54
4.	Results.....	55
4.1	Initial considerations.....	55
4.2	Demographic variables	57
4.3	Characteristics of the multiple-item indicators.....	61
4.4	Summary.....	63
5.	Analysis.....	64
5.1	Exploratory factor analysis	64
5.1.1	Prerequisites of EFA	64
5.1.2	Initial tests	65
5.1.3	Factor extraction.....	67
5.1.4	Appropriateness of the seven factors from the EFA	71
5.2	Analysis methods.....	75
5.3	Testing the hypotheses.....	78
5.3.1	Hypothesis 1	78
5.3.2	Hypothesis 2, 3, and 4	79
5.3.3	Hypothesis 5	83
5.3.4	Hypothesis 6 and 7	84
5.3.5	Hypothesis 8.....	84
5.3.6	Aggregated results of hypothesis testing.....	86
5.4	Other tests	87
5.4.1	Other factors from the EFA.....	87
5.4.2	The importance of the factors for the end-users.....	89
5.4.3	Correlations between the factors	91
5.5	Summary.....	92
6.	Discussion.....	93
6.1	Research questions	96

6.1.1	Implications for the framework.....	99
6.1.2	Implications for St. Olav's	101
6.2	Generalizability	104
6.2.1	Quality criteria affecting generalizability	104
6.2.2	Generalizing to theory	105
6.2.3	Generalizing to the larger population.....	106
6.2.4	Generalizing across populations.....	106
7.	Conclusion	108
8.	Limitations and further research	110
9.	Bibliography	112
10.	Appendices	117
	Appendix A.....	117
	Appendix B	125
	Appendix C.....	127

List of tables

Table 1	– Critical success factors with attributes.....	20
Table 2	– CSFs and attributes to be investigated.....	27
Table 3	– Investigated CSFs (a plus sign means fulfilled and a minus sign otherwise).....	37
Table 4	– Sex distribution of three target groups.....	56
Table 5	– Sex distribution of the respondents by the three target groups.....	56
Table 6	– Distribution of sex	57
Table 7	– Distribution of age	57
Table 8	– Distribution of Visma usage	57
Table 9	– Percentage of orders through Visma.....	58
Table 10	– Respondents projected to use Visma in the future.....	58
Table 11	– Other ordering channels than Visma	59
Table 12	– Reasons for not ordering through Visma.....	59
Table 13	– The four groups by the variable vGroup.....	60
Table 14	– Internal reliability for survey attributes	62
Table 15	– KMO and Bartlett's Test	65
Table 16	– Total Variance Explained - Extraction Method: Principal Axis Factoring.	66
Table 17	– Factor extraction sorted by size	68
Table 18	– KMO and Bartlett's Test for the final factor structure.....	70
Table 19	– Factor Correlation Matrix	70
Table 20	– EFA: End-user training.....	71
Table 21	– EFA: Attitude and perceptions towards EP	72
Table 22	– EFA: Communication to key stakeholders	72

Table 23 – EFA: Supplier relationships	73
Table 24 – EFA: Status quo	73
Table 25 – EFA: Communication	74
Table 26 – EFA: In-house training	74
Table 27 – Named EFA factors	75
Table 28 – Testing of hypothesis H2a, H3a, and H4a	80
Table 29 – Testing of hypothesis H2, H3 and H4 parts b) + c)	82
Table 30 – Testing of hypothesis H5	83
Table 31 – Testing of hypothesis H6 and H7	84
Table 32 – Testing of hypothesis H8 part a) and b).....	85
Table 33 – Aggregated results of hypothesis testing	86
Table 34 – Testing for IHTs effect on the grouping variables.....	88
Table 35 – Testing for EPPERCP’s effect on the grouping variables	89
Table 36 – Descriptives one-sample t-test for dependent variables	90
Table 37 – One-sample t-test with test value 3.5 for dependent variables	90
Table 38 – Spearman’s rho correlation for EFA factors.....	91
Table 39 – The survey.....	124
Table 40 – Descriptive results of the survey.....	126
Table 41 – Pairwise t-tests dependent variables	128

List of figures

Figure 1 – Actors and ordering channels at St. Olav’s	4
Figure 2 – Hypothesized groups of end-users.....	7
Figure 3 – Approach to developing the framework	11
Figure 4 – General procurement process adapted from (Van Weele, 2005, p. 34)	13
Figure 5 – The framework	14
Figure 6 – Methodology chapter overview	31
Figure 7 – Development of question items for the survey.....	37
Figure 8 – Number of individuals in each target group.....	43
Figure 9 – Gantt chart of the survey deployment procedure	43
Figure 10 – Scree plot for factor extraction	66
Figure 11 – Actual distribution of user groups	78
Figure 12 – Process for addressing the main research question	96
Figure 13 – EFA factors’ relation with the framework.....	101

1. Introduction

“A \$500 million manufacturer of steel products rolled out an e-Procurement pilot project to its plants. The company did not have a change management plan or an overarching procurement strategy. Employees resisted the use of unfamiliar tools and the introduction of revised buying procedures. Despite later mandates by company executives, the lack of early focus on change management resulted in poor adoption and therefore no substantial savings.”

ICG Commerce – reasons e-procurement projects fail to achieve their ROI, white paper (ICGCommerce, 2009).

Electronic procurement (hereafter called EP or e-procurement) refers to the use of internet technology in purchasing processes (De Boer et al., 2002). Purchasing can be defined as externally obtaining all goods, services, capabilities, and knowledge which are necessary for running, maintaining, and managing the company’s primary and support activities (Weele, 2005).

Companies both privately and publicly held are increasingly realizing the potential benefits of e-procurement (Croom, 2000, De Boer et al., 2002, Puschmann and Alt, 2005, Reunis et al., 2006, Vaidya et al., 2006). In fact, public sector organizations worldwide have identified e-procurement as a priority and have implemented or are in the process of implementing e-procurement (Vaidya et al., 2006). Previtali (2012) finds that the public procurement of goods and services are strategically important for at least three reasons: 1) The economic impact is between 15-20% of the GDP of European countries. 2) It is relevant for potential improvements in public services. 3) It affects both the competitiveness of nations and the welfare of citizens.

Procurement usually represents one of the largest expense items in a firm’s cost structure (Angeles and Nath, 2007). The Aberdeen Group (2001) finds that the purchase of maintenance, repair, and operations (MRO) goods usually account for 30-60% of a firm’s total spending. MRO goods are products such as office supplies, personal computers, non-manufacturing items, etc. (Angeles and Nath, 2007). Furthermore, Turban et al. (2006) find that corporate buyers tend to waste time on non-value adding activities such as data entry, expediting delivery and correcting errors in paperwork.

The possible benefits from an e-procurement system are, for instance, shorter order cycles, higher compliance with purchasing contracts, and increased negotiation leverage leading to lower costs (Arbin, 2008b, Croom and Brandon-Jones, 2007, Puschmann and Alt, 2005, Reunis et al., 2006). However, such systems are not always implemented successfully. A study by

Boston Consulting Group in 2001 found that only 20% of all e-procurement investments were successful (Caniato et al., 2012). A similar study by Calyptus Consulting Group in 2009 reported this value to be 27% (Caniato et al., 2012). As illustrated by the introductory example as well, it is important to understand what drives a successful implementation in order to reap the gains from investments in e-procurement initiatives. One such driver is adoption of the system by the end-users. Adoption of the system is critical, and without, the gains can evaporate (Arbin, 2008b, Reunis et al., 2006).

1.1 Case company - St. Olav's Hospital

This master's thesis is written for NTNU (Norwegian University of Science and Technology in Trondheim) in collaboration with St. Olav's Hospital. St. Olav's Hospital (hereafter called St. Olav's), the University Hospital for Mid-Norway, and NTNU are integrated. Patient treatment, research, and education are such integrated functions. The hospital covers the population of Sør-Trøndelag with 302,000 inhabitants, but has several regional and national tasks for the population in the three counties of Møre and Romsdal, Sør-Trøndelag and Nord-Trøndelag with a total of 695,000 inhabitants (St.OlavsHospital, 2014a).

St. Olav's is currently in the process of upgrading the electronic ordering system for MRO goods. Electronic ordering (hereafter called e-ordering) is a subset of e-procurement, and can be defined as the use of internet to facilitate operational purchasing processes, including the ordering (requisitioning), order approval, order receipt, and payment process (Reunis, 2007).

Some of the main reasons for St. Olav's upgrading the system are: 1) To conform with Norwegian law, called the Public Procurement Act, requiring that all public purchases are done through supplier contracts. 2) The hospital hopes to reduce maverick buying (off-contract buying) and purchasing costs while increasing the quality of the procurement process. 3) The project is a preface before implementing a new enterprise resource planning (ERP) system in April 2016, and should help the organization to attain valuable insights (St.OlavsHospital, 2014b). The topic of this thesis is of high interest to the case company in order to learn from this introductory project and understand what the critical factors for a successful implementation are.

The project initiated in 2013 as part of a project to reduce costs of goods. During 2012, St. Olav's had a considerable increase in cost of goods purchased, both compared to earlier years and to the budget for 2012. The deviation from the budget was complex, but partially because of the lack of supplier contracts, maverick buying, and the lack of a system to identify current

supplier contracts. The purchasing system was out of date and allowed for little transparency. In addition, many of the routines were manual, and usage of electronic orders was low (Project documents, 2014). To clarify, the e-ordering system at St. Olav's is called Visma Enterprise (hereafter called Visma). This system was already in place before this project, but there were some shortcomings as mentioned above. The new upgrade is implemented in the same system, but will allow new features. Among these, a standardized access point for connecting suppliers and buyers, allowing catalogues from suppliers to be imported to the buyer (DIFI, 2014). In addition, this will allow for standardized ordering processes, where end-users can buy directly from suppliers through the system (Project documents, 2014).

Following is a list of the main groups of actors affected by the new system, briefly summarizing their main tasks and some of the planned changes (Project documents, 2014):

- Central purchasing division: This division handles contract agreements and is responsible for entering and updating the product and supplier agreements in the system. The upgrade will automate some of the previously manual routines and make their work easier.
- Central warehouse: This warehouse receives incoming goods from suppliers and dispatches these to the different departments and end-users. Prior to the upgrade, the warehouse received orders from the end-users via Visma, phone, email, or fax. Orders through phone, email, and fax led to unnecessary manual work, such as checking inventory and matching orders with suppliers. The intention of the upgrade is to ensure that orders are transferred directly to the supplier through the e-ordering system, thus reducing unnecessary work at the warehouse.
- Suppliers: Prior to the upgrade, suppliers could receive orders directly from end-users by either email, phone or fax (potential maverick buying), or from the central warehouse. The upgrade will allow suppliers to receive orders directly from end-users through the e-ordering system, upholding contract agreements.
- End-users: Approximately 800 end-users, located around at different departments at St. Olav's, purchase MRO-goods. These users range from auxiliary nurses and cleaning personnel to physicians, where ordering is often a small part of their duties. Some of these users order via phone, email and fax, directly to suppliers, while others order to the warehouse via Visma or other channels. The project will introduce the system to users not previously familiar with Visma, as well as preventing orders via other channels by enforcing Visma.

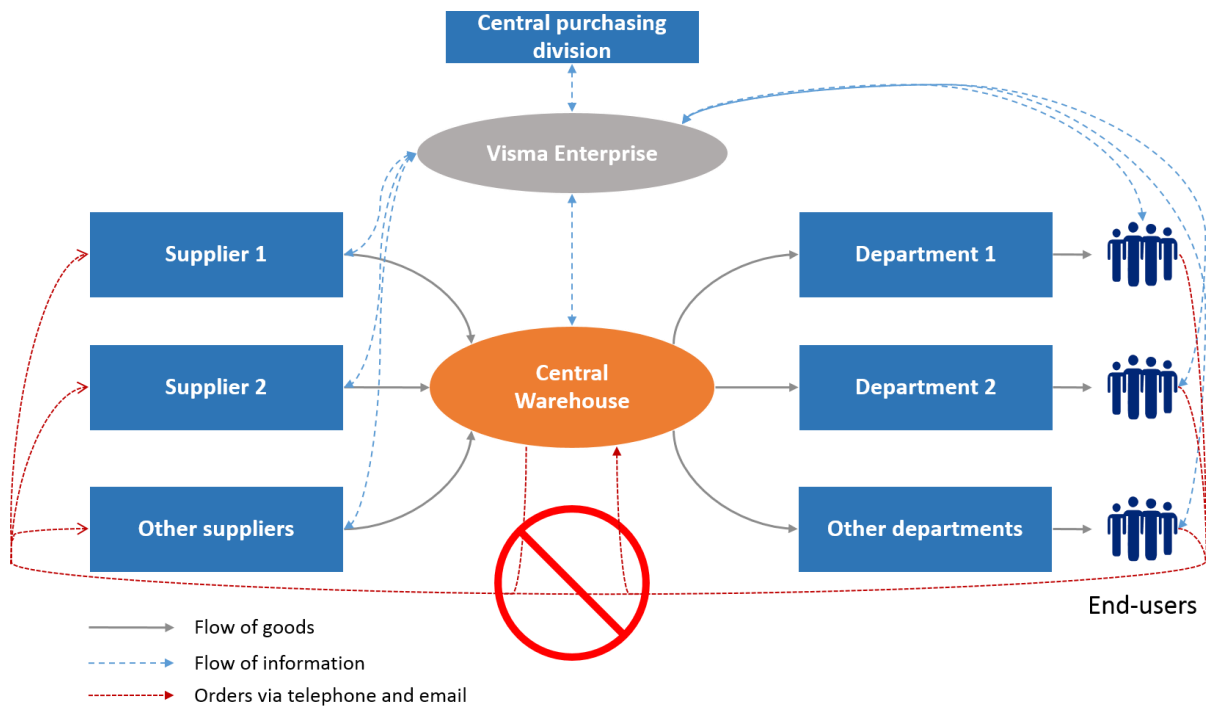


Figure 1 – Actors and ordering channels at St. Olav's

As discussed, the upgrade of the e-ordering system affects multiple actors to a large degree. Furthermore, many end-users have not previously ordered through Visma, but via phone, email, and fax directly to suppliers. As shown in Figure 1, it will not be possible to do so anymore when the e-ordering system is enforced. Although it is technically an upgrade of the system, it is still classified as an e-ordering implementation as many end-users will experience this system for the first time. In addition, as ordering through other channels than Visma will be discontinued, even experienced Visma-users are affected if they previously ordered through such channels. This is further elaborated on in section 1.4 when discussing the scope of this thesis.

1.2 Continuation of pre-diploma thesis

This master's thesis is a continuation of our pre-diploma thesis. During the Fall of 2014, we developed a framework for e-ordering implementation, covering critical success factors (CSFs) of three distinctive project phases in such an implementation. A CSF is simply defined as a factor critical for achieving a successful implementation (Vaidya et al., 2006). The aim was to aid managers in general and St. Olav's in decision-making by presenting a convenient framework with clear managerial implications. To our knowledge, such a framework is lacking in the literature.

Before introducing the research questions in the next section, it is necessary to shortly introduce some of the main findings of the pre-diploma thesis in order to bring the reader up to date. This is especially important since some of the research questions presented in the next section refer to some of these findings.

A framework covering CSFs was developed through combining literature on e-ordering, e-procurement, and change management research. The framework also highlights three typical phases of an implementation, and in which of these CSFs are of importance. The three phases are pre-implementation, implementation, and post implementation. This allows managers to assess what factor they need to consider in each phase, but also learn from what may have been lacking in previous phases, hence increasing managerial implications. We employed an exploratory case study to gain insight into St. Olav's (the case company), refine the research questions, and gather empirical data. After the development of the framework, empirical data was used in a preliminary analysis in order to ascertain the relevance of the framework to the case company. Findings suggested that the framework is helpful for emphasizing important aspects, analyzing whether these aspects have been considered sufficiently, and for aiding the case company through the implementation. Details about the development of the framework and theory employed are elaborated on in chapter 2.

The exploratory case study unveiled two distinct end-user groups at St. Olav's, while a third group were believed to exist. The research questions relate to these groups, so they are shortly introduced here. The groups are as follows:

1. End-users that exclusively order through Visma
2. End-users that order via Visma as well as other channels, such as phone and email
3. End-users that have not yet ordered via Visma

Group 1 is the hypothesized group believed to exist. Our interviews only unveiled users in group 2 and 3 (Interviews central warehouse, 2014), but we believed there was a group of end-users that made all their orders through the system.

As already mentioned above, the scope of the pre-diploma thesis was e-ordering implementation at St. Olav's. Due to limited literature on e-ordering implementation and critical success factors, we decided to turn towards e-procurement literature, a broader field of study, as e-ordering is a subset of e-procurement (Gunasekaran and Ngai, 2008). Furthermore, literature on CSFs within e-procurement was supplemented with two main streams of research on e-ordering implementation by Arbin (2009) and Reunis (2007). This was done in order to ensure that the framework developed would be relevant in an e-ordering implementation, as the one currently happening at St. Olav's. The reason for going into detail on this subject is that the difference between the two terms seemed to be problematic before developing the framework. As e-procurement is a broader topic, we did not know whether all CSFs found in such literature would be relevant in an e-ordering context. It affected both our literature search and development of the framework. The preliminary analysis did however not indicate any CSFs from the e-procurement-literature being irrelevant in an e-ordering context. On the contrary, we were able to relate the empirical data gathered to most of the CSFs. Furthermore, any CSFs from an e-ordering context should be applicable in an e-procurement context, since e-ordering is a subset of e-procurement. It is also reassuring that Arbin (2009) and Reunis (2007) did implicitly not consider this a problem, by mixing the two concepts in their doctoral theses. Based on this notion, we will hereafter use the term e-procurement instead of e-ordering.

1.3 Research questions

Due to time constraints, we were unable to validate the framework developed in the pre-diploma thesis. Limitations of the pre-diploma thesis suggested further validation to ensure the relevance of the framework, which justifies the aim of this master's thesis. For practical reasons, such as access to St. Olav's, this master's thesis will attempt to validate only parts of the framework. A large base of approximately all the end-users with purchasing rights has been made available to us by St. Olav's, which facilitates testing whether the CSFs in the framework are relevant to end-users. This is further elaborated on in chapter 3.

Figure 2 below illustrates groups of end-users hypothesized to exist in the available database of end-users.

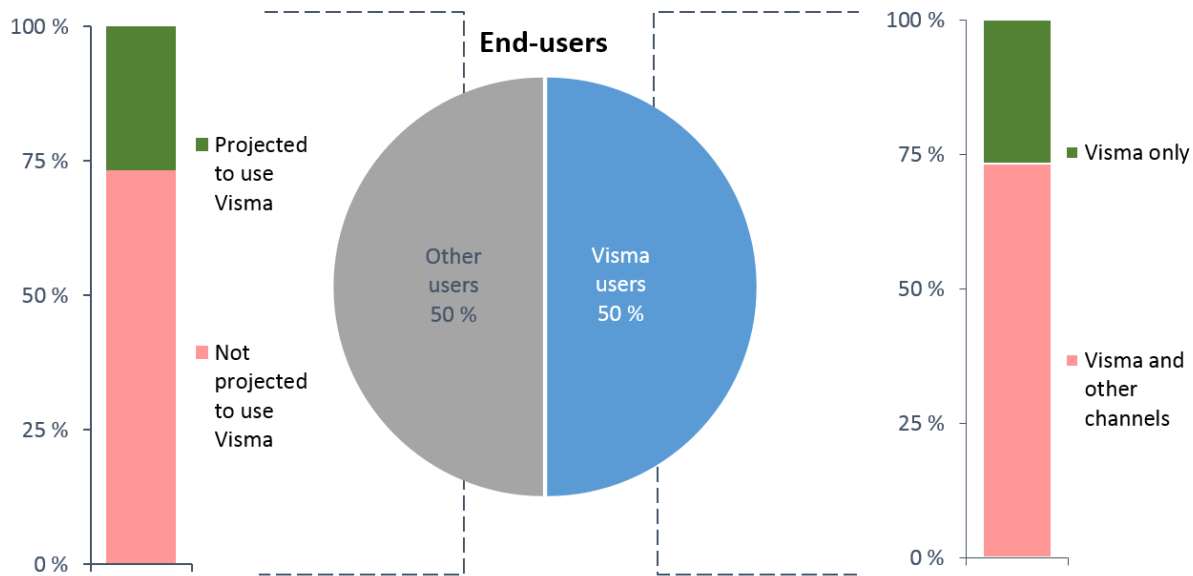


Figure 2 – Hypothesized groups of end-users

End-user group 3 mentioned previously has now been divided into two groups where both are currently not using the system, but one group has plans to use the system while the other not. The reason for the grouping of end-users in Figure 2 is to ascertain the relevance of the CSFs not only to the end-users as a whole, but also to see if there are any differences between the end-users. For example, one such difference could be that end-users already exposed to a new system need less training when being introduced to another system. Such differences may prove valuable for both research and St. Olav's, especially as St. Olav's is planning to introduce a new ERP system called SAP already in 2016. The topic of this master's thesis is therefore in the interest of the case company and other companies in a similar situation, in addition to possibly contributing to research on e-procurement implementation generally.

The main research question is as follows:

Are different CSFs more or less relevant for different types of end-users in an e-procurement implementation context?

To answer this main question, we will address the following research questions:

1. What are critical success factors for e-procurement implementation in the literature?
2. What is likely for the different end users to assess as important?
3. To what degree do different end-user groups in an empirical survey assess the importance of different factors, and are the factors indeed relevant to end-users?
4. Which insights may be gained from this statistical analysis?

1.4 Limitations of the thesis

Some limitations were necessary in order to scope this master's thesis, and to cope with time and resource limitations. First, the sample of end-users are limited to employees with purchasing rights at St. Olav's. The research questions are therefore limited to those CSFs from our framework that are most relatable to the end-users made available to us by St. Olav's. Many of the CSFs in our framework are difficult to relate to for end-users, as they have a managerial and project perspective. Furthermore, with St. Olav's currently being in the implementation phase, the research questions are also limited to CSFs relating to this phase because we lack empirical data for the other phases in the framework. A multiple-case study could be more suitable for attempting to validate the entire framework, allowing researchers to assess the importance of CSFs that end-users cannot relate to directly, and CSFs relating to the other two phases. Specifically, also including project managers in a multiple-case study could enable the investigation of aspects not related to end-users or the implementation phase, such as ROI decisions.

The research questions focus on the case company, St. Olav's, a publicly held hospital in Norway. This is the domain of the empirical data, which can affect generalizability across settings. Generalizability will be further explored in section 6.2. The supply chain perspective, such as what aspects of a successful implementation affects the suppliers, is not considered. Covering the supply chain would result in a more comprehensive study and require access to suppliers in order to gather empirical data. Moreover, direct goods and the sourcing process (supplier selection, negotiation, and contracting) are not considered, as this is not in the scope of the project at St. Olav's. The project entails only MRO goods, and therefore this thesis is limited to such goods. In addition, the sourcing process may be difficult to relate to for the end-users, because St. Olav's has a dedicated purchasing department handling those processes.

1.5 Structure of the thesis

The first part of this master's thesis highlighted the context of the case company, St. Olav's, currently undergoing an e-procurement implementation. The aim of this thesis, as a continuation of the framework developed in the pre-diploma thesis, was then discussed. Furthermore, this led to the research questions, trying to ascertain the relevance of the framework for the end-user population made available to us at the case company. The scope and limits of the thesis was then discussed, to highlight which aspects we decided to exclude from this thesis for various reasons. Following, the next part of the thesis, chapter 2, will concentrate on the contents of the framework, the theoretical basis, and how it was developed. This forms the basis for the hypotheses, which are directly related to CSFs in the framework. Chapter three deals with the research design and method chosen. We elaborate on how the survey was developed and how it was conducted, before relevant quality criteria are discussed. The end of chapter 3 contains a brief assessment of the options and assumptions for statistical analysis of the gathered empirical data. Thereafter, in chapter 4, we present necessary considerations before moving on to demographics of the empirical data and the characteristics of the results. In chapter 5, we discuss challenges of the survey, and the reasons for why an exploratory factor analysis was carried out. Subsequently, analyses are conducted and the hypotheses are tested. Next, in chapter 6, we discuss the results of the survey and the analyses, before addressing research questions 1 to 4. After that, we move on to generalizability of the findings. In chapter 7 we present the conclusion of this thesis and address the main research question, before moving on to further research in chapter 8.

2. Theory, framework, and hypotheses

In this section of the paper, the theoretical basis for this master's thesis is presented in order to address research question 1 - *What are critical success factors for e-procurement implementation in the literature?* First, we will briefly discuss the methodology approach used to form the theoretical basis in the pre-diploma thesis, as this founds the theoretical basis for this master's thesis. Next, we present relevant theory on purchasing regarding e-procurement and CSFs found in literature, before moving the framework. We will decompose the framework for brevity instead of developing it from scratch, which was the case when constructing it. Therefore, we will only briefly explain the contents of the framework and how it was developed. Finally, we present the hypotheses in order to address research question 2 - *What is likely for the different end users to assess as important?*

2.1 Methodology underlying the framework

Before presenting the framework developed in the pre-diploma thesis, the underlying methodology of the development is introduced briefly. When developing the framework, we employed an analytical conceptual approach rather than an empirical one, due to limited access to the case company. Wacker (1998) defines this type of research as aiming to add new insights by logically developing relationships between concepts into an internally consistent theory. Furthermore, the distinction between e-procurement and e-ordering, as discussed in section 1.2, was a key driver to how the literature search and review was conducted. A change in the review question from what is known about critical success factors of e-procurement to entailing e-ordering instead resulted in a shift in the review approach from systematic to narrative. Using the narrative approach via snowball sampling was the only feasible choice to investigate the scarce research on e-ordering. However, aiming to find relationships between concepts from both e-procurement and e-ordering, literature findings from the systematic approach was kept. In addition, relevant theory and concepts of change management were added to the literature search. In addition to acting as a backdrop, change management was included to serve as both as an interface and a supplementary perspective to possible concept relationships. The approach is summarized in Figure 3 below.

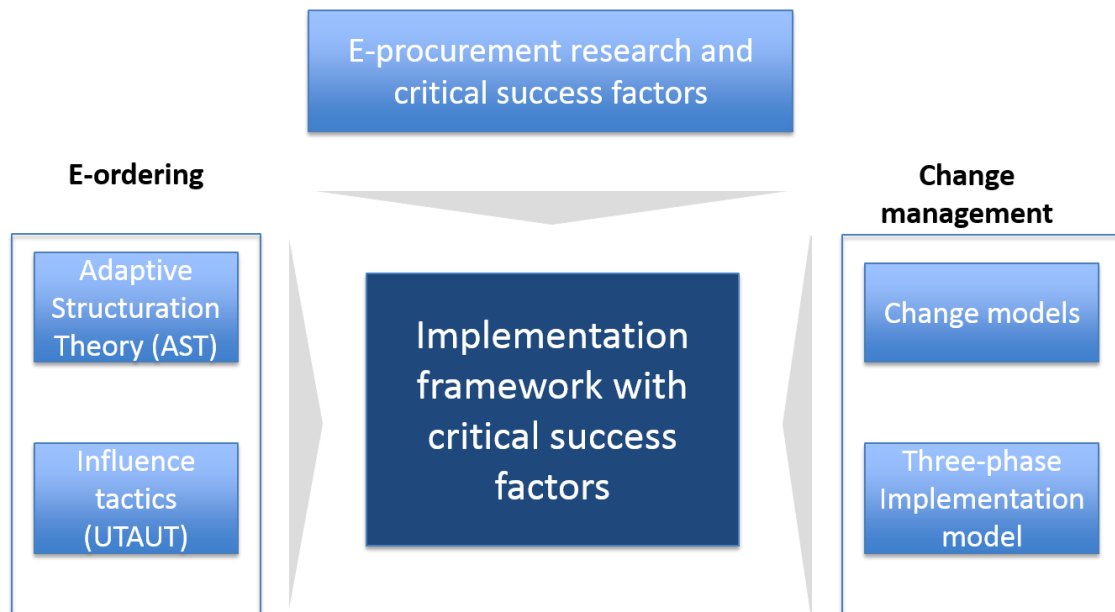


Figure 3 – Approach to developing the framework

As mentioned in section 1.2, we employed an exploratory case study. Wacker (1998) states that case studies are often applied to illustrate examples when employing an analytical conceptual research approach. The empirical data was gathered using semi-structured interviews with seven individuals and by examining a myriad of documents relating to the e-ordering project at St. Olav's. Of the seven interviews, three of the interviewees were end-users, two having used Visma. By collecting opinions and predispositions toward an electronic ordering system, the overall target was to provide examples and insight to the developed framework. Employing only a single exploratory case study did limit generalization of the findings ascertaining the relevance of the framework. Moreover, as only three end-users were interviewed, this inhibited the possibility to generalize the findings to a larger group of end-users at St. Olav's.

Furthermore, regarding limitations, the theory behind the framework, especially the CSFs, were mainly from findings in the public sector although findings from the private sector were used as supplements. As the framework was only applied on St. Olav's, which is a publicly held company, this implied that further validation is necessary, for instance, in case studies where the focal companies are privately held, in order to ascertain if the framework is applicable to the private sector. Another limitation relevant for the findings is in regards to where the theory findings are originally from, i.e. nationality. Although not turning out to be an issue, some care should be included when developing a framework by using theory based from other parts in the world and by applying it on a Norwegian company. However, it may be

stated that frameworks generally are not to be used rigidly, but to have some room for interpretation and adaptation considering each specific instance they are used.

Before continuing the work of the pre-diploma thesis, we performed a new, updated literature search in order to see if there were any new discoveries since the pre-diploma thesis was written the Fall of 2014. This search did not yield any new insight, and thus did not yield any reason to revise the framework before setting out to test the framework in a survey.

2.2 Purchasing and E-procurement research

The importance of purchasing has increased from being an administrative function in the 1970s to becoming a function that drives the strategy of the firm (Cousins, 2008). During the first years of this century, top management began to realize that purchasing is a key contributor to corporate strategy (Axelsson et al., 2005, Puschmann and Alt, 2005). The ratios of purchasing-to-sales is usually in the range of 30-60% for service organization, and even higher in the retailing business. This illustrates the potential impact of purchasing, and a dollar saved in purchasing is a dollar added to the bottom line (Van Weele, 2005).

Purchasing is usually divided into direct and indirect purchasing. Direct purchasing is materials going into the final product, and has been the focus of management for many decades. Indirect purchasing includes purchasing of so-called MRO goods. Moreover, as mentioned in the introduction, indirect purchasing has not received the same amount of attention until the last decade. It has often been done in a decentralized and uncoordinated way (Arbin, 2008a). According to Bechtel and Patterson (1997), there are three possible explanations for this: 1) When reducing further costs in direct purchasing becomes difficult, companies turn to other areas to reduce costs. 2) Little time is spent on strategic MRO issues, since most of the time is spent on day-to-day tactical decisions and processing of routine paperwork. 3) Pressures of global competition incentivize managers to cut costs everywhere.

Based on a survey, Cox et al. (2005) find that MRO purchasing amount for up to 20% of all purchases. This is lower than what the Aberdeen Group, as mentioned in the introduction, found in 2001 (30-60%), which may indicate an increased awareness towards MRO costs. Cox et al. (2005) also find that MRO purchases in organizations still suffer from lack of internal support, maverick buying, and compliance rates as low as 25-50%. A recommended strategy for MRO purchasing is having a centralized purchasing function, using a category management strategy in combination with an e-ordering system functioning as a tool to steer purchasing orders towards suppliers, thus reducing maverick buying (Croom, 2000, Puschmann and Alt,

2005, Van Weele, 2005). As companies experience increased pressures of competition and costs reductions (Christopher, 2011), new tools to reach these goals are developed. One such gaining tool gaining popularity is e-procurement (Ronchi et al., 2010). According to Gunasekaran and Ngai (2008), the use of e-procurement is inevitable in services following the increased use of internet and electronic data interchange (EDI).

Van Weele (2005) depicts the general procurement process, which can be found in Figure 4 below. The figure illustrates the different procurement activities, ranging from the need of an internal customer, to delivery from the supplier.

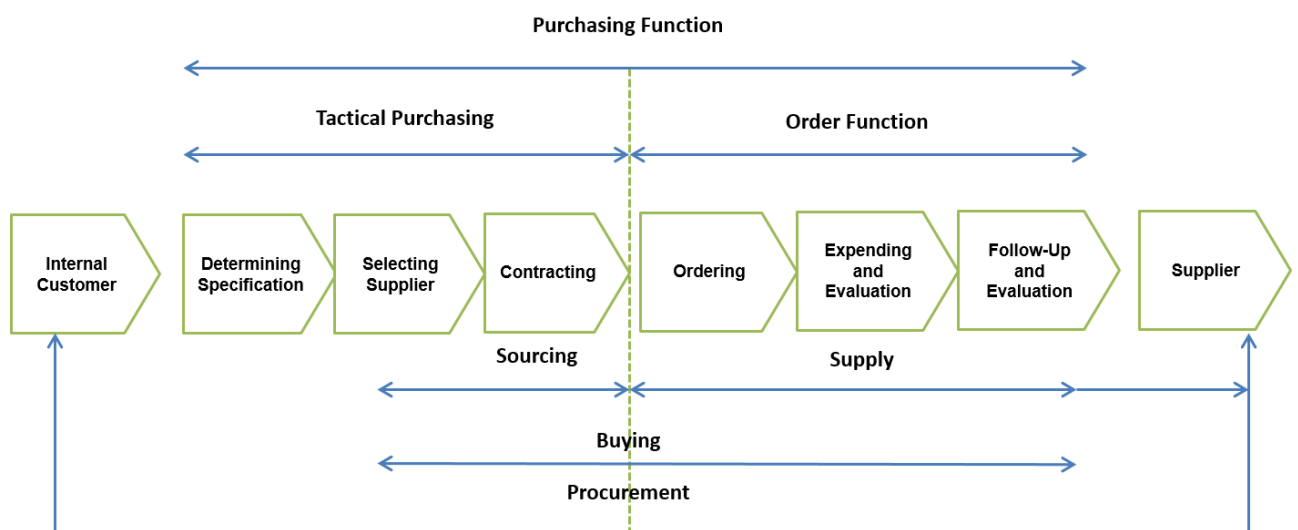


Figure 4 – General procurement process adapted from (Van Weele, 2005, p. 34)

E-procurement refers to the use of internet-based information and communication technologies to carry out stages of the procurement process, as depicted above, such as supplier selection, sourcing, ordering, expediting, and evaluation (Croom and Brandon-Jones, 2005). There are various forms of e-procurement concentrating on the different stages. E-ordering is one such form, but generally, e-procurement is considered as an end-to-end solution. E-procurement is further a subset of e-commerce. E-commerce can be defined as doing business using Internet technology (De Boer et al., 2002).

In order to reap the gains from an e-procurement investment, as mentioned in the introduction, adoption is necessary (Arbin, 2008b, Reunis et al., 2006). Pinto and Slevin (1987, as referred to in Panda and Sahu, 2012) postulated that project success is dependent on addressing CSFs associated with the project. There is a need for a much better understanding of CSF concerning e-procurement implementation and use in the public sector. Without a set of CSFs, it seems impossible to assess the successfulness of e-procurement initiatives in the public sector (Vaidya et al., 2006).

2.3 Presentation of the framework

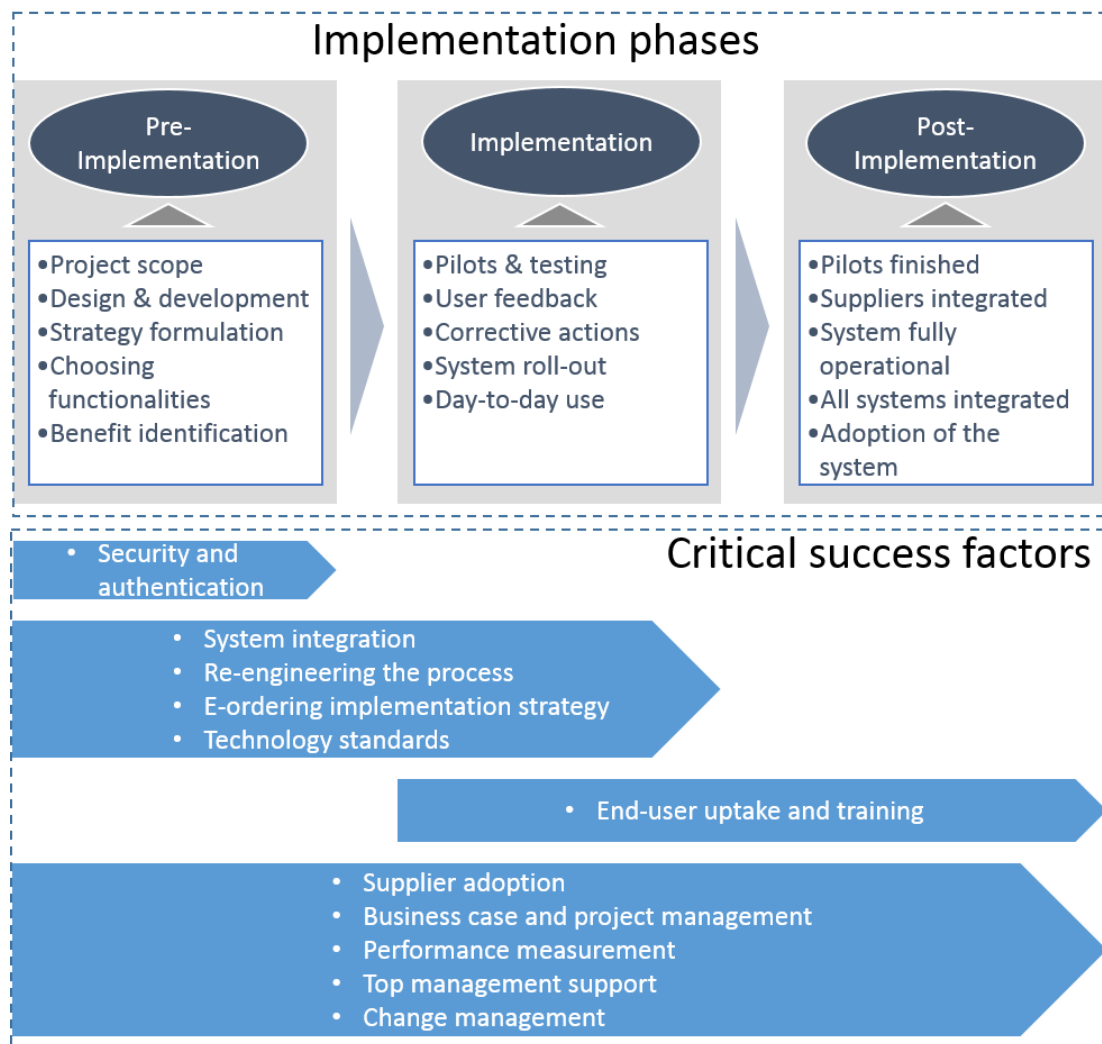


Figure 5 – The framework

Figure 5 above illustrates the framework we developed. The framework contains two main parts; implementation phases and critical success factors. The implementation phases illustrate three distinct phases we argue e-procurement projects typically undergo. These three phases can help the organization to assess which phase they are currently in, in addition to illustrating

typical characteristics of each phase. The theoretical basis and a more thorough explanation of these phases are found in section 2.4 and 2.5. The second part of the framework consists of eleven critical success factors we argue are necessary to consider in order to carry out a successful implementation of an e-procurement initiative. The blue arrows indicate in which phase each factor is most important, to increase managerial implications. Each of the critical success factors has multiple attributes (or “sub-factors”) to help managers understand the substance of the factors. In some sense, an attribute is also a CSF, but on a lower level than the CSFs. A more thorough explanation of the CSFs and their attributes is given in section 2.5. It must be noted that the attributes of the CSFs are omitted from the figure above for illustrative purposes. The framework and the following paragraphs explaining what it consists of addresses research question 1: *What are critical success factors for e-procurement implementation in the literature?*

2.4 Change management and project phases

Chan and Swatman (1998) state that an organizational change process over time is what best describes an information system implementation. Vaidya et al. (2006) elaborate on importance of change management in light of e-procurement, and argue among several aspects that a lack of change management may lead to project failure.

Consequently, we decided to have a two-pronged approach toward developing our framework, which entailed a purchasing perspective and a change management perspective. Change management played an important role in developing the framework, and served as a backdrop for the CSFs included in the framework. As explained earlier, we wanted to include project phases in our framework to increase managerial implications.

A planned change approach was proved being feasible, as such models attempt to capture the evolvement of change processes, usually divided into stages. Weick and Quinn (1999) argue that planned change is categorized as an episodic change, whereas emergent change is of a continuous change nature. We argue that planned change was a suitable approach for us. The process of e-procurement implementation does not relate to continuous change; it belongs in the category episodic event. We also argue that cost intensive e-procurement systems call for a planned change regardless if the idea to implement such a system was of a reactive nature.

The focus on planned change led us to Lewin’s three-stage model. Categorizing change into phases is in general a feasible tactic in order to underline the different aspects during the evolvement of a change process. However, in order to put this rather broad model into

perspective, we turned to other researchers such as Levasseur (2001), who emphasizes important aspects of Lewin's three-stage model. In addition, Cummings and Worley (2015) point out that Lewin's model is comprehensive and that the model has been extended to comprise more stages that are specific. One of these is Kotter's (1995) eight-step model for leading change, which, in combination with Lewin's model and Levasseur's (2001) specific properties of the three stages, was used as a backdrop of relevant change models. However, it was also necessary to put the change models in context of an information system implementation. Caniato et al. (2012) propose an implementation framework to identify main decisions companies should take during e-procurement adoption. On that note, adoption will first be defined due to its importance in information systems implementation in general, the eleven CSFs, and Caniato's et al. (2012) model

Rogers (1995, referred to in Reunis, 2007, p. 17) defines adoption as "the process to make full use of an innovation as the best course of action available". Zaltman (1973, referred to in Reunis, 2007) recognized organizational adoption process using a two-phased approach. The first phase, primary adoption, is the decision to adopt an innovation on an organizational level. The second phase, secondary adoption, is when the organizational members in an implementation stage of the innovation decide to adopt it. This two-phased approach can be viewed in light of change acceptance. Primary adoption relates to imposing a new system on the organization, whereas secondary adoption relates to end-user uptake, which has been discussed in earlier chapters. In the first phase, the important factor is to make the organizational level accept the change. In the second phase, even though potential users have accepted the change, it will not necessarily mean that they will adopt the system.

Returning to Caniato et al. (2012), the framework they propose is divided into three stages: pre-implementation, implementation, and post implementation. The pre-implementation phase addresses the identification of the goals of the project before selecting the appropriate functionalities to ensure achieving those goals. During the implementation phase, the functionalities (e-procurement elements) selected in the previous phase are introduced within the company. Furthermore, Caniato et al. (2012) argue that within this phase, the implementation should be frozen until the desired benefits are verified being achieved. They also highlight that some possible changes in management might be needed in this phase. In the final phase, post-implementation, assessment of whether the desired benefits have been attained or not should be the main focus. The implementation model Caniato et al. (2012)

propose is quite similar to Lewin's change model with a clear diagnosis and an analysis phase, which are Lewin's first two stages.

In order for managers to ascertain easily what phase they are currently in, we elaborate on the three phases and interpret the characteristics of each phase. Furthermore, upon having established the project phases and characteristics in the framework, the process of mapping CSFs onto the phases could begin.

Pre-implementation phase

Caniato et al. (2012) find that during pre-implementation, the objective is to decide what elements the company should focus on, and how these should be managed. This phase will typically include the scope of the project, design and development, strategy formulation, choosing e-procurement functionalities, and benefit identification. For example, among all benefits, the company must understand and choose which of the potential benefits it wants to focus on. Such benefits are automation of non-value-added activities, centralization of tactical and strategic activities, and information sharing (Caniato et al., 2012). This phase relates to work and planning needed before one can initiate pilot projects and system roll-out to end-users. Munns and Bjeirmi (1996) support this notion by stating that sufficient project planning in the early phase is critical to achieve success.

Implementation

During this phase, the system and functionalities selected are introduced to the company in order to be used in the day-by-day activities. Pilot projects will typically be the first step in introducing the system to end-users (Caniato et al., 2012). This will allow for testing of the system and its functionalities, gathering feedback from users, correction of any problems encountered etc., before rolling out the system to all end-users.

According to Caniato et al. (2012), this phase also concerns actions needed to achieve the goals of the implementation. Employee training and people involvement, especially top management commitment, were found to be crucial factors for successful implementation. As mentioned above, the authors suggested freezing the implemented solution to verify achievement of benefits. Adoption has to be measured, and incentives linked to individual targets. Furthermore, the authors noted that change management was an important aspect during this phase.

Post-Implementation

According to Caniato et al. (2012), this phase concerns benefit assessment, quantification of user satisfaction, and possible modification of misaligned functionalities. This phase is reached when pilots are finished, supplier catalogues and systems are integrated, and the system has been rolled out to all end-users for day-to-day use.

Caniato et al. (2012) found that adoption of the system was vital to achieve expected benefits in this phase, and that higher adoption led to larger benefits. This notion is supported other authors as well, such as Arbin (2008a) and Reunis et al. (2006). Although measurement was found to be especially important during this phase, the authors stressed that measurement systems must be implemented already during the pre-implementation phase, and monitor achievements throughout the project to assure success.

After having introduced most of the vital parts concerning the implementation phases of the framework, next section will elaborate on the critical success factors underlying the framework.

2.5 Critical success factors in the framework

Vaidya et al. (2006) state that previous research on CSFs in e-procurement is sparse, and that there is a need for much better understanding of CSFs concerning e-procurement implementation. Subsequently, Vaidya et al. (2006) employ a literature survey of practitioner material to capture practitioners' perceptions of e-procurement practices. The result is eleven identified CSFs, with attributes.

A more recent and extensive study by Panda and Sahu (2012) covers 28 different research-studies identifying CSFs in an e-procurement context. The authors extend the work of Vaidya et al. (2006) by finding literature support for the same eleven CSFs and by adding attributes having impact on e-procurement project outcomes. Furthermore, Panda and Sahu (2012) and Caniato et al. (2012) request a tabulation of the CSFs in terms of e-procurement stages, as the CSFs lack a project execution context like project phases.

Clark et al. (2012) attempt to address a request by Vaidya et al. (2006) for in-depth case studies to validate the eleven CSFs. Clark et al. (2012) use these CSFs to analyze a large and successful e-procurement project for the State of Arizona. More specifically, they analyze how efforts of the project team relate to the different CSFs, what obstacles a focus on the CSFs are able to overcome, and what benefits are realized by using the CSFs. The project initiated incremental implementation of an e-procurement solution in June 2009. By July 2011, all state agencies had successfully implemented all the phases of the project on time and within budget. Some of

the benefits realized was increased transparency and competitiveness, leading to expected cost reduction in the range of 5-20% on a \$6 billion state spend (Clark et al., 2012).

The findings of Panda and Sahu (2012) and Clark et al. (2012) were used to support and enrich those of Vaidya et al. (2006) in our pre-diploma thesis. We chose to enrich the framework of Vaidya et al. (2006) as it was well established. Furthermore, as mentioned introductorily, we explored two different streams of literature regarding individual information system adoption and use. These theories are 1) adaptive structuration theory (AST), and 2) The Unified theory of acceptance and use of technology (UTAUT). Both theories have previously been adapted to analyze e-ordering adoption in case companies. Arbin (2009) adapted AST to an e-ordering context through her Ph.D. dissertation. Reunis (2007) adapted UTAUT to an e-ordering context through his dissertation by looking at how influence tactics affect adoption. These dissertations were analyzed in our pre-diploma thesis, and attributes were extracted from their findings also to enrich our framework.

The eleven CSFs by Vaidya et al. (2006) are explained briefly below in subsections 2.4.1-2.4.11. The approach we employed to enrich the factors by Vaidya et al. (2006) will also be exemplified briefly only for the first factor, *End-user uptake and training*, to illustrate the method we used for the eleven CSFs. The same goes for the approach used to map the CSFs onto the three project phases and using change management as a backdrop. This process was comprehensive and is not the focus of this master's thesis. The result of the enrichment process is for convenience shown in Table 1 below. Each headline is one CSF, coinciding with those in the framework above in Figure 5. The paragraph below each headline contains all the attributes, and increase managerial implications by indicating what each CSF consists of and important aspects to consider.

<p>1. End-user uptake and training -User involvement, user support/communication, user training, in-house training, on-demand training, self-learning, focus on users with low IT-experience, include end-user benefits</p>
<p>2. Supplier adoption -Supplier e-readiness, supplier adoption strategy and communication plan, supplier education and benefits demonstration, compliance to best practices with content and catalogue management</p>
<p>3. Business Case and Project management -Identification of business drivers, business process assessment and requirement, Return on Investment (ROI), Total Cost of Ownership (TCO), risks identification and management, pilot projects</p>
<p>4. System integration -Information matching, sending and receiving of real time information to other information systems, electronic commerce with suppliers</p>
<p>5. Security and authentication -Infrastructure, authentication, authorization, planning before system roll-out, confidentiality and integrity, security requirements</p>
<p>6. Re-engineering the process -Transparency improvement, automated invoice payment and reconciliation, compliance with purchasing procedures and standards, diagnosis, analysis and redesign of processes, integrated supply chain processes, mapping previous order routines and work routine, customization</p>
<p>7. Performance measurement -Goals and targets, Key performance Indicators (KPIs), baseline measurement, progress monitoring, alignment of compensation and rewards with performance evaluation, capturing use behavior in adoption measures</p>
<p>8. Top management support -Management sponsor, allocating sufficient resources, involvement of the steering committee, management support on all levels, investment in organizational change, capabilities and limitations of IT, system alignment with strategy, establishment of an appropriate work culture</p>
<p>9. Change management -Identification, management, and communication to key stakeholders and peers, convincing information exchange, group composition, e-ordering impact assessment, potential barriers to implementation, organizational resistance, communication plan, Irreversible changeover to e-ordering, mandate that considers prerequisites and outcomes</p>
<p>10. E-ordering implementation strategy -Sound procurement practices, opportunities for aggregation, a consistent approach to procurement, investigate preferred suppliers, relationships with suppliers</p>
<p>11. Technology standards -Technical standards, content standards, process and procedural standards, compliance with the standards frameworks, interoperability, user friendly system with minimum effort required</p>

Table 1 – Critical success factors with attributes

2.5.1 *End-user uptake and training*

An e-procurement initiative may introduce changes to the current procurement approach, as well as introducing new technology to users. The need for training staff in the usage of new procurement practices and the technology is critical to the success of the initiative (Panda and Sahu, 2012, Vaidya et al., 2006). For the organization and its end-users to realize benefits from the new system, they must understand the functionalities of system. Training should therefore be given high priority, as well as the need for identifying skills required by staff engaged in procurement activities (Vaidya et al., 2006).

A new technology itself does not ensure successful adoption; the success depends on both users and buyers actually using the new system and processes (Vaidya et al., 2006). For this to happen, the system must attract users so that they prefer using the new system and processes compared to going back to the old ones. Success of the project also hinges on communication to users, as the two major obstacles to increased support by users are their level of acceptance and technological awareness, and their willingness to change long-established internal business processes (Vaidya et al., 2006).

Enriching the CSF End user uptake and training

Firstly, the findings from Panda and Sahu (2012) were added to the framework. They found literature support for almost identical attributes to this CSF as those of Vaidya et al. (2006), but also added that four previous studies found developing an own in-house training as an important attribute. Secondly, the in depth case study by Clark et al. (2012) was analyzed to support the CSFs and attributes by Vaidya et al. (2006) and Panda and Sahu (2012). For example, Clark et al. (2012) found that by implementing a training program that included agency personnel, i.e. in-house training, end-users were able better understand the system and understand how the agency wanted them to operate. This led to rapid improvements and reduced purchase order cycle times of 46%. Some other benefits were increased sense of ownership in the core team and that early adopters could guide late adopters.

Thirdly, findings from case company studies by Reunis (2006) and Arbin (2008a) were extracted into possible attributes. Ten possible attributes were deducted from the findings of Arbin (2008a) and nine possible attributes from the findings of Reunis (2006). Furthermore, these possible attributes were then assorted to respective CSFs through a logical and extensive process following the MECE (mutually collective, collectively exhaustive) principle. Change management theory was also used to support the assigning of attributes to factors. The CSF *end-user uptake and training* could be related to multiple of the attributes we deducted from

the two research streams by Arbin (2008a) and Reunis et al. (2006). Since both of these research streams focus on end-user adoption, it was no surprise that multiple findings related to the CSF *end-user uptake and training*. To illustrate the process, one attribute deducted from Reunis (2006) was *to include end-user benefits in training and communication*. Reunis (2006) found that this was especially important in order to motivate end-users and to ensure that they adopted the system. This attribute is directly related to the CSF *end-user uptake and training*, and was therefore added as an attribute to emphasize the importance of including end user benefits. This way, the work of Vaidya et al. (2006) and Panda and Sahu (2012) was extended by shedding light on aspects not previously covered, as well as increasing managerial implications.

Moreover, from a change management perspective, fear of a bigger workload and fear of the unknown may lead to resistance (Jacobsen, 2012). Levasseur (2001) and Kotter (1995) exemplify that a focus on communication may reduce fear of the unknown. Regarding fear of a bigger workload, communication alone cannot convince an individual to support a change initiative, i.e. implementing e-procurement, if the workload in fact becomes bigger. However, according to Jacobsen (2012), the workload usually becomes bigger only for a short period of time, which includes training and phasing out old work procedures. If benefits are communicated in such a way that the end-users see the benefits for the company as a whole and for themselves, this may lower some of the potential causes for resistance, for instance, fear of a bigger workload and the fear of the unknown. Based on the change management perspective, it made sense to add the attribute. The remaining 18 attributes were assigned to relevant CSFs in a similar fashion.

Mapping the CSF *end-user uptake and training* onto the framework

Referring back to the framework, Figure 5, we have three distinct project phases where the CSFs are assigned according to the phase they are most important in. It should be noted that some of the factors are difficult to place within one phase only, as they may be important in more than one phase, and their attributes may diverge towards more than one phase. In those cases, the CSFs was mapped onto more than one phase. The factor *End user uptake and training* is depicted to be most important in the implementation and post-implementation phase.

In order for the end-users to understand the functionalities of the system, people involvement, especially in terms of training, is important. The functionalities chosen in the pre-implementation do not become apparent for end-users before the system is rolled out or partially by a pilot project. Based on this notion, the importance of end-user uptake and training

are most important in the implementation phase. Caniato et al. (2012) also highlight the importance of end-user training, as well as communication and feedback in this phase. Regarding the post-implementation phase, when requirements of the implementation phase are met, day-to-day use of the system should be in place. The need for further training may however still be required if adoption is lagging behind. For these reasons, this CSF's main emphasis is assigned to be applicable to the implementation phase, but also applicable to the post-implementation phase depending on the specific case. The remaining ten factors have been assigned to their relevant phases in a similar fashion.

2.5.2 Supplier adoption

Suppliers are one of the most important group of stakeholders for assuring the success of an e-procurement system (Panda and Sahu, 2012). Early supplier involvement is closely related to the success, and they must be involved in every step of the implementation. Demonstrating the proposed solution to the suppliers and discussing concerns and issues such as development and maintenance of supplier catalogues are important (Vaidya et al., 2006). Allowing suppliers to offer feedback should be encouraged, and may allow the purchasing department to find areas of improvement and adjust practices accordingly (Panda and Sahu, 2012, Vaidya et al., 2006). Furthermore, suppliers may see the e-procurement initiative as an attempt to force prices down through increased leverage, especially if they are uncertain about the benefits to be gained. Suppliers should therefore be educated on benefits that can be provided to them as early as possible in the project (Vaidya et al., 2006). The e-procurement system should also be simple and effective so that most suppliers can use the system (Panda and Sahu, 2012). The successfulness of the e-procurement initiative may well be related to the electronic readiness of suppliers, and communication with suppliers is therefore important (Vaidya et al., 2006).

2.5.3 Business case and project management

Successful initiation and progression of the e-procurement initiative requires making a strong business case (Panda and Sahu, 2012). Achieving planned benefits is contingent upon users and buyers making use of the program, which requires senior management sponsorship and championing of the project. Ensuring buy-in is particularly important, i.e. any inertia of procurement officers and other people associated with the implementation must be removed. To do this, one must identify business drivers and benefits of the system (Vaidya et al., 2006, Panda and Sahu, 2012). Furthermore, Birks et al (2001, as referred to in Vaidya et al., 2006) suggest that the business case should include understanding the starting point, benefits,

approaches, affordability, risks, and benefit realization. Identification of risks and implementation of risk mitigation plans are crucial in order to deal with any temporary setback the organization might encounter. Furthermore, pilots can highlight benefits of the system and allay all fears, and counter arguments potential opponents of the system may have (Panda and Sahu, 2012).

2.5.4 System integration

It is imperative to determine the level of integration required between the e-Procurement solution and existing information systems. For the e-procurement initiative to be successful, the system has to be well integrated with existing IT systems, especially financial systems (Panda and Sahu, 2012, Vaidya et al., 2006). This facilitates the process of online payment to suppliers. Furthermore, it is important that information is shared to all stakeholders in real-time across systems, and that it is reliable and accurate (Panda and Sahu, 2012). If integration issues are complex, it is likely that underlying business processes within an organization should be changed or adapted (Vaidya et al. 2006).

2.5.5 Security and authentication

Security of data is a critical factor to consider in a governmental context due to the sensitivity of data and the legal nature of order and payments. Mechanisms for identifying and authenticating the user that places an order is crucial so that the supplier knows the order is safe to fulfill (Vaidya et al., 2006), but also due to corruption and fraud considerations (Panda and Sahu, 2012). In order to encourage buyer and supplier use of the system, both parties must have complete confidence of the underlying security infrastructure (Vaidya et al., 2006)

2.5.6 Re-engineering the process

E-procurement can be viewed in means of making the procurement process more efficient, in terms of time, cost, and value for money. It is important that the organization undertake a renewed look at all the procurement processes, as long-established internal business processes may stand in the way for reaching the goals of the initiative, and may require re-engineering (Panda and Sahu, 2012). According to Vaidya et al. (2006), a significant proportion of benefits gained from implementing e-procurement initiatives is due to changes made through business process re-engineering rather than the implementation of the system itself. Such re-engineering should also address supplier relationships and all internal groups affected. New processes may introduce substantial changes to roles and responsibilities, which require staff to adapt to these (Vaidya et al., 2006).

2.5.7 *Performance measurement*

To ensure that the e-procurement initiative yields intended results, it is vital to continuously measure key performance indicators (KPIs). Early identification of KPIs is important as it enables tracking of benefits achieved and distills the business case into measurable targets. In addition, measurement drives behavior, and can encourage stakeholders to work towards their goals (Panda and Sahu, 2012). A lack of measurement capability can result in management having only limited tools for assessing organizational progress. An accurate measurement of process functionality before the change is also important in order to understand the effect of the project and whether the initiative was worth the investment or not (Vaidya et al., 2006).

2.5.8 *Top management support*

It is well recognized that top management support is critical for success in most projects. This is the case for an e-procurement implementation as well (Vaidya et al., 2006). Top management buy-in and political will have been found to be the most important factors in e-procurement implementation (Panda and Sahu, 2012). The management team must involve stakeholders such as project manager, project consultants, and staff in order to develop an implementation strategy and policies necessary for initiation of the e-procurement implementation. Considerable attention and support is needed by management to make sure that the reform is well understood in the organization. Moreover, a vision with goals must be set forth to create a collective commitment for change (Vaidya et al. 2006; Panda and Sahu, 2012).

2.5.9 *Change management*

The speed of adoption of an e-procurement initiative is directly related to change management. Although change management can be the least expensive aspect of an e-procurement implementation, it is critical for the project's success. With change management issues becoming more substantial as stakeholders needs increase, more attention must be given to such issues (Vaidya et al, 2006). A plan for managing change must be in place to ensure a smooth roll-out and consider elements such as training of users and gathering feedback. In addition, help desk systems or call centers, online help and FAQs should be readily available on the online e-procurement portal (Panda and Sahu, 2012).

2.5.10 E-procurement implementation strategy

Once the decision to implement an e-procurement solution has been taken, it is imperative to plan far ahead and create documented and executable strategies. E-procurement strategy, being both procurement and technology driven, must account for major institutional changes from both the procurement perspective as well as the organizational perspective (Vaidya et al., 2006). It should also be based on sound procurement practices, and take into account opportunities the e-procurement system creates, such as demand aggregation and increased leverage. A downside however, is potential hurdle due to vested departmental interests, perceived loss of authority and effort required in achieving it. The need for demand aggregation should therefore be weighed against the degree of decentralization desired for (Panda and Sahu, 2012).

2.5.11 Technology standards

In order for e-procurement to function properly, buyer and supplier systems must exchange information and electronic documents through common standards (Vaidya et al., 2006). Such standards should be open source so that systems can be linked to other systems for interoperability. This can also simplify upgrading of the system (Vaidya et al., 2006). Common concerns that may surface is what standard should be used for formatting the electronic catalogues and what the legal requirements are (Panda and Sahu, 2012, Vaidya et al., 2006). Furthermore, the interface design should be intuitive and easy to use, and require minimal efforts. Adoption of the system depends on the ease of which data can be exchanged both within the organizations and between their supply bases (Vaidya et al., 2006).

2.6 Hypotheses

In this section of the paper, we will develop multiple hypotheses based on CSFs and attributes, that end-users can relate to, from our framework. The selection of these is discussed in subsection 3.2.1. The hypotheses will also be based on relevant theory where possible, and aim to address research question 2: *What is likely for the different end users to assess as important?*

Table 2 below shows which CSFs and attributes we chose to investigate. We note beforehand that not all of these attributes could be included in the hypotheses, due to a lack of theoretical basis for what differences one can expect across end-users. The lack of theory is due to the attributes being specific and not validated before. The survey will still consider all attributes to assess potential differences across end-users.

Investigated CSF	Investigated attributes
End-user uptake and training	<i>User involvement, user support/communication, user training, in-house training, on-demand training, self-learning, focus on users with low IT-experience, and include end-user benefits</i>
Change management	<i>Organizational resistance, communication to key stakeholders and peers, and irreversible changeover to EP</i>
Re-engineering the process	<i>Previous work/order routines, transparency improvements, and compliance with purchasing procedures and standards</i>
E-procurement implementation strategy	<i>Preferred suppliers and relationships with suppliers</i>

Table 2 – CSFs and attributes to be investigated

Recapping the main research question, we want to ascertain if our framework is more or less important for different end-users in the organization. In order to compare our results across groups of end-users, we must first find a suitable grouping of end-users. Based on the work in the pre-diploma thesis, see Figure 2, we found users in group 2 and 3, but were unable to find group 1 and 4. Although interviews uncovered that such groups probably did exist, we did not know for sure, or their size (Interview central warehouse, 2014). On this notion, the first hypothesis is as follows:

H1: *Four distinct groups of end-users can be found at St. Olav's:*

1. *End-users solely using Visma as their purchasing tool*
2. *End-users using Visma in addition to other channels*
3. *End-users currently not using Visma, but are projected to use it*
4. *End-users not using Visma and not projected to use it*

By using the grouping above, we can serve more than one purpose. Firstly, we can try to find useful information that can help St. Olav's in the ongoing implementation of e-procurement by knowing about group differences. In addition, this may also help them in the coming ERP implementation. Secondly, these groups differ in terms of experience, which is defined by Reunis (2007) as having used either the focal system or having used similar systems. In our case the focal system is Visma, and the groups are in ascending order according to the groups' experience with Visma. This grouping is interesting from a theoretical standpoint, because experience has been found to positively affect the users' perspective of how easy the system is to use, and how useful the system is (Taylor and Todd, 1995, Venkatesh and Davis, 2000, Venkatesh and Morris, 2000, Venkatesh et al., 2003).

Furthermore, Eisenhardt (1989) and Hamel et al. (1993) advocate assessing within-group similarities and between-group differences in cross-case analysis by clustering findings. In addition to the grouping suggested in hypothesis H1 above, there are also other possible comparisons if designed into the survey. Two such groups are age and sex, both found by Venkatesh et al. (2003) to affect perceived ease of use and perceived usefulness of IT-systems. More specifically, the effects were more pronounced for females than males and for older than younger individuals. The split between old and young is made at 40 years. Reunis (2007) used this split when investigating, among several concepts, perceived ease of use and perceived usefulness in a case study.

In our pre-diploma thesis, we found that the magnitude of change might have affected what the interviewed subject considered important. More specifically, we found that end users new to Visma considered the CSFs *end user uptake and training* to be especially important compared to end-users experienced in using Visma. Communication and support in terms of follow up was also more important, as well as other differences. We predicted that the magnitude of change experienced was an underlying factor for these differences, as the user experienced in Visma was not undergoing the same change process as the users new to Visma. Although these results were based on only three preliminary interviews, they still indicate possible group differences. The second hypothesis is as follows;

H2: The following attributes are likely to be significantly more important for end users in group 3 and 4 than group 2 and 1

- (a) user training*
- (b) user involvement,*
- (c) user support/communication*

The reasoning behind hypothesis H2 is firstly that the magnitude of change differs as discussed above. Ryan and Harrison (2000) find that as the magnitude of change the end-users experienced increased, the need for user training also increased. These findings support hypothesis H2. Secondly, the degree of experience affects the perceived ease of use for an EP system, also introduced above. The perceived ease of use can further be induced to affect the end-users need for training, because users that believe the system is easy have less need for training and vice versa. As there are also other factors affecting the perceived ease of use than experience, such as age and sex as discussed above, the next two hypotheses are;

H3: The following attributes are likely to be significantly more important for older individuals than younger ones

- (a) *user training*
- (b) *user involvement*
- (c) *user support/communication*

H4: The following attributes are likely to be significantly more important for females than males

- (a) *user training*
- (b) *user involvement*
- (c) *user support/communication*

Furthermore, regarding magnitude of change, Francalanci (2001) finds that as the magnitude increases, the resistance within the organizational generally increase as well. This supports the foundation of the next hypothesis;

H5: Regarding the attribute organizational resistance, end users in group 3 and 4 will be significantly more negative than users in group 1 and 2

As users in groups 1 and 2 have already used the system, the changes they experience are only minor compared to those users of group 3 and 4 will experience eventually. Based on this, it is feasible that users of group 3 and 4 are more resistant to the new system.

Social influence has been found to be more pronounced for older than younger individuals (Venkatesh et al., 2003) and more pronounced for females than males (Morris et al., 2005, Venkatesh et al., 2003). In our framework, the attribute *communication to key stakeholders and peers* was deducted from findings relating to social influence, and captures how workers are affected by other workers. Based on the findings above, it is interesting to see whether these differences exist at St. Olav's. The two next hypotheses are therefore;

H6: The attribute communication to key stakeholders and peers will be significantly more important to females than males

H7: The attribute communication to key stakeholders and peers will be significantly more important for older individuals than younger ones

Kulp et al. (2006) find that in employees in a case company were not motivated to employ electronic procurement tools due to already established relationships with local suppliers. The employees were familiar with particular manufacturers and their products, and did not want to change suppliers, i.e. order from suppliers in the e-procurement system. Finally, the last hypothesis, H8, is based on this notion, and is as follows;

H8: The following attributes will be significantly more important to group 3 and 4 than group 1 and 2

- (a) *Relationships with suppliers*
- (b) *preferred suppliers*

The reasoning behind the last hypothesis, H8, is that users in group 3 and 4 order only through other channels than Visma, such as phone and email. Therefore, they possibly order more from local or preferred suppliers, based on long established relationships. On the other hand, users in group 1 and 2 order through Visma where suppliers are selected by the central purchasing division. These users may therefore have less prominent supplier relationships or preferences regarding suppliers.

2.7 Summary

In this chapter, we have introduced the theoretical basis for this master's thesis. As pointed out, the theoretical basis was developed in the pre-diploma thesis, and consisted of both e-procurement and e-ordering literature, with change management serving as a backdrop. The result was a framework covering eleven CSFs, and specifying in what project phases each CSF is relevant. As the framework constitutes the basis for this master's thesis, it was introduced and its content explained. Last, the hypotheses were developed based theory and on CSFs from the framework deemed to be relevant to end-users.

3. Methodology

In this chapter, we will elaborate on the methodology underlying this thesis. The research questions must be considered for choosing an appropriate research design and research method. Choosing research method is also dependent on the hypotheses. As already indicated by the research questions and hypotheses, an empirical survey is implied as a proper approach to comply with the aim of this master's thesis. As mentioned in chapter 1, this aim includes partially validating the framework by investigating the large base of end-users and their beliefs and attitudes towards e-procurement. The following sections will entail justifications for choosing research design, research method, quality criteria, and finally an elaboration on potential statistical data analyses. An overview of the methodology chapter is illustrated by Figure 6 below.



Figure 6 – Methodology chapter overview

3.1 Research design and research questions

Bryman (2012) defines research design as the framework that guides the process of collection and analysis of data. The decisions and choices made in the research process are reflected in the choice of research design. Moreover, according to Bryman (2012), causal connections between variables, and generalization to a larger group than the often downscaled one which is investigated, are two aspects affected by the choice of research design. The first aspect concerning variables should be addressed more closely. As Bryman (2012, p. 48) defines it, “a variable is simply an attribute on which cases vary”. Often deemed as people, however, cases may also include everything from a household to an entire nation.

Among several research designs, Bryman (2012) highlights case studies as a research design. Yin (2014) elaborates on different components that are important for developing a case study design, and one of these components is concerning the research questions of the study. Furthermore, this component is strongly linked to which method of data collection that should

be used since the research questions steer the researchers somehow to which method to be chosen. Research questions in terms of “how” or “why”, according to Yin (2014), infer that a case study may be appropriate, whereas “who”, “what”, “where”, and “how much/many” infer a survey. By recapping our research questions, they are in the form of “what/which” and “how much”, hence indicating a research design comprising of a survey is favorable.

Survey research comprises, according to Bryman (2012, p. 60), “[...] a cross-sectional design in relation to which data are collected predominantly by questionnaire or by structured interview on more than one case [...] a single point in time in order to collect a body of quantitative or quantifiable data in connection with two or more variables [...], which are then examined to detect patterns of association.” In other words, a survey research design is by Bryman’s (2012) definition is subjected to be a cross-sectional design where data is collected at one occasion. Mills et al. (2009), on the other hand, introduce a research design called case study survey research. The definition of case study survey research design is: “[...] a research design in which a survey is administered to a case, either a small sample or an entire population of individuals, to describe an aspect or characteristic of that population” (Mills et al., 2009, p. 125). The definition is somewhat alike to Bryman’s (2012) definition of survey research. However, the notion of a case distinguishes the definitions. It is appropriate to follow the case study survey research design because the gained insight by the sample of the case is aimed to assist the case company. Furthermore, we argue that St. Olav’s has a central position in thesis, and some of the hypotheses regarding group differences are heavily influenced by the findings of the exploratory case study of the pre-diploma thesis. Mills et al. (2009) also state that case study survey research design also can entail a longitudinal design, which also enables collecting data at more than one point in time if desired. This is appropriate in our case as this thesis is a continuation of the pre-diploma thesis.

The advantages of this research design, according to Mills et al. (2009), are firstly its usefulness by allowing the researcher to sample a large number of people within a case both economically and quickly. Secondly, the design enables sampling individuals that are not necessarily in close proximity to the researcher. Thirdly, it allows information to be gathered either over a period of time or at one point in time.

However, Mills et al. (2009) point out some drawbacks to this research design being that survey data are self-reported information, and therefore only reflect what individuals think, or believe they should report at the specific point of time, for instance, the questionnaire is being completed. Moreover, as there is no manipulation of conditions, the research design cannot be

used to address causality between relationships to any degree of certainty. The individuals' honesty and willingness to participate are two important aspects influencing the validity of the information gathered. On the other hand, there are some steps that can be taken in order to minimize the effect of these drawbacks, and these will be addressed in the following sections.

Although it has only been mentioned implicitly, the research strategy of this thesis is of a quantitative nature. Bryman (2012) underlines other research designs that also allow for quantitative strategy in addition to case study and cross-sectional design, and one of these are comparative designs. A comparative design could also have been employed in this thesis, which allows assessing two or more populations (i.e. multiple companies) at one point in time, according to Bryman (2012). However, given the main focus on St. Olav's and their specific context and current situation in an e-procurement implementation, finding another similar public hospital for reference is challenging. Therefore, employing the case study survey research design is deemed as an appropriate approach.

3.2 Research Method

Mills et al. (2009) state that the most common instruments to collect data when employing a case study survey research design are either questionnaires or interview guides. Considering that interviews with a large group of end-users in order to collect data is out of the question because of time and other resource constraints, a self-completion questionnaire is the relevant method. Administering this instrument to a large sample can be done both efficiently and economically when considering the option of online surveys. Furthermore, this enables gathering a potential large response set that can be employed for testing the hypotheses and provide insight and understanding to the research questions.

3.2.1 Survey development

Before the research instrument could be developed, some decisions had to be made regarding which of the CSFs to address in the questionnaire and which of the implementation phases to focus on. Eleven CSFs are intuitively too many to investigate in one questionnaire as they all consist of several attributes that each demands at least one question item. In addition, it also makes sense that end-users cannot relate to all of the CSFs given that the framework is aimed at managers. Because of the implementation phases in the framework, this potentially gives a second dimension to the investigated CSFs that the respondents need to consider when completing the survey. Nevertheless, St. Olav's is deemed to be in the implementation phase, as mentioned as one of the limitations of this thesis in section 1.4. As the case company cannot

yet be placed in the post-implementation phase since the characteristics (fully integrated, pilots finished, adoption of the system, etc., cf. Figure 5) of this final phase are not yet established at St. Olav's. Including this final phase in the survey is not sensible because the end-users have to relate to potential future events. Moreover, including the first phase, pre-implementation, in the survey is not sensible either as it is easier for the respondents to relate to present issues than previous. However, one may argue that some of the respondents can perceive themselves to be in either one of the three phases, but the phases are mostly aimed at the project as the whole, not on an individual level. Overall, the investigated project phase of the survey is the current implementation phase, leaving the CSF, *security and authentication*, out of the question of being explored. This leaves us ten possible CSFs to be investigated. Considering that the framework is to be validated in terms of the end-users' point of view, CSFs were partially selected in terms of what these respondents may have opinions and beliefs about. Five criteria were employed for selecting which of the CSFs and which of their interdependent attributes to investigate, and are as follows:

1. CSFs that are relatable to end-users.
2. Feasibility, i.e. whether it is possible to measure the CSF and the attributes in a survey.
3. The CSF's possible link to the different user groups (the hypotheses).
4. The number of attributes to be investigated should be at least two for each of the selected CSFs.
5. To which further investigation on a CSF may be useful to St. Olav's.

Being relatable to end-users was an important criterion for the chosen investigated CSFs and attributes. Although this criterion is somewhat similar to the following criterion, feasibility is more specifically tied to the survey itself. For instance, if we were to investigate the CSF business case and project management in the survey, measuring the attributes *return on investment* and *total cost of ownership*, perhaps would not be feasible being explored in a survey. The third criterion was important in order to investigate attributes that could be linked to user groups. Moreover, the fourth criterion emphasizes that at least two attributes should be included of each investigated CSF. This is because indications of a CSF's overall importance are of greater utility when investigating at least two attributes. In addition, usefulness for St. Olav's is greater when more attributes are investigated.

Four CSFs were selected based on these criteria: *end-user uptake and training*, *change management*, *re-engineering the process*, and *e-procurement implementation strategy*. The argumentation of selecting these CSFs based on the criteria is presented below. The remaining CSFs were not seemingly relatable to end-users as opposed to the investigated four.

End-user uptake and training

This CSF and all of its attributes are intuitively clearly relatable to end-users. Moreover, by having opinions and beliefs about matters concerning their role in an e-procurement context, all of the eight attributes can seemingly be captured in question items in a survey with end-users as the respondents. The different user groups are also linked to the hypotheses via the research questions, and by testing the respective hypotheses, insight may be gained regarding this CSF's relevance. Although not all of the attributes are specifically hypothesized (for instance in-house training and on-demand training), the results may be useful to St. Olav's. As already elaborated on in chapter 2, adoption by end-users is critical when it comes to information systems such as e-procurement tools. By investigating this CSF, we aim to aid the case company with knowledge regarding end-user uptake, which may prove to be helpful when St. Olav's is going to implement SAP. In addition, the insight is also likely to be of utility value regarding the current EP implementation.

Change management

In this CSF, most of the attributes are aimed at a managerial perspective. Nevertheless, many of the attributes can be measured indirectly in terms of end-users' point of view. Three attributes of this CSF strike as directly relatable to end-users, and these are *communication to key stakeholders and peers*, *organizational resistance*, and *irreversible changeover to e-procurement*. In other words, it should be possible to translate these attributes to question items. However, as an irreversible changeover to EP is not yet the case at St. Olav's, it may be more cumbersome to capture this attribute in a survey than the first two attributes since the questions will be more of a hypothetical nature. Furthermore, group differences of the first two attributes are captured by hypotheses H5 and H6, whereas the third, *irreversible changeover to EP*, is not hypothesized. Regardless of this, *Irreversible changeover to EP* is of high interest for St. Olav's as the future transition to SAP implies that other channels than ordering using the selected electronic system will be discontinued. Therefore, examining this attribute in addition to *organizational resistance* more in depth may give indications to St. Olav's of how well employees will handle this transition.

Re-engineering the process

As with the previous CSF, most of the attributes of *re-engineering the process* are aimed at a managerial perspective. However, three attributes are seemingly directly relatable to the end-users in addition to being translatable to question items, and these are *previous work/order routines*, *transparency improvements*, and *compliance with purchasing procedures and standards*. We were not able to find theory that supports any group differences among the end-users regarding this CSF. However, investigating the current views and reflections on these topics in lack of hypotheses may still provide useful information to St. Olav's. The reason is that it can aid the consecutive implementation of e-procurement as well as SAP.

E-procurement implementation strategy

For this CSF as well, most of the attributes are more relevant in a managerial setting. However, there are some attributes that may be measured in terms of the end-users' point of view. Two attributes strike as more or less directly relatable to end-users, which are *preferred supplier* and *relationships with suppliers*. In addition, these attributes are linked to possible group differences in hypothesis H8. Investigating end-users' perspective on suppliers may prove useful to St. Olav's when end-user preferences are replaced with supplier contract agreements to a larger extent. By having information on this subject in terms of possible group differences, St. Olav's can be prepared to a greater extent on how to handle this coming transition.

Criterion	End-user uptake and training	Change management	Re-engineering the process	EP implementation strategy
1. Relatable to end-users	+	(+)	(+)	(+)
2. Survey feasibility	+	+	+	+
3. Link to RQ via hypotheses	+	+	-	+
4. At least two attributes	+	+	+	(+)
5. Usefulness to St. Olav's	+	+	+	+

Table 3 – Investigated CSFs (a plus sign means fulfilled and a minus sign otherwise)

As Table 3 above shows, three of the four chosen CSFs to be investigated fulfill all of the selection criteria. However, the parentheses on some of the criteria indicate that the criterion is fulfilled but have some room for improvement. *Re-engineering the process* fulfills all criteria except number 3. As hypotheses were cumbersome to deduce regarding this CSF due to lack of relevant theory related to end-users, any possible relationship found at St. Olav's may only be used in this research context, not to either support or falsify a hypothesis. Including this CSF is therefore done on the background of assessing its relevance in the framework and to provide St. Olav's with useful information.

After having decided which CSFs and which of their respective attributes to investigate, developing question items was the next step in the process. Following a typical iteration process, our approach to developing the questions is depicted in the Figure 7 below. Developing question items ourselves was deemed as a suitable approach as there is lacking research on these CSFs in the specific context of e-procurement. Moreover, Vaidya et al. (2006) express the need for an in-depth study to validate these CSFs in further research, as mentioned in section 2.5. However, to our knowledge, this is yet to be done, and research instruments consisting of question items in this context have not yet surfaced.

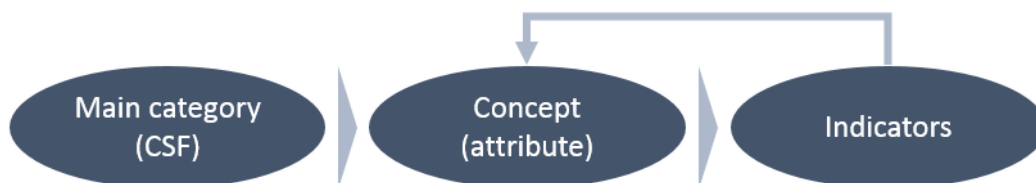


Figure 7 – Development of question items for the survey

When investigating a certain attribute (concept) of a certain CSF (main category), see Figure 7, it was important to decide what relevant indicators were. An example of this is regarding the attribute *user training* for which we needed multiple questions (indicators) to assess. After a question item had been developed, reviewing it in terms of the attribute was crucial in order to assess whether the item truly measures what it is aimed to measure. Other criteria such as ensuring the questions have a clear and easily understood language and avoiding double-barreled questions were also necessary precautions.

By following this approach, all the questions regarding the investigated CSFs were developed in a consistent manner. All of these questions were designed as closed questions (statements) aimed to be measured as multiple-indicators in terms of a Likert scale. According to Malhotra (2006), the Likert scale is one of the most common itemized scale, and that the end-points of the scale usually is in the form of “strongly disagree” and “strongly agree”. Using a Likert scale has many advantages, for instance, constructing and administering the scale are achieved without great effort for the researcher and that the respondents can readily grasp Likert scale questions. However, using a Likert scale implies that the respondents have to read the entire statement, which may result in a more time consuming process as opposed to using another form of itemized rating scales (Malhotra, 2006). Although constructing a Likert scale is quite straightforward as Malhotra (2006) states, selecting the number of response categories, or points for that matter, is necessarily not the same case.

Preston and Colman (2000) state that when using rating scales, the issue of the optimal number of response categories is still unresolved. According to Malhotra (2006), a general point of view is to have seven plus or minus two response categories. Furthermore, the human mind is capable of distinguishing about seven (plus or minus two), and having any higher than nine may be nonproductive, according to Preston and Colman (2000). Dawes (2008) finds that when rescaling a five- and seven-point scale, the two scales produced the same mean score. Furthermore, Dawes (2008) suggests that five- to seven-point scales have higher reliability and validity than scales with less than five categories and that using more than seven categories do not improve validity or reliability further. By keeping the number of categories to a minimum for the end-users and for the aforementioned reasons, the point of departure was to use a five-point Likert-scale. However, this topic was also included in the pilot tests, which is elaborated on in subsection 3.2.3, aiming to ascertain whether a five-point Likert scale is preferred or not.

After making decisions concerning which of the CSFs to investigate and how to assess their importance in terms of statements rated on a five-point Likert scale, the demographic part of the survey was developed. Recalling hypothesis H1, the predicted end-user groups at St. Olav's are:

1. End-users solely using Visma as their purchasing tool
2. End-users using Visma in addition to other channels
3. End-users currently not using Visma, but are projected to use it
4. End-users not using Visma, not projected to use it

Furthermore, the majority of the demographic questions of the survey were intended to assess hypothesis H1 by separating the respondents into the mentioned groups based on their answers. In addition to the Likert scale items, we employed interval/ratio and dichotomous variables. The latter was used in order to assess whether the respondents had used Visma or not in addition to determine gender, while the former was employed in order to assess the respondents' possible other channels of purchasing, frequency of Visma usage, frequency of ordering, etc.

The selected tool for developing the survey is SelectSurvey, created by Class Apps, to which access was granted by the Faculty of Social Sciences and Technology Management, NTNU. Allowing the choice of filtering, correct questions were displayed according to the respondent's previous answers such that, for instance, questions regarding Visma usage were not displayed to respondents who had never used Visma. Another example of filtering is that respondents who would answer that they solely use Visma for ordering would not be questioned about other channels of purchasing.

Although filtering determines some of the displaying of questions, the resulting survey/questionnaire comprises of seven preliminary demographic questions, 20 statements assessing the CSF end-user uptake and training, eight statements assessing change management, six statements assessing re-engineering the process, five statements assessing E-procurement implementation strategy, and three final demographic questions assessing age, sex, and which department the respondent work at. The entire survey can be found in Appendix A. Advised by Bryman (2012), socio-demographic questions should be left until the end of the questionnaire so that the survey ends with questions that are easy to reply to, as preferred by respondents. Following up on this advice, this is why the final question items in the survey to cover these socio-demographic aspects.

3.2.2 *Response bias*

Some of the mentioned drawbacks (in section 3.1) to the chosen research design are that responses may be affected by how the respondents believe they should report and that the validity of the gathered data is contingent upon their honesty. The latter can be linked to a type of response bias named acquiescence. According to Bryman (2012), multiple-indicator measures are especially prone to this type of response bias when respondent systematically report to a series of questions (for instance only use one of the item categories). A way to mitigate this issue is to employ question items that reflect opposite stances for the same concept by reversing some of the items. In the survey found in Appendix A, question items marked with an asterisk are reversed.

The first mentioned drawback can be linked to another type of response bias called social desirability bias. This type of bias occurs when a respondent is influenced by what is perceived as social desirable when completing a survey (Bryman, 2012). Nederhof (1985) states that there are indications that social desirability bias is reduced somewhat when using anonymous self-administration questionnaires. Although the effects of this type of bias can be present, using anonymous questionnaires should reduce it.

3.2.3 *Conducting the survey*

Notification form to NSD

In order for the method of data collection to comply with Norwegian research guidelines, the study was reported to Norwegian Social Science Data Services (NSD). A notification form including the survey, information letter to the respondents, description of what indirect personal identifying questions were to be included, aim of study, etc. was forwarded to NSD February 27th, 2015 and was approved March 10th, 2015. During this period, we conducted pilot tests and survey reviewing.

Pilot tests

Both Yin (2014) and Bryman (2012) stress the importance of a pilot study and pilot testing before the actual data collection of either a case study or survey research. Considering that survey research and case study survey research are somewhat similar, pilot tests are important also for the latter and chosen research design. Van Teijlingen and Hundley (2001, p. 1) underline this: “Pilot studies can be based on quantitative and/or qualitative methods and large-scale studies might employ a number of pilot studies before the main survey is conducted.” Furthermore, van Teijlingen and Hundley (2001) elaborate on pilot studies and that the term is used in two different ways in social science research. The term can refer to pre-testing of a

certain research instrument, such as a questionnaire, but can also refer to feasibility studies. The latter will not be discussed any further, as the relevant in our case is pre-testing of a research instrument. Pilot studies can unveil, for instance, issues with the wording, the order of questions, or the number of response categories (van Teijlingen and Hundley, 2001).

Before conducting pilot tests with individuals in the target population at St. Olav's, the survey was reviewed by two fellow master's degree students, our supervisor, a change manager at St. Olav's, and a member of the e-procurement project group at St. Olav's. The two fellow students also employ a survey as the method for data collection in their theses, and the reviewing with them was conducted individually and comments were noted. For the rest of the reviewers, the survey was deployed via SelectSurvey and text boxes were included on each page of the survey where possible comments could be added.

The major changes resulting from the reviewing were 1) wording revising on several questions, mainly on statements, 2) discarding some questions, mainly redundant questions, and 3) ideas and comments to the information pages within the survey.

Three individuals were chosen for the pilot test, of which two were individuals interviewed for the pre-diploma thesis. Of these two, one was an experienced Visma user and the other new to the system. The last of the three was an end-user who had never used Visma, but utilized other channels for purchasing. The main objectives of why conducting the pilot test were to identify ambiguities in the survey, record the time for completion, check whether sufficient information had been included, and to assess whether the five-point Likert scale was appropriate.

The pilot tests were carried out individually with the pilot subjects, and concerns and ideas they had were noted for following revising. The pilot subjects took the survey as they normally would, but stopped whenever something was unclear or they had ideas that should be taken into consideration. Therefore, the objective of recording the test subjects' time for completion could not be assessed as the pilot tests were conducted more in terms of a dialogue. On the one hand, van Teijlingen and Hundley (2001) suggest that the pilot study should be administered the same way as it would in the main study. This would imply that the ongoing process of feedback from the test subjects should have been omitted. On the other hand, as none of the statements (question items) in the survey had been based on other surveys, the main focus of the pilot tests was on these, especially the wording and possible ambiguities. Waiting until completion of the pilot tests would possibly had led the test subjects to forgetting some of the comments and concerns. In addition, including text boxes on the survey pages or handing the

test subjects a piece of paper and a pen to write their comments would have had the same level of intervening as a dialogue. However, in hindsight, the test subjects could have completed the survey first and then visit the survey once more for addressing issues by engaging in a dialogue.

The revising the pilot tests led to were similar to the initial reviewing: certain questions' wording were revised additionally, information pages were revised only to contain the necessary and the most important information text was changed to bold, and some redundant questions discarded. They also expressed that there should not be any additional statements and that a seven-point scale would be too comprehensive. Considering the latter, the choice of having a five-point rating scale, as elaborated on in section 3.1, was supported.

Target population/sample

Bryman (2012) states that it is unlikely that a researcher is able to send out a survey to the entire target population, and that sampling, especially statistical probability sampling, is a good tool in order to assure generalization. Moreover, there also exists an alternative to the statistical methods of sampling. This is defined by Bryman (2012) as convenience sampling, which is a non-probability sample that is merely available to the researcher by accessibility.

In our case, using convenience sampling, we acquired two email lists, which we were told to cover all end-users with a purchasing right employed at St. Olav's. The first of these was a list covering all registered Visma-users. For this list, we predicted that it would cover both group 1 and 2 as of hypothesis H1, as they are registered in Visma. The other list was an email list covering so-called "level 4 users", predicted to cover the two remaining groups of H1, not using Visma.

Specifically, 541 Visma-users were included in the first email list. The second list included 262 "level 4 leaders". Of these 262 individuals, 95 of them were also registered as Visma-users, and since they were redundancies, they were listed only once in order to avoid double deploying the survey. Therefore, the remaining 167 individuals registered as "level 4 leaders" were subject of belonging in either group 3 or 4 as of H1. The target groups are illustrated in Figure 8 below.

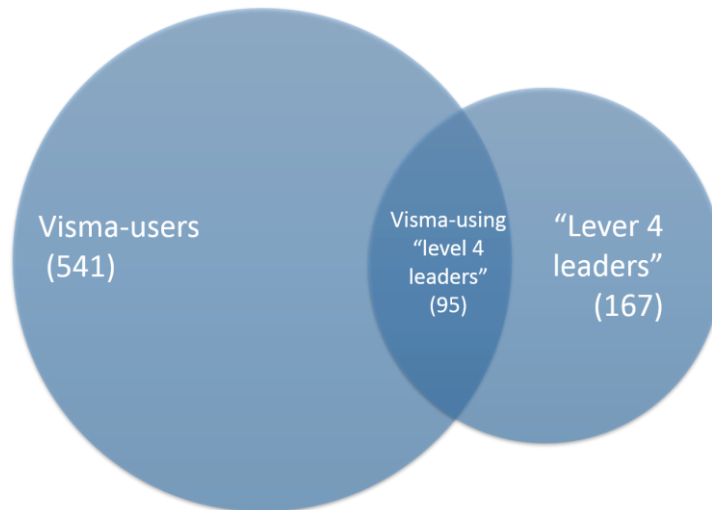


Figure 8 – Number of individuals in each target group

Survey deployment

For the sake of order, the final version of the survey developed in SelectSurvey was divided into three versions with different titles. These were then deployed using the built-in deploy tool in SelectSurvey, each to the corresponding target group (illustrated in Figure 8 above). Each individual received an email with an embedded link to the survey. Also included was an embedded decline link and general information regarding the survey both included in the email and an additional link to an information sheet. The Gantt-chart below in Figure 9 illustrates the time frame of when the survey was active and when follow-ups occurred.

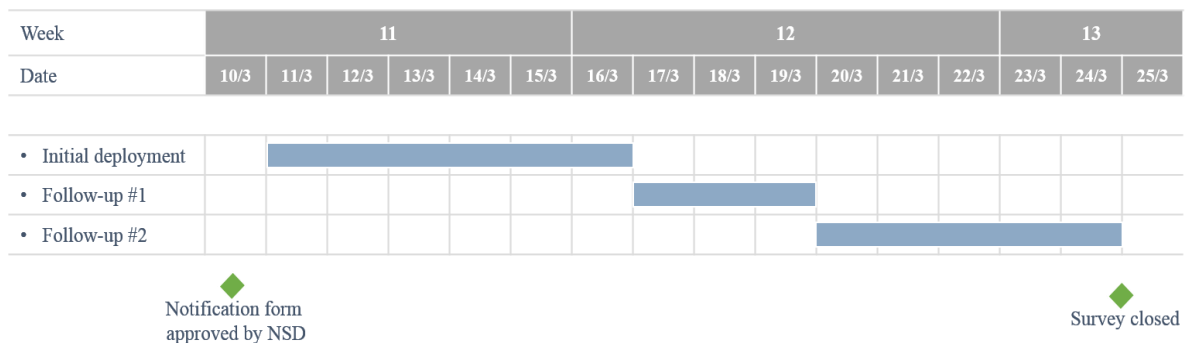


Figure 9 – Gantt chart of the survey deployment procedure

After the initial six days given (see Figure 9 above) of completing the survey and before sending a follow-up email, a total of 170 respondents had completed the survey with a distribution of 115, 24, and 31 of the Visma-user group, Visma-using “level 4 leaders”, and “level 4 leaders” using other channels, respectively. Before follow-up number two, March 20, 253 respondents had completed the survey with a distribution of 173, 36, and 44. When the survey was closed March 24, 2015, 306 respondents had completed the survey with a distribution of 203, 44, and 59. Since the total number individuals that received the survey was 803 (541+95+167), this leaves us with a response rate of 38.1 % (306/803). Incidentally, the response rate per target group, Visma-users, Visma-using “level 4 leaders”, and “level 4 leaders” using other channels, is 37.5 %, 46.3 %, and 35.3 %, respectively.

Asch et al. (1997) investigate survey response rates to mail surveys found in medical journals, and find that among nurses there is an average response rate (\pm standard deviation) of $61\pm 23\%$ and among health-care workers a mean response rate of $56\pm 24\%$. In a more recent study by Baruch and Holtom (2008), the authors investigate survey response rate levels and trends by covering 463 studies between 2000 and 2005 employing a questionnaire. The authors find a mean response rate of 54.7 % with a standard deviation (SD) of 23.9% when using email as the distribution method. When categorizing response rates by sectors, the mean response rate in the health care sector is 53.8 % (SD equal to 20.0 %). The latter average accounts for other distribution methods as well, which tend to have a lower response rate compared to mail surveys (Bryman, 2012). This may explain why the response rate is lower than the averages found by Asch et al. (1997). As our survey was deployed by email, it is more interesting to rest upon the interval found by Baruch and Holtom (2008) using this distribution method ($54.7\pm 23.9\%$). Although scoring below the three means, our response rate lies within the standard deviation by professional title, industry sector, and most importantly in the interval by distribution method. Furthermore, Bryman (2012) also underlines that when having a convenience sample, a low response rate may be of less significance. By considering this in addition to the argument of lying in the aforementioned intervals, a response rate of 38.1 % is satisfactory for the purposes of this master’s thesis.

Furthermore, Deutskens et al. (2004) investigate follow-ups when conducting a survey. The authors find that there is no optimal timing for the follow up. They do however find a slight preference towards sending a follow-up after one week instead of after two, although the results are not conclusive. In addition, follow-ups should be sent out quite early in order to take advantage of the fast turnarounds of online surveys, according to Deutskens et al. (2004). In

our case, the pilot test subjects were also asked about their preference for follow-ups, for which they considered three days or more to be adequate. We also monitored the response rates, which declined rapidly after three days. After six days of the initial deployment, no new responses were registered, so we decided to initiate the first follow up the seventh day, according to Figure 9 above. Three days after the first follow-up, day ten, we initiated the last and final follow-up, due to the considerations mentioned above and only a few new responses. The survey was finally closed five days later, as no new responses were registered two weeks after the initial deployment.

Non- and late-response

According to Bryman (2012), non-response can affect the external validity (which will be addressed in the following section). In a sample, there can be people not willing to participate. Therefore, it is important to address whether the responses of these individuals will affect the results by responding differently or if their opinions and beliefs would conform to the acquired results (Bryman, 2012). Considering that 61.9% decided not to complete the survey is not explored, there can be numerous reasons of why they did not take the survey. The somewhat similar response rate per target group as mentioned above can indicate that the target groups is not a key driver for non-response. A comprehensive follow-up of the non-responses would have been a resource intensive process considering that email was not an effective approach. Calling the non-respondents on the telephone instead could have been a way to investigate reasons for not completing the survey, but we do not have the capacity to do so. Nevertheless, there exists one possible approach to investigate non-response, which is in terms of assessing whether sex can be a driver. This will be further evaluated in section 4.1.

Kanuk and Berenson (1975) state that there is a type of bias involving a potential difference between early and late responses. In our case, late-response can be linked to the procedure of follow-ups, as discussed above. Without employing follow-ups but leaving the survey open for the same amount of time, the rate of non-response would most likely have been higher. It is further explored in section 4.1 whether the late-responses are biased in any way, thus affecting the total results.

3.3 Quality criteria

According to Bryman (2012), three criteria are important in order to assess the quality of quantitative research, namely reliability, validity, and replication. Following, the main types within each quality criteria for evaluating social research are presented. Some of these types will not be further investigated as they either are not relevant or in the scope of this thesis, mostly caused by time and other resource constraints.

3.3.1 *Reliability*

Bryman (2012) refers to reliability as the consistency of a measure of a concept, and states that reliability can be divided into three sub categories for deciding whether a measure is reliable.

Stability is the first of them, and this refers to which degree a measure is stable over time. The aim is to assure that results relating to a measure do not fluctuate over time. By choosing a longitudinal case study survey research design instead of one point in time, it would have been possible to assess stability. However, due to time constraints this was not in the scope of this thesis.

Internal reliability is the second of the sub categories, and addresses to which degree the scale or index that is employed is consistent with the indicators that constitute it. Furthermore, this includes to which extent the respondents' scores on one indicator is related to their scores on other indicators. Internal reliability applies to multiple-item measures, which for instance are surveys consisting of questions measured by a Likert scale. Generally, when using a Likert scale, each respondent's answers aggregated form an overall score. When this is the case, there exists a possibility that the indicators do not relate to the same thing, i.e. lack coherence. It is necessary that the indicators are related to each other to prevent that some of the items are indicative of something else (Bryman, 2012). A split-half method is commonly employed as a test to assess the internal reliability, where Cronbach's alpha is the most used test. Moreover, a computed alpha value of 1 denotes perfect internal reliability and 0 denotes no internal reliability. However, the alpha value may also be negative, which may, for instance, imply wrongly coding of one or more items. A rule of thumb is to aim for an alpha value higher than 0.8, according to Bryman (2012). In a meta-analysis by Peterson (1994), numerous recommended reliability levels are assessed of which below 0.6 is regarded as an unacceptable level. However, the range between 0.5 and 0.6 can also be regarded as fitting for preliminary research (Nunnally, 1967, as referred to in Peterson, 1994). Moreover, George and Mallery (2003) suggest that a Cronbach's alpha in the range 0.6-0.7 is questionable. Categorizing the

research in this master's thesis as preliminary is reasonable considering this is the first attempt to validate our framework. Therefore, the lowest acceptable value of Cronbach's alpha is set to 0.6, although in the lower region. Furthermore, an important property of the reliability measure Cronbach's alpha is that the number of items that are included in the calculation affect the alpha value positively (Peterson, 1994). In other words, wrongly including too many items in a reliability analysis may give an incorrect high alpha value.

Whether our survey exhibits internal reliability or not is determined by performing reliability analyses on the sets of question items of the four investigated CSFs' attributes. If the Cronbach's alpha value is lower than 0.6, actions must be taken in order to assure internal reliability of the attribute(s) in question.

Inter-observer consistency is the third of the sub categories for assessing reliability. As the name may indicate, this applies to research methods such as structured observations or structured interviews when there is more than one observer (Bryman, 2012). In order to assess that the observers interpret mutually and consistently, different actions and precautions can be made. However, as our method of data collection is a survey, inter-observer consistency only applies to the interpretation of the question items of the survey to which the respondents' responses are in free text (for instance, SRCord and ReasonNot in Appendix A). Ruling out that these responses are to be included in the testing of the hypotheses, inter-observer consistency is not assessed in this master's thesis. The responses to the free text question items, on the other hand, are included to give a demographic overview.

3.3.2 *Validity*

Validity is primarily concerned with the integrity of the conclusions a piece of research yields (Bryman, 2012). There are different types of validity that address this quality criterion. Following, the relevant types will be presented.

Face validity concerns whether the measure reflects the content of the concept or not, according to Bryman (2012). Establishing face validity may be done by asking people, preferably with experience, to act as referees on the matter whether the concept is reflected by the measure (Bryman, 2012).

In our case, assuring face validity is reflected upon the pilot test period, where both third parties and end-users at St. Olav's were included. The level of face validity may be considered medium, as there were some flaws to the pilot testing. Firstly, as we are attempting to measure concepts (the attributes of the four CSFs), knowledge on this matter is important to have a

saying whether the question items of an attribute are consistent. The group of third parties, except our supervisor, was lacking this knowledge, and input on whether question items are consistent did not surface. Despite this, wording on numerous question items were reviewed, redundancies removed, and the content of the information pages was revised, all because of the reviewing with the third parties. Secondly, the pilot test with the three test subjects could have been done differently, as mentioned in subsection (3.2.3). Allowing the individuals to complete the survey once without any form of interaction before taking it again, but the second time engaging in a dialogue with them, could have resulted in some improvements. For instance, we would have been able to track their time for completion. In addition, the setting of completing the survey the first time would have been more natural to them than the approach we chose. Still, our approach proved to be of good value as the test subject could share their insights and concerns at any point they themselves felt appropriate.

Concurrent and predictive validity deal with employing a criterion that is relevant to the concept in question (Bryman, 2012). Typically, a golden-standard constitutes this criterion to which the results of the measure are compared with. For example, if a new test designed to measure mathematical aptitude was given to a group of students, a typical way to evaluate concurrent validity is to correlate the new test with an already established math test developed by professors. If the criterion, the established math test, is employed contemporary, it is to determine concurrent validity. However, if the criterion is employed after a certain point of time, it is to determine predictive validity, i.e. whether the new measure has a predictive ability (Bryman, 2012). Addressing concurrent and predictive validity, on the other hand, is not in the scope of this thesis. One of the reasons for this is that finding an established test (a golden-standard) that contains the same CSFs and attributes is not possible, which inhibits the opportunity to address these types of validity.

Construct validity concerns whether a measure that is supposed to denote a concept reflects the concept in question or not (Bryman, 2012). Deducing hypotheses from theory that are relevant to the concepts is often how researchers are encouraged to estimate construct validity. However, determining construct validity by a single study is not possible, according to Peter (1981). Furthermore, in order to determine even a tentative construct validity, it is required that a series of validity studies and reliability studies are conducted (Peter, 1981). Considering this is out of the scope of this thesis, determining construct validity is left to potential further research.

External validity concerns generalizability, i.e. whether the results of a study can be generalized beyond the context of the research (Bryman, 2012). Lucas (2003) underlines that external validity concerns generalizability to theory, and hence external validity is closely tied to construct validity. This can be illustrated by the quote: “Because theoretical concepts are never measured directly, and because generalization can only occur through applying findings to theoretical concepts, measures in any study must relate to each other consistently with theoretically derived hypotheses for the study to have external validity.”(Lucas, 2003, p. 248). Therefore, in order to ensure external validity, the level of construct validity plays a crucial role. In addition, according to Lucas (2003), the situation designed to test the theory must be relevant to the theory, the level of replicability (see subsection 3.3.3), and consistency between the theory and the observations (i.e. internal reliability) are important conditions for ensuring external validity. Moreover, choosing either convenience and other non-probability sampling or probability sampling does not affect the ability of generalizing across either populations or settings. However, Lucas (2003) states that if generalizing to the specified larger population is the intention, probability sampling can increase the external validity. As mentioned in subsection 3.2.3, also non-response can influence the level of external validity. Therefore, addressing potential bias due to non-response must be done in order ensure that generalizability is not negatively influenced.

As external validity is dependent of multiple aspects from reliability, validity, and replication, the discussion of this quality criterion is left until section 6.2. However, it is worth mentioning that omitting to address construct validity affects the external validity of this thesis negatively. Nevertheless, considerations regarding generalizability may still be made, but with some caution.

3.3.3 *Replication*

Replication, and inherently replicability, deal with the research’s capacity of being replicated (Bryman, 2012). In order to eliminate the chance that a research is undermined due to the researchers’ potential biases, researchers tend to be highly explicit about their procedures, according to Bryman (2012). Accordingly, objectivity is a key ingredient, especially in social research as the researcher’s values and expectations can influence and color the results of a study.

Regarding the replicability of this thesis, it is attempted to keep the level of detail to a high degree. Specifically, by reporting relevant underlying statistics of the analysis, important decision-making, and considerations of the data, a replication of this study should be possible in the future. As with external validity, we will return to replication in section 6.2.

3.4 Statistical analysis

In this section, we will present briefly some approaches to the statistical analyses in this thesis. However, in order to prevent going into too much detail in this section, some further description of the chosen approaches is left until section 4.1 and 5.2. Following, the desired data analysis approaches for hypothesis testing are presented before moving on to a description of the alternative approach, namely exploratory factor analysis.

3.4.1 Data analysis

The chosen tool for statistical analysis is IBM SPSS Statistics version 22. Using this rather advanced software allows for managing responses and variables, analyzing data by, for instance, performing reliability tests, analyses of variance, etc. in a systematic way. Before moving on to the results acquired by using SPSS and discussing these in light of the hypotheses in chapter 4 and 5, some steps have to be addressed first.

The data for the three versions of the survey were exported from SelectSurvey as readable files to be used in SPSS. The three data files were merged into one with all of the 306 responses. In order to have the ability to assess their origin, a new population variable was added to track each case's (response) originating data file. As all of the variables were named vN (where N is an integer starting at 1), all variables were assigned new names in order to readily understand each variable's nature either it was demographic or in connection to either one of the investigated CSFs. The label of each variable was also changed to contain the English translation to prevent including a second language in the analysis. Moreover, since some variables were constructed as reverse items, cf. subsection 3.2.2, recoding of these variables was necessary in order to make sense of the scoring on these variables. For instance, scoring 5 (strongly agree) was changed to 1, 4 to 2, and so on. The variable names and labels are included in Appendix A in addition to the original Norwegian question items. In addition, all recoded variables are marked with an asterisk before the variable label.

In order to assess potential group differences, a new variable, *vGroup*, was added to categorize each case in one of the predicted four groups of H1 based on their response to question item *PRCNTvis* and *FUTuseVis* (see Appendix A). Due to filtering, *PRCNTvis* was only displayed to *Visma*-users and *FUTuseVis* only to respondents that had not used *Visma*. If a respondent replied that he or she only used *Visma* to complete orders (100% *Visma* use), he or she would end up in group 1 of the four groups. If the respondent replied less than 100% but more than 0% to the same question item, he or she would be categorized in group 2. Furthermore, if a respondent replied that he or she is projected to use *Visma* in the nearest future, he or she was categorized in group 3. If the reply to this question item was “no”, or that the respondent did not know whether he or she was going to make use of *Visma*, the respondent was categorized in group 4. The reasoning behind categorizing respondents that do not know whether they are going to make use of *Visma* in group 4 is that we find it rather unlikely that they are projected to use *Visma* if they have not received any information. We had to include this response category for potential respondents that do not know what *Visma* is. This means that group 3 may potentially be a bit larger while group 4 a bit smaller.

The desired analysis approaches consist of descriptive statistics to show the distribution and demographics of the respondents, analysis of variance (ANOVA) to assess differences across more than two groups, independent t-test to test differences in means when there are only two groups, and running different tests in order to assure that the requirements of the different analysis approaches were met.

The aim of the one-way ANOVA is, according to Ho (2014), to assess whether the means of more than two independent groups differ. By this notion, it is therefore a relevant statistical approach to test the hypotheses concerning either the user groups or age. One-way ANOVA is also known as an extension of the t-test (only two independent groups), and the statistical assumptions of the two approaches are similar. Ho (2014) states that the requirements for ANOVA are that there can only be one independent variables, which should consist of more than two groups, and that only one dependent variable can be tested in each analysis. Furthermore, there are two underlying assumptions of ANOVA. The first one is normality, i.e. that the dependent variable is normally distributed and the second is homogeneity of variance, i.e. that the groups have close to equal variance on the dependent variable (Ho, 2014).

Normality may be checked by Kolmogorov-Smirnov and Shapiro-Wilk tests, by statistical z value for the skewness and kurtosis, and visually by normal and detrended normal Q-Q plots, according to Ho (2014). However, Field (2013) advocates that because of the central limit theorem when a sample is large, significance tests such as Kolmogorov-Smirnov, Shapiro-Wilk, and z-tests will not be of much use since they will almost exclusively show a trend towards a normal distribution. In fact, according to Field (2013), if normality regardless is to be assessed using significance tests, the normality plots ought to be addressed as well. Visually, the normal Q-Q plot should show the data distribution falling more or less on the diagonal and the detrended Q-Q plot should show clustering of the distribution around the straight line with no pattern for normality to be assumed (Ho, 2014).

Homogeneity of variance is typically tested by the Levene statistic (an F-value) with the corresponding level of significance (Ho, 2014). As with the flaw pointed out for the significance tests to check normality, Field's (2013) argument also applies for homogeneity of variance, meaning that large samples almost exclusively do not violate homogeneity of variance. However, when the group sizes are not equal, homogeneity of variance should be assessed (Field, 2013).

3.4.2 *Alternative data analysis*

Using ANOVA to analyze group differences on the different attributes, i.e. test the hypotheses, is contingent upon the attributes having a satisfactory level of Cronbach's alpha, i.e. being internally reliable, in order for the analysis to make sense. Since we developed the survey ourselves, there is chance that the question items are more relatable to other concepts (covered by our survey as either included attributes or concepts that are not addressed). Therefore, addressing this potential issue is important for hypothesis testing. Exploratory factor analysis (EFA) is a promising candidate if the attributes fail to fulfill the requirement level of Cronbach's alpha at 0.6. The reasons of why performing an EFA, according to Field (2013), is to, for instance, either understand the structure of a set of variables by the variance they commonly share or to reduce a data set to a more manageable size without losing too much of the original information. A combination of these two causes constitute a good reason for why choosing EFA as a contingency plan. However, using EFA as a contingency plan has some drawbacks. Turning to explorative factor analysis may involve taking some steps away from the initial aim of validating the framework. Additionally, it may involve that some of the specific hypotheses regarding the attributes cannot be tested. This is because it is rather unlikely for the possible extracted factors to consist of the same variables as the attributes if the

attributes show a tendency towards low alpha values. On the contrary, the extracted factors may be used to assess group differences and to provide some understanding to the underlying questions of the hypotheses presented in section 2.6. However, the aforementioned relies heavily on the explorative factor analysis resulting in reasonable factors all aiming to measure seemingly appropriate variables.

Among of the requirements for EFA, Ho (2014) states that the sample size should be at least 100 and have at least five times as many cases as variables included in the EFA, although having ten times as many cases than variables is more in the acceptable range. Beavers et al. (2013) underline that there is no consensus on this matter and that, for instance, Gorusch (1983, as referred to in Beavers et al., 2013) and Norušis (2005, as referred to in Beavers et al., 2013) advocate at least 200 and at least 300 cases, respectively, regardless of the ratio of cases to variables. Field (2013) states that if the sample size is greater than 300, factors with few and low loadings (less than 0.40) may still be interpreted. Furthermore, a typical approach to assessing the sample adequacy for EFA is Kaiser-Meyer-Olkin measure of sampling adequacy (KMO). A KMO statistic under 0.5 is not acceptable, a value in the 0.70s is in the middling range, and in the 0.90s in the marvelous range (Field, 2013).

According to Ho (2014), the statistical assumptions of EFA consist of normality and sufficient significant correlations in data matrix. The latter refers to a test called Bartlett's Test of Sphericity, which can be used to test for the adequacy of the correlation matrix (Ho, 2014). The test is satisfactory if the result is significant ($p < 0.05$). Bartlett's Test of Sphericity is often tested along with the KMO.

3.5 Summary

In this chapter, we have discussed that case study survey as the research design for this master's thesis conforms well to the data collection method as a survey. Exploring the trends in the population concerning attitudes and beliefs about e-procurement through a self-completion questionnaire at one occasion may yield data that can be used to investigate certain relationships, primarily those hypothesized in section 2.6. In order being able to do this, we selected CSFs and their attributes using five criteria, which resulted in four of the CSFs covered by the implementation phase in the framework. Unable to find previous research instruments that have explored the CSFs in the same e-procurement implementation context, we developed the survey ourselves with question items related to the attributes. After presenting how we developed the survey, how it was conducted was elaborated on. This included pilot testing before we deployed the survey to all of the end-users with a purchasing right employed at St. Olav's. In addition, we discussed different kinds of bias and some of the approaches we included to lower their effect. Thereafter, we presented quality criteria important to quantitative research and evaluated the relevant ones. However, some of these criteria will be further evaluated in section 6.2 when addressing generalizability. Following, we presented the desired statistical approaches to test the hypotheses using ANOVA and independent t-tests. Additionally, we also presented an alternative data analysis approach concerning an exploratory factor analysis in case our survey did not exhibit internal reliability.

4. Results

In this chapter, we will first present some initial considerations such as late-response and outliers, before moving on to the demographics of the survey in section 4.2. Following, in section 4.3, we introduce the descriptive results and the internal reliability of the variables measuring the attributes of the four investigated CSFs.

We note that the survey included more demographic variables than the ones that are presented in this chapter, but not all of these proved useful in this thesis. Some variables that were seemingly appropriate at first when the survey was developed turned out to irrelevant as the research process progressed.

4.1 Initial considerations

Two initial tests were performed in order to check for potential outliers and late-response bias. In order to ascertain that follow-ups did not affect the overall results, it was necessary to check whether late-responses were significantly different from the earlier results. This was executed by assigning values to the respondents based on their time of completion, and then by comparing means to check whether the overall results were affected by time of completion. There were no indications that this was the case, hence late-responses did not affect the overall results.

According to Field (2013), outliers have an effect on the normal distribution of variables, and suggest that variables with a z-score of ± 3.29 are likely outliers and can be considered extreme. Furthermore, a z-score of ± 2.58 represent probable outliers and a z-score of ± 1.96 are potential outliers (Field, 2013). We computed z-scores for all of the 39 attribute variables by using SPSS and checked for outliers in both ends of the scale. Intuitively, using both the floor and roof value of a five-point Likert-scale should be acceptable, as outliers in multiple-indicator measures such as a Likert-scale could represent actual variance in the sample, which is favorable. In addition, there is disagreement in research whether or not to remove outliers, where many considerations can be taken. However, there is an overall tendency of agreement towards removing extreme outliers (Osborne and Overbay, 2004). Therefore, we removed the likely outliers of the data set (16 in total), as the probable and potential outliers did not seem to affect the normal distribution of their respective variables.

Non-response bias can affect the level of external validity, and as discussed in 3.2.3, due to time and other resource constraints we will only explore whether sex can be a driver for non-response. Below, Table 4 show the distribution of sex cross-tabulated with the three target groups.

Population	SEX			Total
	Male	Female	Male %	
Level4Visma	17	78	17.9%	95
Level4NoVisma	58	109	34.7%	167
VismaUsers	95	446	17.6%	541
Total	170	633	21.2%	803

Table 4 – Sex distribution of three target groups

Population	SEX			Total
	Male	Female	Male %	
Level4Visma	9	35	20.5%	44
Level4NoVisma	21	38	35.6%	59
VismaUsers	38	161	19.1%	199
Total	68	234	22.5%	302

Table 5 – Sex distribution of the respondents by the three target groups

Table 5 above is similar to Table 4, but shows the respondents of the survey instead of the entire populations. Comparing the percentage of males between the tables, it shows a difference of less than three percentage points for the three target groups. This indicates that sex is not a driver for non-response.

As stated in subsection 3.2.2, acquiescence bias can relate to a respondent systematically responding to a series of questions. We used reversed wording on some of the question items to prevent and spot this kind of bias, and checked all of the respondents' scoring on the 39 attribute variables to assess potential occurrence of this. None of the reversed items were equally responded to as their opposed non-reversed items except when both of the items were assigned the neutral value of 3. In addition, there were no occurrences of respondents systematically responding to a set of question items. Based on this notion, we argue that acquiescence bias was not apparent in our results.

4.2 Demographic variables

SEX	Frequency	Percent	Valid Percent	Cumulative Percent
Male	68	22.2	22.5	22.5
Female	234	76.5	77.5	100.0
Total	302	98.7	100.0	
Missing	4	1.3		
Total	306	100.0		

Table 6 – Distribution of sex

In Table 6 above, the distribution by sex of the respondents is shown, indicating a representation of females at 77.5 valid percent and a representation of males at 22.5 valid percent. We note that males are underrepresented among the respondents.

AGE	Frequency	Percent	Valid Percent	Cumulative Percent
1. 18-28 years old	3	1.0	1.0	1.0
2. 29-38 years old	45	14.7	14.8	15.8
3. 39-48 years old	92	30.1	30.3	46.1
4. 49-58 years old	114	37.3	37.5	83.6
5. 59-68 years old	50	16.3	16.4	100.0
Total	304	99.3	100.0	
Missing	2	.7		
Total	306	100.0		

Table 7 – Distribution of age

Table 7 above shows the distribution of the respondents by age. Referring to the split among young and old, introduced in section 2.6, we see that young people (group 1 and 2) represent only 15.8 percent of the sample. Furthermore, we note that none of the respondents are older than 68 years of age, as this option had zero respondents and was therefore omitted from the table above.

VISMAuse - Have you placed orders using Visma?	Frequency	Percent	Valid Percent	Cumulative Percent
1. No	128	41.8	41.8	41.8
2. Yes	178	58.2	58.2	100.0
Total	306	100.0	100.0	

Table 8 – Distribution of Visma usage

Distribution by the variable VISMAuse, whether respondents have placed orders through Visma or not, is shown in Table 8 above. The valid percent of respondents not having used Visma to place order and the ones that have counts to 41.8% and 58.2%, respectively. These numbers are further commented on when assessing hypothesis H1 in section 5.3.

PRCNTvis - Please estimate how many of your orders in total are conducted through Visma today.	Frequency	Percent	Valid Percent	Cumulative Percent
1. All of my orders (100%)	23	7.5	12.9	12.9
2. Most of my orders (76-99%)	67	21.9	37.6	50.6
3. Many (51-75%)	24	7.8	13.5	64.0
4. Some (26-50%)	20	6.5	11.2	75.3
5. Few (1-25%)	39	12.7	21.9	97.2
6. None (0%)	5	1.6	2.8	100.0
Total	178	58.2	100.0	
Missing	128	41.8		
Total	306	100.0		

Table 9 – Percentage of orders through Visma

Table 9 above show the characteristics of Visma-usage, only for the respondents replying 'yes' to the variable VISMAuse due to a filter. For the variable PRCNTvis, the distribution in valid percent shows a tendency towards users placing most of their orders through Visma. The valid percent of the respondents who only use Visma as their channel of purchase counts to only 12.9%. We note that almost one quarter (24.7 %) of the users place less than or equal to a quarter of their orders through Visma, while 2.8 valid percent of the respondents do not use Visma as of the time the survey was taken.

FUTuseVis - Is it projected that you are going to make use of Visma in the nearest future?	Frequency	Percent	Valid Percent	Cumulative Percent
1. Yes	42	13.7	33.9	33.9
2. No	24	7.8	19.4	53.2
3. Not that I know of	58	19.0	46.8	100.0
Total	124	40.5	100.0	
Missing	182	59.5		
Total	306	100.0		

Table 10 – Respondents projected to use Visma in the future

Table 10 above shows the distribution of the variable FUTuseVis of the respondents replying 'no' to VISMAuse. Of the respondents (in valid percent), 33.9% are projected to be using Visma in the nearest future, 19.4 % are not projected, and the remaining 46.8 % do not know whether they are projected to be using Visma or not.

SRCord - Which others channels for ordering are you using (besides Visma)?	Responses		Percent of Cases
	N	Percent	
1. Telephone	122	25.6%	44.0%
2. Email	214	44.9%	77.3%
3. Fax	51	10.7%	18.4%
4. Other	90	18.9%	32.5%
Total	477	100.0%	172.2%

Table 11 – Other ordering channels than Visma

Table 11 above shows the distribution of other ordering channels used by the respondents besides Visma. As the variable SRCord is a multiple response variable, this explains the total number of 477 responses to this variable, which exceeds the total number of responses to the survey. The results indicate that email is the most frequent used channel for ordering besides Visma, accounting for 44.9%, while 18.4 % still use fax to order. The response category “other” allowed respondents to specify with text, and consisted mainly of different web solutions and other persons ordering instead of the respondents.

ReasonNot – Reason for not ordering through Visma	Responses		Percent of Cases
	N	Percent	
1. I have not been given the offer to use Visma	47	35.1%	37.9%
2. I do not think it coincides with the supplies I am purchasing	19	14.2%	15.3%
3. I do not believe Visma will lead to any improvements	2	1.5%	1.6%
4. I believe that training will take too long time	3	2.2%	2.4%
5. Other	63	47.0%	50.8%
Total	134	100.0%	108.1%

Table 12 – Reasons for not ordering through Visma

The distribution of the responses to the multiple response variable ReasonNot is shown in Table 12 above. Besides not having been given the offer to use Visma (the first response category) accounting for 35.1%, the most frequent response to this question item is of other reasons, which account for 47.0 % of the responses. Almost everyone (61 out of 63) specified what other was. Roughly half of these responded that another person is responsible for ordering than the respondent herself, while one third responded that it was not appropriate for different reasons. The remaining responded other reasons such as lack of access, information, and training. One respondent replied that the system was too requiring and difficult to use, while not having problems with other systems.

As mentioned in subsection 3.4.1, the grouping variable vGroup was developed using the responses of PRCNTvis and FUTuseVis. Below, in Table 13, the distribution of respondents by vGroup is shown. Group 2 is the largest, users that use Visma but not exclusively, whereas the smallest is group 1, users that exclusively use Visma. Both these groups account for the respondents using Visma with a cumulative percent of 58.9. The remaining respondents belong in group 3 and 4, with the projected Visma group as being the smallest.

vGroup	Frequency	Percent	Valid Percent	Cumulative Percent
1. 100percentVisma	23	7.5	7.6	7.6
2. VismaButLessThan100percent	155	50.7	51.3	58.9
3. NonVismaProjected	42	13.7	13.9	72.8
4. NonVismaNonProjected	82	26.8	27.2	100.0
Total	302	98.7	100.0	
Missing	4	1.3		
Total	306	100.0		

Table 13 – The four groups by the variable vGroup

4.3 Characteristics of the multiple-item indicators

After presenting the characteristics and demographic distribution of the respondents in section 4.2, the results from the 39 multiple-item indicators regarding the investigated CSFs and their attributes will now be discussed.

In Table 40 in Appendix B, the number of responses, means, standard deviation, and maximum and minimum responses to all of the 39 variables are shown. As mentioned in subsection 3.4.2, normality can almost exclusively be assumed when the sample size is large. However, a visual check of the normality using Q-Q plots was conducted for each variable, indicating that the responses to all of the 39 variables were normally distributed.

The entire range of the Likert-scale is used by the respondents for almost all of the variables, except for three of them, according to Table 40 in Appendix B. Furthermore, the mean for each variable show a trend toward lying around 3 and the standard deviation between 0.5 and 1. Of 306 respondents, 266 (the valid N) of them replied to all of the variables. This is reflected upon the sum of responses to the variables where $N < 306$.

Reliability analysis, shown in Table 14 below, was performed on all attributes consisting of more than one variable. For those attributes consisting of only one variable, a Cronbach's alpha cannot be computed, and these are therefore indicated with 'N/A'. Only the attributes *user training*, *communication to key stakeholders and peers*, and *relationships with suppliers* show a Cronbach's alpha of higher than 0.6, with $\alpha=0.830$, $\alpha=0.605$, and $\alpha=0.610$, respectively. The remaining attributes show $\alpha < 0.6$. For some attributes alpha is negative, indicating that one or more variable should have been recoded. Recoding these did however not yield sufficiently high alpha. On the other hand, two of the attributes meet the criterion of $\alpha > 0.6$ when exploring the option "scale if deleted" in SPSS reliability analysis. This option evaluates the potential alpha value if one or more variables were omitted. This is the case for the second item in *in-house training* and the first in *organizational resistance*, in which the overall alpha value is adjusted to 0.748 and 0.793, respectively, if these items are omitted. Performing reliability analyses on the entire CSFs did yield higher alpha values, however, as mentioned in subsection 3.3.1, this is an improper approach as too many variables are included thus yielding an artificially high alpha.

CSF (abbreviation)	Attributes (abbreviation)	Variable	Cronbach's Alpha	Scale if item deleted
End-user uptake and training (EUA)	User involvement (ui)	EUAui1	-0.470	-0.223
		EUAui2		0.161
		EUAui3		-0.140
		EUAui4		-0.002
	User support/communication (usc)	EUAusc1	0.533	N/A
		EUAusc2		
	User training (ut)	EUAut1	0.830	0.783
		EUAut2		0.758
		EUAut3		0.750
	In-house training (iht)	EUAiht1	0.456	-0.032
		EUAiht2		0.748
		EUAiht3		0.125
	On-demand training (odt)	EUAodt	N/A	N/A
Self-learning (sl)	EUAsl	N/A	N/A	
Focus on users with low IT-experience (itX)	EUAitX	N/A	N/A	
End-user benefits (eub)	EUAeub1	0.501	0.403	
	EUAeub2		0.523	
	EUAeub3		0.388	
	EUAeub4		0.346	
	EUAeub5		0.543	
Change management (CHM)	Organizational resistance (or)	CHMor1	0.444	0.793
		CHMor2		-0.364
		CHMor3		-0.039
	Communication to key stakeholders and peers (cks)	CHMcks1	0.605	N/A
		CHMcks2		
	Irreversible changeover to EP (icep)	CHMicep1	-0.063	-0.168
		CHMicep2		-0.389
CHMicep3		0.316		
Re-engineering the process (RENG)	Previous work/order routines (pwo)	RENGpwo1	0.307	N/A
		RENGpwo2		
	Transparency improvements (ti)	RENGti1	0.186	0.409
		RENGti2		-0.456
		RENGti3		0.254
Compliance with purchasing procedures and standards (comp)	RENGcomp	N/A	N/A	
E-procurement implementation strategy (EPIS)	Relationships with suppliers (rel)	EPISrel1	0.610	N/A
		EPISrel2		
	Preferred suppliers (ps)	EPISps1	0.529	0.226
		EPISps2		-1.597
EPISps3		-0.573		

Table 14 – Internal reliability for survey attributes

4.4 Summary

In this chapter, we have presented the initial considerations of the gathered survey data, which entailed assessing different types of bias and their potential effect. Late-response did not show an effect on the overall results, sex could not be considered as a driver for non-response, and acquiescence bias was not found to be apparent in the data. In addition, outliers were addressed, but only the extreme values were removed.

Next, several demographics variables were presented giving insight into the sample. The sample was also grouped into four different groups for the variable vGroup, showing how many of the respondents use Visma, too which degree, and whether they are project to use Visma in the future or not. This variable will be used throughout section 5.3 where we test our hypotheses. Last, characteristics of the survey, such as internal reliability, was shortly discussed as it forms the basis for the exploratory factor analysis in the next chapter.

5. Analysis

In this chapter, an exploratory factor analysis (EFA) is carried out and explained in detail. In section 5.3, the results of the EFA are used to analyze group differences in order to address the hypotheses from section 2.6. In section 5.4, results of relevant other tests are presented.

5.1 Exploratory factor analysis

The results from the reliability analysis of the attributes in Table 14 show that most of the attributes are seemingly consisting of variables not related to the same concept (the investigated attributes). *User training* and *communication to key stakeholders and peers*, and *relationships with supplier* exceeded the minimum criteria of $\alpha > 0.6$, and may be used to test relevant hypotheses. Furthermore, if scaling the two attributes *in-house training* and *organizational resistance*, these may also be used for relevant hypotheses. However, hypotheses H2, H3, H4, and partially H8 are affected by insufficient alpha values for their respective attributes, indicating that another approach to analyzing the data should be considered.

Exploratory factor analysis was presented in subsection 3.4.2 as an alternative data analysis approach. The alternative approach was included in order to mitigate the event of low Cronbach's alpha values for the attributes, as the overall results in Table 14 show.

5.1.1 Prerequisites of EFA

The total number of respondents that have responded to all of the 39 multiple-indicator variables is 266 (Valid N in Table 40, Appendix B). This is close to the requirement by Field (2013) of 300 samples presented in subsection 3.4.2, and still exceeding Ho's (2014) and Gorusch' (1983, as referred to in Beavers et al., 2013) requirement of 100 and 200 samples, respectively. The samples-to-variables ratio is 6.82 (266/39), which is above the minimum requirement of 5 but less than the more acceptable range of 10. Overall, we consider this prerequisite as fulfilled since the sample size of 266 is sufficient for running EFA, although not optimal.

As stated in section 4.3, all of the responses to the 39 variables were normally distributed, hence the assumption normality may be assumed. The assumption of sufficient correlations in data matrix is addressed in the following section using Bartlett's Test of Sphericity. The adequacy test of KMO is also included. These assumptions were introduced in subsection 3.4.2.

5.1.2 Initial tests

An initial EFA using principal axis factoring was run on all the 39 variables in order to assess sampling and correlation matrix adequacy by testing for KMO and Bartlett's test of Sphericity. Table 15 below shows Bartlett's test of Sphericity of being significant at $p < 0.05$, which indicates that sufficient correlations in the data matrix may be assumed. A KMO value of 0.725 is in the middling range and well above the minimum of 0.5, as introduced in subsection 3.4.2. Therefore, sample adequacy may be assumed. However, using the option of anti-image in SPSS, the sampling adequacy of each variable may also be assessed. According to Field (2013), the elements on the diagonal should, as with the KMO, be greater than 0.5. By investigating the anti-image correlation matrix, two variables do not fulfill the requirement of 0.5. These are EUAui4 and EUAeub2 with values of 0.497 and 0.465, respectively, which indicates that these two should be omitted in factor extraction.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.725
Bartlett's Test of Sphericity	Approx. Chi-Square	3185.228
	df	741
	Sig.	.000

Table 15 – KMO and Bartlett's Test

Moreover, choosing how many factors to extract is a key question of EFA. Table 16 below shows a truncated output of the total variance explained using EFA. All 39 variables are listed, but only the first 13 have an eigenvalue greater than 1. This is an indication of that up to 13 extracted factors is appropriate. The two rightmost columns under *Extraction Sums of Squared Loadings* further indicate their appropriateness in terms of percent of variance and cumulative percentage. As the column *Cumulative %* indicates, the payoff of adding an extra factor is incrementally declining, leaving the total possible variance explained at 49.193% for 13 factors.

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.625	14.422	14.422	5.189	13.305	13.305
2	3.309	8.483	22.906	2.820	7.232	20.537
3	2.372	6.082	28.987	1.903	4.879	25.416
4	2.086	5.349	34.336	1.628	4.174	29.590
5	1.843	4.726	39.062	1.419	3.638	33.228
6	1.783	4.572	43.634	1.329	3.408	36.636
7	1.532	3.928	47.563	1.018	2.611	39.247
8	1.335	3.422	50.985	.827	2.120	41.367
9	1.322	3.389	54.374	.766	1.964	43.330
10	1.159	2.973	57.346	.676	1.732	45.063
11	1.116	2.861	60.208	.622	1.595	46.658
12	1.061	2.722	62.930	.520	1.333	47.991
13	1.018	2.611	65.540	.469	1.202	49.193
14	.930	2.383	67.924			
...			
39	.182	.467	100.000			

Table 16 – Total Variance Explained - Extraction Method: Principal Axis Factoring.

Figure 10 below shows a scree plot with the number of factors plotted on the x-axis and the eigenvalue on the y-axis. A scree plot may also be used to assess how many factors to extract (Field, 2013). Two points of inflection can be found in the plot, indicated with the blue and red lines. By assessing the corresponding factor number of these two points, it may be indicated that the total number of factors to be extracted is between five and nine factors.

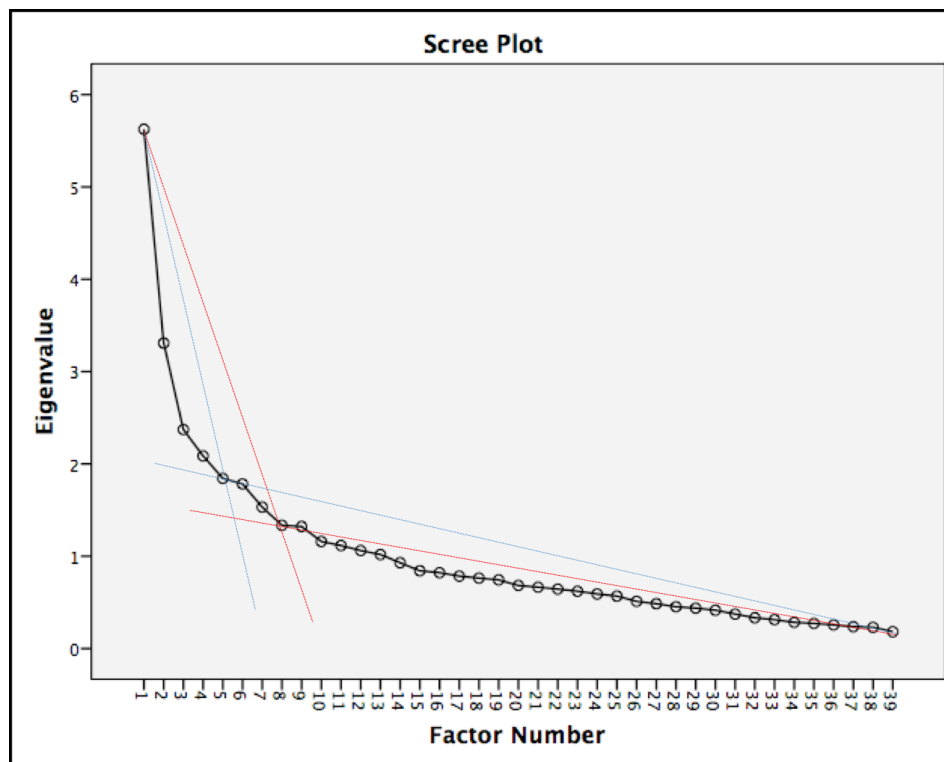


Figure 10 – Scree plot for factor extraction

5.1.3 Factor extraction

Extracting the factors can be described as an iteration process, aiming to find an appropriate factor structure while at the same time preserving as much of the total variance explained as possible (Field, 2013). Even though the number of factors advised by the scree plot is between five and nine, the maximum of 13 factors was also addressed. In other words, exploratory factor analyses were run for all of the possible outcomes with the set number of factors from five to thirteen. For each respective number of factors, different solutions were tested, attempting to find the most meaningful factor structure.

The cutoff level of minimum factor loading for all possible outcomes was set to 0.32, as advised by Yong and Pearce (2013) and Beavers et al. (2013), which corresponds to a variable at least explaining 10% of the variance in the factor. Occurrences of cross-loadings (i.e. variables with a loading on more than one factor) should be omitted if they cannot be explained to be appropriate, as suggested by Yong and Pearce (2013). We considered this for each respective number of factors to be extracted. Moreover, the chosen rotation method was the oblique method instead of orthogonal, as the orthogonal method assumes that all factors are non-correlated (Field, 2013). The assumption of non-correlated factors is not suitable for the CSFs' and attributes' contextual nature, because all variables measure end-users' perceptions and it is likely that these are somewhat correlated. In addition, we did a preliminary test with orthogonal rotation which confirmed our assumption, the resulting factor structure did not make sense and had too many cross-loadings. Therefore, the *direct oblimin* rotation method, which corresponds to oblique, was chosen. Missing values were excluded using the option of *exclude cases listwise*, in which case a respondent with missing data for any of the variables is excluded. Also, finding a factor consisting of loadings not exclusively being either positive or negative was mitigated by recoding one or more variables in the factor as long as this made sense, otherwise it would indicate that one or more variables are inversely correlated to the rest of the variables in a factor. The final result of the factor extraction can be found in Table 17 below.

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
EUAut2	.778						
EUAut1	.708						
EUAsl	.705						
EUAut3	.662						
EUaitX	.609						
EUAodt	.479						
EUAusc1	.454						
CHMor1		.720					
EUAeub5		.614					
CHMicep1		.553					
RENGcomp		.467					
RENGti1		.449					
RENGti2		.334					
CHMcks1			.646				
CHMcks2			.608				
EPISrel1				.593			
EPISps1		-.335		.592			
EPISrel2				.521			
EPISps2				.456			
EPISps3				.394			
CHMor3					.813		
CHMor2					.809		
RENGpwo2							
EUAeub4						-.757	
EUAeub3						-.677	
RENGpwo1						-.431	
EUAeub1						-.332	
EUAut1						-.325	
EUaiht3							.747
EUaiht1							.688
Eigenvalues	4.575	2.516	1.621	1.350	1.139	1.063	.746
% of variance	15.248	8.387	5.403	4.499	3.797	3.544	2.487
% cumulative variance	15.248	23.636	29.038	33.537	37.335	40.879	43.366
Cronbach's α	0.826	0.683	0.605	0.644	0.790	0.681	0.748

Table 17 – Factor extraction sorted by size

Table 17 shows the factor structure of the 30 variables included in the final EFA, covering seven promising factors. The nine missing variables were omitted because of unexplainable cross-loadings, too low factor loadings, and anti-image requirements. For each of the seven factors, the eigenvalue, percentage of variance, cumulative percentage of variance, and Cronbach's alpha measure of internal reliability are included.

Some properties of the chosen EFA structure above must be elaborated on. Firstly, all of the alpha values exceed the minimum set criterion of 0.6, which we consider as satisfactory, see subsection 3.3.1. Secondly, the total variance explained of 43.366 % must be addressed. Tinsley and Tinsley (1987) note that less than 50% of the total variance explained by a factor solution occurs frequently. However, the authors also note that the range of 30-40% is alarming. In a meta-analysis of variance accounted for in EFA by Peterson (2000), the author

notes that there is no consensus of what is an acceptable measure of total variance explained. In the 568 articles analyzed in the meta-analysis, the mean of total variance explained is 56.6%. Even though the total variance explained of our EFA is about 13 percent points below this, it is still above the alarming range, and we argue that the variance the EFA explains is bearable, especially considering the lack of consensus. Furthermore, Tinsley and Tinsley (1987) note that factors that only explain a small percentage of the variance (around 1%) are not realistically prone to be of either practical or theoretical significance. Although being the factor that explain the least variance, factor 7 accounts for 2.487 %, which is still within the limits of Tinsley and Tinsley (1987).

A third property of the factor structure shown in Table 17 is the number of variables in each factor. Factor 3, 5, and 7 only consist of two variables, whereas factor 1, 2, 4, and 6 consist of five or more variables. Yong and Pearce (2013) suggest that a factor consisting of only two variables should be kept if the two variables in the factor has a higher correlation than 0.7. However, this should be considered as a guidance, and some room for interpretation is allowed. The variables within factor 3, 5, and 7 correlate 0.435, 0.653, and 0.597, respectively. As can be seen, factor 5 and 7 is closer to the acceptable level than factor 3. Although none of the factors met the guiding criterion, we will include them in the testing of hypotheses if deemed reasonable. This will be further discussed in the next subsection when determining the appropriateness of the factors.

The last property is the sign of the factor loadings. According to Field (2013), all the items within a factor should have the same sign. In our case, we note that all the items in factor 6 are negative, which is indicates consistent coding of the variables.

The results of the adequacy tests, KMO and Bartlett's Test of Sphericity, are presented in Table 18 below. The KMO is now a bit higher than in the initial test and the Bartlett's test of Sphericity is significant at $p < 0.0001$. Both results are adequate.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.743
Bartlett's Test of Sphericity	Approx. Chi-Square	2519.419
	df	435
	Sig.	.000

Table 18 – KMO and Bartlett's Test for the final factor structure

The factor correlation matrix of the final EFA is presented in Table 19 below. Although most of the coefficients are low, there still exist some correlation coefficients that show signs of interdependency between the factors, for instance factor 1 and 6 ($r = -0.373$).

Factor	1	2	3	4	5	6	7
1	1.000	.044	-.064	.172	-.055	-.373	.245
2	.044	1.000	.007	-.103	-.176	-.117	-.075
3	-.064	.007	1.000	.080	.084	.000	.180
4	.172	-.103	.080	1.000	.084	-.149	.035
5	-.055	-.176	.084	.084	1.000	-.017	.046
6	-.373	-.117	.000	-.149	-.017	1.000	-.055
7	.245	-.075	.180	.035	.046	-.055	1.000

Extraction Method: Principal Axis Factoring.

Rotation Method: Oblimin with Kaiser Normalization.

Table 19 – Factor Correlation Matrix

5.1.4 Appropriateness of the seven factors from the EFA

The resulting seven factors from Table 17 above need to be interpreted and named for the new factor structure to have any value. Moreover, factors and their relations to the framework will also be discussed in section 5.3 in order to address the hypotheses.

Factor 1: End-user training

The first factor we name *end-user training*, because all seven items the factor contains, see Table 20 below, relate to how the end-user feel about training and what they prefer. By studying this factor, we see that it only contains elements from the original CSF *end-user uptake and training*, which is somewhat reassuring because it indicates that the deviation from the original framework structure is not that high. Furthermore, compared to the original CSF, this factor contains items from five of a total of eight attributes. Most predominant is the attribute *user training* which have all three items from the original CSF structure present in this factor, but it is still evident that this factor is more comprehensive than the attribute *user training*. Interestingly, we note that all items relating to the attribute *in-house training* has been removed although they explicitly mention training. This could indicate that *in-house training* is a different kind of preference as elaborated on when discussing factor seven. The fact that an item from the attribute *user support and communication*, EUAusc1, is present although not mentioning training explicitly is not that surprising because this factor is also closely related to training. For example, when a person is being trained, they need a contact person and a support network to get going. The items of factor 1 are shown in Table 20 below.

Item	Statement
EUAut1	Training is necessary when a new ordering system is introduced
EUAut2	I have a need for training regardless whether I perceive the ordering system as easy to use or not
EUAut3	Adapted training could increase my willingness to use a new ordering system
EUAst	*In order to teach me to use a new ordering system, self-learning is favorable rather than training with a competent person
EUAitX	I have had the need for training every time a new electronic system has been introduced
EUAodt	*If I have the opportunity to be given training on demand, I will prefer this over planned training
EUAusc1	It is of high importance to me to have a contact person/support network available when a new ordering system is introduced

Table 20 – EFA: End-user training

Factor 2: Attitude and perceptions towards EP

By a similar fashion as the previous, factor 2 is named *attitude and perceptions towards EP*, because all items in the factor relate to how end-users perceive an e-procurement system and what attitudes they have toward such systems. Compared to the original CSF framework, this factor is a combination of both *change management*, *re-engineering* and *end-user uptake and training*, supporting the fact that the many of the attributes from the original framework had too low internal reliability. The items of factor 2 are shown in Table 21 below.

Item	Statement
CHmor1	It is advantageous for St. Olav's Hospital that purchases are done electronically rather than using other methods such as email, phone or fax
EUAeub5	I believe an electronic ordering system can simplify my workday
CHMicep1	St. Olav's Hospital has the need for <u>one</u> universal electronic ordering system, which does not allow other methods for ordering such as email and phone, given that all purchases can be done in this system
RENGti1	*I believe an electronic ordering system will impact my purchasing overview negatively
RENGti2	It is of high importance to me that a potential new electronic ordering system gives a better history overview such as <u>previous purchases and consumption</u>
RENGcomp	It is important to me that all of my purchases comply with the current regulations for public purchases, even if this means that I have to do all of my purchases in an electronic ordering system

Table 21 – EFA: Attitude and perceptions towards EP

Factor 3: Communication to key stakeholders

Factor 3 is identical to the attribute *communication to key stakeholders* (CHMcks) from the original framework, and does therefore not need to be renamed. It measures to which degree other workers influence the respondents, and tries to capture the importance of communicating to key stakeholders that can influence other peers. Although this factor only contains two items, which is the lower bound as mentioned in the previous section, the factor is obviously kept, because it is identical with the attribute *communication to key stakeholders* from the original framework. The items of factor three are shown in Table 22 below.

Item	Statement
CHMcks1	If a colleague recommends using a new system, this would have a greater impact on me than a recommendation from top management
CHMcks2	If I know that several other colleagues are using a new ordering system, this would influence me positively to be using the system myself

Table 22 – EFA: Communication to key stakeholders

Factor 4: Supplier relationships

Factor 4 is named *supplier relationships*, because it captures to which degree current relationships are important to the respondents. Compared to the original CSF framework, we see that all the items relate to the two attributes *preferred suppliers* and *relationships with suppliers*. This may indicate that the two concepts we tried to measure is in fact the same concept, supplier relationships. In Table 17, we see that the items EPISps1 has a cross-loading

where the factor analysis shows that it can also belong to factor two, *attitude and perception towards EP*. As mentioned in subsection 5.1.3, it is permissible to have cross-loadings as long as these can be explained. In this case, we see that the factor loading is higher towards this factor, 0.592 compared to -0.335. In addition, the item does not mention electronic ordering explicitly such as all the other items in factor two, although one can argue that acting more on general agreements is a result of electronic ordering. Still, the relation is weaker than for the others. Based on this notion, we argue that the EPISps1 belongs to this factor. The items of factor 4 are shown in Table 23 below.

Item	Statement
EPISps1	*It is more important that St. Olav's as an entity acts more on general agreements than that I can continue to order from today's suppliers
EPISps2	It is important to me that in the future, I have the opportunity to order the same supplies as I do today
EPISps3	*I am open to order other supplies which are equivalent to those I order today if St. Olav's hospital has better agreements on these
EPISrel1	I have a close relationship to one or more suppliers (for example if you have a good communicational relationship to some suppliers)
EPISrel2	I perceive it negatively if the present relationships to suppliers are affected as a result of an introduction of electronic ordering

Table 23 – EFA: Supplier relationships

Factor 5: *Status quo*

The fifth factor is named *status quo* because it measures whether the respondents is satisfied with status quo or not. Compared to the original framework, this factor aligns well with the attribute *organizational resistance*. One item however, CHmor1, has been removed from attribute and is included in factor two instead, *Attitude and perceptions towards EP*. In other words, *status quo*, is identical to the attribute *organizational resistance* if we were to delete CHmor1 as suggested by the column “scale item if deleted” in Table 14. In hindsight, this makes sense because CHmor1 measures perception towards EP instead of measuring whether the respondent is satisfied with status quo or not. Furthermore, it is worth mentioning that the item RENGpwo2 is also placed within the *status quo* in the factor analysis, according to Table 17, but is suppressed because its factor loading is less than 0.32. However, the content of this item is in line with that of *status quo*. The items of factor 5 are shown in Table 24 below.

Item	Statement
CHmor2	I am satisfied with today's ordering system
CHmor3	I do not see the need for a new ordering system

Table 24 – EFA: Status quo

Factor 6: Communication

Factor 6 is named *communication*, as the items measure how important communication is to the respondents. More specifically, the factor entails both two-way communication and communication of benefits. Although the item RENGpwo1 may not seem to fit in at first sight, we argue that it is related to EUAeub1. If communication of benefits is important for a respondent, it should also be important for the respondents that their work routines are considered. That way, the benefits the respondents can achieve can be communicated and may be more trustworthy. Furthermore, communication is also necessary in order to map the previous work routines of the respondents. The items of factor 6 are shown in Table 25 below.

Item	Statement
EUAeub1	It is of high importance to me that a new ordering system simplifies my workday
EUAeub3	Which benefits I can experience must be communicated to me prior to introducing a new ordering system
EUAeub4	If benefits are communicated to me, this could have a considerable positive impact on my attitude towards the new system
EUAui1	It is of high importance to me that I have the opportunity to provide input if a new system for electronic ordering were to be introduced
RENGpwo1	It is important that my previous work routines are mapped prior to the introduction of a new ordering system (for example returns, rush orders, etc., I am supposed to follow up)

Table 25 – EFA: Communication

Factor 7: In-house training

Finally, the last factor is named *in-house training*, because it measures whether the respondents prefer in-house training as opposed to external training. Compared to the framework, this factor is very close to the original attribute *in-house training* although the item EUAiht2 is not included. This item tried to measure whether respondents felt in-house was safer than external training, but is evidently not linked directly to their preference of in-house training. As with factor 5, *status quo*, this factor is the same as the result we would get if we were to scale the attribute *in-house training*, according to Table 14. Moreover, as mentioned when discussing the first factor, *end-user training*, we noticed that *in-house training* was not included although mentioning training explicitly. We argue the reason is that the two factors measure different things. Whereas the first measures the need for end-user training, this last factor measures the preference for in-house training versus external training. The items of factor 7 are shown in Table 26.

Item	Statement
EUAiht1	I prefer that training is given by employees at St. Olav's rather than external personnel
EUAiht3	Employees at St. Olav's may understand my needs for training better than external personnel

Table 26 – EFA: In-house training

The seven factors obtained from the EFA are summarized in Table 27 below.

Factor	Name	Abbreviation
1	<i>End-user training</i>	<i>EUT</i>
2	<i>Attitude and perceptions towards EP</i>	<i>EPPERCP</i>
3	<i>Communication to key stakeholders</i>	<i>CHMcks</i>
4	<i>Relationships with suppliers</i>	<i>SUPREL</i>
5	<i>Status quo</i>	<i>STATQUO</i>
6	<i>Communication</i>	<i>COMM</i>
7	<i>In-house training</i>	<i>IHT</i>

Table 27 – Named EFA factors

5.2 Analysis methods

In order to run ANOVAs, mean variables were computed and added to the data set in SPSS, not only for the EFA factors, but also for the three attributes, EUAut, CHMcks, and EPISrel, for which $\alpha > 0.6$. Only cases that had provided responses to at least 25% of the respective variables for each mean were added. For example, when adding the mean variable for the factor SUPREL, a response ratio of four out of five question items was set as the criterion. For mean variables consisting of only two or three items, answers for all of the items were required. According to Walpole (2012), the mean of two or more normally distributed variables is also normally distributed. By this notion, we treat all of the computed mean variables as normally distributed, as their input variables are not violating the assumption of normality.

In addition to normality, homogeneity of variance is the second statistical assumption underlying ANOVA, and this second assumption was assessed for each ANOVA test. Field (2013) advocates that homogeneity of variance is especially important to address when group sizes are not equal. Considering the relevant tables presented in section 4.2, equal group sizes is not the case. Furthermore, if homogeneity of variance was violated ($p < 0.05$), a robust test of equality of means using Welch was applied. If the Welch test yields a significant p-value ($p < 0.05$), this implies that further tests can be performed.

As mentioned in subsection 3.4.1, ANOVA is an extension of the independent t-test. Independent t-tests were used to assess differences across sex since the number of independent groups are only two (males and females). Furthermore, the statistical assumptions of independent t-tests are the same as for ANOVA.

As using ANOVA only gives indications that there potentially are group differences, a method for addressing where the potential differences stem from is needed. In order to assess the specific group differences for hypotheses regarding age or vGroup as the independent

variables, planned contrasts for these respective ANOVAs were applied to assess the specific hypothesized differences. Field (2013) states that planned contrasts can be regarded as independent t-tests in which only two independent groups are tested. As mentioned in section 2.6, the split for age was made at 40 years for categorizing whether an individual is considered young or old. By using the existing response categories to the variable (cf. Table 6), response category 1 (18-28 years) and 2 (29-38 years) is set to correspond as ‘young’ whereas the remaining categories correspond to ‘old’. Our split is only one year skewed in favor of the ‘old’ group, which is not a big deviation from the recommended split. Contrasts were applied to assess a potential significant difference between age group 1 and group 3, 4, 5, and 6 (Contrast₁) and to assess a potential significant difference between age group 2 and group 3, 4, 5, and 6 (Contrast₂). Furthermore, for the hypotheses concerning vGroup as the independent variable, contrasts were also applied. This was done by testing for a difference between, cf. Table 13, group 1 and group 3 and 4 (Contrast₁) and between group 2 and group 3 and 4 (Contrast₂).

A significant difference in both contrasts and that the directions are not contradicting the hypothesis would imply that the respective hypothesis is supported. A non-significant difference imply that the test results are inconclusive because a *p*-value greater than 0.05 means the alternative hypothesis is false, but not that the null-hypothesis is true (Field, 2013). Specifically, this is because a non-significant result means that the effect is not big enough to be found, but not that is equal to zero.

Following is a presentation of different methods to calculate effect sizes of the statistical tests. As Field (2013) suggests, effect sizes should be presented even though the test is non-significant.

In order to calculate the effect sizes of the independent t-tests for the hypotheses regarding sex as the independent variable, Cohen’s *d* is used. Field (2013) states that a Cohen’s *d* of 0.2, 0.5, and 0.8 correspond to a small, medium, and large effect, respectively. The underlying calculation of Cohen’s *d* is as follows:

$$d = \frac{M_1 - M_2}{\sqrt{\frac{SD_1^2 + SD_2^2}{2}}}$$

where *M* is the mean and *SD* is the standard deviation.

In order to measure the overall effect size for the ANOVAs, omega squared (ω^2) was used instead of the traditional R^2 . As Field (2013) suggests, R^2 tends to be biased since it does not reflect the effect size in the population. ω^2 , on the other hand, also accounts for the degrees of freedom in the model (df_M) and the means square of the residual (MS_R), which gives a better understanding of the effect size. According to Field (2013), a value of ω^2 of 0.01, 0.06, and 0.14 correspond to a small, medium, and large effect, respectively. The underlying calculation of ω^2 is as follows:

$$\omega^2 = \frac{SS_M - (df_M)MS_R}{SS_T + MS_R}$$

For calculating the effect sizes for the results of planned comparisons using contrasts, Field (2013) suggests using an alternative to Cohen's d for independent t-test since these planned comparisons are merely t-tests. The effect coefficient is named $r_{contrast}$, and an r-value of 0.10, 0.30, and 0.50 correspond to a small, medium, and large effect, respectively. The underlying calculation of $r_{contrast}$ is as follows:

$$r_{contrast} = \sqrt{\frac{t^2}{t^2 + df}}$$

where t is the t-value of the test and df is the corresponding degrees of freedom.

The aforementioned relations and elaborations on the ANOVA and the independent t-test are employed in the following section, which includes testing of the hypotheses.

5.3 Testing the hypotheses

5.3.1 Hypothesis 1

The first hypothesis, H1, can readily be addressed by looking at Table 13 in section 4.2. All the four groups hypothesized to exist at St. Olav's were found, and the hypothesis is therefore confirmed. We note that the group solely using Visma is quite small, accounting for only 7.6 % of the users. This is not that surprising as we were uncertain whether any users did in fact only use Visma for their purchases. Furthermore, we also see that the group projected to use Visma is also quite small at 13.7 %. However, we were only able to find two such users when conducting interviews for the pre-diploma thesis. Figure 11 below is updated with the actual numbers, based on whether the respondents have used Visma to order or not. We note that according to Figure 8, in subsection 3.2.3, 76 % (541/803) of the respondents are registered as Visma users. Figure 11 below, however, indicate that many registered users have not yet placed a single order. We argue that Figure 11 below is a more realistic measure of how many use Visma, compared to registered users.

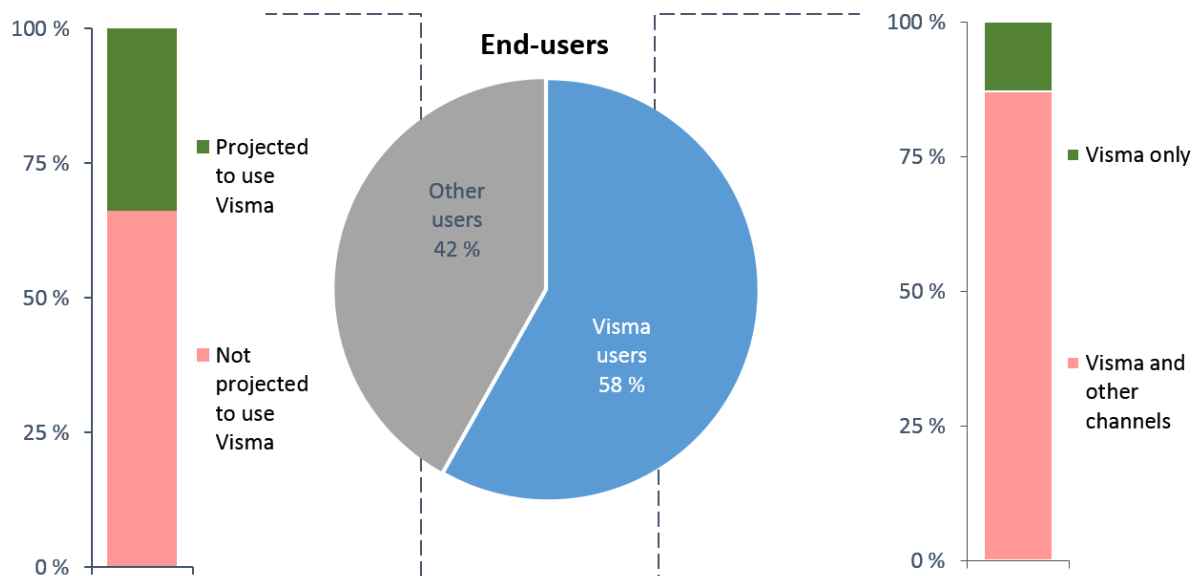


Figure 11 – Actual distribution of user groups

5.3.2 Hypothesis 2, 3, and 4

Hypotheses H2, H3 and H4, contain the three attributes a) *user training*, b) *user involvement*, and c) *user support/communication*. Part a) of the hypotheses is easiest to address because we can use the attribute *user training* from the original framework directly as its $\alpha > 0.6$. We also found a similar factor from the EFA called *end-user training*, although this factor was more comprehensive. We can use this factor from the EFA as well, and compare the results to those of the attribute *user training* to see if this may give any insight. The remaining parts of the hypotheses, b) and c), cannot formally be addressed because the relating attributes show $\alpha < 0.6$. However, we did find a similar factor in the EFA that can be used to discuss what the results could have been. Specifically, the factor called *communication*, as discussed in subsection 5.1.4, contain elements of both attributes *communication/support* and *user involvement*. Thus, we can employ this factor to see whether the results do in fact vary across the groupings of users.

Hypothesis (Type of test)	Dependent variable	Independent variable	Assumptions	Results
H2a (ANOVA)	EUAut (Attribute)	vGroup	Normality assumed and homogeneity of variance (F=0.402, p=0.752) not violated.	There was no significant effect of EUAut on the different user groups (vGroup), F(3, 295)=1.104, p=0.348, $\omega^2=0.001$. Although not significant, applying contrasts: Contrast ₁ : t(295)= -1.067, p=0.287; $r_{contrast_1} = 0.06$ (r<small). Contrast ₂ : t(295)= -1.622, p=0.106; $r_{contrast_2} = 0.09$ (r<small).
	EUT (EFA factor)	vGroup	Normality assumed and homogeneity of variance assumed (F=0.526, p=0.665)	There was no significant effect of EUT on the different user groups (vGroup), F(3, 298)=1.015, p=0.386, $\omega^2 \approx 0$. Although not significant, applying contrasts: Contrast ₁ : t(298)=-1.179, p=0.239; $r_{contrast_1} = 0.07$ (r<small).. Contrast ₂ : t(298)= -1.580, p=0.115; $r_{contrast_2} = 0.09$ (r<small).
H3a (ANOVA)	EUAut (Attribute)	AGE	Normality assumed and homogeneity of variance (F=2.215, p=0.067) assumed	There was a significant effect of EUAut on age, F(4, 296)=4.349, p=0.002, $\omega^2 = 0.043$. Applying contrasts: Contrast ₁ : t(296)=-2.109, p=0.036; $r_{contrast_1} = 0.13$. (small<r<medium). Contrast ₂ : t(296)=-1.904, p=0.058; $r_{contrast_2} = 0.11$ (small<r<medium).
	EUT (EFA factor)	AGE	Homogeneity of variance assumed (F=1.116, p=0.349) and normality assumed.	There was a significant effect of EUT on age, F(4, 299)=7.454, p=0.000010, with a medium effect of $\omega^2 = 0.078$. Applying contrasts: Contrast ₁ : t(299)=-3.053, p=0.002; $r_{contrast_1} = 0.17$ (small<r<medium). Contrast ₂ : t(299)=-2.326, p=0.021; $r_{contrast_2} = 0.13$ (small<r<medium).
H4a (Independent t-test)	EUAut (Attribute)	SEX	Normality assumed and homogeneity of variance (F=0.078, p=0.780) assumed	On average, males scored lower on the variable EUAut (M=4.0245, SE=0.1006) than females (M=4.0909, SE=0.0491). This difference, -0.0664, BCa 95% CI [-0.2744, 0.1416], was, however, not significant t(297)= -0.628, p=0.530; d=0.08 (d<small).
	EUT (EFA factor)	SEX	Normality not violated and homogeneity of variance assumed	On average, males scored lower on the variable EUT (M=3.674, SE=0.079) than females (M=3.864, SE=0.039). This difference, - 0.189, BCa 95% CI [-0.3549, -0.0240], was significant

Table 28 – Testing of hypothesis H2a, H3a, and H4a

As Table 28 above shows, hypothesis H2a is inconclusive both when testing the effects of the dependent variables EUAut (attribute from the framework) and EUT (from the EFA) on the independent variable, vGroup. None of these tests were significant and both tests expressed an effect size close to zero. However, when applying contrasts for EUAut, both of the t-statistics is negative, which indicates that the means of group 1 and 2 are lower than group 3 and 4. Although not being significant or exceeding the small effect size threshold of 0.10, this complies with the direction hypothesized in H2a. The very same indication can be seen when applying contrasts for EUT, the direction complies with H2a, but it is non-significant and the effect size lower than 0.10.

Furthermore, regarding H3a, EUAut showed an overall significant effect on age with an effect size closer to a medium effect. When applying contrasts to assess the group differences, both t-statistics are negative and of statistical significance, which indicate that the direction of the hypothesis is supported. However, contrast₁ is of a slightly higher significance and a slightly higher effect size than contrast₂ ($p=0.036$ vs. $p=0.058$ and $r=0.13$ vs. $r=0.11$, respectively). Nevertheless, H3a is supported when testing for EUAut's effect on age. When it comes to EUT's effect on age, the results show the same characteristics as of EUAut. The overall test is significant with a medium effect size, the contrasts revealed both contrasts being significant in the same direction as the hypothesis, and contrast₁ showed a higher significance and effect size than contrast₂ although both in the range between a small to a medium effect. Overall, H3a is supported when testing for both EUAut and EUT's effect on age.

Moreover, hypothesis H4a was inconclusive when testing for EUAut's effect on sex. The direction of the difference in means (user training being more important to females than to males), however, did comply with H4a, although the effect size did not exceed the small-threshold. On the other hand, when testing EUT's effect on sex, the results show a significant difference between males and females with a small to medium effect. The difference of means (females scoring higher than males) does also comply with the direction of H4a. Overall, H4a is supported when using EUT as the dependent variable.

Testing of the hypotheses H2bc, H3bc, and H4bc will, as stated, be tested using the *communication* factor from the EFA as the dependent variable. The findings are listed in Table 29 below.

Hypothesis (Type of test)	Dependent variable	Independent variable	Assumptions	Results
H2 b) + c) (ANOVA)	COMM (EFA factor)	vGroup	With normality assumed and homogeneity of variance assumed (F=1.055, p=0.368)	There was no significant effect of COMM on the different user groups (vGroup), F(3, 297)=1.273, p=0.284, $\omega^2 = 0.002$. Although not significant, applying contrasts: t(297)= -0.547, p=0.585; $r_{contrast_1} = 0.03$ (r<small). Contrast ₂ : t(297)= -0.993, p=0.321; $r_{contrast_2} = 0.06$ (r<small)
H3 b) +c) (ANOVA)	COMM (EFA factor)	AGE	Normality and homogeneity of variance assumed (F=1.016, p=0.399).	There was no significant effect of COMM on age, F(4, 298)=0.768, p=0.547; $\omega^2 = 0.003$. Although not significant, applying contrasts: Contrast ₁ : t(298)= -0.176, p=0.860; $r_{contrast_1} = 0.01$ (r<small). Contrast ₂ : t(298)= -1.397, p=0.163; $r_{contrast_2} = 0.08$ (r<small).
H4 b) + c) (Independent t-test)	COMM (EFA factor)	SEX	Normality assumed and homogeneity of variance assumed (F=2.322, p=0.129).	On average males scored higher on COMM (M=3.8706, SE=0.0592) than females (M=3.8260, SE=0.0346). This difference, 0.0446, BCa 95% CI [-0.0964, 0.1857], was, however, not significant t(299)=0.623, p=0.534; d=0.09 (d<small).

Table 29 – Testing of hypothesis H2, H3 and H4 parts b) + c)

H2b and c and H3b and c are inconclusive because the overall COMM's effect on vGroup and AGE are non-significant. In addition, the overall effect size of the tests is below the small-threshold. However, when applying contrasts, the negative t-statistics express that the difference between the user groups and age groups comply with the direction stated in the hypotheses. On the other hand, the effect sizes none of the contrasts did exceed the small-threshold.

Furthermore, H4b and c) are also inconclusive since COMM's effect on sex as the independent variable of the t-test is not significant. This is exemplified by the low difference in mean between males and females in addition to the measure of effect size being below the small-threshold. In addition, the direction of the hypotheses is violated as the results indicate that *communication* is slightly more important to males than females.

5.3.3 Hypothesis 5

Regarding hypothesis H5, the original attribute *organizational resistance* had $\alpha < 0.6$ and cannot therefore be used to address the hypothesis directly. However, if we scaled it by omitting the first item of the attribute, the alpha would be 0.79, which is equivalent to using the factor *status quo* from the EFA. Therefore, we address hypothesis H5 by testing the factor *status quo*'s effect on the user groups. These findings are listed in Table 30 below.

Hypothesis (Type of test)	Dependent variable	Independent variable	Assumptions	Results
H5 (ANOVA)	STATQUO (EFA factor)	vGroup	Normality assumed, but the homogeneity of variance test violated ($F=2.774$, $p=0.042 < 0.05$). Welch's robust test yielded $F(3, 76.986)=3.745$, $p=0.014$, which indicates that the adjusted F ratio is significant, allowing for further tests.	There was a significant effect of STATQUO on the user groups (vGroup), $F(3, 295)=3.772$, $p=0.011$; $\omega^2 = 0.027$. Applying contrasts: Contrast ₁ $t(31.283)=2.973$, $p=0.006$; $r_{contrast_1} = 0.47$ (medium < r < large). Contrast ₂ $t(178.263)=2.398$, $p=0.018$; $r_{contrast_2} = 0.18$ (small < r < medium). (None of the contrast tests assume equal variances due to the significant Levene's test.)

Table 30 – Testing of hypothesis H5

STATQUO's overall effect on the user groups, with vGroup as the independent variable, is statistically significant, according to the results found in Table 30. The overall effect size was in the range of small to medium. Both of the contrasts showed a significant difference between the respective groups. For contrast₁, the effect size is close to large when testing for the difference between group 1 and group 3 and 4, whereas the effect for contrast₂ is between small and medium for the difference between group 2 and group 3 and 4. More importantly, the positive t-statistics for the contrasts imply that the direction stated in hypothesis H5 is supported. Overall, hypothesis H5 is supported.

5.3.4 Hypothesis 6 and 7

The hypotheses H6 and H7 can be tested directly with the attribute *communication to key stakeholders and peers* since its $\alpha > 0.6$ (cf. Table 14). The results are listed in Table 31 below.

Hypothesis (Type of test)	Dependent variable	Independent variable	Assumptions	Results
H6 (Independent t-test)	CHMcks (Attribute)	SEX	Normality and homogeneity of variance assumed (F=0.078, $p=0.780$).	On average, males scored higher on the variable CHMcks (M=3.2794, SE=0.0780) than females (M=3.1048, SE=0.0459). This difference, 0.1746, BCa 95% CI [-0.0111, 0.3603], was however not significant $t(295)=1.851$, $p=0.065$; $d=0.26$ (small<d<medium) .
H7 (ANOVA)	CHMcks (Attribute)	AGE	Normality and homogeneity of variance (F=1.252, $p=0.289$) assumed.	There was no significant effect of CHMcks on age, $F(4, 294)=0.650$, $p=0.627$; $\omega^2 = 0.005$. Although not significant, applying contrasts: Contrast ₁ : $t(294)=0.066$, $p=0.947$; $r_{contrast_1} = 0.004$ (r<small). Contrast ₂ : $t(294)= -1.123$, $p=0.262$; $r_{contrast_2} = 0.07$ (r<small).

Table 31 – Testing of hypothesis H6 and H7

Even though the difference between males and females showed a small effect, H6 was inconclusive because CHMcks' effect on sex was non-significant. We note that the difference in means contradicts the stated direction of H6 as CHMcks was more important to males than females. Moreover, H7 is also inconclusive since CHMcks' overall effect on age was not significant. In addition, the overall effect size was below the small-threshold. The different signs of the t-statistics for the contrasts imply that contrast₁ (difference between group 1 and group 3 and 4) contradicts the stated direction of H7 whereas contrast₂ supports it. However, as none of the contrast tests were significant nor exceeded the small-threshold, these results can be regarded as a merely random occurrence.

5.3.5 Hypothesis 8

Hypotheses H8 a) and b) contain the two attributes *relationships with suppliers* and *preferred suppliers*, respectively. We found that the attribute *relationships with suppliers* had $\alpha > 0.6$, which is sufficient to address hypothesis H8a directly. From the EFA, we found the factor *supplier relationship*, which is more comprehensive than the attribute from the original framework although they are almost identical in name. The factor, as discussed in 5.4.1, consists of both attributes *relationships with suppliers* and *preferred suppliers*. We can therefore only address the whole hypothesis, part a) and b) together, with the factor. It does not

make sense to address part b) when the EFA showed that the items for the two attributes are perhaps measuring only one concept, *supplier relationships*. However, we can still compare the results of the factor and the attribute across the user groups. The results are listed in Table 32 below.

Hypothesis (Type of test)	Dependent variable	Independent variable	Assumptions	Results
H8a) (ANOVA)	EPISrel (Attribute)	vGroup	Normality and homogeneity of variance assumed (F=2.297, p=0.078)	There was no significant effect of EPISrel on the user groups (vGroup), F(3, 297)=1.077, p=0.359; $\omega^2 = 0.0008$. Although not significant, applying contrasts: Contrast ₁ : t(297)=0.324, p=0.746; $r_{contrast_1} = 0.02$ (r<small). Contrast ₂ : t(297)=1.158, p=0.248; $r_{contrast_2} = 0.07$ (r<small).
H8 a) + b) (ANOVA)	SUPREL (EFA factor)	vGroup	Normality and homogeneity of variance assumed (F=1.336, p=0.263)	There was no significant effect of SUPREL on the user groups (vGroup), F(3, 298)=0.755, p=0.520; $\omega^2 = 0.0002$. Although not significant, applying contrasts: Contrast ₁ : t(298)=0.807, p=0.420; $r_{contrast_1} = 0.05$ (r<small). Contrast ₂ : t(298)=1.455, p=0.147; $r_{contrast_2} = 0.08$ (r<small).

Table 32 – Testing of hypothesis H8 part a) and b)

H8a was inconclusive, as EPISrel's overall effect on the user groups was non-significant. In addition, the overall effect size of the test was close to zero. Although not exceeding the small-threshold for the effect sizes nor being statistically significant, both contrasts showed a positive t-statistics, which contradicts the direction stated in H8a.

When testing for the hypothesis in its entirety, H8 a)+b), using the factor SUPREL, the overall effect of SUPREL on the user groups was non-significant with a very low overall effect size. As for the results of H7a using EPISrel as the dependent variable, both contrasts showed a positive t-statistics, which contradicts the direction stated in H8a. Overall, H8 a)+b) was inconclusive.

5.3.6 Aggregated results of hypothesis testing

The results of the hypothesis tests, with their effect size, are summarized in the Table 33 below. When stating the effect size of the contrasts for the respective hypotheses, we will present the

mean of the two contrasts: $r_{mean} = \frac{r_{contrast_1} + r_{contrast_2}}{2}$

H#	Hypothesis	Effect size	Supported (√) or Inconclusive (X)
H1	Four distinct groups of end-users can be found at St. Olav's	N/A	√
H2a	<u>User training</u> is likely to be significantly more important to end-users in group 3 and 4 than group 2 and 1	Overall r<small for both EUAut ($r_{mean} = 0.08$) and EUT ($r_{mean} = 0.08$)	X
H3a	<u>User training</u> is likely to be significantly more important to older individuals than younger ones	Overall small<r<medium for both EUAut ($r_{mean} = 0.12$) and EUT ($r_{mean} = 0.15$)	√
H4a	<u>User training</u> is likely to be significantly more important to females than males	d<small for EUAut (d=0.08), but small<d<medium for EUT (d=0.30)	√
H2b) +c)	<u>User involvement</u> and <u>user support/communication</u> are likely to be significantly more important to end-users in group 3 and 4 than group 2 and 1	Overall r<small ($r_{mean} = 0.05$)	X
H3b) +c)	<u>User involvement</u> and <u>user support/communication</u> are likely to be more important to older individuals than younger ones	Overall r<small ($r_{mean} = 0.05$)	X
H4b) +c)	<u>User involvement</u> and <u>user support/communication</u> are likely to be significantly more important for females than males	d<small (d=0.09)	X
H5	Regarding the attribute <u>organizational resistance</u> , end users in group 3 and 4 will be significantly more negative than users in group 1 and 2	Overall medium<r<large ($r_{mean} = 0.32$)	√
H6	The attribute <u>communication to key stakeholders and peers</u> will be significantly more important to females than males	small<d<medium (d=0.26)	X
H7	The attribute <u>communication to key stakeholders and peers</u> will be significantly more important for older individuals than younger ones	Overall r<small ($r_{mean} = 0.04$)	X
H8a	<u>Relationships with suppliers</u> will be significantly more important to group 3 and 4 than group 1 and 2	Overall r<small ($r_{mean} = 0.045$)	X
H8a) +b)	<u>Relationships with suppliers</u> and <u>preferred suppliers</u> will be significantly more important to group 3 and 4 than group 1 and 2	Overall r<small ($r_{mean} = 0.07$)	X

Table 33 – Aggregated results of hypothesis testing

5.4 Other tests

5.4.1 Other factors from the EFA

Although not having any specific hypotheses regarding the two factors *in-house training* (IHT) and *attitude and perceptions towards EP* (EPPERCP) from the EFA, it is appropriate to investigate these as they can give valuable insight. In a similar fashion as for the tested hypotheses, we will assess potential differences across age, sex, and the categorization of end-user groups using vGroup. However, planned contrasts when testing for differences in either age or vGroup are not appropriate due to lack of hypotheses. Instead, a post hoc test in ANOVA using Bonferroni is employed to assess potential significant differences. As Field (2013) suggests, the post hoc procedure of Bonferroni is the preferable test to use when no specific predictions are made concerning the data. However, there is a stricter level of significance when using Bonferroni compared to an independent t-test (as the planned contrast can be regarded as). This strictness is related to the number of tests in the ANOVA. Rice (1989) states that in order for a difference in means to be significant using Bonferroni, the requirement is $\frac{\alpha}{k}$, where α is the significance level and k is the number of tests. Since vGroup consists of four groups, this gives $k=6$ (1-2, 1-3, 1-4, 2-3, 2-4, and 3-4) and a requirement of $\frac{0.05}{6} = 0.00833$, which is quite strict compared to a t-test where the requirement would be 0.05. Similarly, as there are five age groups (not including group 6, >69 years, as no end-user were in this group), this gives $k=10$ and a requirement of $\frac{0.05}{10} = 0.005$ in order for the difference in means being significant. If the overall ANOVA shows a significant difference between the respective groups, but Bonferroni does not yield any significant difference, the test will not be further evaluated. This is because we cannot say anything about the direction of the difference without further testing with an independent t-test, and this is not relevant since the potential difference is not hypothesized. On the other hand, when exploring potential differences across sex for the dependent variables IHT and EPPERCP, independent t-tests are used. This is because Bonferroni for only two groups (male and female) is not possible. As for when testing across sex in the previous section, the independent t-tests yield a direction of the difference and the effect size can be calculated. If finding significant differences across sex for IHT and EPPERCP, the findings will be further discussed.

Factor (Type of test)	Independent variable	Assumptions	Results
IHT (ANOVA)	vGroup	Normality and homogeneity of variance assumed (F=0.393, p=0.758)	There was no significant effect of IHT on the user groups (vGroup), F(3, 294)=0.125, p=0.945; $\omega^2 = -0.008$. Using the Bonferroni post hoc test revealed no significant differences in means of IHT across the four user groups.
IHT (ANOVA)	AGE	Normality and homogeneity of variance assumed (F=1.299, p=0.271)	There was a significant effect of IHT on the age groups, F(4, 295)=2.830, p=0.025; $\omega^2 = 0.024$. Using the Bonferroni post hoc test revealed, however, that there were no significant differences in means of IHT across the age groups.
IHT (Independent t-test)	SEX	Normality assumed, but homogeneity of variance violated (F=4.937, p=0.027). Using t-test value for “equal variances not assumed” instead.	On average, males scored higher on the variable IHT (M=3.3507, SE=0.0896) than females (M=3.0974, SE=0.0434). This difference, 0.2533, BCa 95% CI [-0.0558, 0.4509], was significant t(99.026)=2.545, p=0.012; d=0.36 (small<d<medium).

Table 34 – Testing for IHT's effect on the grouping variables

Table 34 above shows the effect of IHT on the independent variables vGroup, AGE, and SEX. As there was no significant effect of IHT, and consequently no significant difference in means when investigating the post hoc test, there is no reason to believe that the respondents' preferences of *in-house training* vary across the end-user groups, vGroup. The effect size did not exceed the small-threshold ($\omega^2 = 0.01$). On the other hand, there was a significant effect of IHT on age with a small to medium effect size. Since the requirement of statistical significance is a tenth of 0.05, no significance across the five age groups were found using Bonferroni. In addition, as discussed, further evaluation of this relationship is not included due to a lack of a specific hypothesis. More importantly, there was significant difference in means between males and females with a small to medium effect size. In our sample, this can indicate that the preference of having internal personnel giving training is more important to males than to females.

Factor (Type of test)	Independent variable	Assumptions	Results
EPPERCP (ANOVA)	vGroup	Normality and homogeneity of variance assumed (F=0.433, p=0.730)	There was no significant effect of EPPERCP on the user groups (vGroup), F(3, 298)=1.049, p=0.371; $\omega^2 = 0.0005$. Using the Bonferroni post hoc test revealed no significant differences in means of IHT across the four user groups.
EPPERCP (ANOVA)	AGE	Normality and homogeneity of variance assumed (F=0.763, p=0.550)	There was no significant effect of EPPERCP on the age groups, F(4, 299)=1.563, p=0.184; $\omega^2 = 0.007$. Using the Bonferroni post hoc test revealed that there were no significant differences in means of EPPERCP across the age groups.
EPPERCP (Independent t-test)	SEX	Normality and homogeneity of variance assumed (F=0.012, p=0.912).	On average, males scored higher on the variable EPPERCP (M=4.0417, SE=0.0599) than females (M=3.8339, SE=0.0312). This difference, 0.2078, BCa 95% CI [-0.0773, 0.3382], was significant t(300)=3.134, p=0.002; d=0.43 (small<d<medium).

Table 35 – Testing for EPPERCP’s effect on the grouping variables

Table 35 above shows the effect of EPPERCP on the independent variables vGroup, AGE, and SEX. When testing EPPERCP’s effect on vGroup and AGE, there was no significant effect on either, and consequently no significant difference in means when investigating the post hoc test. Neither of the effect sizes exceeded the small-threshold. There is no reason to believe that *attitude and perceptions towards EP* vary across either the user or age groups. On the other hand, EPPERCP’s effect on sex was significant, expressing that males score higher on the factor than females. In addition, the effect size was close to being a medium. Therefore, there is a reason to believe that *attitude and perceptions towards EP* vary across sex for our sample.

5.4.2 The importance of the factors for the end-users

As suggested by multiple authors, it is possible to assess the importance of the factors by comparing the means in an one-sample t-test (Brotherton, 2004, Liu et al., 2010, McGaughey and Roach, 2001). If the means are significantly higher than the test value, one can infer that the factor are important for the respondents. McGaughey and Roach (2001) suggest that using four as test value on a six-point scale, whereas we will use 3.5 as the test value because we use a five-point scale. Considering the ratios $\frac{4}{6} = 0.67 < \frac{3.5}{5} = 0.70$, our test is a bit stricter. In Table 36 and Table 37 below, we see the results of comparing all the factors from the EFA including the attributes for which alpha was greater than 0.6.

	N	Mean	Std. Deviation	Std. Error Mean
EUT	306	3.82215	.612216	.034998
EPPERCP	306	3.8790	.48535	.02775
CHMcks	301	3.1545	.68694	.03959
SUPREL	306	2.9328	.58162	.03325
STATQUO	303	2.8053	.88496	.05084
COMM	305	3.8362	.51759	.02964
IHT	302	3.1556	.68123	.03920
EUAut	303	4.0781	.76108	.04372
EPISrel	305	3.0311	.79151	.04532

Table 36 – Descriptives one-sample t-test for dependent variables

	Test Value = 3.5					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
EUT	9.205	305	.000	.322148	.25328	.39102
EPPERCP	13.659	305	.000	.37898	.3244	.4336
CHMcks	-8.726	300	.000	-.34551	-.4234	-.2676
SUPREL	-17.058	305	.000	-.56716	-.6326	-.5017
STATQUO	-13.665	302	.000	-.69472	-.7948	-.5947
COMM	11.345	304	.000	.33623	.2779	.3945
IHT	-8.785	301	.000	-.34437	-.4215	-.2672
EUAut	13.222	302	.000	.57811	.4921	.6641
EPISrel	-10.345	304	.000	-.46885	-.5580	-.3797

Table 37 – One-sample t-test with test value 3.5 for dependent variables

The results in Table 37 above show us that three factors and one attribute are significantly important for end-users. These factors are EUT, EPPERCP, COMM, and the attribute EUAut. The reasoning behind this is that these four have a positive mean difference, which correspond to the respective means shown in Table 36. Regarding all the others, we see that they are significantly lower than 3.5. This does not mean that we can infer that they are not important to end-users. To do so, we would need another test with a test value of 2.5. This is no point however, as all the means are above 2.5. In other words, we cannot certainly say if these factors are important or not with our gathered data.

These results above have to be handled with care however, because acquiescence bias and social desirability, as mentioned in 3.2.2, can affect the results. In our case however, as pointed out in section 4.1, the respondents have not scored systematically high for all factors, indicating that acquiescence bias is probably not a problem. Also, while a high mean rating on a factor can imply that it is important for end users, it can also be a result of a poorly developed questionnaire where everyone agrees. We argue that this is unlikely in our case as we have more than three items in all the factors found to be significant, and the means for the other factors are significantly lower. Last, we point out that one-sample t-tests are mostly used in

conjunction with specific hypotheses, unlike in our case, which is a weakness that must be noted.

Furthermore, multiple authors suggest that one can assess the relative importance of factors, i.e. which factor is most important compared to the other factors (Aksorn and Hadikusumo, 2008, McGaughey and Roach, 2001, Ng et al., 2009). To do this, a paired sample t-test is used, found in Appendix C, Table 41, due to the length of this table. The table shows tests for significant differences in all the possible combinations of the factors and attributes, 34 in total. Table 41 shows us that among those that were significant and relatively important in the test above, we can only say that EUAut is significantly more important than EUT, EPPERCP, and COMM. This is however not that surprising, as EUAut only contains three items, while the other five or more. The more items in a factor, the less likely it is that all end-users have scored high on all the factors. Regarding the other factors, we do not see any reason for discussing which of them that was more important than the other, as we cannot say whether they were overall important.

5.4.3 Correlations between the factors

According to Field (2014), one cannot make any conclusion about causality from a correlation. However, we can square the correlations to see how much variance the factors share. Table 38 below show Spearman's rho correlations for all the factors, which is the appropriate measure to use for ordinal factors, according to Bryman (2012). A correlation of 0.1 is considered a small correlation, while 0.3 is medium, and 0.5 is consider large (Field, 2013). Furthermore, a correlation of 0.5 means that the factors share $0.5^2 = 25\%$ of the variance. As we cannot make any conclusions about causality, the results from the correlation matrix is not important for this thesis. However, we will comment on those correlations that are medium or higher in effect size, as these factors share a considerable amount of variance.

	EUT	EPPERCP	CHMcks	SUPREL	STATQUO	COMM	IHT	EUAut	EPISrel
EUT	1.000	.098	.025	.147*	-.124*	.529**	.194**	.878**	.191**
EPPERCP	.098	1.000	.030	-.141*	-.157**	.161**	.032	.106	.031
SOCINFLU	.025	.030	1.000	.144*	.111	.061	.183**	.084	.177**
SUPREL	.147*	-.141*	.144*	1.000	.072	.068	.140*	.098	.760**
STATQUO	-.124*	-.157**	.111	.072	1.000	-.099	.037	-.090	-.004
COMM	.529**	.161**	.061	.068	-.099	1.000	.240**	.545**	.092
IHT	.194**	.032	.183**	.140*	.037	.240**	1.000	.225**	.189**
EUAut	.878**	.106	.084	.098	-.090	.545**	.225**	1.000	.164**
EPISrel	.191**	.031	.177**	.760**	-.004	.092	.189**	.164**	1.000

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).

Table 38 – Spearman's rho correlation for EFA factors

In the Table 38 above, the correlations higher than 0.3 are highlighted, of which there are four cases in total. The first, EUT and COMM share $0.529^2 = 28\%$ of the variance. The next one, EUT and EUAut share 77 % of the variance, but this is no surprise as EUT contains all the items of EUAut, and these account for three sevenths of the factor. The same argument goes for SUPREL and EPISrel that share 57 % of the variance. The last, COMM and EUAut share 30 % of the variance, but this is likely because EUT and COMM share 28 % variance, while EUT and EUAut share 77 % of the variance. In other words, the only interesting finding is that EUT and COMM share 28 % of the variance. We cannot say in which direction this relationship works, or if there is a relationship at all, i.e. that other factor(s) cause this variation.

5.5 Summary

In this chapter, we have explained in detail how the EFA was run, and the results of the factor extraction that led to seven factors. These seven factors, together with the three attributes from the framework with alpha larger than 0.6, were then used to test the hypotheses. Hypothesis H1, H3a, H4a, and H5 were confirmed, while the rest were inconclusive due to a lack of significance. In addition, some other tests were performed although lacking specific hypotheses, because they could provide useful insight. These tests uncovered significant differences for the factors *in-house training* and *attitude and perceptions towards EP* across sex. More specifically, males scored significantly higher than females on both factors. The additional tests also uncovered significant importance of three factors and one attribute, *End-user training, communication, attitude and perceptions towards EP*, and *user training (EUAut)*, but were unable to ascertain the relative importance of these factors. The only finding was that EUAut was relatively more important than the other three factors, but this did not yield much insight.

6. Discussion

In this chapter, we will first consider the motive for this thesis, before moving on to St. Olav's as a case company. Thereafter, we will discuss the weaknesses of the survey before discussing the results of the exploratory factor analysis. In section 6.1, we address research questions 1 to 4 and the implications for our framework and St. Olav's. In section 6.2, we explore generalizability, before moving on to the conclusion, in chapter 7, where we address the main research question.

As pointed out in section 1.2, the need for empirical validation was one of the limitations of the pre-diploma thesis. The pre-diploma thesis did however still indicate the relevance of the framework through the exploratory case study that was carried out. Recommendations for further research in the pre-diploma thesis proposed that an empirical survey could assess the relevance of the framework, as well as contribute with valuable information for St. Olav's. In cooperation with St. Olav's and our supervisor, we determined that further testing of the framework was both feasible and appropriate. Furthermore, this led to the aim of this master's thesis, to assess the relevance of the framework.

As the case company, St. Olav's proved valuable and helpful in providing both interview subjects and a suitable end-user population for surveying. Still, having St. Olav's as our case company did lead to some limitations. The fact that St. Olav's was already in the implementation phase of the project, instead of pre-implementation or post-implementation, did not affect the choice of what CSFs to validate much, as only one relevant factor was excluded, cf. 3.2.1. On the other hand, it was difficult to assess the relevance of CSFs that end-users cannot easily relate to. This last aspect heavily influenced what CSFs from the framework we could validate. Although it could be possible to explore more CSFs than the five chosen for the survey, we had to limit the length of the survey. In fact, according to the pilot tests, we were already pushing the attention span of the end-users with our 50 question items, which is understandable as ordering is only a minor part of the end-users' responsibilities.

In hindsight, it is clear that the survey had some weaknesses. Although using the iteration method described in subsection 3.2.1 when developing the question items and conducting a pilot test with the third parties for assessing face validity, most of the attributes showed a non-satisfactory Cronbach's alpha value. This is best illustrated by the results in Table 14 that show insufficient internal reliability for all but three of the sixteen attributes. Regardless that further two of the attributes (EUAiht and CHMor) could be rescaled by omitting one of the variables,

five out of the sixteen attributes having a sufficient internal reliability is still a considerable drawback. Face validity was regarded to be of a medium degree, cf. 3.3.2., as either the pilot test with the end-users or the pilot test with the third parties could provide input regarding consistency of the question items. A larger group of end-users for the pilot tests could have been an option for assessing internal reliability in greater detail, but was not viable due to busy work schedules. Another option could have been to reach out to a larger group of experienced third parties for evaluating the developed survey and ensuring internal consistency. However, as parties who possess such experience with surveys and the underlying theory are hard to come by, we did not have such an option. As we have experienced, developing a survey is a comprehensive process. Field (2013) suggests that exploratory factor analysis is typically used as an iteration approach to increase the quality of the research instrument. On this notion, our survey may be considered to be in the early process of developing a sound research instrument. The flaws of the survey support this reasoning, by, for instance, considering the low internal reliability. The optimal approach would have been to develop the survey using multiple iterations as suggested above. However, this was not possible in our case due to time and resource limitations.

Despite the weaknesses of the survey discussed above, the survey did yield valuable data. The demographic variables proved useful to group the end-users for the testing of the hypotheses. In addition, we believe these results will prove very useful for St. Olav's. They do not currently have any such data, at least not that we were able to unveil through interviews with the project group in the exploratory case study. As the internal reliability for the survey was insufficient, we decided to do an EFA to find a new factor structure. The output of the EFA was seven factors with sufficient internal reliability. Many of these factors were similar to attributes of the framework, and some identical, which was somewhat reassuring. However, with only seven factors compared to the fifteen we attempted to measure, it is probable that not all the items measured the concepts (the attributes) we initially wanted them to measure. Another explanation can be that the questions were complicated or difficult for the end-users to relate to, and that this caused inconsistent answers. The pilot tests did not raise this issue, but it is possible that the interviewees did not want to admit that some questions were difficult to relate to.

In subsection 5.1.4, we named and discussed the substance of the seven factors from the EFA in relation to our framework. But what do the factors actually mean for our framework? From a theoretical standpoint, the factors from an EFA are hypothetical constructs used to represent latent variables (Yong and Pearce, 2013). One of the main uses of the EFA, in addition to the two reasons stated in subsection 3.4.2, is to measure the underlying variable, by finding the group of items that represent this underlying variable, presented as a factor (Field, 2013). In other words, the factors indicate that groups of items, from our survey, together measure the underlying latent variable. Furthermore, this means that the factors should be good measures for the underlying concepts we want to measure, unless they measure something else. It comes down to whether we have interpreted what they measure correctly, i.e. named the concept correctly. Santos and Clegg (1999) state that interpreting the factors is an uncomplicated process where the researcher intellectually synthesizes and describe the common thread that binds all the variables involved in each construct. We argue that we were able to do so, because many of the factors were similar to the framework, and interpreting their substance was not too challenging. This means that it was appropriate to test hypotheses by using seven factors, as long as the factors measure the same attributes we wanted to test in the hypotheses. As discussed when testing the hypotheses, it turned out that we were able to test most of the hypotheses by using the factors with some small modifications. For example, we tested hypotheses H2-4 b)+c) with only one factor instead the two attributes included in the hypotheses. Upon having established the factors as a means for testing the hypotheses, we can move on to the research questions in the following section.

6.1 Research questions

In this section, we will first attempt to address research questions 1 through 4. These four questions together will form the basis for addressing the main research question. This process is depicted below in Figure 12.

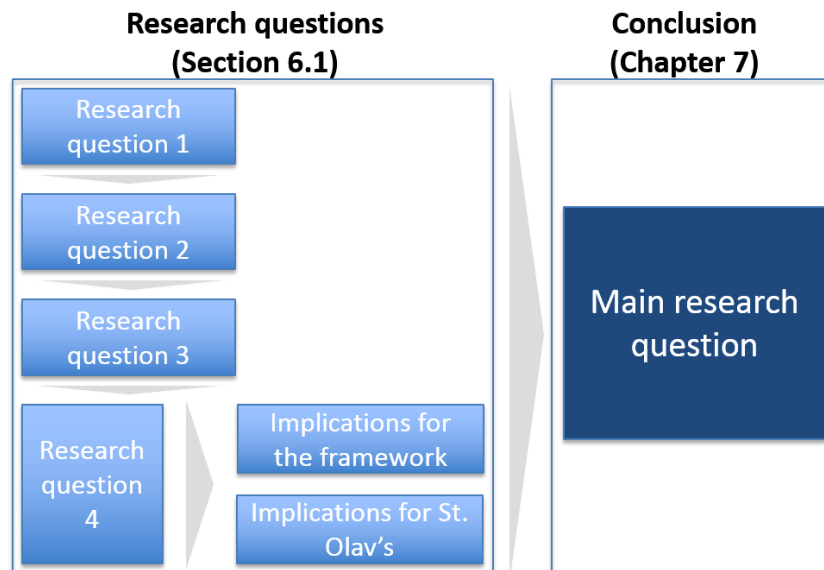


Figure 12 – Process for addressing the main research question

Research question 1 “*What are critical success factors for e-procurement implementation in the literature?*” was implicitly addressed in chapter 2 when introducing relevant theory and the framework. We pointed out that literature on e-procurement CSFs is sparse, as noted by Vaidya et al. (2006). Through both a systematic and narrative literature search on CSFs, we argue that we were able to gather most, if not all, of the relevant literature on e-procurement CSFs. These sparse findings further led us to synthesize the work regarding CSFs by multiple authors in the pre-diploma thesis. The result was a framework covering eleven CSFs with multiple attributes, tabulated according to which project phases they were deemed most important. In addition, we also did a new literature search at the beginning of this master’s thesis to see whether there were any new findings regarding the subject that had to be considered before continuing. As pointed out, this did not yield any results. Based on this notion, we argue that our framework is the answer for research question 1. The framework has clear managerial implications by being mapped onto phases, and should prove useful for St. Olav’s in their project, as well as other actors in similar situations.

Research question 2 “*What is likely for the different end users to assess as important?*” was also implicitly addressed in chapter 2. The hypotheses, presented in section 2.6, aim to address what different end-users probably assess as important. The first hypothesis, H1, founds the basis for being able to compare groups of end-users, other than across age and sex, in the remaining hypotheses. The remaining hypotheses, H2-H8, were formed based on the notion that the end-users with different experience in Visma assessed the importance of CSFs differently, as suggested by findings in the pre-diploma thesis. Relevant theory was used to support the hypotheses, but also to form hypotheses regarding differences not indicated by any of the findings in the pre-diploma thesis, such as across age and sex, as these were not looked into. Furthermore, relevant theory was used so that the results can be compared to those suggested in theory. It should be noted that the hypotheses are not exhaustive at all, but selected on the basis of what we believed to be feasible to measure, what implications they would have for St. Olav’s and research in general, and on the findings of other authors that surfaced in the literature search.

Research question 3 consists of two parts, and will be answered subsequently throughout this paragraph. The first part “*To what degree do different end-user groups in an empirical survey assess the importance of different factors?*” is addressed based on the results of the testing of hypotheses. The second part “*are the factors indeed relevant to end-users?*” is addressed based on the tests of importance in Table 37. Referring to Table 33, we first point out that the first hypothesis, H1, was confirmed, which enabled us to continue testing of the remaining hypotheses regarding differences across the end-user groups (the vGroup variable). Regarding H2a, we did not find significant differences for either *end-user training* (from the EFA) or *user training* (from the framework) across the four groups. Although the difference in means were in favor for the hypothesis, it was not sufficient for statistical significance. This was somewhat surprising, because it indicates that experienced Visma users consider training to be equally important as users never having used Visma. Both the pre-diploma findings and theoretical findings suggested that such a difference would exist. On the other hand, we found that both the factor *end-user training* and the attribute *user-training* are significantly important to end-users. Moving on, hypothesis H3a was supported, meaning there is a significant difference across age. Both tests showed that training is significantly more important for older people than younger people. Hypothesis H4a was also supported, but only for the factor *end-user training*. This means that there is in fact a significant difference between males and females in regards to training, specifically that females deem training as more important than males. As discussed,

parts b) and c) of hypotheses H2-4 were tested with one factor, *communication*. However, there was no significant difference across any of the groups (vGroup, age, and sex), thus hypotheses H2-4, part b) and c), were all inconclusive. On the other hand, we did find that the factor *communication* was significantly important for the end-users as a whole using. The next hypothesis, H5, was supported; end-users in group 3 and 4, never having used Visma, are significantly more negative for the factor *status quo* than end-users of group 1 and 2, experienced with Visma. This may indicate that end-users that are using the old ways to order are more resistant to change. However, we could not confirm the overall importance of the factor *status quo* for end-users as a whole. Hypotheses H6 and H7 were both tested using the attribute *communication to key stakeholders and peers* as the internal reliability of this attribute was sufficiently high. Neither the hypotheses nor the overall importance of the factor in other tests could be confirmed. Last, parts a) and b) of hypothesis H8 were inconclusive when testing with the factor *communication*. The overall importance of the factor *supplier relationships* and the attribute *relationships with suppliers* could not be established. Although these were the planned tests for research question 3, we did also analyze the last two factors emerging from the EFA even though these did not have specific hypotheses. For the factor *in-house training*, we found a significant difference across sex, indicating that males preferred in-house training to a greater extent compared to females. Similarly, the factor *attitude and perceptions towards EP* was found to be significantly higher for males, indicating that males have a more positive attitude towards EP than females. Additionally, we were able to establish an overall importance of this factor, i.e. indicating that end-users are generally positive towards EP.

Research question 4 – “*Which insights may be gained from this statistical analysis?*” - will contain two perspectives. First, what the implications are for our framework will be addressed in subsection 6.1.1. Second, what the implications are for the case company will be addressed in subsection 6.1.2.

6.1.1 *Implications for the framework*

There are mainly three aspects in this thesis that have implications for our framework. These are the emerging factors from the EFA, the results of the hypothesis testing, and the other tests in section 5.4.

The seven factors that emerged from the EFA are all relatable to our framework. Of all the seven factors, only one factor, *attitude and perceptions towards EP*, is not mentioned directly or indirectly in the framework. Being able to establish an overall importance of this factor, we argue that it is relevant to our framework. As pointed out, the factor was overall important, which means end-users in general have positive attitudes and perception towards e-procurement systems. This is not necessarily equivalent to the factor being critical for a successful implementation. However, also finding significant differences among males and females supports the notion of the factor's relevance, as it indicates that there are also differences between groups of individuals. Although we did not find significant differences among the other groups, this may be a coincidence in our sample. Still, we argue that the differences we did find are enough to consider adding the factor to our framework as an attribute under the CSF *change management*. It highlights the importance of measuring end-user attitudes, as negative attitudes can be interpreted as a form of resistance. By doing so, management, for instance, can target these users by providing additional information and training to remedy such attitudes. Regarding the phases of implementation, measuring *attitude and perceptions towards EP* continuously may unveil changes in attitudes and give insight, as well as continuously enabling management to focus on minimizing negative attitudes. This is in line with the CSF change management and the attribute *organizational resistance*. The CSF change management is mapped in all of the phases of the framework, according to Figure 5. Although the factor, *attitude and perceptions towards EP*, can be considered as a measure for organizational resistance, we argue that explicitly pointing this out can increase the managerial implications and establish a clear connection between organizational resistance and e-procurement systems. We do, however, emphasize that adding *attitude and perceptions towards EP* as an attribute under change management should be further researched. A case study can perhaps establish the causality of including this attribute in an e-procurement implementation. Having established the relevance of the one factor not mentioned in the framework, implications of the other findings will be discussed. The remaining six factors will, however, not be discussed to the same degree because they are already included in the framework.

Although only three of the seven factors emerging from the EFA contain identical items as the attributes we tried to measure in the survey, we saw that the correlations were quite high with the attributes for the remaining factors. For example, the factor *end-user training* from the EFA had a correlation of 0.878 with the attribute *user training*, cf. Table 38. We argue that these minor differences in items and names are not important at all for the framework, because this is probably only a result of us not being able to measure the attributes sufficiently. Another example is the factor *supplier relationships*, which contained items believed to measure both attributes *relationships with suppliers* and *preferred suppliers*. The EFA showed that all items of both attributes ended up in the same factor, *supplier relationships*. This probably means that all the items we developed was a better measure for the first attribute than the latter. Therefore, we will not discuss these differences any further, but rather discuss what the factors represent in comparison to the attributes in our framework.

Continuing with the results, we pointed out that two more factors were established as important, *end-user training* and *communication*. As both are already covered in the framework, these results support the importance of our framework. *Communication* supports the importance of the attribute *end-user support/communication* in the framework, while *end-user training* supports the attribute *user training*. Furthermore, *in house training*, and *communication to key stakeholders* are identical to the attributes in the framework, while *status quo* is identical to the attribute *organizational resistance*. Although we were unable to establish significant importance of the other factors than the three pointed out, we argue that significant differences for the factor *status quo* may indicate importance of the attribute *organizational resistance*. We pointed out that experienced Visma users (group 1 and 2) were significantly less resistant to new e-procurement systems. Therefore, we argue that this aspect is important to consider from a change management perspective because such resistance may be one of the obstacles hindering adoption of a new e-procurement system, and is likely to exist in any organization.

For the remaining factors, we cannot say whether they are important or not from our findings, we argue, however, they are important because they are deducted from similar cases where they were found to be important. Validating the importance of the remaining factors for certain, will however require further research.

Figure 13 below summarizes how the seven factors from the EFA relate to our framework.

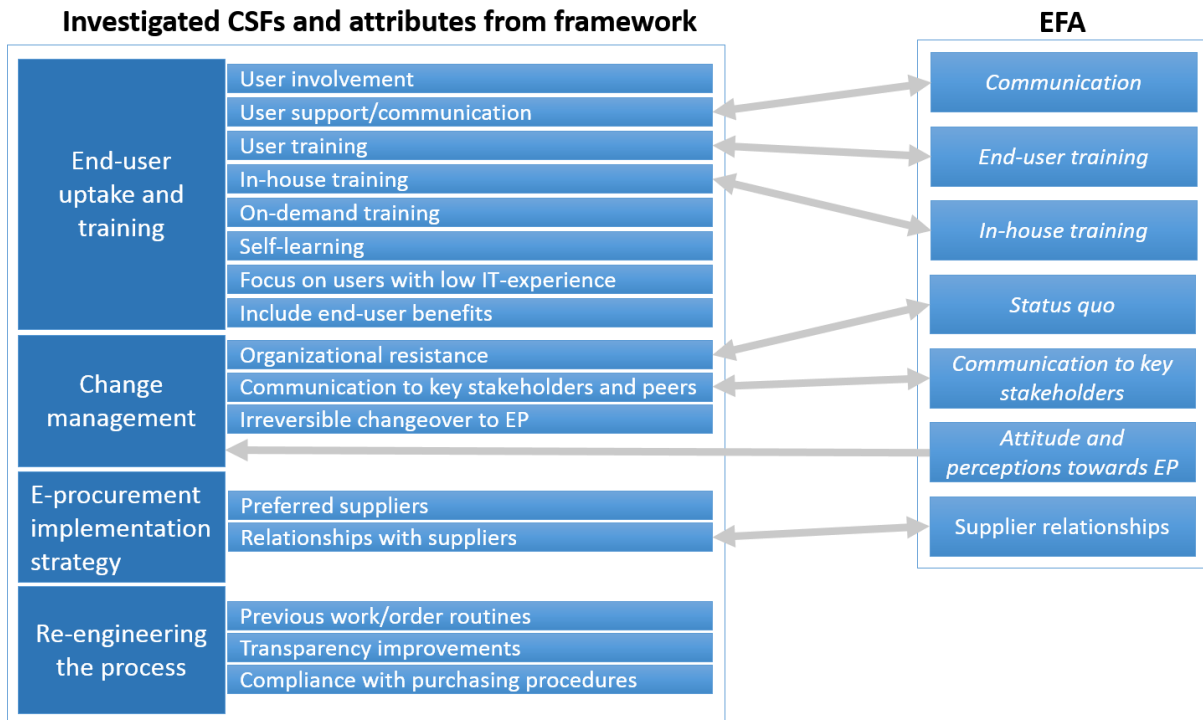


Figure 13 – EFA factors' relation with the framework

6.1.2 Implications for St. Olav's

The main aspects in this thesis that have implications for St. Olav's are the demographic variables and characteristics found in chapter 4, the significant differences found when testing the hypotheses in chapter 5, and the other tests from section 5.4.

Starting with the demographic variables, the first interesting observation is found by combining the results of Table 8 and Table 9, cf. section 4.2. 178 end-users (58.2 %) responded that they have placed orders through Visma. Of these, 44 users (24.7 %) reply that they conduct less than or equal to a quarter of all their orders through Visma. Apart from possible misunderstandings, it may be worthwhile for St. Olav's of looking into these low numbers. Possible explanations can be; a) end-users do not want to use the system for unknown reasons, i.e. showing resistance towards the system, b) the needs of these users cannot be fulfilled due to the system not being comprehensive enough or under-developed, and c) other ordering channels, such as phone and fax, are still available, and users prefer these. As pointed out in the introduction, St. Olav's plans to enforce Visma in the near future, and by doing so we argue it is critical to monitor and collect feedback from these users to ensure adoption of the system. Either way, we recommend investigating these end-users further.

Next, Table 10 illustrates that almost half of all the respondents (58 of 124) do not know whether or not they are going to use Visma in the nearest future. This number may point to a possible problem, as it indicates a lack of information directed at these end-users. However, this may be a deliberate action by St. Olav's for reasons unknown to us, or it may be a result of a slow progress in rolling out the system. If not, St. Olav's should consider informing these end-users. Table 11 gives an overview of what channels end-users order through other than Visma. While email accounts for 77.3 % is not that surprising, it is more alarming that almost 20 % still use fax to order. The end-users that are relying on an outdated ordering method may be less prone to using e-procurement than other users, and therefore require more training and information.

Table 12 lists the reasons for not using Visma among the users that have not yet used Visma. We note that most of the users (37.9 %) responded "I have not been given the offer to use Visma", category 1, Table 12. Not having a response category for users that delegate orders to others is a weakness with the survey in the sense that some of these respondents might have instead used response category 1. However, as pointed out in section 4.2, roughly 25 % responded that they delegate orders through specified text when using the response category "other reasons", indicating that perhaps not many of the 37.9 % mentioned above are users that delegate orders. Nevertheless, we cannot conclude on this. We address this because, if these are relevant users that do not delegate their orders, St. Olav's should identify this group and give them the option to use Visma. This may bring St. Olav's closer to its overall goal of reducing maverick buying. Furthermore, with over 15 % responding that they do not believe the system will coincide with their purchasing routines, the question is whether St. Olav's has done a good enough job of informing these users, or if the system is in fact too limited for these respondents needs. Although few, almost 4 % responded negatively considering that they do not believe Visma will lead to any improvement or that training will take too long time. Social desirability bias could have had an impact meaning the actual numbers could be higher. However, it is probably difficult, if not impossible, for St. Olav's to find these users as our survey was anonymous and that the gathered data only is to be used for this thesis, according to the guidelines given by NSD. Therefore, the implication of this is limited. Finally, we found that of the roughly 50 % that responded "other reasons", a few responded lack of information and training to be the reason. However, considering that St. Olav's is a large and complex organization with many different types of people, the results are arguable rather acceptable.

The distribution of sex and age in Tables 4 and 5, respectively, are worth mentioning. In itself, these two variables, sex and age, may not have any direct implications for St. Olav's, but the results of tests across these groups were significant. The two first significant results found are relating to *user training*. User training being found to be significantly more important for older than younger people suggests that St. Olav's can target these users with extra training. Furthermore, finding user training to be more important for females than males can also be considered, however one need to be careful not to discriminate. The same goes for age, as this can be insulting for older people. Final, the only significant difference found across the groups of Visma use (vGroup) was for organizational resistance. We found that experienced users were more positive towards e-procurement system than the unexperienced users. This can imply that St. Olav's should consider targeting unexperienced users with extra communication and training.

As mentioned in 6.1.1, we were able to establish a significant importance of the three attributes *user training*, *communication* and *attitude and perceptions towards EP*. The first two simply implicate that these were important for all the end-users at St. Olav's. If St. Olav's only were to consider two attributes from the framework, these would be the ones we recommend. The last, simply indicates that end-users are generally positive towards e-procurement systems, which is good news for St. Olav's. We did find a difference among males and females regarding attitude towards e-procurement systems, and a preference for in-house training for males, but the implications of these two are unclear.

Although not directly related to the findings in this thesis, our final recommendation is that St. Olav's consider all the CSFs (with their attributes) in our framework, as these have been proven useful in other studies, and as some of them have been proven as relevant to end-users at St. Olav's. For convenience, any employee of St. Olav's reading this thesis should also consider chapter 2 where the CSFs are explained, as well as the conclusion.

6.2 Generalizability

In this section, we will present quality criteria affecting generalizability, as previously stated in section 3.3. The quality criterion external validity concerns generalizability. As mentioned, generalizability can be divided into generalizing to theory, generalizing to the specific larger population, and generalizing across settings or populations. Because external validity depends on other quality criteria such as replication, internal reliability, and construct validity, this connection must be borne in mind when evaluating generalizability.

6.2.1 *Quality criteria affecting generalizability*

The capacity of this thesis being replicated can be said to be medium. This is reflected upon that reaching out to the same sample as we did may be cumbersome in a potential future replication. Although using a convenience sample may be viewed as a disadvantage for replication, we did, however, deploy the survey to the entire population of end-users (as stated in section 3.2.3), not a sample of this. Nevertheless, the email lists we acquired in order to target the population of end-users may be outdated from time to time and are prone to change over time. As individuals either change jobs, their job description is changed, quit their job, etc., the characteristics of the sample may change. Although we have no data on how often this occurs, if an end-user starts working in another department at the hospital, the job description of this individual is likely to be changed. Considering that each department has different purchasing needs, this example supports that the end-user groups are also prone to change. Therefore, finding the same results as we did may be challenging in terms of replication. However, we argue that the level of replicability of this thesis is medium due to the high level of detail through the thesis, especially in the sections attending to the statistical analysis.

The internal reliability of the survey was not satisfactory until extracting the factors from the exploratory factor analysis. Nevertheless, all the hypotheses were able to be tested using a combination of factors and the satisfactory original attributes from the framework, which all had sufficient internal reliability. Hence, we do not treat internal reliability as an issue regarding generalizability. Moreover, there exists a link between internal reliability and replication when considering the exploratory factor analysis. Not finding evidence for most of our groups of question items relating to the same concepts, EFA was a feasible tactic for finding groups of items that did. However, a future replication of this EFA may yield other factors as the implementation process may have progressed at St. Olav's and therefore end-users' attitudes may have changed.

As stated in subsection 3.3.2, in order to determine construct validity, a single study is inadequate. Therefore, in this thesis, construct validity was omitted being addressed, which was mentioned to affect the level of external validity. This is sensible since the developed research instrument was the first, to our knowledge, aiming to measure the specific attributes in the specific context of e-procurement. Ensuring construct validity must therefore be addressed in further research.

6.2.2 *Generalizing to theory*

Generalizing to theory should be dealt with care as replicability and construct validity negatively affect this ability. Although internal reliability is deemed not to be an issue, our findings lack further determination of construct validity. Therefore, we need to be careful by stating that our results can supplement theory by supporting previous research findings or indicating that others' research findings may be wrong. However, generalizing to our framework is possible by considering the implications for our framework as discussed above. For instance, a potential attribute, *attitude and perceptions towards EP*, was found to be relevant to the CSF change management. In addition, the exploratory factor analysis, hypothesis testing, and other tests were useful to provide information on the importance and relevance of some of the factors, but not all of them. On the other hand, as stated in the limitations of this thesis, only end-users were included in the sample, hence the attempt to validate the relevance of the framework is limited to the end-user perspective. The framework mostly aims at managers, considering the fact that few of the CSFs are relatable to by end-users. Generalizing to the framework is therefore somewhat limited due to the end-user perspective in this thesis, although we argue that end-users constitute a stakeholder of high relevance in an e-procurement implementation. As with the limited ability to generalize to the theory underlying the hypotheses, generalizing to the theory underlying the framework should also be dealt with care. As pointed out earlier, this is mainly because this thesis is a single study and that construct validity remains being established.

6.2.3 *Generalizing to the larger population*

As stated in subsection 3.3.3, the ability to generalize to the specific larger population, which in our case concerns all of the end-users employed at St. Olav's, may be negatively affected by the choice of a non-probability sample. However, as the survey was deployed to all of the registered end-users with the responsibility of purchasing at St. Olav's, and not merely a sample of these, the choice of using a convenience sample does not impact the ability to generalize as negatively as first indicated. Concerning the sample and being told that the 803 individuals cover more or less all of the end-users with a purchasing right, we can assume that we have reached out to approximately all of these end-users employed at St. Olav's. In order to have representativeness of our gathered data, there are three criteria. These criteria are that the distribution of sex, age, and vGroup of our respondents need to be similar to the whole population of 803 end-users with a purchasing right employed at St. Olav's. The only data we have for the entire population is the distribution of sex. This was assessed in section 4.1 in relation to non-response bias, which indicated that sex was not a driver for non-response. Therefore, our data is representative in terms of distribution of sex. We do not have data for the remaining two criteria and therefore cannot guarantee representativeness of our data as a whole. Moreover, there can be a myriad of reasons for non-response bias. Since we did not have the capacity to investigate possible drivers other than sex, we cannot guarantee that there is no non-response bias present in our sample. Overall, we are careful by generalizing to the entire population of end-users with purchasing rights because of the aspects considered in this subsection.

6.2.4 *Generalizing across populations*

The next question is whether the findings can be generalized to other settings or populations. As mentioned in the limitations of this thesis, the domain affects the ability to generalize across settings. As the setting of this thesis only includes one publicly held, Norwegian hospital undergoing an e-procurement implementation, generalizing the findings to other settings than this domain cannot be done. This is because that there are differences between publicly and privately held organizations and that there are cultural differences across countries. Nevertheless, it can be stated that cultural differences between the Scandinavian countries are not as apparent as opposed to differences between Norway and, for instance, the United States. However, discussing whether our results may be applicable in either Sweden, Denmark or Finland is not possible without investigating, for instance, regulatory differences of purchasing across the countries. Moreover, since some of the findings are related to the different user

groups, characterized by vGroup, this is a rather specific property of St. Olav's, which may not be the case at other hospitals undergoing an e-procurement implementation. However, the findings related to age, sex, and findings relating to the importance of the CSFs and their attributes can be important to other populations in the same setting as St. Olav's.

In summary, generalizing to theory underlying the hypotheses and the framework is negatively affected the medium level of replicability and by this study only focusing on one case and therefore lacking construct validity. However, generalizing to the framework was considered to be feasible. Furthermore, since we could not guarantee representativeness and no non-response bias present in our data, generalizing to the specified larger population concerning all the end-users with a purchasing right cannot be done. In addition, generalizing to other settings is feasible, but considering the narrowness of the setting, it concerns quite few other populations. Nevertheless, none of the findings suggest that our framework cannot be applied to other EP implementations and settings, but as this was not in scope of this thesis, it needs to be researched further.

7. Conclusion

In this master's thesis, we have seen the trend of increased importance of the purchasing function and awareness towards MRO costs. Furthermore, this has led companies to utilize electronic procurement tools in order to make indirect purchasing more efficient and effective. To ensure a successful implementation of these tools, it is advantageous to consider critical success factors (CSFs) capturing important aspects of the implementation. Adoption of the system is one such aspect, critical for achieving planned benefits.

Previous research on e-procurement have covered such CSFs to some degree, but this research was not unified into one framework, covering both CSFs and relevant project phases. Based on this notion, as well as being requested by other researchers, we set out to construct such a framework in the pre-diploma thesis. The resulting framework formed the basis for this thesis, and it was introduced in chapter 2. One of the caveats of the framework was a lack of validation. Other researchers contributing with the theories that make up our framework, also called for further validation. The aim of this master's thesis was to remedy this lack of validation, and to assess its relevance for end-users. This was a feasible option considering St. Olav's as the case company, and the available sample of end-users.

Referring back to the main research question “*Are different CSFs more or less relevant for different types of end-users in an e-procurement implementation context?*”. The answer to this is complicated. As mentioned, we set out to validate CSFs from an end-user perspective. Of the four CSFs we attempted to validate, only the CSF *End-user uptake and training* had all attributes relevant for end-users. The three remaining CSFs *Change management*, *Re-engineering the process*, and *E-procurement implementation strategy* had only a few attributes seemingly relevant for end-users. Therefore, the aim was to indirectly validate these CSFs from an end-user perspective by validating the CSFs' attributes relevant to end-users. Furthermore, as mentioned in section 2.6, we were not able to hypothesize differences for all these attributes as they lacked being previously explored in theory. This led us to develop a survey attempting to find differences regarding the attributes not hypothesized as well. As it turned out, our survey had too low internal reliability to validate all of these attributes. Figure 13 in subsection 6.1.1 provides an overview of what attributes we were able to measure by using the extracted factors of the EFA. Using these attributes, we were able to find significant differences across sex, age, and the four groups relating to use of Visma. Although we were unable to test the CSFs to the

degree we set out to do, we argue that these differences for attributes are enough to infer that the different CSFs are more relevant for some end-users than others. Based on this notion, the answer to the research question is therefore yes.

However, we emphasize the weakness of not finding differences for all the CSFs' attributes, and that the differences found may only be relevant for our specific case company. On the other hand, one may be able to find more differences when looking at other companies. Furthermore, not finding differences can be a strength of the framework, indicating that factors are more or less equally important to all end-users. This can facilitate the implementation as managers do not need to consider end-user differences for all CSFs. We also point out that the framework is comprehensive, as illustrated by the fact that we were only able to look at a small portion of the framework. It was not possible to consider the framework in its entirety due to resource and time limitations, as elaborated on in further research.

Overall, by finding an additional attribute to be added we have seen that our framework has room for improvements. We did also find differences among the end-users as well as well as being able to establish the overall importance of some of the attributes measured. Although many of the hypotheses were inconclusive, we found nothing indicating that any parts of the framework are irrelevant. On the contrary, our findings indicate that the framework is relevant, at least to end-users, and that further research on the framework is appropriate.

8. Limitations and further research

A further in-depth validation of the CSFs underlying our framework, as mentioned, was urged by Vaidya et al. (2006). The same can be stated for our pre-diploma thesis, in which further research was expressed as a requirement in order for validation of the framework. In this thesis, we have attempted to follow-up on this by investigating some of the CSFs in terms of end-users' opinions and beliefs. However, additional further research is both encouraged and needed in order to validate the rather comprehensive framework.

One of the aspects to potentially research further is to establish construct validity of a research instrument aimed to measure the CSFs and their relevance. Using our survey as a point of departure, may be feasible, especially when considering the factors from the exploratory factor analysis. In general, this thesis has only explored some of the CSF and their attributes relevant to end-users. Although we were able to test all of the hypotheses, some of the attributes being explored in the survey, such as the ones relating to the CSF *re-engineering the process*, were not tested in the hypothesis testing. In order to validate the framework in its entirety, including in which of the project phases the CSFs are important, more comprehensive research is needed, which does not only take into account end-users' opinions and beliefs, but also the managerial perspective. A longitudinal research design, which employs a mix between quantitative and qualitative research methods, should be a feasible tactic. More specifically, multiple-case studies can be used to investigate qualitatively the cause and effect of the CSFs for managers, the project group, and end-users in several e-procurement settings. In addition, by employing case study survey research, which allows for a longitudinal design, one can investigate findings from the multiple-case studies quantitatively. By doing so, generalizing to theory, which was a limitation for this thesis, is conceivable. As opposed to this thesis, the suggested research design also enables generalizing to the framework to a higher degree. Due to the potential relevance in any e-procurement implementation setting, further investigation should also explore settings not included in the domain of this thesis, such as privately held companies and companies based in other countries than Norway.

Although change management theory was employed to serve as a backdrop when developing the framework, we encourage further research on change management aspects in the context of e-procurement implementation. Vaidya et al. (2006) suggest that changes are required and directly related to the rate of e-procurement adoption. As stated, adoption can be viewed in light of change acceptance and since one of the overall goals of e-procurement implementation

is to achieve a high level of end-user uptake, change management's importance is apparent in this context. Only three attributes of the CSF change management were investigated and only two of these were used for hypothesis testing. In addition, as change management related aspects should be assessed over time, our survey did not suffice. By using the aforementioned proposed research design, it enables the change management aspects of this context to be explored over time. This can, for instance, result in supplementary attributes, such as the potential attribute, *attitude and perceptions towards EP*, suggested to be added to the CSF change management in subsection 6.1.2. However, further research must establish whether this is appropriate by further investigating the causality and the reciprocal relationship with other CSFs and attributes.

Regardless of the fact that most of the hypothesis tests were inconclusive, we are satisfied with the initial validation of our e-procurement implementation framework. Our work has established the relevance of the framework from an end-user perspective, in addition to suggesting appropriate methods to continue validating the framework. We also encourage further revision of the framework in order to capture all relevant aspects. As research on electronic procurement has been attended to the last two decades without establishing a comprehensive framework with clear managerial implications, we hope that further research can establish ours.

9. Bibliography

- ABERDEENGROUP 2001. Best Practices in e-Procurement: The Abridged Report. Boston, MA.
- AKSORN, T. & HADIKUSUMO, B. 2008. Critical success factors influencing safety program performance in Thai construction projects. *Safety Science*, 46, 709-727.
- ANGELES, R. & NATH, R. 2007. Business-to-business e-procurement: success factors and challenges to implementation. *Supply Chain Management: An International Journal*, 12, 104-115.
- ARBIN, K. 2008a. The road towards successful e-ordering implementation: success factors and barriers. *International Journal of Procurement Management*, 1, 415-429.
- ARBIN, K. 2008b. The structure of determinants of individual adoption and use of e-ordering systems. *Human Systems Management*, 27, 143-159.
- ARBIN, K. 2009. *Individual information system acceptance behaviour: an electronic ordering system case*. Ph.D., Stockholm School of Economics.
- ASCH, D. A., JEDRZIEWSKI, M. K. & CHRISTAKIS, N. A. 1997. Response rates to mail surveys published in medical journals. *Journal of clinical epidemiology*, 50, 1129-1136.
- AXELSSON, B., ROZEMEIJER, F. & WYNSTRA, F. 2005. *Developing sourcing capabilities: Creating strategic change in purchasing and supply management*, Wiley.
- BARUCH, Y. & HOLTOM, B. C. 2008. Survey response rate levels and trends in organizational research. *Human Relations*, 61, 1139-1160.
- BEAVERS, A. S., LOUNSBURY, J. W., RICHARDS, J. K., HUCK, S. W., SKOLITS, G. J. & ESQUIVEL, S. L. 2013. Practical considerations for using exploratory factor analysis in educational research. *Practical assessment, research & evaluation*, 18, 1-13.
- BECHTEL, C. & PATTERSON, J. L. 1997. MRO partnerships: a case study. *International Journal of Purchasing and Materials Management*, 33, 18-23.
- BROTHERTON, B. 2004. Critical success factors in UK budget hotel operations. *International Journal of Operations & Production Management*, 24, 944-969.
- BRYMAN, A. 2012. *Social research methods*, Oxford university press.
- CANIATO, F., LONGONI, A. & MORETTO, A. 2012. Effective eProcurement implementation process. *Production Planning & Control*, 23, 935-949.
- CHAN, C. & SWATMAN, P. M. EDI implementation: a broader perspective. Bled'98"-11th International Conference on Electronic Commerce, 1998. 8-10.
- CHRISTOPHER, M. 2011. *Logistics & Supply Chain Management*, Great Britain, Pearson.
- CLARK, J., KENNEDY, C., SCHMITT, T. & WALTERS, J. Critical factors that influenced e-Procurement implementation success in the State of Arizona: the Procureaz Project. 5th International Public Procurement Conference, Seattle, USA, 2012.
- COUSINS, P. 2008. *Strategic supply management: principles, theories and practice*, Harlow, Financial Times/Prentice Hall.

- COX, A., CHICKSAND, D., IRELAND, P. & DAVIES, T. 2005. Sourcing Indirect Spend: A Survey of Current Internal and External Strategies for Non - Revenue - Generating Goods and Services. *Journal of Supply Chain Management*, 41, 39-51.
- CROOM, S. & BRANDON-JONES, A. 2007. Impact of e-procurement: experiences from implementation in the UK public sector. *Journal of Purchasing and Supply Management*, 13, 294-303.
- CROOM, S. R. 2000. The Impact of Web - Based Procurement on the Management of Operating Resources Supply. *Journal of Supply Chain Management*, 36, 4-13.
- CROOM, S. R. & BRANDON-JONES, A. 2005. Key issues in e-procurement: procurement implementation and operation in the public sector. *Journal of Public Procurement*, 5, 367.
- CUMMINGS, T. G. & WORLEY, C. G. 2015. *Organization development & change*, Australia, Cengage Learning.
- DAWES, J. G. 2008. Do data characteristics change according to the number of scale points used? An experiment using 5 point, 7 point and 10 point scales. *International journal of market research*, 51.
- DE BOER, L., HARINK, J. & HEIJBOER, G. 2002. A conceptual model for assessing the impact of electronic procurement. *European Journal of Purchasing & Supply Management*, 8, 25-33.
- DEUTSKENS, E., DE RUYTER, K., WETZELS, M. & OOSTERVELD, P. 2004. Response rate and response quality of internet-based surveys: An experimental study. *Marketing letters*, 15, 21-36.
- DIFI. 2014. *Elektronisk handelsformat (EHF)* [Online]. Available: <http://standard.difi.no/forvaltningsstandarder/standard/elektronisk-handelsformat-ehf/elektronisk-handelsformat-ehf> [Accessed 01.10 2014].
- EISENHARDT, K. M. 1989. Building theories from case study research. *Academy of management review*, 14, 532-550.
- FIELD, A. 2013. *Discovering statistics using IBM SPSS statistics: and sex and drugs and rock 'n' roll*, Los Angeles, SAGE.
- FRANCALANCI, C. 2001. Predicting the implementation effort of ERP projects: empirical evidence on SAP/R3. *Journal of information technology*, 16, 33-48.
- GEORGE, D. & MALLERY, P. 2003. *SPSS for Windows step by step: a simple guide and reference, 11.0 update*, Boston, Allyn and Bacon.
- GUNASEKARAN, A. & NGAI, E. W. 2008. Adoption of e-procurement in Hong Kong: an empirical research. *International Journal of Production Economics*, 113, 159-175.
- HAMEL, J., DUFOUR, S. & FORTIN, D. 1993. *Case study methods*, Sage.
- HO, R. 2014. *Handbook of univariate and multivariate data analysis with IBM SPSS*, Boca Raton, Fl., Taylor & Francis.
- ICGCOMMERCE. 2009. Reasons e-Procurement Projects Fail to Achieve Their ROI. IDG TechNetwork.
- INTERVIEW CENTRAL WAREHOUSE 2014. Interview with former leader of the central warehouse. Conducted on November 11, 2014, at St. Olavs Hospital, Trondheim
- JACOBSEN, D. I. 2012. *Organisasjonsendringer og endringsledelse*, Bergen, Fagbokforl.

- KANUK, L. & BERENSON, C. 1975. Mail surveys and response rates: A literature review. *Journal of Marketing Research*, 440-453.
- KOTTER, J. P. 1995. Leading change: Why transformation efforts fail. *Harvard business review*, 73, 59-67.
- KULP, S. L., RANDALL, T., BRANDYBERRY, G. & POTTS, K. 2006. Using organizational control mechanisms to enhance procurement efficiency: how GlaxoSmithKline improved the effectiveness of e-procurement. *Interfaces*, 36, 209-219.
- LEVASSEUR, R. E. 2001. People skills: Change management tools—Lewin's change model. *Interfaces*, 31, 71-73.
- LIU, X., GRANT, D. B., MCKINNON, A. C. & FENG, Y. 2010. An empirical examination of the contribution of capabilities to the competitiveness of logistics service providers: A perspective from China. *International Journal of Physical Distribution & Logistics Management*, 40, 847-866.
- LUCAS, J. W. 2003. Theory - testing, generalization, and the problem of external validity. *Sociological Theory*, 21, 236-253.
- MALHOTRA, N. K. 2006. Questionnaire design and scale development. *The handbook of marketing research: Uses, misuses, and future advances*, 176-202.
- MCGAUGHEY, R. E. & ROACH, D. W. 2001. CIM planning: an exploratory study of factors practitioners perceive important to CIM planning success. *International Journal of Computer Integrated Manufacturing*, 14, 353-366.
- MILLS, A. J., DUREPOS, G. & WIEBE, E. 2009. *Encyclopedia of case study research*, Sage Publications.
- MORRIS, M. G., VENKATESH, V. & ACKERMAN, P. L. 2005. Gender and age differences in employee decisions about new technology: An extension to the theory of planned behavior. *Engineering Management, IEEE Transactions on*, 52, 69-84.
- MUNNS, A. & BJEIRMI, B. F. 1996. The role of project management in achieving project success. *International journal of project management*, 14, 81-87.
- NEDERHOF, A. J. 1985. Methods of coping with social desirability bias: A review. *European journal of social psychology*, 15, 263-280.
- NG, S. T., TANG, Z. & PALANEESWARAN, E. 2009. Factors contributing to the success of equipment-intensive subcontractors in construction. *International Journal of Project Management*, 27, 736-744.
- OSBORNE, J. W. & OVERBAY, A. 2004. The power of outliers (and why researchers should always check for them). *Practical assessment, research & evaluation*, 9, 1-12.
- PANDA, P. & SAHU, G. 2012. e-Procurement implementation: Critical analysis of the impact of success factors on project outcome. *The IUP Journal of Supply Chain Management*, 9, 44-72.
- PETER, J. P. 1981. Construct validity: a review of basic issues and marketing practices. *Journal of marketing research*, 133-145.
- PETERSON, R. A. 1994. A meta-analysis of Cronbach's coefficient alpha. *Journal of consumer research*, 381-391.

- PETERSON, R. A. 2000. A meta-analysis of variance accounted for and factor loadings in exploratory factor analysis. *Marketing Letters*, 11, 261-275.
- PRESTON, C. C. & COLMAN, A. M. 2000. Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta psychologica*, 104, 1-15.
- PREVITALI, P. 2012. European Public E-Procurement: The Italian Experience. *Manoharan A. & Holzer M., Active Citizen Participation in E-Government: A Global Perspective, Igi Global*.
- PROJECT DOCUMENTS 2014. Various internal project related documents on the e-ordering implementation. Various authors. Trondheim, 2013-14.
- PUSCHMANN, T. & ALT, R. 2005. Successful use of e-procurement in supply chains. *Supply Chain Management: An International Journal*, 10, 122-133.
- REUNIS, M. R., SANTEMA, S. C. & HARINK, J. H. 2006. Increasing e-ordering adoption: A case study. *Journal of Purchasing and Supply Management*, 12, 322-331.
- REUNIS, M. R. B. 2007. *The effect of influence tactics on e-procurement adoption cognitions*, PhD thesis. Technische Universiteit Delft, Delft.
- RICE, W. R. 1989. Analyzing tables of statistical tests. *Evolution*, 223-225.
- RONCHI, S., BRUN, A., GOLINI, R. & FAN, X. 2010. What is the value of an IT e-procurement system? *Journal of Purchasing and Supply Management*, 16, 131-140.
- RYAN, S. D. & HARRISON, D. A. 2000. Considering social subsystem costs and benefits in information technology investment decisions: A view from the field on anticipated payoffs. *Journal of Management Information Systems*, 11-40.
- SANTOS, J. R. A. & CLEGG, M. D. 1999. Factor analysis adds new dimension to extension surveys. *Journal of Extension*, 37, 5RIB6.
- ST.OLAVSHOSPITAL. 2014a. *About* [Online]. Available: <http://www.stolav.no/en/About-the-hospital/83713/> [Accessed 22.10 2014].
- ST.OLAVSHOSPITAL. 2014b. *St. Olav handler på e* [Online]. Available: <http://www.stolav.no/no/Nyheter/St-Olav-handler-pa-e/130362/> [Accessed 10.10 2014].
- TAYLOR, S. & TODD, P. A. 1995. Understanding information technology usage: A test of competing models. *Information systems research*, 6, 144-176.
- TINSLEY, H. E. & TINSLEY, D. J. 1987. Uses of factor analysis in counseling psychology research. *Journal of counseling psychology*, 34, 414.
- TURBAN, E., KING, D., VIEHLAND, D. & LEE, J. 2006. A managerial perspective. *Electronic commerce*.
- VAIDYA, K., SAJEEV, A. & CALLENDER, G. 2006. Critical factors that influence e-procurement implementation success in the public sector. *Journal of public procurement*, 6, 70.
- VAN TEIJLINGEN, E. & HUNDLEY, V. 2001. The importance of pilot studies. *Social research update*, 1-4.
- VAN WEELE, A. J. 2005. Purchasing & supply chain management: analysis, strategy, planning and practice.
- VENKATESH, V. & DAVIS, F. D. 2000. A theoretical extension of the technology acceptance model: four longitudinal field studies. *Management science*, 46, 186-204.

- VENKATESH, V. & MORRIS, M. G. 2000. Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. *MIS quarterly*, 115-139.
- VENKATESH, V., MORRIS, M. G., DAVIS, G. B. & DAVIS, F. D. 2003. User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.
- WACKER, J. G. 1998. A definition of theory: research guidelines for different theory-building research methods in operations management. *Journal of operations management*, 16, 361-385.
- WALPOLE, R. E. 2012. *Probability & statistics for engineers & scientists*, Boston, Mass., Pearson.
- WEELE, A. J. V. 2005. *Purchasing & supply chain management: analysis, strategy, planning and practice*, London, Cengage Learning.
- WEICK, K. E. & QUINN, R. E. 1999. Organizational change and development. *Annual review of psychology*, 50, 361-386.
- YIN, R. K. 2014. *Case study research: Design and methods*, Sage publications.
- YONG, A. G. & PEARCE, S. 2013. A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, 9, 79-94.

10. Appendices

Appendix A

Question Item	Norwegian (original)	English translation	Filtering	Item categories	Type
VISMAuse	Har du gjennomført bestillinger i Visma?	Have you placed orders using Visma?	Displayed to all	1) Yes 2) No	Dichotomous
DurVis	Hvor lenge har du benyttet Visma?	For how long have you used Visma?	Displayed if VISMAuse=Yes	1)Less than a month, 2)Between 1 month and three months 3)Between three months and six, 4)More than six months but less than a year 5)More than a year	Ordinal
PRCNTvis	Vennligst anslå hvor stor andel av alle dine bestillinger du gjennomfører via Visma per I dag	Please estimate how many of your orders in total are conducted through Visma today.	Displayed if VISMAuse =Yes	1)All of my orders (100%) 2)Most of them (76-99%) 3)Many of them (51-75%) 4)Some of them (26-50%) 5)Few (1-25%) 6) None (0%)	Ordinal
OftenVis	Hvor ofte benytter du Visma til å gjennomføre bestillinger?	How often do you use Visma to place orders?	Displayed if VISMAuse =Yes	1) Daily 2) Once a week 3) Once every two weeks 4) Once a month 5) Rarer than once a month	Ordinal
SRCord	Hvilke andre bestillingskanaler (annet enn Visma) benytter du?	Which others channels for ordering are you using (besides Visma)?	Displayed to all except PRCNTvis=100%	Email, fax, phone, and others	Multiple response + text box
FUTuseVis	Er det planlagt at du skal ta i bruk Visma i nærmeste framtid?	Is it projected that you are going to make use of Visma in the nearest future?	Displayed if VISMAuse =no	1)Yes 2)No 3)Not that I know of	Nominal
ReasonNot	Hva er grunnen til at du ikke har benyttet Visma?	Why is it that you are not using Visma?	Displayed if VISMAuse =no	1) I have not been given the offer to use Visma	Multiple response

				2) I do not think it coincides with the supplies I am purchasing 3) I do not believe Visma will lead to any improvements 4) I believe that training will take too long time 5) Other	
VisChg	Har du i løpet av det siste opplevd noen endringer I bestillingssystemet Visma?	Have you experienced any changes in Visma the last 12 months?	Displayed if VISMAuse =Yes +if DurVis≠1&2	No, Yes If yes, please specify	Dichotomous + text box
EUAui1	Det er svært viktig for meg at jeg har mulighet til å komme med innspill dersom et nytt system for elektroniske bestillinger introduseres	It is of high importance to me that I have the opportunity to provide input if a new system for electronic ordering were to be introduced	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUAui2	*Jeg foretrekker å komme med innspill på eget initiativ, fremfor en oppfordring til å komme med innspill	*I prefer to provide input on my own initiative rather than being urged to do it	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUAui3	Jeg foretrekker møter hvor jeg kan komme med innspill fremfor andre kanaler som telefon og e-post	I prefer meetings where I can provide input rather than other channels such as phone and email	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUAui4	Jeg føler mine innspill sjelden blir hørt	I feel that my input rarely is taken into account	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUAusc1	Det er svært viktig for meg å ha en kontaktperson /støtteapparat tilgjengelig når et nytt bestillingssystem introduseres	It is of high importance to me to have a contact person/support network available when a new ordering system is introduced	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal

EUAus2	Jeg ønsker at informasjon gis i forkant av en endring som innebærer et nytt bestillingssystem	I want information prior to a change that involves a new ordering system	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUAut1	Opplæring er nødvendig når et nytt bestillingssystem introduseres	Training is necessary when a new ordering system is introduced	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUAut2	Jeg har behov for opplæring uavhengig om jeg oppfatter bestillingssystemet som lett å bruke eller ikke	I have a need for training regardless whether I perceive the ordering system as easy to use or not	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUAut3	Tilrettelagt opplæring vil kunne øke min vilje til å benytte et nytt bestillingssystem	Adapted training could increase my willingness to use a new ordering system	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUaiht1	Jeg foretrekker at opplæring utføres av ansatte ved St. Olavs fremfor eksternt personell	I prefer that training is given by employees at St. Olav's rather than external personnel	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUaiht2	*Opplæring utført av eksternt personell vil føles tryggere enn opplæring utført av ansatte ved St. Olavs	*Training given by external personnel will feel safer than training given by employees at St. Olav's	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUaiht3	Ansatte ved St. Olavs vil kunne forstå mine behov for opplæring bedre enn eksternt personell	Employees at St. Olav's may understand my needs for training better than external personnel	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUaodt	*Dersom jeg har muligheten til å få opplæring kun ved behov, vil jeg foretrekke dette fremfor planlagt opplæring	*If I have the opportunity to be given training on demand, I will prefer this over planned training	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUasl	*For å lære meg å bruke et nytt bestillingssystem er selv-læring fordelaktig fremfor opplæring med en annen kyndig person	*In order to teach me to use a new ordering system, self-learning is favorable rather	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal

		than training with a competent person			
EUAtX	Hver gang et elektronisk system har blitt innført, har jeg hatt behov for opplæring	I have had the need for training every time a new electronic system has been introduced	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUAeub1	Det er svært viktig at et nytt bestillingssystem bidrar til en enklere arbeidshverdag for meg	It is of high importance to me that a new ordering system simplifies my workday	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUAeub2	*Min arbeidshverdag trenger ikke bli enklere så lenge det nye systemet gagnar St. Olav som en helhet	*My workday does not necessarily need to become simpler as long as the new system benefits St. Olav's as a whole	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUAeub3	Det må kommuniseres til meg i forkant hvilke fordeler jeg kan oppleve ved et nytt bestillingssystem	Which benefits I can experience must be communicated to me prior to introducing a new ordering system	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUAeub4	Dersom fordeler kommuniseres til meg, vil dette kunne ha en betydelig positiv innvirkning på min holdning til det nye systemet	If benefits are communicated to me, this could have a considerable positive impact on my attitude towards the new system	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EUAeub5	Jeg tror et elektronisk bestillingssystem kan gi meg en enklere arbeidshverdag	I believe an electronic ordering system can simplify my workday	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
CHMor1	Det er fordelaktig for St. Olavs Hospital at bestillinger gjennomføres elektronisk fremfor andre metoder, slik som e-post, telefon og faks	It is advantageous for St. Olav's Hospital that purchases are done electronically rather than using other methods such as email, phone or fax	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal

CHMor2	Jeg er fornøyd med dagens bestillingssystem for å gjennomføre bestillinger	I am satisfied with today's ordering system	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
CHMor3	Jeg ser ikke et behov for et nytt bestillingssystem	I do not see the need for a new ordering system	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
CHMcks1	Anbefaling fra en kollega om å bruke et nytt system har større påvirkning på meg enn en anbefaling fra ledelsen	If a colleague recommends using a new system, this would have a greater impact on me than a recommendation from top management	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
CHMicks2	Dersom jeg vet at flere andre kollegaer benytter et nytt bestillingssystem, vil dette påvirke meg positivt til å ta systemet i bruk	If I know that several other colleagues are using a new ordering system, this would influence me positively to be using the system myself	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
CHMicep1	St. Olavs Hospital har behov for <u>ett</u> universelt elektronisk bestillingssystem som ikke tillater andre metoder for bestilling slik som e-post og telefon, såfremt alle bestillinger kan gjøres i dette systemet	St. Olav's Hospital has the need for <u>one</u> universal electronic ordering system, which does not allow other methods for ordering such as email and phone, given that all purchases can be done in this system	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
CHMicep2	Mine nåværende bestillingsrutiner vil ikke påvirkes i noen særlig grad dersom andre metoder for bestillinger enn gjennom et elektronisk system fases ut	My present ordering routines will not be affected to any great extent if other methods for ordering than using an electronic system are discontinued	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
CHMicep3	Gitt at alle bestillinger kan gjøres i det nye bestillingssystemet, er det likevel viktig for meg at utfasing	Given that all purchases can be done in the new ordering system, it is still important to me that other purchasing	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal

	av andre bestillingskanaler skjer gradvis, fremfor over natten	channels are discontinued gradually rather than immediately			
RENGpwo1	Det er viktig at mine bestillingsbehov kartlegges i forkant av et nytt bestillingssystem (eksempelvis hvor mange returer, hasteordre, etc. jeg skal følge opp)	It is important that my previous work routines are mapped prior to the introduction of a new ordering system (for example returns, rush orders, etc., I am supposed to follow up)	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
RENGpwo2	Det er svært viktig at et nytt bestillingssystem ikke fører til nye bestillingsrutiner innad i avdelingen (eksempelvis hvem i avdelingen som samler inn og gjennomfører bestillinger)	It is of high importance that a new ordering system does not lead to new ordering routines within the department (for example changing who in the department that is responsible for purchasing)	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
RENGti1	*Jeg tror et elektronisk bestillingssystem vil gi meg mindre oversikt over mine bestillinger	*I believe an electronic ordering system will impact my purchasing overview negatively	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
RENGti2	Det er svært viktig for meg at et eventuelt nytt elektronisk bestillingssystem gir økt oversikt over historikk, som tidligere bestillinger og forbruk	It is of high importance to me that a potential new electronic ordering system gives a better history overview such as previous purchases and consumption	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
RENGti3	Mine bestillingsrutinene per i dag gir meg ikke tilstrekkelig oversikt over mine bestillinger	My present ordering routines do not give me a sufficient overview of my purchases	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
RENGcomp	For meg er det viktig at alle kjøp jeg gjør følger gjeldende regelverk for offentlige bestillinger, selv om det betyr at jeg må gjøre alle mine	It is important to me that all of my purchases comply with the current regulations for public purchases, even if	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal

	bestillinger i et elektronisk bestillingssystem	this means that I have to do all of my purchases in an electronic ordering system			
EPISrel1	Jeg har et tett forhold til én eller flere leverandører (eksempelvis at du har et godt kommunikasjonsforhold med enkelte leverandører)	I have a close relationship to one or more suppliers (for example if you have a good communicational relationship to some suppliers)	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EPISrel2	Jeg oppfatter det som negativt dersom dagens forhold til leverandørene skulle påvirkes som følge av innføring av elektronisk bestilling	I perceive it negatively if the present relationships to suppliers are affected as a result of an introduction of electronic ordering	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EPISps1	*Det er viktigere at St. Olavs som helhet kan handle mer gjennom rammeavtaler, enn at jeg kan fortsette med å bestille hos dagens leverandører	*It is more important that St. Olav's as an entity acts more on general agreements than that I can continue to order from today's suppliers	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EPISps2	Det er viktig for meg at jeg har muligheten i fremtiden til å bestille de samme varene jeg bestiller i dag	It is important to me that in the future, I have the opportunity to order the same supplies as I do today	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
EPISps3	*Jeg er åpen for å bestille andre varer som tilsvarer de jeg gjør i dag dersom St. Olavs Hospital har bedre avtaler på disse	*I am open to order other supplies which are equivalent to those I order today if St. Olav's hospital has better agreements on these	Displayed to all	Likert: 1)Strongly disagree, 2)Disagree, 3)Neutral, 4)Agree, 5)Strongly agree	Ordinal
AGE	Vennligst angi din alder i én av de følgende intervaller	Please specify your age in one of the following intervals	Displayed to all	1) 18-28 years old 2) 29-38 years old 3) 39-48 years old 4) 49-58 years old 5) 59-68 years old 6) 69 years or older	Interval/ratio
SEX	Vennligst angi kjønn	Please specify your sex	Displayed to all	1)Female	Dichotomous

				2)Male	
DEPARTM	Vennligst angi hvilken avdeling du jobber i	Please specify which department you belong in	Displayed to all	Text box	Open

Table 39 – The survey

Appendix B

Descriptives	N	Min	Max	Mean	Std. Deviation
EUAui1 It is of high importance to me that I have the opportunity to provide input if a new system for electronic ordering were to be introduced	306	1	5	3.73	.843
EUAui2 *I prefer to provide input on my own initiative rather than being urged to do it	302	1	5	3.17	.902
EUAui3 I prefer meetings where I can provide input rather than other channels such as phone and email	306	1	5	3.01	1.021
EUAui4 I feel that my input rarely is taken into account	305	1	5	2.94	.703
EUAusc1 It is of high importance to me to have a contact person/support network available when a new ordering system is introduced	305	1	5	4.18	.863
EUAusc2 I want information prior to a change that involves a new ordering system	305	2	5	4.48	.591
EUAut1 Training is necessary when a new ordering system is introduced	306	1	5	4.42	.712
EUAut2 I have a need for training regardless whether I perceive the ordering system as easy to use or not	305	1	5	3.82	1.007
EUAut3 Adapted training could increase my willingness to use a new ordering system	304	1	5	4.00	.894
EUAiht1 I prefer that training is given by employees at St. Olav's rather than external personnel	305	1	5	3.10	.758
EUAiht2 *Training given by external personnel will feel safer than training given by employees at St. Olav's	305	1	5	3.30	.745
EUAiht3 Employees at St. Olav's may understand my needs for training better than external personnel	303	1	5	3.20	.766
EUAodt *If I have the opportunity to be given training on demand, I will prefer this over planned training	301	1	5	2.99	1.054
EUAsl *In order to teach me to use a new ordering system, self-learning is favorable rather than training with a competent person	306	1	5	3.52	.910
EUAitX I have had the need for training every time a new electronic system has been introduced	306	1	5	3.17	1.053
EUAeub1 It is of high importance to me that a new ordering system simplifies my workday	305	2	5	4.41	.674
EUAeub2 *My workday does not necessarily need to become simpler as long as the new system benefits St. Olav's as a whole	305	1	5	3.17	1.008
EUAeub3 Which benefits I can experience must be communicated to me prior to introducing a new ordering system	305	1	5	3.63	.805
EUAeub4 If benefits are communicated to me, this could have a considerable positive impact on my attitude towards the new system	304	1	5	3.87	.757
EUAeub5 I believe an electronic ordering system can simplify my workday	306	2	5	3.94	.715
CHMor1 It is advantageous for St. Olav's Hospital that purchases are done electronically rather than using other methods such as email, phone or fax	302	2	5	4.14	.694
CHMor2 I am satisfied with today's ordering system	306	1	5	2.94	.973
CHMor3 I do not see the need for a new ordering system	303	1	5	2.67	.971
CHMcks1 If a colleague recommends using a new system, this would have a greater impact on me than a recommendation from top management	303	1	5	2.83	.828

CHMcks2 If I know that several other colleagues are using a new ordering system, this would influence me positively to be using the system myself	304	1	5	3.49	.792
CHMicep1 St. Olav's Hospital has the need for one universal electronic ordering system, which does not allow other methods for ordering such as email and phone, given that all purchases can be done in this system	305	1	5	3.46	.976
CHMicep2 My present ordering routines will not be affected to any great extent if other methods for ordering than using an electronic system are discontinued	305	1	5	2.95	.920
CHMicep3 Given that all purchases can be done in the new ordering system, it is still important to me that other purchasing channels are discontinued gradually rather than immediately	306	1	5	3.30	.972
RENGpwo1 It is important that my previous work routines are mapped prior to the introduction of a new ordering system (for example returns, rush orders, etc., I am supposed to follow up)	304	1	5	3.55	.823
RENGpwo2 It is of high importance that a new ordering system does not lead to new ordering routines within the department (for example changing who in the department that is responsible for purchasing)	305	1	5	3.36	.926
RENGti1 *I believe an electronic ordering system will impact my purchasing overview negatively	303	1	5	3.81	.773
RENGti2 It is of high importance to me that a potential new electronic ordering system gives a better history overview such as previous purchases and consumption	306	2	5	4.18	.701
RENGti3 My present ordering routines do not give me a sufficient overview of my purchases	305	1	5	3.26	1.053
RENGcomp It is important to me that all of my purchases comply with the current regulations for public purchases, even if this means that I have to do all of my purchases in an electronic ordering system	305	1	5	3.74	.804
EPISrel1 I have a close relationship to one or more suppliers (for example if you have a good communicational relationship to some suppliers)	305	1	5	2.99	.987
EPISrel2 I perceive it negatively if the present relationships to suppliers are affected as a result of an introduction of electronic ordering	306	1	5	3.06	.883
EPISps1 *It is more important that St. Olav's as an entity acts more on general agreements than that I can continue to order from today's suppliers	302	1	5	2.50	.939
EPISps2 It is important to me that in the future, I have the opportunity to order the same supplies as I do today	302	1	5	3.90	.817
EPISps3 *I am open to order other supplies which are equivalent to those I order today if St. Olav's hospital has better agreements on these	305	1	5	2.22	.884
Valid N (listwise)	266				

Table 40 – Descriptive results of the survey

Appendix C

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	EUT - EPPERCP	-.056828	.753823	.043093	-.141625	.027970	-1.319	305	.188
Pair 2	EUT - CHMcks	.668802	.900337	.051895	.566679	.770926	12.888	300	.000
Pair 3	EUT - SUPREL	.889305	.748729	.042802	.805081	.973530	20.777	305	.000
Pair 4	EUT - STATQUO	1.019645	1.127448	.064770	.892187	1.147103	15.742	302	.000
Pair 5	EUT - COMM	-.011386	.561347	.032143	-.074636	.051865	-.354	304	.723
Pair 6	EUT - IHT	.664853	.792252	.045589	.575140	.754567	14.584	301	.000
Pair 7	EUT - EUAut	-.260864	.350494	.020135	-.300487	-.221240	-12.956	302	.000
Pair 8	EUT - EPISrel	.790827	.891765	.051062	.690347	.891308	15.487	304	.000
Pair 9	EPPERCP - CHMCKS	.73023	.83055	.04787	.63603	.82444	15.254	300	.000
Pair 10	EPPERCP - SUPREL	.94613	.82009	.04688	.85388	1.03838	20.181	305	.000
Pair 11	EPPERCP - STATQUO	1.07624	1.07472	.06174	.95474	1.19773	17.431	302	.000
Pair 12	EPPERCP - COMM	.04432	.65779	.03766	-.02980	.11843	1.177	304	.240
Pair 13	EPPERCP - IHT	.72064	.83172	.04786	.62646	.81482	15.057	301	.000
Pair 14	EPPERCP - EUAut	-.19648	.87152	.05007	-.29500	-.09795	-3.924	302	.000
Pair 15	EPPERCP - EPISrel	.84907	.93268	.05341	.74398	.95416	15.899	304	.000
Pair 16	CHMCKS - SUPREL	.22259	.82753	.04770	.12873	.31646	4.667	300	.000
Pair 17	CHMCKS - STATQUO	.35570	1.06093	.06146	.23476	.47665	5.788	297	.000
Pair 18	CHMCKS - COMM	-.68183	.81777	.04721	-.77475	-.58892	-14.441	299	.000
Pair 19	CHMCKS - IHT	-.00842	.86452	.05016	-.10714	.09031	-.168	296	.867
Pair 20	CHMCKS - EUAut	-.92754	.97782	.05655	-1.03882	-.81625	-16.402	298	.000
Pair 21	CHMCKS - EPISrel	.12167	.94905	.05479	.01384	.22950	2.220	299	.027
Pair 22	SUPREL - STATQUO	.12822	1.04107	.05981	.01052	.24591	2.144	302	.033
Pair 23	SUPREL - COMM	-.90230	.74343	.04257	-.98606	-.81853	-21.196	304	.000
Pair 24	SUPREL - IHT	-.22103	.81028	.04663	-.31278	-.12927	-4.740	301	.000

Pair 25	SUPREL - EUAut	-1.14642	.88817	.05102	-1.24683	-1.04602	-22.468	302	.000
Pair 26	SUPREL - EPISrel	-.09525	.48000	.02748	-.14933	-.04116	-3.465	304	.001
Pair 27	STATQUO - COMM	-1.03593	1.06517	.06129	-1.15655	-.91531	-16.901	301	.000
Pair 28	STATQUO - IHT	-.35117	1.09319	.06322	-.47559	-.22676	-5.555	298	.000
Pair 29	STATQUO - EUAut	-1.28000	1.20460	.06955	-1.41687	-1.14313	-18.405	299	.000
Pair 30	STATQUO - EPISrel	-.23179	1.19074	.06852	-.36663	-.09695	-3.383	301	.001
Pair 31	COMM - IHT	.67990	.74586	.04299	.59530	.76450	15.815	300	.000
Pair 32	COMM - EUAut	-.24906	.66113	.03804	-.32393	-.17420	-6.547	301	.000
Pair 33	COMM - EPISrel	.80707	.90901	.05214	.70448	.90967	15.480	303	.000
Pair 34	EUAut - EPISrel	1.05022	.99767	.05741	.93725	1.16320	18.294	301	.000

Table 41 – Pairwise t-tests dependent variables