

## MASTER

### Democracy in business process management

validation of process models created through the Plural method : a framework for decentralized business process modeling through modularization

van den Hurk, H.M.M.

*Award date:*  
2014

[Link to publication](#)

#### **Disclaimer**

This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

#### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

Eindhoven, July 2014

**Democracy in Business Process  
Management:**

**Validation of process models created  
through the Plural method; a framework  
for decentralized business process  
modeling through modularization**

By

H.M.M. (Henk) van den Hurk

BSc. Innovation Sciences

Student identity number: 06518564

in partial fulfillment of the requirements for the degree of

**Master of Science  
In Innovation Management**

Supervisors:

Dr. O. Türetken, TU/e, IS

Professor Dr. Ir. H.A. Reijers, TU/e, AIS

Dr. J. van Moll, Senior Quality & Regulatory Manager Quality Systems Philips Healthcare

TUE. School of Industrial Engineering.  
Series Master Theses Innovation Management

Subject headings: BUSINESS PROCESSES; REENGINEERING, Business process management, process modeling, BPMN, subject-oriented business process modeling, Modularity

*“Essentially, all models are wrong, but some are useful”*

George. E.P. Box (1919-2013)

## **Abstract**

In a common setting, business process models are designed by modeling experts who create their processes in a top-down fashion. However, this may not be the best way to model business processes. Current research studies the Plural method, a framework for bottom-up process modeling where process participants model their own process. Guided by a coordinator, 11 process participants modeled 4 of their own processes. These models were compared with classical models using an online experiment involving 51 participants. Analyses showed that though the application of the Plural method resulted in more complete process models, they were harder to understand and maintain than their classical counterparts. This can be attributed to the lack of control flow and larger size, which may hamper understanding. It has been shown that the Plural method is a powerful tool for process discovery, but an improvement on its models is needed to obtain full value of the framework.

## Preface

This section is the final part of this thesis to write and maybe it does not feel as glorious and grand as it should. This thesis concludes my life as a student and rings the bell to a new phase in my life. Though I have never spent this much time on a project before, it still does not feel like the grand finale to taking part of 20 years of the Dutch educational system. It will probably sink in to me in retrospect when I am applying for a job in the near future. However, what I can say about this thesis is that it has been, without a doubt, the most interesting educational experience of the previous 20 years. The ability to do research in an interesting company, on an interesting topic surrounded by interesting people felt unlike I have ever experienced before.

Before I can conclude this thesis, I do have to provide my thanks to the giants who's metaphorical shoulders my work rests on. First and foremost my gratitude goes out to Oktay Türetken, who has provided nothing less than excellent mentorship during the past one and a half year. He is a very busy man, guiding many students other than myself besides his other duties at TU/e. Regardless; he found time for thought provoking discussions, feedback and overall great counseling. Furthermore, I'd like to express my deepest thanks to the gentlemen who shared an office with me at Philips Healthcare: Jan van Moll, Zouhair Bedawi and Godfried Webers. These men have taken their time to aide and facilitate my research in the best possible way. This was done all while their regular duties kept going on during hectic times of audits and inspections. I would also like to say thanks to professor Reijers for asking the right questions at the right time to foster the quality of my research. His feedback has helped me greatly in reaching the point I am at right now.

Lastly, I would like to say thanks to my girlfriend Eva, with whom I have had many great discussions about our respective master's theses and my parents, who have provided me with outstanding support throughout my entire life.

I owe thanks to all these people and all other people who have added to my education and life. Thank you, I could not have had this great experience without you.

Henk van den Hurk  
Eindhoven, 2014

## Executive Summary

For any business, it is important to maintain clear overview of what is going on inside the organization. A frequently used paradigm for obtaining such oversight is Business Process Management. Business Process Management (BPM) is a body of methods, techniques and tools to discover, analyze, redesign, execute and monitor business processes (Dumas, La Rosa, Mendling, & Reijers, 2013). These business processes are defined as *“a collection of inter-related events, activities and decision points that involve a number of actors and objects, and that collectively lead to an outcome that is of value to at least one customer”* (Dumas, La Rosa, Mendling, & Reijers, 2013). The value of applying BPM in business has been shown frequently.

Currently, processes are often modeled by external consultants who have a high level of modeling knowledge, but are not Subject Matter Experts (SME's). This results in a set of business process model which have been made by so-called white-collars, who tell the blue-collars what to do (Fleischman, Raß, & Singer, 2013). This may pose a problem in the way process participants (who do the actual work) follow the described workflow of the process model. This may even more so be the case for knowledge workers, who have a high level of curiosity and creativity (Reinhardt, Schmidt, Sloep, & Drachsler, 2011). Therefore, it would be useful to include these people in the modeling effort of their own process.

Enter the Plural method. The Plural method is a framework for having process participants model their own processes (Türetken & Demirors, 2011) (Türetken & Demirors, 2013). During execution of the Plural method, the context is first defined. It is discussed with stakeholders what processes should be in the scope. Then, so-called operations are defined. In these operations, a process participant (now known as an 'agent') states what goes on in this specific task. The agent states what input is required for this task, what is being done with it and what the output is. If all operations of a certain process are defined in this way, they can be connected through their inputs and output to obtain a complete process model. If this is not the case, a conflict has occurred. E.g. one operation requires input no-one else offers. In such cases, the agents can discuss how the conflict should be resolved. After the process models are complete, they can be integrated for potential improvements. Due to the use of operations, a process model created through the Plural method becomes highly modular. This results in several potential benefits over classical models, such as easier understanding and maintenance. The framework of the Plural method is depicted in figure 1.

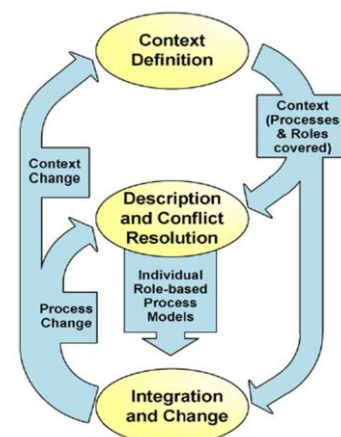


Figure 1: the Plural method

To assert the value of the process models created through the Plural method, the research question for this thesis was:

To assert the value of the process models created through the Plural method, the research question for this thesis was:

### **To what extent do the models created through the Plural Method outperform those created through classical process modeling techniques?**

Due to the highly modular nature of the process models, three hypotheses were coined, underpinned by literature:

- Hypothesis 1: A process model acquired through the Plural Method is easier to adapt to changes than a conventional process model.
- Hypothesis 2: A process model acquired through the Plural Method is easier to understand than a conventional process model.
- Hypothesis 3: A person is more willing to use the models created through the Plural Method, rather than using conventional business process models.

To assess these hypotheses, the Plural method had to be applied in a business setting. It was made possible to conduct this research in the Quality & Regulatory department of Philips Healthcare MRI. There, four processes were identified as part of the scope: Corrective Action & Preventive Action (CAPA), Complaint Handling, Field Change Order (FCO) and Risk Management. In the Quality Management System (QMS) of this business unit, textual descriptions are present for each of these, called ‘procedures’. These procedures sometimes contain a flowchart, but no formal process model. What is interesting about these procedures is that they often have been implemented in a top-down fashion, where a business layer higher up decides on the contents and makes the knowledge workers if the business units use them. These procedures were translated to a BPMN model in the classical method, where the researcher acted as a consultant. For each of these models, the process owner (the responsible employee with intimate knowledge) validated the models. These classical models served as a control group and can be found in the appendices of this thesis.

After the models were made in the classical method, participants were invited to model the same process, but this time from their perspective. The four processes were modeled by 11 experts and in all cases came up with correct models. The only process which did not seem to fit for the Plural method was the FCO process, which turned out to be in the middle of a redesign effort, causing some confusion. The Plural models showed quite some differences in content from the classical models, as they contained the same flow but with more exceptions and feedback loops. Furthermore, the creation of these Plural models pointed at some inconsistencies and unclear points in the classical models. This identification pointed at a value of the Plural method over the Classical method outside of the research scope.

The three hypotheses were asserted using an online experiment, where participants were invited to a questionnaire. In this questionnaire, participants were informed of what was expected of them and subsequently presented with two process models. For the experiment, only two out of four processes were selected to keep group sizes as large as possible. Each participant saw one of each processes and always saw a Plural and a classical model. For each model, the participant was asked eight questions. Six of those questions were to assert understanding based on control flow, data flow and organization aspects (van der Aalst, 2000). Furthermore, the participants were asked two questions to assert their ability to update the model. The questions were all validated by a Philips SME and designed similar to previous research (Mendling, Strembeck, & Recker, 2012). Furthermore, a total of eleven items was asked to

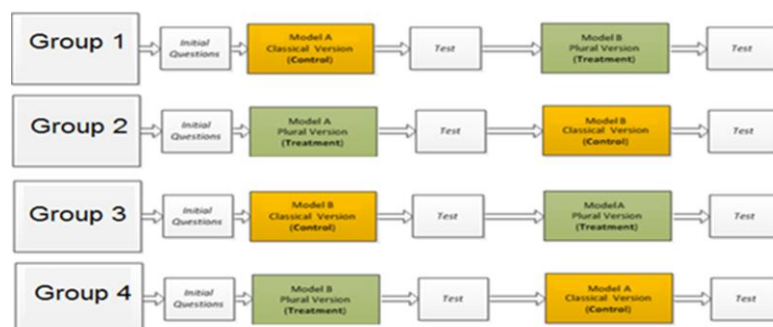


Figure 2: experiment design



assert the participant's intention to use, perceived usefulness, perceived ease of use and information retrieval, common to the Technology Acceptance model (TAM) (Venkatesh & Davis, 2000) (Seddon & Yip, 1992). This was to assert hypothesis 3. The experiment design is depicted in figure 2.

51 participants filled out the questionnaire, of which 10 people were involved in modeling sessions, 13 people were random Philips Healthcare employees and 28 people were students. Their performance for the questionnaire was asserted based on the number of correct questions, the time it took to answer these questions and the efficiency of obtaining correct questions (score divided by time). For the assertion of the hypotheses, both parametric and non-parametric tests had to be conducted as not all dependent variables possessed the parametric properties. The most common issue was the non-normality of some of the dependent variables. When transformations were not suitable, the non-parametric Mann-Whitney test was conducted, rather than parametric tests like the student t-test or the ANOVA. The result of the assertion of hypotheses is depicted in table 1.

Hypothesis	Result	#	Process	Significant results (order of model)
<i>H1: A process model acquired through the Plural Method is easier to adapt to changes than a conventional process model.</i>	Partially rejected	1	Complaint	Score for updating favors classical (1 <sup>st</sup> )
		2	Complaint	Efficiency update 1 favors classical (1 <sup>st</sup> )
		3	CAPA	Efficiency update 2 favors classical (2 <sup>nd</sup> )
<i>H2: A process model acquired through the Plural Method is easier to understand than a conventional process model.</i>	Partially rejected	4	Complaint	Efficiency understanding organizational aspects favors classical (both)
		5	Complaint	Time Q5 favors classical (both)
		6	Both	Efficiency Q3 favors classical (both)
		7	CAPA	Efficiency Q3 favors classical (2 <sup>nd</sup> )
		8	CAPA	Efficiency Q4 favors classical (both)
		9	CAPA	Efficiency Q4 favors classical (1 <sup>st</sup> )
		10	Complaint	Efficiency Q5 favors classical (1 <sup>st</sup> )
<i>H3: A person is more willing to use the models created through the Plural Method, rather than using conventional business process models.</i>	No support found	11	Both	None of the hypothesized differences among groups were significant

Table 1: results of the hypotheses

It was found that in most cases, the classical models outperformed the Plural models, which was equal parts surprising and interesting. The rejection of hypothesis 1 and 2 suggests that a completely modular model, like a Plural model, is not automatically easier to understand or maintain than a model with less modularity, like a classical model. It was found that besides differences in modularity between the Complaint Handling classical and Plural model, there was a large size difference. When completely expanded, the Plural model (183 nodes) was far larger than the classical model (97 nodes). This implies that the alleged beneficial effect of modularity can be counteracted by an increase in size.

Furthermore, in the CAPA process there was only a little difference in modularity between the Plural method (19 sub-processes) and the classical model (14 sub-processes). Still, the CAPA classical model was easier to understand for participants. This can be explained by the addition of some control flow

elements on the top level of the classical model. This points to the larger conclusion that 100% modularity is not the best way to make a process model understandable. Understandability seems to be fostered best when a large amount of modularity is combined with some control elements. The insignificance of hypothesis 3 can be attributed to the possibility that participants did not see the difference between Plural and Classical models. Where they would give significant different answers between the CAPA and the Complaint handling models, they did not respond significantly different for either of the two modeling methods (Plural versus Classical).

This research has answered the main question to some extent by showing that the Plural models were harder to understand than those created by the classical method. Still, the value of the Plural method cannot be dismissed as the process models created through the Plural method were more complete and correct than the classical models. The results of this study indicate that the Plural method outperforms the classical method in process discovery, but not yet in process understanding. This shows that the Plural method could be used as a process discovery tool right now. However, with more research on the use of modularity in the Plural models it could reach its full potential as a complete process modeling framework.

## List of abbreviations & Definitions

**CSF:** Critical Success Factors

**BPM:** Business Process Management

**BPMN:** Business Process Modeling and Notation

**BPR:** Business Process Reengineering

**CAPA:** Corrective Action & Preventive Action

**DFD:** Data Flow Diagram

**EPC:** Event-driven Process Chains

**FCO:** Field Change Order

**FDA:** Food & Drug Administration

**GoM:** Guidelines of Modeling

**PH:** Philips Healthcare

**Q&R:** Quality & Regulatory

**S-BPM:** Subject-oriented Business Process Management

**TRA:** Theory of Reasoned Action

**TAM:** Technology Acceptance Model

**Business Process:** a collection of inter-related events, activities and decision points that involve a number of actors and objects, and that collectively lead to an outcome that is of value to at least one customer.

**Business Process Management:** a body of methods, techniques and tools to discover, analyze, redesign, execute and monitor business processes.

**Business Process Modeling and Notation:** a process modeling notation with widespread support and attention. A large number of elements is present to model process very accurately.

**Business Process Redesign:** using analysis techniques to change an as-is model in a to-be model.

**Business Process Reengineering:** analyzing a business, start with a blank sheet of paper and completely create the business from scratch.

**Complaint Handling:** a business process in Philips Healthcare Magnetic Resonance unit. Entails receiving, analyzing, investigating and potentially resolving incoming complaints.

**Corrective Action & Preventive Action:** A business process in Philips Healthcare Magnetic Resonance unit. Entails the analysis, removal and monitoring of the root cause to a business problem.

**Critical Success Factor:** An influential power, paramount to successfully completing a project.

**Data Flow Diagram:** a tool for modeling data in a process or system.

**Encapsulation:** notion in modularization: adapting elements of one sub-process doesn't affect others

**Event-driven Process Chains:** a process modeling notation containing events, activities and decision points.

**Field Change Order:** a business process in Philips Healthcare Magnetic Resonance unit. Entails collecting small problems in the field, their collective solution and the rollout of that solution.

**Flowchart:** simplistic map of showing a process through activities and decision points.

**Food & Drug Administration:** governing authority of Food and Drug regulation in the United States of America.

**Information hiding:** hiding information in a system from a reader to prevent overburdening the reader. Facilitates understanding.

**Modularity:** The design principle of composing a large system into smaller, connected but separately manageable sub-systems

**Operation:** a vital construct of the Plural method. An operation consists of input received from another role, all the tasks the current role does with that input and the provided output.

**Petri nets:** a process modeling notation containing only two elements: states and transitions.

**Quality & Regulatory:** The business function in Philips Healthcare responsible for quality management and regulation management.

**Risk Management:** a business process in Philips Healthcare Magnetic Resonance unit. Entails management of all risks associated with a Philips Healthcare market, both prior to and after market release

**Subject-oriented Business Process Management:** a paradigm of business process management where the people performing the business process are placed in a central position.

**Sub-Process:** a part of a business process model which contains a self-contained process on a lower level.

**Plural Method:** A framework for having a process owner model his/her own process.

**Technology Acceptance Model:** a conceptual model to explain what factors drive acceptance of technology, based on TRA. TAM was extended to Technology Acceptance Model 2 by adding more factors.

**Theory of Reason Action:** the theory that human behavior by the person's belief towards that behavior along with the norm regarding that behavior.

## List of Figures

Figure 1: the phases of the Plural Method.....	3
Figure 2: a process in the Plural Method .....	4
Figure 3: The operation “Review Project plan” expanded .....	4
Figure 4: Research Design of this study.....	8
Figure 5: The BPM life-cycle .....	10
Figure 6: Example DFD of webshop.....	11
Figure 7: BPMN example .....	12
Figure 8: The SIQ model .....	14
Figure 9: expanding ‘Cut Hair’ in the barbershop process .....	15
Figure 10: Technology Acceptance Model .....	17
Figure 11: Experiment Design .....	23
Figure 12: Sub-Process Expansion .....	25
Figure 13: Demographics of respondents .....	29
Figure 14: Perceived difficulties of CAPA and Complaint Handling models.....	30
Figure 15: Boxplot for Scoring on Plural questions .....	31
Figure 16: Perceived difficulties per group .....	31
Figure 17: Total score per group .....	31
Figure 18: Score for Updating.....	32
Figure 19: Logarithm of time to update .....	33
Figure 20: Proposed new Plural framework.....	46

## List of Tables

Table 1: Overview of questions in experiment .....	25
Table 2: Mann-Whitney results for updating the Complaint Handling process (score).....	32
Table 3: Mann-Whitney results for updating the CAPA process (score).....	32
Table 4: Conducted tests for hypothesis 1 .....	34
Table 5: Conducted tests on total set of questions.....	35
Table 6: Mann-Whitney results for understanding control flow of the Complaint Handling process (score).....	35
Table 7: Mann-Whitney results for understanding control flow of the CAPA process (score).....	35
Table 8: Conducted tests on three types of understanding.....	36
Table 9: Conducted tests on all six individual questions.....	38
Table 10: Descriptives of TAM-related questions per model and process type .....	39
Table 11: Results of asserting hypotheses .....	40
Table 12: Metrics of researched process models.....	40

# Table of Contents

- Abstract ..... i
- Preface.....ii
- Executive Summary .....iii
- List of abbreviations & Definitions .....vii
- List of Figures.....ix
- List of Tables .....ix
- 1. Introduction..... 1
  - 1.1 The Research proposition..... 2
  - 1.2 The Plural Method..... 3
- 2. Research Questions..... 6
  - 2.1 Sub questions ..... 6
  - 2.2 Hypotheses ..... 6
- 3. Research Design ..... 7
  - 3.1 Research Context & project drivers ..... 8
- 4. Literature Review ..... 9
  - 4.1 A brief history of BPM in literature ..... 9
  - 4.2 The Business Process Management life-cycle..... 10
  - 4.3 Process Modeling ..... 11
  - 4.4 Understandability of Business Process Models..... 13
  - 4.5 Quality of process models ..... 14
  - 4.6 Sub-processes..... 15
  - 4.7 Critical Success Factors (CSF’s) ..... 15
    - 4.7.1 End-user participation ..... 15
    - 4.7.2 Employee Empowerment ..... 16
    - 4.7.3 Communication ..... 16
    - 4.7.4 Process ownership..... 17
  - 4.8 Adoption of Information Systems ..... 17
  - 4.9 TAM ..... 17
- 5. Trial Experiment ..... 18
  - 5.1 Scope of processes ..... 18
  - 5.2 Models of Trial Experiment ..... 18
- 6. Conducting Plural Modeling Sessions..... 19
  - 6.1 The Complaint Handling process..... 19

6.2 The FCO process .....	20
6.3 The Risk Management process.....	21
6.4 Reflection on the Plural Method .....	21
6.5 Adapting the Classical models.....	22
7. Experiment Construction.....	23
7.1 Question Construction .....	24
7.2 Creating the online experiment .....	25
7.3 Sampling .....	26
7.4 Internal validity & Survey errors .....	26
7.5 Rationale for made decisions .....	27
7.5.1 Experiment Design.....	27
7.5.2 Experiment setting .....	28
8. Results of Experiment.....	29
8.1 Descriptive statistics.....	29
8.2 Control variables .....	30
8.2.1 Learning effects .....	30
8.2.2 Differences between CAPA and Complaint Handling.....	30
8.2.3 Differences in modeling experience.....	31
8.2.4 Differences between groups .....	31
8.3 Hypothesis testing.....	32
8.3.1 Hypothesis 1 .....	32
8.3.2 Hypothesis 2 .....	35
8.3.3 Hypothesis 3 .....	38
9. Discussion.....	39
9.1 Discussion of experiment .....	39
9.2 Discussion of modeling sessions .....	43
10 . Limitations .....	44
11. Implications .....	45
11.1 Implications for Research.....	45
11.2 Implications for Managers.....	46
12. Future Research.....	46
13. Conclusion .....	48
References.....	xix
Appendices .....	xxiv

Appendix A: the CAPA process .....	xxiv
A1: CAPA classical.....	xxiv
A2: CAPA Plural.....	xxvii
A3: CAPA Classical adapted for experiment.....	xxxiii
Appendix B: the Complaint Handling process .....	xxxiii
B1: Complaint Handling Classical .....	xxxiii
B2: Complaint Handling Plural.....	xxxiv
B3: Complaint Handling Classical adapted for experiment.....	xxxix
Appendix C: the FCO process .....	xxxix
C1: FCO Classical.....	xxxix
C2: FCO Plural (incomplete) .....	xli
Appendix D: Post Market Risk Management .....	xliii
D1: Risk Management Classical .....	xliii
D2: Risk Management Plural .....	xliv
Appendix E: Questions and Questionnaire .....	xlvi
Appendix E1: constructed questions.....	xlvi
Appendix E2: TAM (and related) questions.....	xlvi
Appendix E3: other questions .....	xlvi
Appendix E4: Relevant screenshots of questionnaire.....	xlvi
Appendix F: Descriptive results experiment .....	lii
Appendix G: data and data processing.....	liv



## 1. Introduction

Businesses, large international ones in particular, constantly struggle to improve their competitive position. They try to achieve this by increasing effectiveness as an organization while maintaining internal efficiency. Though it might be possible for small business owners to have a clear overview of their company, large business are not as transparent. Small businesses may only have a single segment of customers, only a few employees and all overview is still present in the general manager. However, once the corporation increases in size, this overview cannot be maintained anymore. Therefore, it is important to incorporate a vision which helps maintain this overview and allow for control over the way your business functions. A well-researched, established and generally accepted vision is Business Process Management. Business Process Management (BPM) is a body of methods, techniques and tools to discover, analyze, redesign, execute and monitor business processes (Dumas, La Rosa, Mendling, & Reijers, 2013).

A business process can intuitively be defined as ‘one of the things your business does’. However, there should be more to an organizational effort before it can be called a business process. There are several important elements before the term business process applies. The first important thing is that business process is not a business function. Business functions can be any of operations, sales, marketing, support, human resources and such. However, business processes should be *cross-functional*. Furthermore, a business process should provide some value to a customer, be that an internal or external one. Therefore, a business process is defined as such: “*a collection of inter-related events, activities and decision points that involve a number of actors and objects, and that collectively lead to an outcome that is of value to at least one customer*” (Dumas, La Rosa, Mendling, & Reijers, 2013). The value of applying BPM in a company has been shown frequently (Jeston & Nelis, 2014). Process orientation has been linked to organizational performance indicators like increase customer satisfaction, cost reduction and speed improvement (Kohlbacher, 2010) (McCormack, 2001). As the definition of a business process requires value to the customer, it is not odd that improving business processes has been linked to increasing customer satisfaction (Kumar, Smart, Maddern, & Maull, 2008).

In order to understand and improve ones business processes, it is important to model them. The act of process modeling means to capture all elements of a business process in a graphical way. Once the current ‘as-is’ situation is modeled, the process model can be used for any redesign effort. In a redesign effort, a business process is not viewed the way it is right now, but the way it should be in the future (Davenport & Short, 1990). Modeling the as-is processes, often referred to as Process Discovery, requires a lot of effort; there is a separation in know-how. People who are intimately known with the contents of a certain business process (i.e. the people who execute it) are rarely trained in modeling techniques. However, the people who are modeling experts might lack domain-specific knowledge.

In BPM projects (a BPM Initiative), Process Discovery is usually mentioned as the most time-consuming phase with about 40% of total time spent (Wolf & Harmon, 2006). There are some ways to mitigate the separation of know-how. People who execute the process (process participants) can be trained in modeling techniques, which allows them to model their own process directly. This is a costly and time-consuming effort. Another option is to get the modeling expert familiar with the

domain. This is a more common approach, as there are several techniques to achieve this. The modeling expert can perform interviews, workshops or document analysis (Dumas, La Rosa, Mendling, & Reijers, 2013). A large body of knowledge is attributed to performing process discovery through several automated techniques like process mining (van der Werf, van Dongen, Hurkens, & Serebrenik, 2008) (van Dongen, De Medeiros, & Wen, 2009).

However, domain-knowledge is not the only factor which limits the success of a BPM Initiative. There are numerous matters which play an important role for the BPM Initiative to be successful. These matters, known as Critical Success Factors (CSF's) have been widely discussed in BPM literature (Parkes, 2002) (Mutschler, Reichert, & Bumiller, 2007) (zur Muehlen & Ho, 2006). It turns out that the most frequently cited CSF's often concern informing and including end-users in the BPM Initiative. This shows that these end-users take a crucial role in the progress of a BPM Initiative.

## 1.1 The Research proposition

In a classical setup, modeling business processes is often done by an external consultant, or by an internal process modeling expert (Fleischman, Raß, & Singer, 2013). These people might observe the processes and perform workshops and/or interviews with the employees to find out how a process is executed. They are often skilled in process modeling, but are not familiar or experienced with the execution of the process. The result is a set of business process models, *as perceived by the external consultant*. When the modeled process is a production process like an assembly line, this does not pose any problem. However, this proves to be a problem when the modeled process entails knowledge workers. Knowledge workers have a high level of curiosity and creativity (Reinhardt, Schmidt, Sloep, & Drachsler, 2011). This is among the reasons why they do not like it when external people tell them how they should do their job (i.e. the "white collars" tell the "blue collars" how to work) (Fleischman, Raß, & Singer, 2013). The full value of a process model is not utilized as it could be, which could be attributed to the way these models come about.

An important facet of Business Process Models lies with their complexity. A lot of BPM literature has been devoted to research how end users' understanding of process models comes about (Reijers, Mendling, & Recker, 2010) (Mendling, Strembeck, & Recker, 2012) (Reijers & Mendling, 2011). A gap of knowledge lies between the creators of business process models (i.e. the "white collars") and the people who actually use them (i.e. "the blue collars"), as the former is extensively trained in business process modeling and the latter is not. However, it is important that the users of a process model are able to understand them. Therefore, process models should be made as intuitive as possible to facilitate end user understanding.

A final important aspect of business process modeling is the adaption to changes. A business process model is a snapshot of the way the process looks at the moment it is being modeled (Lindsay, Downs, & Lunn, 2003). However, it is rarely the case that an organization, especially a large one at that, does not change its structure for an extended period of time. According to Saidani and Nurcan that can be attributed to the ever changing demands of customers, which is why business process models should be able to adapt to changes in context (Saidani & Nurcan, 2007). Therefore, it is important that a business process model allows for easy updating and maintenance.

## 1.2 The Plural Method

The previous section addressed several aspects of business process modeling which are crucial to current research. To deal with these aspects, the Plural Method was developed by Türetken et al. (Türetken & Demirors, 2011) (Türetken & Demirors, 2013). This method allows participants to model their own processes. The result is a set of models which resembles the process best, according to the people who actually perform these processes. Furthermore, the structure of the models will be different from a classical process model, which facilitates both understanding and maintenance through modularity. The Plural Method places a heavy focus on the inputs and outputs of process participants.

In the Plural Method, a process modeling expert *guides the modeling effort*, rather than executing it. The framework for this method can be seen in Figure 1. The actual modeling of processes is left to participants. The process modeling expert, known as the coordinator, helps stakeholders define the scope of the BPM Initiative. This first phase is called the context definition. During definition of the context the roles in a process are identified, along with the responsibilities of each role. This set of roles and their responsibilities acts as an architecture to ensure a structured modeling effort. Afterwards, so-called *operations* can be defined. An operation is the amount of work a single role performs before he/she needs new input. This means that an operation always consists of a piece of input, one or more tasks (executed without needing new input) and one or more pieces of output. These operations will be modeled by experts: people who perform these responsibilities in their work every day. These people are from now on referred to as agents. This is part of the Description and conflict Resolution Phase.

Once all operations have been defined, inputs and outputs can be linked to create one big process model. This is possible since the output of an operation should serve as input to another operation. However, it might be that case that not all inputs and outputs are accounted for. People might offer output which no one (inside or outside the process) uses, or input may be required that no one offers. Often, this will be attributed to the fact that people have different terminology for output or input. Where one role calls a specific document 'Product Design Review', another role might call it a 'Design Specification File'. They mean the same file while giving it a different name. In the Plural Method, these differences in terminology immediately surface as a process model is not finished until all inputs and outputs have been accounted for. This phenomenon is called 'Conflict Resolution'. Once these conflicts have been resolved, these individual models can be fed to the *Integration and Change* phase. In this phase, the models can be integrated to create complete overview models. These can be used to perform a process improvement cycle (Türetken & Demirors, 2011).

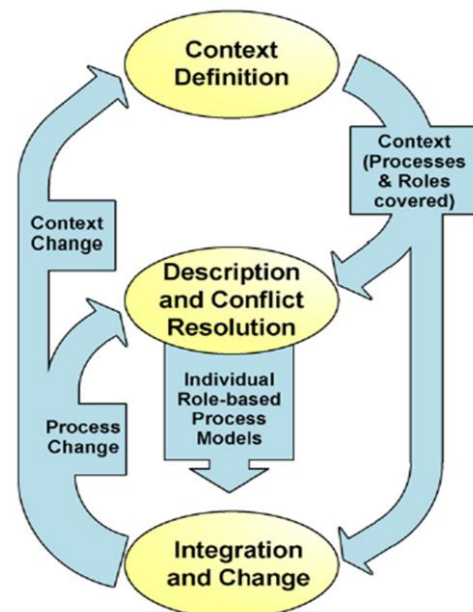


Figure 1: the phases of the Plural Method

The use of operations allows people to focus on the parts they are responsible for rather than being mentally overloaded with the entire process. The concept of shielding users from irrelevant information is known as ‘information hiding’. This concept is well-known in literature of BPM and computer science (Reijers & Mendling, 2008) (Parnas, 1972). Information hiding is applied in Plural through the use of operations, as mentioned earlier. Each operation will be modeled as a sub-process, in which the agent can define his/her activities. To facilitate information hiding, the roles of others are shown as a ‘black box’. Something comes out of the black box (input needed for the current operation), activities are executed with this information and consequently output is sent to another black box. Both in computer science and BPM, information hiding has been linked to increased understanding.

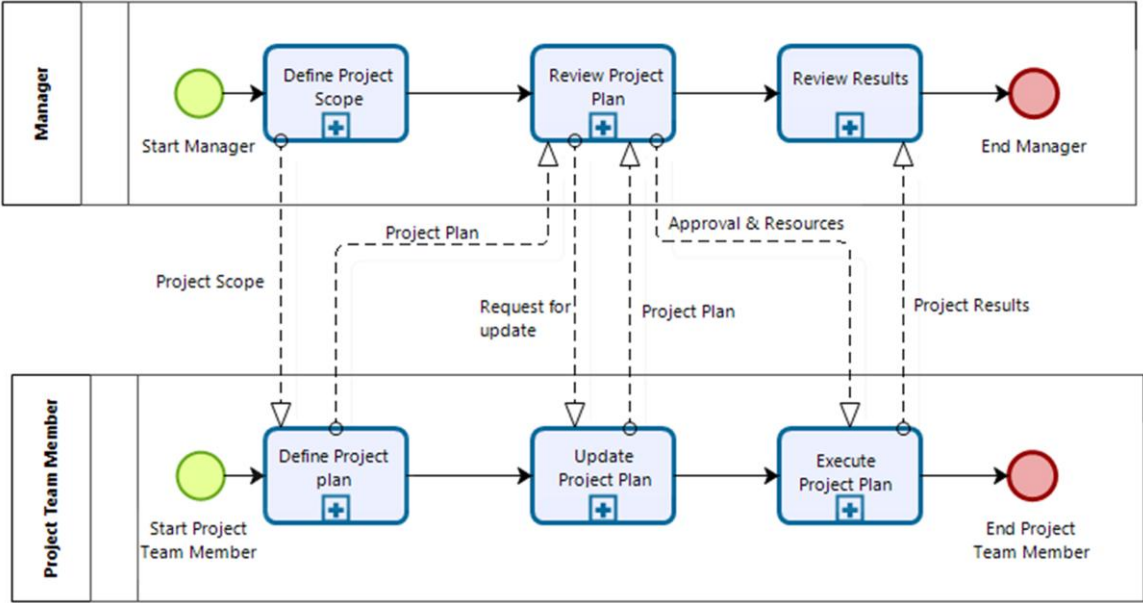


Figure 2: a process in the Plural Method

For clarification of this concept, please review figure 2. It shows a top level process of defining and executing a project, modeled in the Plural Method. On the top level only roles, their operations and the interface is shown. The complete internal behavior is modeled in sub-processes. This allows for quick understanding of the responsibilities and interface of each role. If someone is interested in the exact workings of an operation, it can be expanded to show its behavior.

For instance the operation ‘Review Project Plan’ can be seen in figure 3 when expanded. As

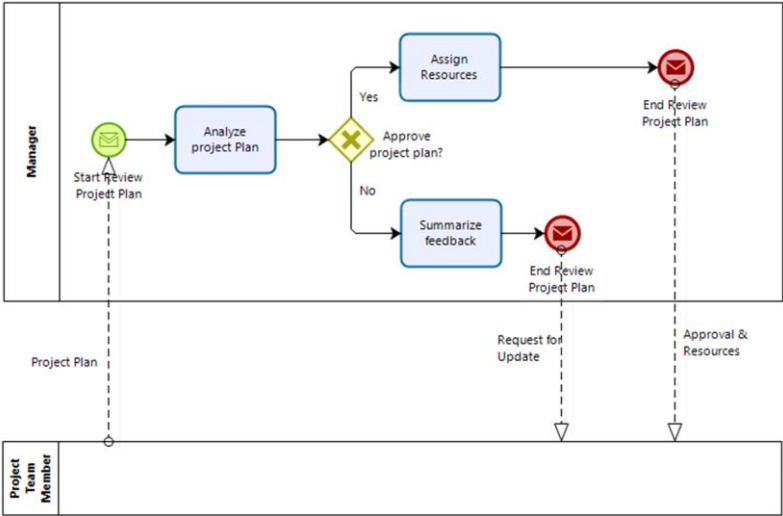


Figure 3: The operation “Review Project plan” expanded

can be seen, the actual behavior of the role 'Project Team Member' is not shown. When the agent who performs the role 'Manager' starts modeling this operation, he/she is presented with these two empty pools. The agent can model the behavior of 'Review Project Plan' and state what input is needed from other roles and what output is presented. Once the agent has indicated what exchange of information is necessary to perform his task, this can be added to the interface on the top level.

This way of obtaining the final process model, causes it *to be structurally different from a classical process model*. Where a classical process model shows more behavior of all roles on the highest level, a process model created through the Plural Method will focus on showing the global responsibilities (the operations) and the interface (input and output) on the top level. This different structure has several forms of impact on the usability of a process model. The first difference over a classical process model is its high level of modularity; Plural Models are completely modularized. A higher level of modularity has been frequently linked to an increase in process model understanding through the notion of information hiding (Mendling, Reijers, & Cardoso, 2007) (Reijers & Mendling, 2008). The notion of modularity is also a very well-researched topic in computer science, for instance in object-oriented programming. This type of computer programs has a high level of modularity (like a process model created through the Plural Method has), which causes it to be easier to maintain (Henry & Humphrey, 1988). This increased maintainability is linked to the notion of encapsulation. Encapsulation means that you can change the content of a sub-process, without disturbing the rest of the process. This allows for easier maintenance of a process which should adapt to the ever-changing context of a business.

Though the use of modularity and the related notions of information and encapsulation hiding have frequently been linked to increased understandability and maintenance, Plural Models show very little of the control flow on the top level. This could be perceived as a downside to the method, as control flow will only be apparent once operations will be expanded. However, there seems to exist a trade-off. As fewer control flow elements are displayed on the top level, it appears to be easier to identify data flow between operations. This, in turn, would increase understandability.

One crucial part of the Plural Method is its independence of any specific notation. In the original papers by Türetken et al. the notation of choice was a modified version of Event-driven Process Chains (EPC) (Türetken & Demirors, 2007). At the time, this was a smart choice as the notation was rather popular. However in recent years the notation Business Process Management & Notation (BPMN) has developed new and improved specifications, called BPMN 2.0. This notation provides rich expressive power, combined with quite a few intuitive business elements. Therefore current research will be the very first study which applies the Plural Method through BPMN 2.0. According to a report by Wolf and Harmon the notation of BPMN is popular among both scientists and practitioners (Wolf & Harmon, 2012). Since BPMN allows for many options in the use of sub-processes, it could fit perfectly for modeling operations. However, from figure 3 one can observe that a Pool element is being used in a sub-process. This is not allowed in regular BPMN 2.0. Therefore, a small update to the notation is proposed to be able to express the source and destination of input/output in an operation. Another powerful advantage of using BPMN is the high availability of modeling tools. The widespread popularity of BPMN might be facilitated by this large offer of tools.

Employing the Plural Method should result in significant advantages over conventional process modeling techniques (Türetken & Demirors, 2011). The current research is about validation of the

models created through the Plural Method by conducting an experiment. Using this experiment, the value of the Plural Method could be revealed. The main focus lies on validation of advantages of the resulting process models, as well as the preference of process participants to use these models.

## 2. Research Questions

The research proposition mentioned how several important aspects in business process modeling could be addressed using the Plural Method. As these claims are relative in nature, this was the main question of this research:

**To what extent do the models created through the Plural Method outperform those created through classical process modeling techniques?**

### 2.1 Sub questions

To determine the relative performance of the Plural Method, several sub questions were defined. From the original papers, two main aspects were reported about the advantages of the Plural Method. The first aspect stated that since the Plural Method allows people to have power and influence over the process models, they will feel empowered and prefer these process models over conventional process models. The second aspect states that the models created through the Plural Method will have a higher level of understandability and maintainability through the use of modularity (Türetken & Demirors, 2013). This research was concerned with addressing this second aspect of the Plural Method. To guide answering the research question, the following sub-questions were defined:

- 1.) *Is a process model acquired through the Plural Method easier to adapt to changes than a conventional process model?*
- 2.) *Is a process model acquired through the Plural Method easier to understand than a conventional process model?*

In order to address the first aspect of the Plural Method as well, this questions is asserted in addition:

- 3.) *Are people more willing to use the models created through the Plural Method, rather than using conventional business process models?*

### 2.2 Hypotheses

The previous section states the research questions regarding validation of the Plural Method. Hypotheses were made for each of the three sub questions. They are reported below.

The process models created through the Plural Method are highly modular. This allows users to easily see what compartmentalized parts of the models are affected once other parts are updated. Furthermore, the use of modularization has been linked to a higher level of maintainability in software development (Henry & Humphrey, 1988) (Mayer, 1988). Another important aspect of updating a process model is identifying the dependent parts. In the Plural Models, if an update does not change anything about its interface, it can quickly be implemented. If updating an operation does change the interface, it immediately becomes clear inside that operation which other roles are affected by this update. The occurrence of this effect is called encapsulation. These notions lead to the first hypothesis:

*Hypothesis 1: A process model acquired through the Plural Method is easier to adapt to changes than a conventional process model.*

The process models created through the Plural Method resemble the structure of Object Oriented programming (Türetken & Demirors, 2011). The process models are created through modeling a set of operations for every role and define how they interact. As these modular operations can be expanded and collapsed, irrelevant information can be easily hidden. Smaller models with a modular structure are easier to understand, rather than models with a larger amount of elements (Mendling, Reijers, & van der Aalst, 2010) (Reijers & Mendling, 2011). Therefore the second hypothesis ensues:

*Hypothesis 2: A process model acquired through the Plural Method is easier to understand than a conventional process model.*

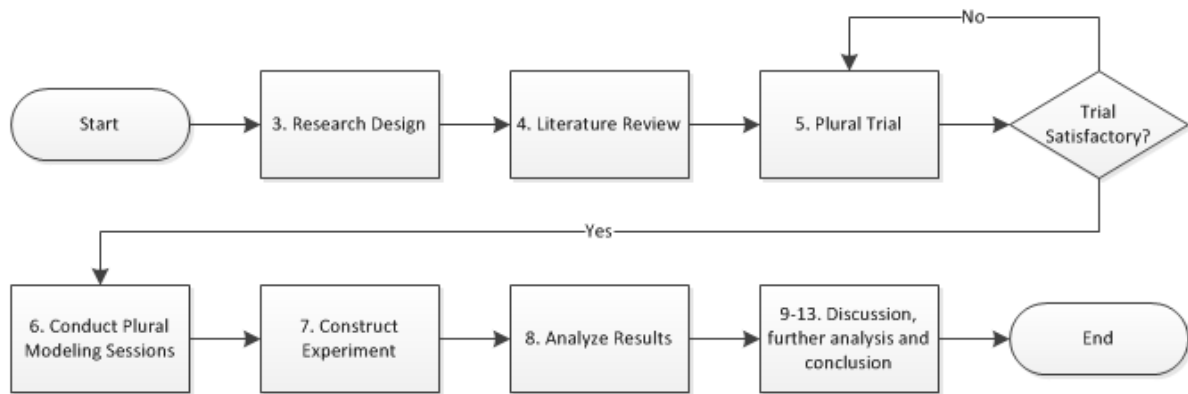
The driver for the final hypothesis lies in the fact that participants model their own process. From the field of subject-oriented process modeling, it is known that people do not like it when an external consultant tells them how they do and or should do their jobs (Fleischman, Raß, & Singer, 2013). As the Plural Method employs very similar concepts, it is likely that people have a high stake in their process and are allowed to determine what the process model looks like, so they are presented with responsibility over the process. This responsibility is a part of process ownership, which increases the intention to use a new technology (Karahanna, Straub, & Chervany, 1999). Moreover, it has been shown that intention to use can be predicted by the constructs of perceived usefulness and the perceived ease of use (Venkatesh & Davis, 2000). As the previous hypotheses stated that the Plural models are easier to use than their classical counterparts, it could be hypothesized that people prefer working with the Plural models. This dictates the third and final hypothesis:

*Hypothesis 3: A person is more willing to use the models created through the Plural Method, rather than using conventional business process models.*

This wraps up the research questions and the accompanying hypotheses. The rest of this thesis is about finding a method to answer these questions and assert the hypotheses, executing that method and analyzing the results.

### **3. Research Design**

To answer the questions which were stated in the previous section, a Research Design was outlined. This Research Design describes in detail what was done in this research and how the rest of this report is structured. To foster the clarity of this report, the structure of the rest of this research is depicted in figure 4. This figure outlines the order of steps taken in this research. Each of the steps depicted in this figure represents a chapter; starting with the current chapter.



**Figure 4: Research Design of this study**

The flowchart of figure 4 starts with the current chapter, the Research Design. After outlining the Research Design, a Literature Review was conducted. This Literature Review was based on a previously written report and can be found in chapter 4. Chapter 5 describes the execution of a short Plural Trial; it was done in order to make sure the used ‘flavor’ of the Plural method was sufficient for current research. Since this was the case, the rest of the research could continue. Chapter 6 describes the conducted Plural modeling sessions in detail and chapter 7 describes how the experiment to assert the hypotheses was constructed. Chapter 8 outlines the statistical analyses used to assess the hypotheses. Chapters 9 through 13 contain all important sections after the results: discussion (chapter 9), limitations (chapter 10), implications (chapter 11), future research (chapter 12) and finally the conclusion to this research (chapter 13).

### 3.1 Research Context & project drivers

The research proposition of section 1.1 mentions how several important aspects in business process modeling can be addressed. Validating the Plural Method requires application and research in an actual company. Such research should be executed in a company which, preferably, already has defined some of their processes. Some of those processes could be selected in the scope and the Plural Method can be applied. Experiments can be conducted to measure the quality of the new models compared to the existing ones.

During the initial phases of this project a suitable company was selected: Philips Healthcare. Philips Healthcare is the largest division of Philips and employs over 3000 people in their headquarters alone. The headquarters of Philips Healthcare are located in Best, a town near Eindhoven. Philips Healthcare is market leader in several different medical systems, like MRI-machines and CT-scans (Philips Healthcare, 2013). A quality manager of the Philips Healthcare MRI division approved of using some of their processes as a field setting for the experiments. This business unit has already defined many of their core processes, which would provide a suitable environment for research.

Philips Healthcare is a company which is subject to strict regulations, due to the nature of their products. This results in several high-impact checks by control agencies. A check of a very influential organization is pending in the near future: the Food & Drug Administration (FDA). The FDA determines whether a firm complies with a set of high standards in order to guarantee the safety of US citizens (FDA - What we do, 2013). A negative result could mean that Philips Healthcare products cannot be shipped over US borders. The results of such a drastic decision would be disastrous for Philips Healthcare and should be avoided at all times. Among other things, the FDA is keen on



checking out the processes of a company and the improvements they have made. The upcoming check of the FDA increases the tension in some of the departments, but it also brings about a clear interest in improving processes.

This combination of factors makes Philips Healthcare a very suitable environment for research. They have both powerful internal and external motivation for looking into process improvements. Their processes are already defined, which adds to its suitability. Please note that these processes are defined, but not yet modeled. Their processes have been captured in textual documents, called procedures and work instructions. What is particularly interesting about these procedures, is that they in fact have been forced upon the process owners and participants by higher layer in the business structure. This results in differences in implementation among different business units and a general skeptical attitude towards the procedures. Furthermore, Philips Healthcare is driven by knowledge and the firm has a high percentage of knowledge workers. It has also been established that the Plural Method is most applicable in a knowledge-centric organization (Türetken & Demirors, 2013) (Türetken & Demirors, 2011).

## 4. Literature Review

To support the current research, a literature review was conducted. This literature review contains an analysis of all relevant pieces of BPM literature to support current research. It starts with a short history of BPM and then continues to address all relevant topics one by one.

### 4.1 A brief history of BPM in literature

Where the first industrial revolution mainly revolutionized the way we use technology for production, the second industrial revolution aimed to revolutionize human labor. A pioneer in this area was Frederick W. Taylor, a mechanical engineer born in the 18<sup>th</sup> century. With his groundbreaking work on human labor, he could achieve large improvements in productivity (Taylor, 1911). Taylor's methods included division of labor, specialization and scientific observation of the workers. Though this method achieved vast improvements, a downside occurred. The focus of this approach was too much centered on internal efficiency, rather than pleasing the customer. Customer dissatisfaction occurred and so-called Taylorism did not seem to be the solution anymore.

A new wave of industrial engineers rose during the late eighties, early nineties. Their way of thinking was called "*Process Thinking*" or "*Business Process Reengineering*". This wave was pioneered by several seminal works, among which the works of Davenport & Short and of Hammer & Champy (Davenport & Short, 1990) (Hammer & Champy, 1993). The focus of this movement was to improve business processes, but in a rather revolutionary way. Instead of taking the present form of the organization and trying to adapt it, the idea is to take a blank sheet of paper and try to "*reengineer*" the company to a desirable new state. The idea was to improve customer satisfaction by improving processes using Information Technology (IT). Both works focus on IT as being a key enabler of Business Process Reengineering. The following quote can be read:

*"We say that in reengineering, information technology acts as an essential enabler. Without information technology, the process cannot be reengineered."* (Hammer & Champy, 1993)

This way of thinking established a large amount of remarkable results. Application of Business Process Redesign in Rank Xerox U.K caused a reduction of 27 days in delivery time and halted a long period of financial stagnation to a 20% growth in revenue (Davenport & Short, 1990). In another case, implementation of this method reduced a 7 day throughput time at IBM Credit Corporation to a grand total of 4 hours per case (Hammer & Champy, 1993). Halfway through the nineties of the last century, over 60% of the Fortune 500 companies claimed to incorporate some type of Business Process Reengineering (Hamscher, 1994).

These tangible results showed the merit of Business Process Reengineering, but there were downfalls lurking about. Besides the great success stories of BPR, there seems to be a number of pitfalls in implementation. In a Business Process Reengineering project at Pacific Bell’s Centrex, criticism of the basic assumptions of BPR is expressed and modified assumptions are presented (Stoddard, Jarvenpaa, & Littlejohn, 1996). They found that the basic assumptions of BPR were too radical. They stated that the redesign should be performed radical, in a large scale and in a top-down fashion. However, implementation of that reengineering should be done in an incremental way, on a more specific scale and owned bottom-up. These problems with the traditional view of BPR are confirmed in the book “*Fundamentals of Business Process Management*” (Dumas, La Rosa, Mendling, & Reijers, Fundamentals of Business Process Management, 2013). They stated that the problems of BPR were Concept Misuse (1), Over-Radicalism (2) and Support Immaturity (3).

However, two main events can be credited for the rise of BPM. Empirical research showed that firms with a process orientation performed better than firms that did not have such an orientation (Kohlbacher, 2010) (McCormack, 2001). Furthermore, the usage of Information Systems in companies advanced; most notably, the rise of Enterprise Resource Planning systems (ERP) and Workflow Management Systems (WfMSs). BPM seems to be a revival of the process thinking of the nineties, but with a larger scope. Where BPR focusses on breaking down the current organization and reengineer it, BPM follows a more cyclical approach where the Business Process Reengineering is just a part of the entire paradigm.

## 4.2 The Business Process Management life-cycle

The paradigm of BPM can be expressed in the business process life-cycle, a model which can be seen in figure 5 (Dumas, La Rosa, Mendling, & Reijers, 2013). This model shows a cyclic process, where it starts all over again after a full iteration. In the first phase, *Process Identification* (1), a business problem is analyzed and the relevant business processes are identified. Those relevant processes are modeled in the *Process Discovery* (2) phase. This step is often aided by the use of as-is models from the company. In the step of *Process Analysis* (3), issues with the relevant business models are identified and listed in order of severity or ease of change. In *Process Redesign* (4) solutions to the previously found issues are proposed. The steps of *Process Analysis* and *Process Redesign* are very much related, as solutions will be analyzed immediately as well. The best solutions are grouped together and often depicted in a new process model. The result of this step is a to-be process model.

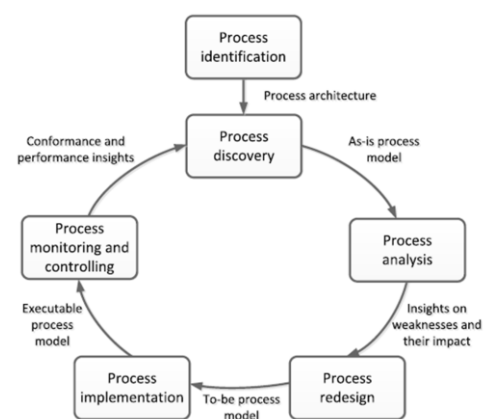


Figure 5: The BPM life-cycle

The fifth step is *Process Implementation (5)*, where the coined solutions are actually executed in the business. This means the steps are taken to change the process from the as-is state to the to-be state. This step consists of two parts: Organizational change management and process automation. Though both parts are of equal importance, Dumas et.al (2013) mainly focuses on the part of process automation. This refers to the use of Information Systems to support the to-be process. The final step that completes the cycle is *Process Monitoring and Controlling (6)*. In this step, relevant data is gathered to assess whether the new situation works as it was intended. Once deviations from the intended effect are measured, a corrective action can be undertaken. When new issues arise, the cycle can enter a new iteration. The BPM life-cycle provides a structured overview of the different aspects of current day BPM. In short, it is the current conclusion to over a hundred years of business process related research.

### 4.3 Process Modeling

Process modeling is the act of capturing the information of a business process in a graphical representation. This may include a number of modeling techniques, many of which have a rigorous syntax. Usually, a process model defines a flow of objects, documents or other entities through the process. This thing flowing through the model will from now on be referred to as a flow object. Up to the resurgence point of 'process thinking' in the 90's, several techniques of modeling were used, though the goal was developing software rather than modeling business processes (Curtis, Kellner, & Over, 1992). Though there are several important known notations for modeling processes, only two will be described in more detail here: Data Flow Diagrams (DFD's) and Business Process Modeling & Notation (BPMN).

A Data Flow Diagram (DFD) was already used quite a lot in practice and has been a software developing staple for many years (Yourdon, 1989). A DFD is a model which places large focus on the manipulation of data through a process. A DFD contains two main elements: circles and arcs. A circle represents an operation on a piece of data (some kind of process) and the arcs represent the flow of this data. A DFD shows very clearly how data changes throughout a certain process. An example DFD is depicted in figure 6.

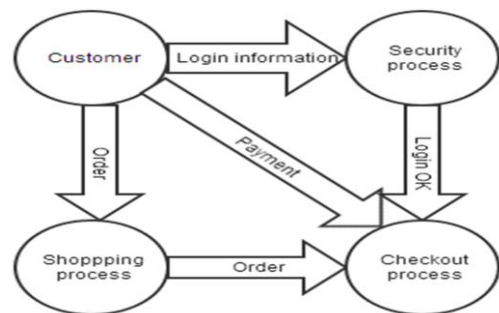


Figure 6: Example DFD of webshop

In spite of its widespread adoption, DFD's could not be used for modeling business processes. This can be attributed to two major limitations. First, DFD's are mostly concerned with the modeling of data and do not support detailed modeling of people or other business constructs. Second, DFD's do not support decision points and ordering of steps (i.e., no control flow) (Giaglis, 2001). As can be seen later, a process model acquired through the Plural method resembles a DFD quite a lot. This attribute should facilitate understanding data flow in a Plural model.

One of the most commonly used methods of process modeling is so-called Business Process Modeling and Notation (BPMN). This might be attributed to its rich notation, the active support by the governing organization and the large availability of supporting computer tools (OMG.org, 2011). BPMN was originally developed by an organization called Business Process Management Initiative (BPMI). Their first specification was released in 2004, called BPMN 1.0. After a merger with the Object Modeling Group (OMG), the first OMG regulated specification of BPMN 1.0 was released

(OMG, 2006). Over the course of the following years, several new versions were released and the notation matured. Through version 1.1 (2007), version 1.2 (2008) and two betas of version 2.0 (2009 and 2010) the final specification were released under version 2.0 (OMG.org, 2011). Business analysts from practice began to notice since the original launch and BPMN increased in usage. Since OMG actively supports and updates BPMN using versions, the usefulness of the notation increases over time (Recker J. , 2010). In a 2012 survey about the current state of Business Process Management, it was indicated that 60% of the respondents from practice had the desire to use BPMN for their process modeling activities (Wolf & Harmon, 2012).

BPMN consists of four types of elements: Flow elements, connecting elements, swimlanes and artifacts. A flow element is an event, an activity or a gateway. They stipulate what is happening throughout the model. A connecting element connects flow elements to each other. Examples are simple arrows (to indicate a sequential relationship) or a message (to indicate communication). Swimlanes provide structure to show which elements are executed by the same party. There are several kinds of artifacts. The most common ones are data elements, grouping with a dotted line or textual annotation. An example of a basic BPMN process is included in Figure 7.

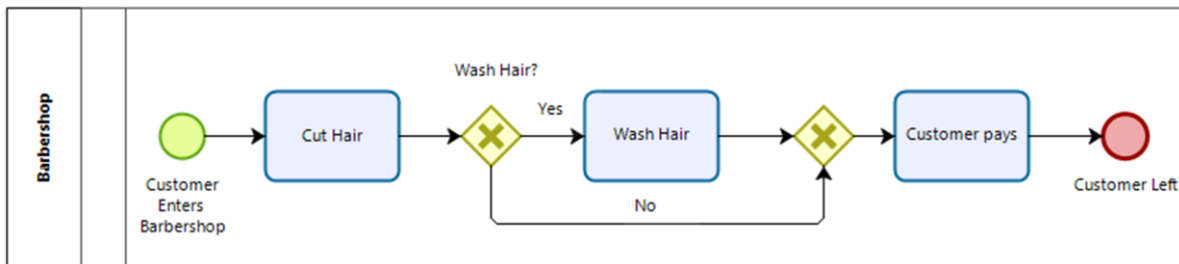


Figure 7: BPMN example

The example shows a simple business process. A customer enters the barbershop, the hairs are cut and now a choice can be made to wash the hair or not. Afterwards, the customer pays and leaves the shop.

Though the example of Figure 7 is very basic, a lot of different additional elements can be used to specify business processes. BPMN grew from already 48 elements at BPMN 1.0 to a total of 116 elements in BPMN 2.0 beta. The extensive use of BPMN in practice and research along with its powerful support allow BPMN 2.0 to be a useful tool for current research. The power of BPMN lies, among other things, in its ability to very clearly specify different kinds of behavior. BPMN allows for a whole lot of different types of events to graphically indicate what you mean. This includes events to show a specific amount of time has lapsed, a message is received, an error occurred and so on and so forth.

Besides plain process modeling, BPMN lends itself very well to other important features of BPM. BPMN 2.0 has a very rich notation which allows people to simulate and even automate their business. Especially the development of BPMN 2.0 over the older versions allow for better execution. The execution semantics of constructs was completely formalized (Chinosi & Trombetta, 2012). In a survey among practitioners, nearly 40% of the respondents indicated to use BPMN 2.0, and 30% still used BPMN 1.2 (Chinosi & Trombetta, 2012). Another important aspect of BPMN is the extensive amount of tools available for creating BPMN models (Yan, Reijers, & Dijkman, 2010) (Recker J. , 2012).

#### 4.4 Understandability of Business Process Models

Process models are often complex diagrams and are not easy to understand. However, the importance of the process models lies, among others, in communication among stakeholders (Reijers, Mendling, & Recker, 2010). Therefore, it is important to know how the understandability of a process model works, and how this understandability can be improved. A lot of research has been done on the comprehension factors of process models.

It has been shown that process models with textual labels are harder to understand syntactically than when these labels are plain numbers (Mendling, Strembeck, & Recker, 2012). This can be attributed to the cognitive loading theory, where it is thought that people have a finite amount of cognitive resources and that the textual labels of process models use some of those resources which could otherwise be used for syntactical comprehension. It was also shown that theoretical knowledge of the respondents improved their understanding of the model, while general modeling intensity and general model experience did not improve this. The same research also showed there seems to be a tradeoff between researching semantic understanding and syntactic understanding (Mendling, Strembeck, & Recker, 2012).

More research has been done about the labeling of process model elements. It has been shown that the best way to label a process model activity is in the style of verb-object (Mendling, Reijers, & Recker, 2010). The verb-object style refers to the practice of labeling activities in an unambiguous way using a very and an object. Once again referring to the barbershop example, the first activity is called "Cut hair". This clearly illustrates what is done in that activity. However, the behavior of this activity would be unclear if the label would say "cut", or "hair", or even "haircut". Even though most process modelers prefer the verb-object style, situations do not always allow for this style.

Experience with process modeling is another important factor for comprehension of process models. Several studies have been performed regarding the experience of process modelers. It has been shown that people, who have had training with process modeling in a specific notation, perform quite well at comprehension tasks in an unknown other notation (Recker & Dreiling, 2011). They also found that domain knowledge helps novice process modelers in understanding process models. This means that someone untrained in process modeling has a higher comprehension of the barbershop example if that person is an actual barber, rather than someone without this domain knowledge. People with BPM experience are even better at understanding these process models, which can again be explained by cognitive loading. People with experience in the domain of BPM require less of their cognitive load to understand business process models, which makes them better and quicker at it. This finding confirms general knowledge about experience and performance (Ericsson & Lehmann, 1996).

A common way of assessing understandability of process models is having a participant answer questions about the model. These questions can be mainly divided into three categories: questions about control flow, questions about data flow and questions about organizational aspects (van der Aalst, 2000). These three categories refer to the uses of process models. If a participant understands the control flow, the data flow and which roles are responsible for certain activities, it could be concluded that the person understands the process model.

## 4.5 Quality of process models

Regardless of the notation of a specific process model, there are several ways to indicate the quality of a process model. To define what makes a process model 'good', a framework is discussed. This framework is the so-called SIQ Model, first introduced in the Handbook on Business Process Management (Reijers, Mendling, & Recker, 2010). The SIQ model is depicted in Figure 8. At the basis of the SIQ model, three intrinsic parts of quality are shown. There is semantic quality, pragmatic quality and syntactic quality. Each of these types of quality will attribute to a process model which is 'good'. Syntactic quality defines whether the model conforms to the rules of the notation. This means that the model is not allowed to show behavior which is not supported by the modeling notation. The importance of syntactical correctness is paramount according to Reijers et. al (2010). They explain it like this: *"Although you may be able to understand the meaning of a word that is not spelled correctly, you may be in doubt sometimes whether it is the actual word the writer intended."* (Reijers, Mendling, & Recker, Business Process Quality Management, 2010). The BPMN model in Figure 6 is syntactically correct. However, BPMN models do not allow you to start with an end event (i.e. if the element called 'Customer enters barbershop' would be red rather than green). If the model contained that error, it would be called *syntactically incorrect*.

Semantic quality consists of two important parts: does the model make true statements about the world (validity) and are all correct statements about the world in the model (completeness) (Lindland, Sindre, & Solvberg, 1994)? In relation to the barbershop example again; if the total amount of services a barbershop can perform consists of cutting hair and washing hair, the model is valid and complete. However if the real barbershop would perform dyeing hair as well, the models would not be complete anymore (they'd still perform correct behavior, just not *all* correct behavior). If, for instance, the barbershop only cuts hair and provides no other services, the model would be called invalid. It would allow for a service (washing hair) which is not present in the real world and thus make a false statement.

Reijers et. al. also talked about pragmatic quality (2010). Pragmatic quality is, according to the authors, a very poorly understood and researched topic. It relates to whether the model can be understood by people. Even though the barbershop models are probably semantically incorrect due to incompleteness (most barbershops offer more services than what the models suggest), they can be understood rather easily. This implies the model is pragmatically correct, while it might not be semantically correct. These three types of correctness are the core of a 'good' process model. of the model, please refer to the original paper (Reijers, Mendling, & Recker, 2010).

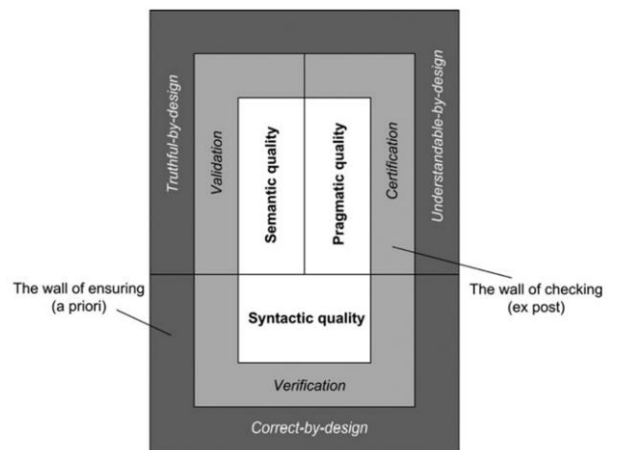


Figure 8: The SIQ model

## 4.6 Sub-processes

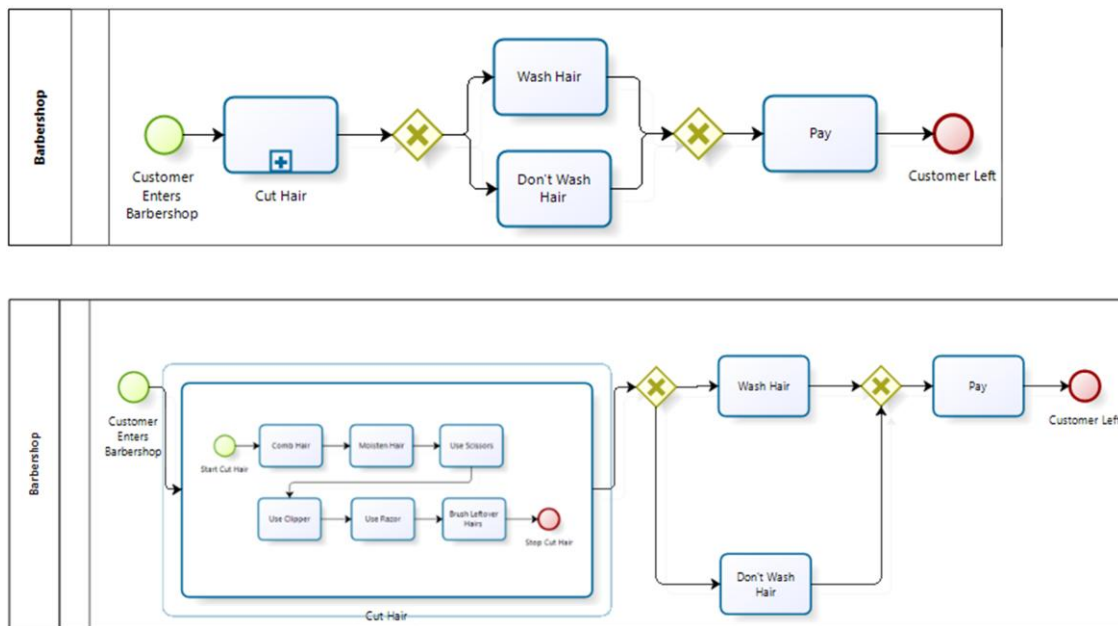


Figure 9: expanding 'Cut Hair' in the barbershop process

One thing that is very relevant to current research is a process modeling language that allows expressing sub-processes. A sub-process is a modularized part of the process, where the internal details are hidden from the user. The sub-process can be expanded for more detail, but the detail is usually hidden from the user. The use of sub-processes (called 'modularity') is an often mentioned predictor of process model comprehension (Reijers & Mendling, 2008) (Mendling, Rijers, & Cardoso, 2007). An example of a sub-process is provided in figure 9. The barbershop example is extended; the task 'Cut Hair' is a sub-process this time (top image). By clicking on the "+", the sub-process is expanded and the internal behavior of 'Cut Hair' is displayed. It is shown that there are a whole lot of activities which are used in cutting hair. The average user of the system might not be interested in this internal behavior, which is why it is placed in a sub-process. Furthermore, displaying all these extra details to cutting hair may make the model harder to read due to the extra information to inspect.

## 4.7 Critical Success Factors (CSF's)

Once a company initiates BPM activities, many different forces act. Incorporating BPM activities in an organization - from now on referred to as the BPM Initiative - is a very costly, time-consuming and risky effort (Melenovsky, 2005). The BPM initiative is usually subject to a large number of threats since stakes are often rather high (zur Muehlen & Ho, 2006). A BPM initiative is more than just a few days' worth of work and a few euros spent; 'failure' is very expensive. To increase the potential success of a BPM Initiative, several factors need to be accounted for. These factors are called Critical Success Factors (CSF). There are several common CSF's to be found in any BPM project. Among the most important CSF's are End-user participation and Communication (Parkes, 2002). For each of these CSF's some more elaboration is provided. With respect to the current research, a more in depth look will be taken at both End-user participation and End-user ownership.

### 4.7.1 End-user participation

The CSF End-user participation entails the inclusion of end-users in the BPM initiative. The importance of this CSF is also well-documented in literature (Parkes, 2002) (Mutschler, Reichert, &

Bumiller, 2007). End-user participation is so important since these people are, in most cases, at the very core of the processes. They are the people who perform some, most or all activities in a process. This means that these are the people who know a lot about how the process currently performs, what is wrong with it and maybe how it should be fixed. As these people have vast knowledge, including them will increase the potential of correct modeling. Furthermore, once changes will be incorporated based on this BPM Initiative, the end-users are the people who are influenced most. If these people have not had the chance to express their opinions, chances are their preferences are not met. This might lead to the undesirable result of rejection of the newly implemented changes.

There are several cognitive and emotional constructs at work in the end-user, when it comes to change. There are plenty of reasons people might resist change and are unwilling to cooperate. An often-mentioned reason for resistance of end-users is fear. People have a tendency to fear what is new, what is different and what is unknown. People might resist change of their jobs due to the fear of becoming obsolete. In a survey among practitioners, it was found that the vast majority (74.28 %) indicated end-user fears to be very critical or critical in the BPM Initiative (Mutschler, Reichert, & Rinderle, 2007). These fears can be mitigated by making the unknown known. By having end-users participate in the BPM Initiative, their fears might be reduced and the success of the BPM Initiative becomes more likely.

One potential way of ensuring end-user participation is the use of a modeling tool which allows for joint modeling. There are quite a few of these tools available like Lombardi and SAP. However, there are two very real problems which are not solved by this method (Rito Silva & Rosemann, 2012). Rito Silva and Rosemann identified that these tools usually do not allow participants to update the model. This means that they are still not empowered to make decisions regarding the model; they can only provide input to the actual business process analyst. The second problem is that these tools still function on the expressive but abstract modeling notations. This means that there is a large bridge between the practical instances of business process models the participants have in their heads, and what can be drawn in a process modeling tool. This might cause end-user to resist the actual tool and participation.

#### **4.7.2 Employee Empowerment**

Employee empowerment is an important factor in the BPM Initiative. In employee empowerment, employees are granted work-related decision-making power to enhance performance (Menon, 2001). The concept has been frequently cited as a powerful management style for increasing motivation and decreasing the feeling of powerlessness (Conger & Kanungo, 1988) (Thomas & Velthouse, 1990). The nature of the BPM Initiative might cause people to believe it is intrusive or a nuisance. During a BPM Initiative, a critical view is used to assess how people do their jobs. People might fear they have to radically change their jobs or even become obsolete. This can cause the feeling of powerlessness. Therefore, employee empowerment is an important CSF for the BPM Initiative as it can help mitigate these fears. This increase in employee empowerment can be achieved by incorporating a process modeling method where employees are asked to do the modeling by themselves (Türetken & Demirors, 2013).

#### **4.7.3 Communication**

As mentioned earlier, communication is a key CSF to the BPM Initiative. By communicating honest, open and frequently, you can inform people about the current state of the project. In a survey



among Dutch practitioners, it was found that the most important success factor for BPM system implementation is Communication with stakeholders (Ravesteyn & Batenburg, 2010). This important CSF is not limited to the BPM Initiative; a large amount of similar fields of information systems implementation seem to be concerned about communication (Galivan & Keil, 2003).

A very potent form of communication with stakeholders is performing training sessions. During such training, people can be educated about BPM and its use (Kappelman & Gunes, 1995). This form of communication provides people with additional information about what is going on. Providing training might also help empower employees, which is very much related to the previous CSF. What makes communication even more so a CSF in BPM, is the fact that often the goal of process modeling lies in facilitating communication (Reijers, Mendling, & Recker, 2010).

#### 4.7.4 Process ownership

The final important CSF is the level of process ownership (Hammer, 2003). This CSF resembles end-user participation, but is different on several levels. The level of process ownership depicts the way the end-user feels he or she has had influence over the way their process is being modeled and updated. If people feel a degree of process ownership, they are also more likely to adopt technology regarding this process (Karahanna, Straub, & Chervany, 1999). Another aspect of process ownership concerns the amount of responsibility people have and display regarding their process (Coombs, Doherty, & Loan-Clarke, 2001). It seems to be that people need to have a certain amount of control over their life, which can be increased with proper End-user empowerment (Kappelman & Gunes, 1995). According to Kappelman & Gunes, providing proper training can be a powerful way of empowering end-users. A large problem in the ownership problem is a difference in knowledge. End users are often very aware of their processes, but not of the technologies that are being used to model them. The result is that people become estranged from their own processes due to unknown technologies. Using training can mitigate this problem.

### 4.8 Adoption of Information Systems

An important aspect of any information systems is the adoption by its users, as an information system is intended to be used to increase productivity. It has been argued before that the reason for disappointing increases in productivity can be attributed to lack of information system adoption (Sichel, 1997). Furthermore, it has been found that a return on investment on information technology is heavily influenced by usage (Devaraj & Kohli, 2003). Therefore, it is important to investigate whether people are inclined to actually use a system. A lot of research has been dedicated to models that explain the usage and adoption of information technology. The common denominator of these models is that they are often based on the principles of Theory of Reasoned Action (Fishbein & Ajzen, 1975). This states that human behavior is influenced by the person's attitude towards behavior, along with the norm regarding that behavior. These two types of factors influence the intention to perform behavior, which in turn influence actual behavior.

### 4.9 TAM

The most familiar model of technology acceptance is the so-called Technology Acceptance Model (TAM). This model was first proposed in 1989 and is based on the Theory of Reasoned Action (Davis, 1989). TAM states a set of external variables determine the perceived ease (E) of use and the perceived usefulness (U), which in turn both influence the attitude toward

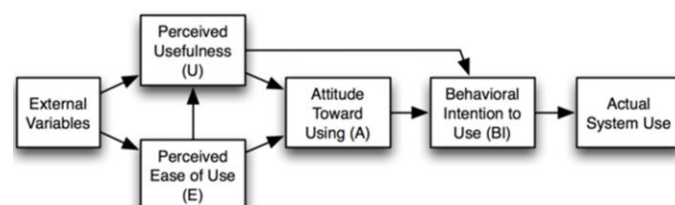


Figure 10: Technology Acceptance Model

using (A). This influences the behavioral intention to use (BI) which in the end predicts actual system use. This model is depicted in figure 10 for clarity. TAM has been validated extensively in studies and has typically been able to explain about 40% of the variance in actual system use (Davis, 1993) (Davis & Venkatesh, 1995). The items that were used to assess the different constructs of TAM have been validated frequently (King & He, 2006) (Lee, Kenneth, & Larsen, 2003) (Legris, Ingham, & Collette, 2003). In current research, TAM can be used as a scale to indicate whether people are willing to use the presented process models in their practice.

## **5. Trial Experiment**

To validate the tools used in this project, a trial study was used. The trial study consisted of the application of the Plural Method on a small scale and in a selected, controlled environment. The main goal of this trial was to ensure that the current ‘flavor’ of the plural method was suitable for application in Philips Healthcare. If this was the case, further analysis could resume.

### **5.1 Scope of processes**

The first important step for the execution of this project was to identify a set of processes for the scope. Together with several people from the Quality and Regulatory department, a set of four processes was defined: The Corrective & Preventive Action process (CAPA), the Field Change Order process (FCO), the Complaint Handling process and the Risk Management process. These four were selected due to their strong connection with the Q&R department. Furthermore, these four processes were common topics of investigation during audits and FDA inspections. Of the four processes, the CAPA process was selected to be investigated first as a trial experiment. Since most of its participants were actual members of the Q&R department, the CAPA process was well-suited for testing the current flavor of the Plural method. This was crucial as the Plural method had never before been applied using BPMN. The following section describes the trial experiment in detail.

### **5.2 Models of Trial Experiment**

A CAPA is a very costly and time-consuming process which is initiated to solve a problem of large impact for the business or the customer. This process is costly as it takes quite a lot of effort and resources to perform a single instance. Furthermore, the CAPA process is often regarded as one of the most important points of investigation for an internal audit or FDA inspection. Having a good CAPA process allows a business to adequately respond to crucial problems in a product or process. The CAPA process is not only used to ‘patch the symptoms’, but also to address the underlying problem of those symptoms, called the ‘root cause’.

The CAPA process can be initiated by several different triggers; for instance when a certain KPI has been above its allowed limit for some time or when a certain product has consistently been functioning subpar. This implies something is wrong and a root cause analysis should be performed. A requestor gathers all initial information and subsequently does a CAPA request. A CAPA review board meets weekly to decide on CAPA requests and monitor progress of current CAPA’s. Once a CAPA request is approved, an SME is appointed as a dedicated CAPA owner. This CAPA owner is now responsible for performing the root cause analysis and come up with actions to mitigate or remove the root cause. At certain controlled points, the CAPA review board provides a decision to either go to the next phase of the CAPA or to ask for rework. After sufficient evidence has been provided that a

certain solution solved the root cause of the problem, the CAPA can be closed. As many people confirmed, a CAPA owner rarely welcomes being appointed due to the extra workload.

For modeling the CAPA process, the procedure from the Quality Management System (QMS) was analyzed. To obtain more information regarding the process, the SME was invited for an information session of an hour. Afterwards, the procedure was transferred into a BPMN 2.0 model. The resulting model was presented to the SME and after incorporating the feedback it was validated by this SME and subsequently 'frozen'. The resulting model can be found in Appendix A1. For each of the identified roles an agent was selected and operations were identified. In the end, seven roles were identified (divided over four agents) with a total of 22 operations. This took a total of 208 minutes, spread over 4 sessions. Initial reactions regarding the Plural Method were positive.

Most feedback comments included positive claims towards the simplicity of the tool (BiZagi suite), the intuitive nature of the notation (BPMN 2.0) and the general usefulness of the method. During one session, a previously unknown feedback loop was found in the behavior of the process. One participant indicated this as *"it is remarkable that you found such a flaw in just under 30 minutes. The [models created in a previous BPM initiative by external experts] were expensively made with a lot of man hours, where you find new issues in a few hours. Remarkable."*

Only one agent indicated reservations regarding the method by stating *"I am interested in the project, but I don't feel my problems will be solved by this method."* This does lead to the tentative conclusion that the Plural Method should be viewed for what it is: a useful framework for process discovery, rather than a catch-all problem solver. From the trial experiment, it was learned that the initial list of identified operations may be insufficient. During most sessions either an additional operation was identified, or another one is found to be superfluous. This is not an issue as long the interface of input/output between roles remains constant. This successful application of the Plural Method in the CAPA process gave the green light to continue the project with the three other selected processes. The following section discusses the results of these sessions. The Plural model of the CAPA process can be found in Appendix A2.

## **6. Conducting Plural Modeling Sessions**

After the trial experiment showed the current setup of the Plural Method would suffice, the three other identified processes were modeled for this research. The following sections describe each process and provide an account of the modeling efforts.

### **6.1 The Complaint Handling process**

The complaint handling process entails identifying, classifying, investigating and if necessary: resolving complaints. Like any other business PH MR receives a great amount of complaints each day and the Complaint Handling Unit (CHU) considers each of them. Based on impact and frequency of occurrence, it is determined whether these complaints will be resolved. The Complaint Handling process is largely performed by the Complaint Handling Unit (CHU). Customers do not directly complain to the CHU, they complain to local PH MR representatives, who report the complaint to the CHU. There, a Complaint Handling Administrator performs some initial checks. The complaint is then passed on to the Complaint Handling Specialist who decides for the need of a main investigator from development. If no such expert is needed, the Specialist investigates the complaint. The classical

model was created as a direct translation of the procedure from the QMS. This model was validated by an SME and can be found in Appendix B1.

For modeling the Complaint Handling process in plural, four roles were identified with a total of 14 operations. These were modeled by three agents who needed a collective 131 minutes to do so. Modeling sessions with a Complaint Handling Administrator (2 operations) and a Main Investigator (3 operations) were relatively simple. However, the Complaint Handling Specialist identified several additional operations and merged some others. All these issues were found and solved rather quickly. Due to the high level of expertise of each of these agents, the sessions went quick. During the modeling sessions, several instances of unclear behavior were identified. There turned out to be a conflict between two agents about the place of a certain task in the workflow. This conflict may have caused duplicate effort for quite some time. The occurrence of this conflict is noted in the report for the responsible people of this process. The Plural model of the complaint handling process can be found in Appendix B2.

## 6.2 The FCO process

The Field Change Order (FCO) process entails an update of systems in the field. The FCO may be the result of a CAPA or complaint. Whenever a system does not perform the way it is supposed to, a so-called Problem Report (PR) is made. When a sufficient amount of PR's are gathered (or when a single PR has sufficient impact), an FCO is initiated. In an FCO, someone from development goes through the required steps to engineer a solution to the problem. This solution can be software-based or hard-ware based. When there is sufficient testing, validation and verification evidence of a correct solution, the Q&R department ensures the required documentation is present and a dedicated FCO manager prepares the FCO for launch. If the FCO concerns some safety-related issue, it is determined whether this FCO should be reported to authorities like the FDA. For each FCO, it is determined what its severity is. If an FCO is mandatory, the entire installed base should update the system. In all other cases it is recommended rather than mandatory. When an FCO is announced, the update can be incorporated. In the case of a software-FCO a service-pack can be downloaded. In the case of a hardware-FCO, an FCO-kit might be necessary. An FCO-kit can consist of some extra parts or tools which solve the original problem. The classical model was created as a direct translation of the procedure from QMS. The FCO manager validated the correctness of this model and Plural sessions could start. The FCO manager did mention that the FCO process would be changed in a few months. However, he stated that the model he just validated actually resembles the way work is being done right now. It is displayed in Appendix C1.

For the plural sessions, four roles with a total of 13 operations were identified. Three out of four roles contained 8 operations and were modeled in 102 minutes by two agents. These sessions showed for the first time that the Plural Method is not applicable to all types of processes. From the get go these sessions were slightly awkward. The first session was with a very experienced Philips employee. Where agents from previous processes immediately could relate to the architecture of the identified roles and operations, this employee had a really hard time doing so. After an interesting session, we finally came up with a model the agent was satisfied with. Still, this was rather different from what was originally validated by the FCO manager. This same confusion occurred the second time, as the next agent did not approve of the structure which the FCO manager proposed. Again, this was a very experienced agent who was able to come up with a satisfactory model in the end. After this session, it was identified that the FCO process was not as suitable for the Plural method as

initially intended. This may be attributed to the fact it is in a redesign effort right now, but other causes might be present as well. It may also have been the case that the agents had different opinions about important details in the model due to a difference in backgrounds and responsibilities. Due to the fact that the FCO process was not deemed suitable to the Plural method anymore, the final session of the Plural method was canceled. This means this model is incomplete and could not be used in the experiment. This incomplete Plural model is displayed in Appendix C2.

### **6.3 The Risk Management process**

The Risk management process is one of the most important processes from the Quality Management System (QMS) of Philips Healthcare, as it is connected to all other business functions of the unit. Where Complaint Handling, CAPA and FCO are mainly quality related processes which are not very salient to people from other business functions, Risk Management is a process that lurks in the shadows of each Philips employee. In essence, Risk Management is the continuing cycle of identifying risks, classifying these risk and, if need be, mitigating them. The Risk management process is split up into two parts: pre-market risk management and post-market risk management. The cutoff point when a product leaves the first to go to the latter is the point where the product is in the so-called 'volume-phase': the moment when the product will be sold to all interested customers.

This process contains many important documents, but there are several which are paramount. Each product has a so-called Risk Management File (RMF), a file which contains all records produced by the Risk Management process. The Risk Management Matrix (RMM) is a matrix which contains all known potential risks of a certain product. For each of these risks, an occurrence estimation is made and the severity is estimated. These calculations are used to assert the overall risk a product carries along with it. During a context definition meeting with the process owner, it was identified that pre-market risk management would not be very suitable to the Plural method. However, the post-market risk management process would be. The classical version of Post-Market Risk Assessment was validated by the process owner and can be found in Appendix D1. In post-market risk management, a member of the complaint handling unit receives a complaint and identifies if the complaint contains a hazard. If this is the case, it is checked whether this hazard was foreseen in the RMM. Unforeseen hazards require a Post-Market Risk Assessment (PMRA), where a risk management specialist investigates the occurred hazard. Afterwards, the investigation results are sent back to the complaint handling unit where it is decided what to do with this information.

Post-market risk management was a small process with only two roles and three operations. Together with the process owner two agents were identified and subsequently invited. In two separate sessions, these two agents required a total 29 minutes to model the entire process. Out of all Plural sessions, these went 'smoothest'. This can be attributed to the high level of stability this process offered. This process was very sequential, with clear steps and control flow. This allowed the participants to quickly catch on what was expected of them and model their work. Both participants were experienced in their work and longtime Philips veterans (19+ years of experience). The resulting process model had no conflicts and found no significant structural differences with the classical model. It is displayed in Appendix D2.

### **6.4 Reflection on the Plural Method**

Modeling efforts have resulted in two models per process; one model for the classical method and one for the Plural Method. All process models can be found in Appendix A, B, C & D. The sessions

went well, as all participants were happy to talk about their work. Most participants called the tool useful and saw the power of BPMN 2.0 over flowcharts. Currently, the procedures in Philips Healthcare are accompanied by a flowchart. These flowcharts are rather rudimentary due to the limited functionality and expressive power. Once participants could work with the new notation of BPMN 2.0, they were enthusiastic. It is important to note, however, that none of the participants actually used the notation. In all sessions the coordinator handled the modeling through the tool, while the participant stated his/her work.

A trending phenomenon during modeling these processes was the discovery/drop of certain operations. It was often the case that an agent identified one or more operations should be added while others should be removed. This proved to be no problem. Another remarkable phenomenon was the frequent occurrence of 'venting' regarding problems with the process or the way work is done. In each session, the agent told valuable information regarding inefficiencies or lack of quality in the workflow. Each of these issues has been carefully documented, as large value lies in these remarks. Even though it is possible middle management is aware of these issues, after the project is finished a report will be written with mention of each of these issues. More on this can be found in the later sections. Only one session was too short for modeling the planned operations. This can be attributed to some confusion about the process. This was the FCO process. The agent was not content with the division of operations among the different roles, which caused a very interesting discussion. However, all troubles were mitigated in the end and all the agent's operations were modeled.

For feedback purposes, an online questionnaire was sent to all plural agents. This questionnaire contained open-ended questions about the Plural sessions they had. The feedback was positive towards BPMN, the used software and the method itself. One of the Plural agents had this to say about the notation: *"It gave good and detailed insight in the actual process and work flow. Graphs were almost self-explaining."* The same agent said this about the Plural method itself: *"I think this kind of business process modeling is really helpful to get the process well described, but also as an interactive tool during doing the work or during training of new employees."* This notion was rather interesting, as the possibility of using the Plural method for employee training purposes has not been coined before. Another Plural agent provided the following feedback about the modeling session: *"It showed me how complex we are working ;-)"*. Though the smiley at the end of the quote indicates a light-hearted nature, the comment does point towards a powerful notion of the Plural method: it helps employees obtain oversight of the process they are a part of.

## 6.5 Adapting the Classical models

As can be seen from appendices A and B, there was a very distinct difference in content between the classical models and the Plural models. This made it difficult to compare them, as this meant any discovered difference might be attributed to the difference in content, rather than structure. Therefore, the original classical models which already have been validated by Process owners were somewhat adapted. This was mostly done by changing labels to obtain comparable items (i.e. if both models had a task which was about checking some report, but one was named 'check report', while the other was named 'review report', the classical model was changed). The Plural model was never changed, as then it would not be completely made by participants anymore. Furthermore, the original classical models contained pools with only one swim lane. To obtain models which most reflected actual classical business process models, some pools were merged into a single pool with

multiple swim lanes, wherever that was possible. The adapted models, as they were displayed in the final questionnaire, are displayed in Appendix A3 and B3 for CAPA and Complaint Handling respectively.

### 7. Experiment Construction

To assert the hypotheses of this study, an experiment was set up. This experiment had to be able to assert whether the Plural models were more understandable, easier to maintain and more liked than their classical counterparts. A way to assert these things was by having people answer the same questions about both types of models and analyze which of the types yielded the best answers. For this experiment, it was identified that not all four processes should be used. The results of different processes cannot be compared, as it may be that case that one process favors a certain method while another process does not. Therefore, group sizes for statistical tests would be too small if all four processes were used in the experiment. Out of the four processes, the CAPA and the Complaint Handling process were selected for experimen.

This study employed a repeated measures design, where participants were asked to answer questions about a classical model and about a Plural model (Field, 2011). It was tested which of these models could best be used to answer a set of questions, based on correct answers, time to obtain those answers and efficiency (number of correct answers divided by the time to do so). Please note that this was a within subjects design, rather than a between subjects design. The research was set up to test the models which were the result of the Plural Method rather than individual differences between the participants.

Participants were assigned to one of four conditions: there were two processes (A&B) and two modeling styles (classical and Plural). This resulted in four possible process models, of which each participant was presented with two. Participants were always presented with one model of process A and one of process B, and were always presented with one classical model and one Plural model. They were assigned to conditions based on participant number; every first, second, third and fourth participant was assigned to a separate condition. A graphical representation of the experiment design is depicted in figure 11 for additional clarification.

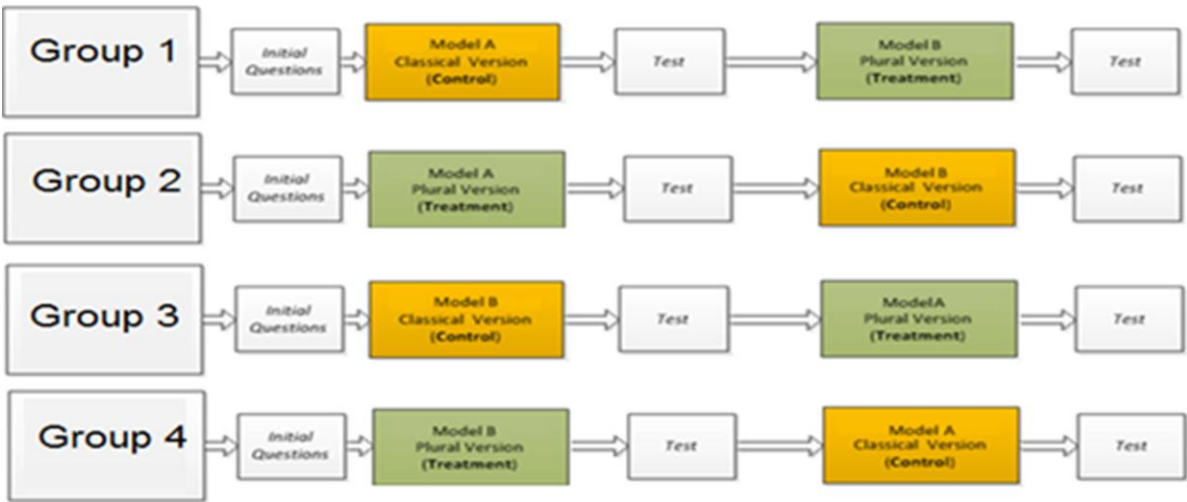


Figure 11: Experiment Design

Such an experiment design requires a carefully constructed set of questions, a proper channel to ask these questions and a set of people to ask these questions to. For each of these three parts, a subsection is written below to address it.

**7.1 Question Construction**

Per process (CAPA and Complaint Handling) two questions were constructed about updating the model. In these questions, a certain update was textually described and a multiple choice question was asked about what other parts of the model would be influenced after such an update. The setting of these questions was made as meaningful as possible. One of the questions stated that the FDA required some extra security check; the other question stated that the manager of the process wanted some improvement. The potential answers were always “yes”, “no” or “don’t know”. The answer “don’t know” was added to prevent participants from guessing. The questions were made to assert hypothesis 1. Here follows an example of the maintenance questions: *“Suppose that the manager of this process wants to improve the throughput time. This should be achieved by skipping the steps “Pre-Check Investigation” and “Move to Implement”. Would this change the deliverables which are required from the CAPA Owner?”* The rest of the questions is displayed in Appendix E1.

For each process, six questions were developed to assert understanding of the model. These questions consisted of three categories with two questions each: control flow, data flow and organizational aspects (van der Aalst, 2000). These questions were used to assert hypothesis 2 and were designed similar to previous research (Mending, Strembeck, & Recker, 2012). To ensure that these questions were meaningful to Philips employees, all questions were designed in such a way that they might be asked by new employees. Again the same answer categories were possible. This is an example of one of the control flow questions for the CAPA process: *“Is it possible to already start executing the action plan before the CAPA Review Board has approved it?”* As one can see, this question asserts understanding of the order of steps; the control flow. Furthermore, an eager CAPA owner could ask such a question to one of the experts when he tries to understand the process. The rest of these questions is displayed in Appendix E1.

To assert hypothesis 3, a set of items from the Technology Acceptance Model (TAM) was used to assess the willingness to use a specific model (Venkatesh & Davis, 2000). These are statements the participant could express their level of agreement with on a 7-point Likert scale, ranging from 1 (completely disagree) to 7 (completely agree). The items were all modified to accommodate this research. Though the initial questions are phrased as such: *“Use of the information system...”*, the items in current research are stated as *“Use of this type of process models...”*. Though TAM is a widely validated measurement tool, some additional questions were asked to assess whether people could use the models for extracting information. Those questions were adopted from previous research (Seddon & Yip, 1992). This entire set of 11 items is depicted in Appendix E2.

The final questions in this research consisted of control questions about age, gender and other demographics. These can be found in Appendix E3. An overview of the asked questions is in table 1.

<b>Questions for online Experiment</b>	<b>Introduction</b>	<i>2 questions about modeling experience</i>				
	<b>19 questions for each model</b>	<i>Hypothesis 1</i>	<i>2x maintenance</i>			
		<i>Hypothesis 2</i>	<i>2x control flow</i>	<i>2x data flow</i>	<i>2x organization</i>	
		<i>Hypothesis 3</i>	<i>2x intention to use</i>	<i>3x perceived usefulness</i>	<i>3x perceived ease of use</i>	<i>3x information retrieval</i>



Table 1: Overview of questions in experiment

## 7.2 Creating the online experiment

The entire questionnaire was created using SSI Web 6.2 by Sawtooth. This Perl-based questionnaire software allows users to create questionnaires for online studies. The tool provides support for HTML (which allows for displaying images) and measuring time spent on a question. Furthermore, SSI Web provides the option of outputting the data to formats suitable for statistical packages like STATA or SPSS. This combination of factors allowed the online survey to be successful.

The participant was first instructed about what was expected of him/her. Then, the participant was asked questions about experience with process models. These questions were taken from a previous study, to ensure a higher validity compared to coming up with new questions (Mendling, Strembeck, & Recker, 2012). To ensure the participant was aware of the notation that was used, a short BPMN 2.0 tutorial was provided. This tutorial can be found in Appendix E4. Then, the participant was explained how the process models would be depicted. The image of a process model was shown, where the participant could have a look at all the sub-processes. It was important to ensure the process models would be displayed in a user friendly way. After some experimenting with proper display techniques, the use of HTML tooltips was selected. This means that if the cursor 'hovered' over the active area (i.e. the sub-process), an image was immediately revealed which displays the contents of the sub-

process. Figure 12 can clarify how this happened in practice. After these initial explanations, the experiment started.

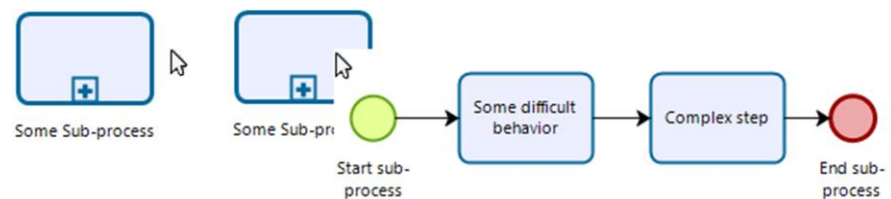


Figure 12: Sub-Process Expansion

Like stated before, every first, second, third and fourth participant would be assigned to a different condition. Skip-logic was used to ensure participants saw only the appropriate questions. This skip-logic is a set of logical rules to determine what question a certain participant is allowed to see and in what order. This made sure the participants would correctly be assigned to their respective conditions like stated in figure 11.

The setup of the experiment was as such. Participants were shown a process model and received a short explanation about it. This explanation outlined, in very broad strokes, what happened in the process. They were told to scan through the process and specifically to look inside all the sub-processes. Participants were also instructed not to spend too much time on analyzing this model, as they had access to the model during the entire set of questions. This way, they did not have to memorize the entire model. In total, the participant was shown two of these process models and was asked 19 questions per process model: six about understanding, two about maintenance and 11 about preference (like stated in the previous section). After answering these 19 questions for a single model, each participant would be presented with the exact same exercise for another model. Afterwards, the control questions were asked.

To ensure the quality of the questionnaire, it was validated in two ways. First, a sanity-check on the questions with a Subject-Matter Expert (SME) was done, using an interview. He was asked whether the questions were well-constructed and generally 'made sense' to practical users of a process. To

prevent a bias, this SME was excluded from participation in the final experiment. The second method of validation included a short trial. The same SME was asked to sit down with the apparatus and try to solve the questions. This session also formed as an indication of the time it would take to perform a single experiment. To prevent loss of interest due to fatigue or boredom, a survey should not take too long. This SME could finish the questionnaire within the desired time, but found a few textual mistakes. He pointed them out and they were subsequently resolved. This SME confirmed that the questions were practical enough in nature and that Philips employees could ask the very same type of questions when they are trying to understand a process. He specifically liked the questions about maintenance, as they were very practical. He stated that these types of updates frequently occur in the business unit.

The survey was verified by testing it on several computers, different web-browsers and with simultaneous experiments. It was specifically checked whether the models would be displayed correctly, if the assignment of conditions worked properly and if the resulting data files could be read by the statistical software. After validation and verification efforts, the survey was uploaded to the domain [www.henkvandenhurk.com](http://www.henkvandenhurk.com). This website is owned and managed by the researcher. The final questionnaire could be reached through <http://henkvandenhurk.com/2014logn.htm>.

### 7.3 Sampling

To obtain a correct sample size, as many participants as possible were invited. First of all, each person involved in the actual modeling of the processes was invited. In fact, whether or not people were involved in the Plural Method itself served as a control variable. Only including these people would seriously limit the sample size, though. The sample size would hardly exceed 15 people, which would be limited at best. Instead, a contact within Philips indicated the willingness to provide a list of email addresses within the business unit. This list included about 750 Philips MR employees. From those, a selection was made based upon availability and suitability to the questionnaire. From the remaining list a pseudo-random sample of sufficient size was drawn to invite. To further increase the sample size, a number of students was asked to fill out the questionnaire as well.

In this experiment, participants received a link to the questionnaire per e-mail. People could log in using a provided username and password. This shielded the questionnaire from potential uninvited users. To ensure plenty of people could enter the survey, a hundred accounts were created; each username would be of the form ID[XX] and each password was a random generated 4-digit integer.

### 7.4 Internal validity & Survey errors

There were several ways the internal validity of the questionnaire was maintained. The following list of precautions was taken to ensure a higher internal validity.

- 1.) Participants were specifically asked to answer the questions based on the model, rather their own expertise.
- 2.) Participants were assigned to a group in a first come, first serve fashion. This means that every first person to start the survey was assigned to the first condition, every second person to the second condition, the third to the third condition, and the fourth to the fourth condition and then the fifth person to the first condition again. This ensured an equal distribution among conditions, while maintaining a pseudo-random assignment
- 3.) The four paths were created to check and control for ordering and learning effects

- 4.) Since the survey was performed online, little control over the environment of the participant could be exerted. However, participants were specifically instructed to perform the entire experiment in one go. The suggestion was made to go get a cup of coffee before performing the questionnaire, along with making sufficient free time in their schedule

Furthermore, as this experiment entailed a survey, there may have been any of the following four forms of survey error present in the setup: sampling error, measurement error, coverage error and non-response error (Dillman, 2000). The sampling error was mediated by drawing a pseudo-random sample from the list of suitable people. This increased the chance that the measurement of the sample said something meaningful about the population. The measurement error was mediated to some extent by expert validation. This validation was used to remove wording issues and general incorrect questions from the questionnaire. The coverage error was limited by including people from different groups: Plural participants, random Philips employees and students.

Finally, the non-response error was mediated by inviting the random Philips employees conform the tailored design method (Dillman, 2000). They were first approached to inform them of the pending experiment. It was asked if these people would approve the researcher to send an invitation of the questionnaire. The pre-approach was done face to face. The idea was that if people got to see the researcher, shake hands, have a chat and ask questions, they were more inclined to participate rather than when another random e-mail popped up in their inbox. People could decline participation if they wanted to, but no one did. Afterwards, all approached people received an invitation to the questionnaire. They got some extra explanation in the email, along with the URL of the study and their account details. To further diminish the non-response error, the people who did not fill out the questionnaire after a week were reminded through another email.

## **7.5 Rationale for made decisions**

There were several possible ways to assert the hypotheses for this research and there were several considered alternatives. This section outlines a rationale for the decisions made to obtain the current setup. For more information about the limitations which reside in the selected design and setup, please see section 10.

### **7.5.1 Experiment Design**

The use of the repeated measures design has both a set of advantages and disadvantages. The large benefit is that the treatment can be applied to all participants, while using the same people as a control group. This means that factors such as IQ, age and gender remain the same for control and treatment groups. This would not necessarily be the case when a regular control group would be used (Field, 2011). The repeated measures occur directly after each other, which might induce an improvement effect or ordering effect. A potential threat to the repeated measures design is the possibility that participants get bored after several measures. This was somewhat counteracted by using the crossover design where all participants were part of the control group and the treatment group (Pan, Shell, & Schleifer, 1994). The main weakness of a crossover study is the threat of carry over effects, where results of the first test influence the second. This was somewhat mitigated by having both orders and always use a different model for the treatment and the control. It was concluded that in this research design the benefits outweighed the limitations.

### 7.5.2 Experiment setting

The decision to conduct the experiment online brought some implications with it. First and foremost, the problem of the online questionnaire was lack of control. When one conducts an online experiment, the researcher cannot directly control the behavior of the respondent. For instance, a participant could be distracted by other people in his/her presence, or the participant could decide to go get a cup of coffee in the middle of the experiment (and thus skewing the time measurement). This limitation was addressed to some extent by informing the participants of this problem. In the first page, participants were explained that this study could only be performed in one session and without a break. Therefore, they should get a coffee right now if they wanted to. See appendix E4 for the way the questionnaire was shown to participants.

Another important consideration was the way the models were depicted. In a face to face setup, participants would receive a big paper printout of the process models. However, this was not possible in the online setup. Participants would have to read the models of their screen. Models were displayed on the screen, but were rather big. Therefore participants would have to scroll. As a control variable, participants were asked the size of their computer screen. This way it was possible to control for screen size in the analyses.

The advantages of the online questionnaire were abundant. First of all, the online questionnaire allowed for parallel experiments. This means that multiple participants could perform the same experiment simultaneously. Furthermore, participants would not need to make an appointment for a face-to-face session. They could start the questionnaire whenever it suited them. This significantly lowered the amount of commitment required from the participants; thus increasing the sample size. This big advantage helped in 'selling' the questionnaire to participants, as they were pleased with this option. The online questionnaire would also allow for very accurate measurement of time and randomization. Using the SSI Web software, each click of the participant could be registered to ensure a very precise measurement of time. The final big advantage to the online questionnaire was the fact that a large amount of participants could be reached. If the experiment would have been conducted in a face to face setting, only a limited amount of people could be approached. However, the online setup would take away that barrier. In fact, the online questionnaire helped reach other groups of people, like students. These considerable advantages caused the researcher to choose for an online setup.

Before the conclusion of performing the experiment online was made, two other methods of performing the experiment were considered. The first alternative was a face-to-face setup with a single participant per session. The same participants could be used, but this time they would be invited for a session in a separate room of a Philips or TU/e building. There, these participants would receive a few big paper printouts like mentioned before and could start answering the questions; either on paper or on a provided laptop. This setup would allow for more control in the way the models would be displayed, along with more control about the way the participants answered the questions. However, there would be serious logistic problems in setting up plenty of sessions with participants. Many potential participants would have a hard time scheduling a certain timeslot.

A way to deal with the logistic challenges while keeping the advantages of control was by organizing group sessions. A timeslot would be reserved for a maximum of 10 people who perform the experiment in a single room. If there were any unclear questions, they could ask the researcher for

more information. This setup would provide a bottleneck problem, however. As all the participants are colleagues, it would be hard to prevent them from influencing each other. They would, for instance, start to chat or try to answer the questions together. These considerations resulted in the setup as it was performed.

## 8. Results of Experiment

After having performed the experiments for this study, a large set of data was gathered. This section addresses all relevant information about this data, as well as the statistical analyses performed to assess the hypotheses. As this section provides a vital backbone in the underpinning of this thesis, it is structured as well as possible. This section starts with an overview of the data, continues with a section about the performed control checks and ends up with the assertion of the hypotheses. In order for other researchers to be able to replicate the results obtained in this study, Appendix G provides a link to a public Google drive folder where interested people can find the dataset, as well as the script used to analyze this data. The statistical package STATA was used for all analyses.

### 8.1 Descriptive statistics

After the questionnaire was made available online, 77 people were invited. These 77 people consisted of 15 randomly drawn Philips MR employees, 13 plural agents and 49 students. Initially, 23 people filled out the questionnaire. By sending out a reminder a week later, another 28 people responded to obtain a final sample size of N=51. The distributions of group, age and education are depicted in figure 13.

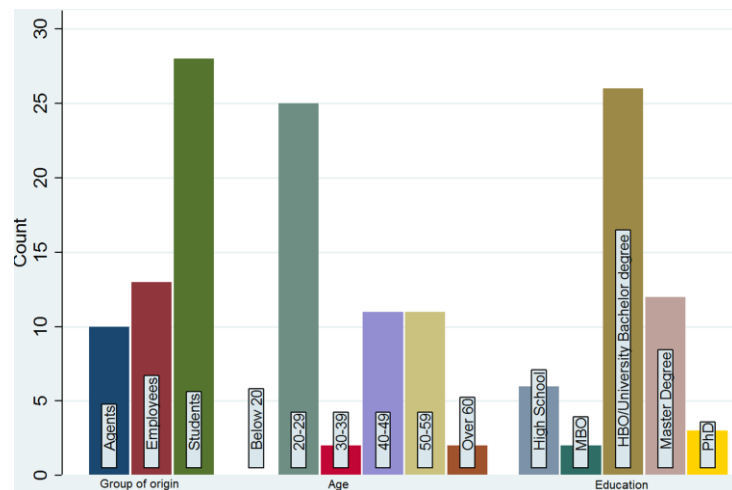


Figure 13: Demographics of respondents

To protect the analyses from being skewed by unmotivated participants, some data were removed. First, each observation was split into two observations as they answered questions for two models. Due to several checks, it could be asserted whether people really took an effort to fill out the questionnaire. If they did not, the results could not be used. First, it was checked how often people indicated they did not know the answer. People answered “don’t know” 1.12 times on average ( $S=1.45$ ). If someone answered four times “don’t know”, or more often, this case was deleted. This led to the deletion of five cases (note: not five participants, but five cases!). Furthermore, if someone answered more than one question in less than 10 seconds, this case would be deleted. This was never the case. If someone took more than 450 seconds to answer a single question (mean + twice the standard deviation), this entire case would be deleted. This could be attributed to the possibility they went for a coffee break, which would skew the results. Eight cases were deleted as such. This resulted in a final sample size of 89 cases, who needed an average 44.71 minutes ( $S=43.83$ ) to complete the total questionnaire. They needed an average 10.62 minutes ( $S=5.56$ ) to complete a single set of eight questions.

To get an idea of how the participants performed the different questions, three tables are placed in Appendix F. These tables represent the performance of the participants for different (sets of) questions, based on score, time and efficiency.

## 8.2 Control variables

### 8.2.1 Learning effects

Some checks should be performed before the analyses can be conducted (and trusted). The first is to try and find out if there's some kind of learning effect. This is asserted by comparing both the times and scores of the first and second time someone answers questions of a plural model and for a classical model. This is done through student t-tests. First it was checked whether there were differences when no control was made for the type of model (Plural versus Classical). There's no significant difference in scores of the first and second model ( $P > 0.05$ ,  $t(87) = -0.0197$ ). To test the time obtain these scores, a transformation of the data was required as the total time was not normally distributed. After this transformation, a significant difference in times to obtain the scores was found ( $P < 0.05$ ,  $t(87) = 2.5372$ ). When checking for the interaction between ordering (first versus second) and model type (Plural versus Classical), it turned out that questions about classical models were answered significantly better if they were for the second model presented rather than the first one ( $P < 0.05$ ,  $t(41) = 1.6284$ ). Moreover, the questions about the Plural models were answered significantly faster when they were for the second model presented rather than the first one ( $P < 0.05$ ,  $t(44) = 2.7899$ ). This leads to the conclusion there was some kind of learning effect where people seem to be more efficient and just as effective for the second model. When controlling that effect for model type, it turns out that in the case of Classical models there is a learning effect for effectiveness and in the case of Plural models there is a learning effect for efficiency. This means the ordering should be included in the ANOVA models as a control variable and interaction variable.

### 8.2.2 Differences between CAPA and Complaint Handling

The next check is to see whether people perceived one process as more difficult than another (CAPA versus Complaint Handling). Participants were asked whether they found the process model they just saw difficult on a five-point likert scale. Since this does not have parametric properties, Mann-Whitney is used to assert differences (Field, 2011).

It was found that Plural models were not perceived as significantly more difficult than classical models ( $P > 0.05$ ). It was found that Complaint Handling models were perceived as significantly more difficult than CAPA models ( $P < 0.05$ ).

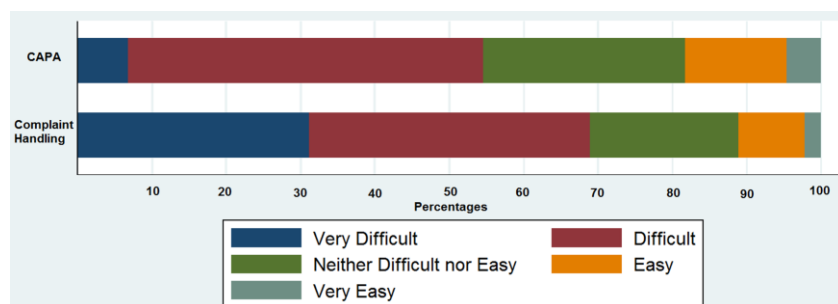


Figure 14: Perceived difficulties of CAPA and Complaint Handling models

Specifically for the classical models, CAPA and Complaint Handling were perceived as equally difficult ( $P > 0.05$ ). However, in the case of the Plural models, Complaint Handling was perceived as more difficult than CAPA ( $P < 0.05$ ). The perceived difficulties are depicted in figure 14.

To further corroborate this control variable, the same was checked for the total scores. As these are normally distributed ratio variables, a student t-test is once again allowed. Using those tests, it was found that for the plural models, the Complaint Handling questions were answered worse than the CAPA questions ( $P < 0.05$ ,  $t(44) = 2.2479$ ) (Figure 15). It turned out to be

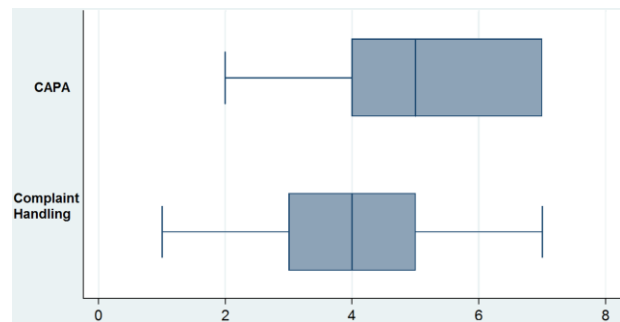


Figure 15: Boxplot for Scoring on Plural questions

that for the plural models the Complaint Handling questions were not significantly answered quicker than the CAPA questions ( $P > 0.05$ ,  $t(44) = 1.5675$ ) (again obtained after a transformation of the data). For the Classical models, there was no significant difference between the Complaint Handling process and the CAPA process, based on score ( $P > 0.05$ ,  $t(41) = -1.0058$ ) and time ( $P > 0.05$ ,  $t(41) = -0.7945$ ) (transformed data for the time). These results seem to indicate that the questions of the Complaint Handling process were perceived as more difficult than the CAPA questions in the case of Plural models. This means that it is meaningful to control for the types of processes in the analyses (and that results from the CAPA models cannot meaningfully be compared to the Complaint Handling processes).

### 8.2.3 Differences in modeling experience

The next control variable was to check whether people with previous modeling experience performed significantly different from inexperienced people. People were asked how long ago they first encountered a process model and how often they used process models in their work. These two questions were combined in a binary value, whether people would only be classified as 'experienced' if they had seen process models at least longer than a year ago and use them at least once a month in their work. Then, using the Mann-Whitney test, it was found that 'experienced' people did not perceive the models as significantly easier or harder ( $P > 0.05$ ). Student t-tests showed that there was no significant difference between experienced and inexperienced participants, based on score ( $P > 0.05$ ,  $t(87) = -0.3062$ ) and time ( $P > 0.05$ ,  $t(87) = -0.6731$ ). This seems to indicate that there is no performance difference in experienced and inexperienced participants, which would mean that it does not have to be included as a control variable in ANOVA analyses.

### 8.2.4 Differences between groups

The final control variable in this study was the group participants belonged to. Mann-Whitney tests showed that students did not find the tasks significantly more difficult than other groups ( $P > 0.05$ ), but it turned out that random Philips employees found them more difficult than others ( $P < 0.05$ ) and Plural agents found them easier than others ( $P < 0.05$ ), as depicted in figure 16. Using student t-tests it was found that students scored worse than other groups

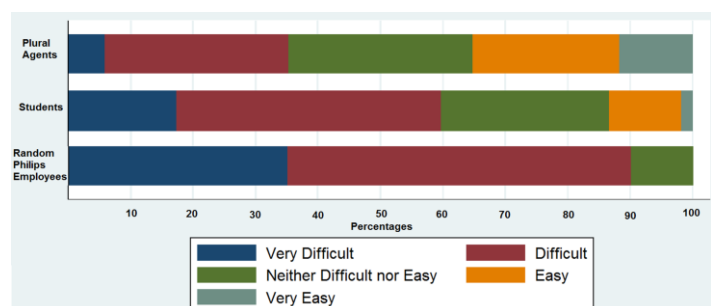


Figure 16: Perceived difficulties per group

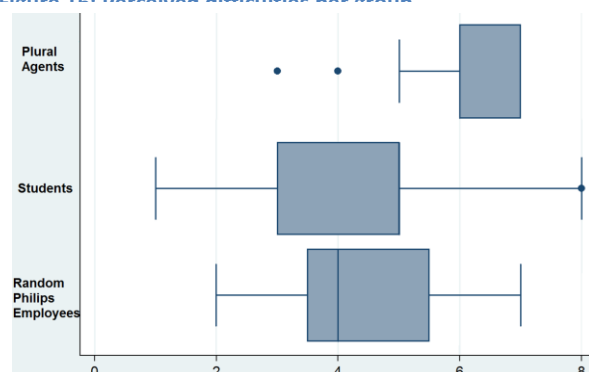


Figure 17: Total score per group

( $P < 0.05$ ,  $t(87) = 2.2268$ ). Random Philips employees did not perform significantly different from the others ( $P > 0.05$ ,  $t(87) = 0.8377$ ), while Plural Agents scored better than others ( $P < 0.001$ ,  $t(87) = 3.9041$ ). These differences are depicted in the boxplot of Figure 17. Based on time, neither students ( $P > 0.05$ ,  $t(87) = -0.8528$ ), nor random employees ( $P > 0.05$ ,  $t(87) = -0.5767$ ) or Plural agents ( $P > 0.05$ ,  $t(87) = -0.4542$ ) performed better than others. This does lead to the conclusion that the group participants came from should be considered in hypothesis testing.

### 8.3 Hypothesis testing

After the control checks were done, the hypotheses were asserted. This was done by using both parametric and non-parametric tests, as not all dependent variables passed the parametric properties (Field, 2009). Some of the dependent variables were either non-normally distributed or not of the ratio or interval format. The frequent occurrence of non-normality could be explained as such: though the scores are ratio variables (i.e. there is an absolute zero point and the difference between two adjacent values is constant), they often have only three categories (i.e. zero, one or two correct values). Such a score variable will only be normally distributed if most people answered one question correct, while the number of people who answered zero questions correct is equal to the number of two correct questions. This is not always likely, especially if the questions were hard to answer (i.e. the median is zero rather than one). Please see the histogram in figure 18 for one of the dependent variables of hypothesis 1.

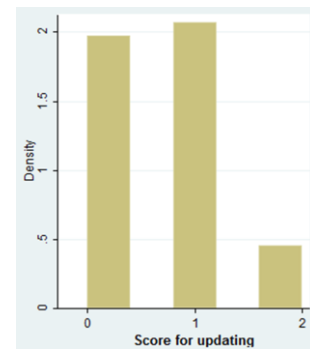


Figure 18: Score for Updating

If this turned out to be the case, no transformations of the data would make it normal. Fortunately, using non-parametric tests like the Mann-Whitney test solved that problem. It has a lower power to detect significant effects than parametric tests, which means the chance of type-II error increases. However, some of the dependent variables did pass the parametric properties after a potential transformation. In those cases, conventional tests like student t-tests and ANOVA's could be conducted, as the F-statistic (based on assumptions of normality) could be trusted.

#### 8.3.1 Hypothesis 1

As the score for updating the models (two questions) did not match the parametric properties, the difference between groups for that variable was asserted using the Mann-Whitney test. The tables below report their outcomes. ClassicalxFirst should be read as: the values for classical models if they are the first model shown.

Group 1 (Mean; Median)	Group 2 (Mean; Median)	Test ( $W_s$ )	z-value	Sig.	Size (r)
Classical (1;1)	Plural (0.56;0)	481.5	2.384	$P < 0.05$	0.36
ClassicalxFirst(1.2;1)	PluralxFirst(0.64;1)	142	2.247	$P < 0.025$	0.46
ClassicalxSecond(0.8;1)	PluralxSecond(0.45;0)	104	1.327	N.S.	(-)

Table 2: Mann-Whitney results for updating the Complaint Handling process (score)

Group 1 (Mean; Median)	Group 2 (Mean; Median)	Test ( $W_s$ )	z-value	Sig.	Size (r)
Classical (0.52;0)	Plural (0.62;1)	493	-0.539	N.S.	(-)
ClassicalxFirst(0.33;0)	PluralxFirst(0.83;1)	78	-1.654	N.S.	(-)
ClassicalxSecond(0.64;0.5)	PluralxSecond(0.33;1)	94.5	0.964	N.S.	(-)

Table 3: Mann-Whitney results for updating the CAPA process (score)



Please note that to assert the interaction between modeling type (classic and Plural) and ordering (first or second model), two concurrent tests had to be executed. This may have inflated the Type-I error, which increases the chance something seems significant while in fact it is not. To correct for this, the Bonferoni Correction was applied (Field, 2009). This means in order to accept the significance of some effect, the significance level  $\alpha$  should be divided by the amount of tests performed to assert this effect. Since two consecutive tests were used to assert this difference, the significance level should be set at  $\alpha=0.025$ . As can be seen from tables 2 and 3, the only significant effects existed for the Complaint Handling process, where it turned out that the questions were significantly harder to answer for Plural models than for the Classical models. There seemed to be an interaction effect with the order of models, as this effect was only present when analyzing the first model shown. This allowed for partial rejection of Hypothesis 1. To understand more, the times to obtain these scores were analyzed as well.

Though a Shapiro-Wilks test showed that the time for the understanding questions is not normally distributed ( $P<0.05$ ), it turned out that its transformation through a logarithm is normally distributed ( $P>0.05$ , figure 19). An ANOVA analysis was subsequently performed. The main effect of the Plural method on the time it took to update the model was not significant,  $F(1,70)=0.43$ ,  $P=0.515$ ,  $\omega^2=-0.006$ . No interaction effect was found between the type of process and modeling method, as there was no significant effect for the Plural method on the time it took to update the Complaint Handling process,

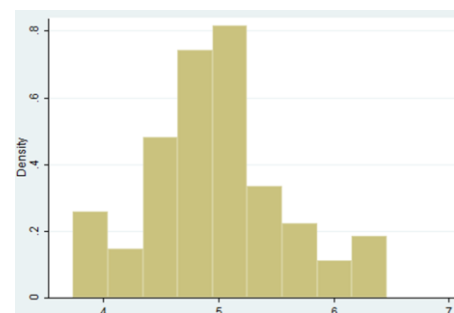


Figure 19: Logarithm of time to update

$F(1,70)=3.16$ ,  $P=0.0796$ ,  $\omega^2=0.024$ . No evidence was found to invalidate the partial rejection of Hypothesis 1.

For further support of this finding, the ratio between the score of updating the models and the time to do so was calculated. This represents the efficiency to obtain correct answers. The difference in efficiency was checked among different conditions as well, using several Mann-Whitney tests. Mann-Whitney tests were selected as this efficiency ratio was not normally distributed, nor could it be transformed into normally distributed data using any of the attempted ladder of powers (Tukey, 1977). The Mann-Whitney tests showed no significant differences among these efficiency ratios in updating the model ( $P>0.05$ ).

For final assertion of Hypothesis 1, the performance for each individual question was asserted. The ability to update the models was asserted through asking two questions per model. For each question the differences in score, the time to do so and the efficiency is asserted among different groups. The score for the first question of updating was asserted through a logistic regression, which is useful in assertion of binary dependent variables (as the question could only be right or wrong). This analysis showed a Pseudo  $R^2$  of 0.26, but none of the independent variables of hypothesis 1, nor the interactions with control variables showed a significant effect on likelihood to correctly answer this question ( $P>0.05$ ).

The time to answer the first question about updating the model was not normally distributed, and none of the attempted ladder of powers could transform the variable to normal. Therefore, a series of Mann-Whitney test was conducted to assert the hypothesis of a significant difference among

groups for the time to answer the first question of updating. No such significant difference existed ( $P>0.05$ ). Finally, the efficiency to answer this question correctly was asserted. The score for the question was divided by the time to answer it, resulting in a non-normally distributed dependent variable. A series of Mann-Whitney tests asserted the differences among groups. A significant difference was found in efficiency for the Complaint Handling process when it was the first model shown. In that case, this question was answered significantly more efficient for a classical model than for a Plural model, even after the Bonferoni correction for two consecutive test ( $W_s=138$ ,  $z=2.581$ ,  $P<0.01$ ,  $r=0.52$ ). The effect size  $r$  indicates the presence of a large effect (Rosenthal, 1991). The same effect did not occur when the same model was the second model shown ( $P>0.05$ ), indicating a learning effect.

For the second question about updating the model, tests were executed to assert the differences among groups based on score, time to obtain that score and efficiency in obtaining correct answers. As score was again a binary value (the question was either right or wrong), logistic regression was applied with the score for the second updating question as a dependent variable and the model type, process type and control effects as predictors. The model turned out to predict the likelihood of a score poorly, with a pseudo  $R^2=0.046$ . Furthermore, none of the predictors significantly influenced the likelihood of correctly answering the question. The time to obtain this correct question was not normally distributed (Shapiro-Wilk's test:  $P=0.00$ ), but a transformation corrected for this. The inverse of the square root of the time was normally distributed, which allowed for a parametric test. A two-way ANOVA was conducted to assert the differences among groups based on the time to answer the last question about updating. Not a single set of groups included in the ANOVA showed a significant difference in time to answer the question ( $P>0.05$ ).

Finally, the efficiency of correctly answering the second updating question was checked for differences among groups. As neither this dependent variable nor any member of the ladder of powers was normally distributed (Shapiro-Wilk's test:  $P<0.05$ ), a series of Mann-Whitney tests was used to assert the differences among groups. A borderline significant effect was found in the CAPA process, where it turned out the question was answered significantly better for the second model if it was a Classical model rather than a Plural model ( $W_s=81$ ,  $z=2.202$ ,  $P=0.027$ ,  $r=0.45$ ). Though the  $P$ -value was well below the regular 0.05 cutoff point, Bonferoni correction should be applied as two consecutive tests were conducted to assert this effect. This would cause the cutoff point to be 0.025 and the test would be insignificant. However, many researchers have stated the Bonferoni correction is too strict which is why this effect can be regarded as significant (Perneger, 1998) (Nakagawa, 2004). A similar test showed this difference did not occur when it was the first model shown ( $P>0.05$ ). This result again seems to support the (partial) rejection of hypothesis 1. All conducted tests during the analysis of hypothesis 1 are summarized in table 4.

	Score		Time		Efficiency	
	Test	Result	Test	Results	Test	Results
Q1-Q2	Mann-Whitney	<b>Significant</b>	ANOVA	Not Significant	Mann-Whitney	Not Significant
Q1	Logistic Regression	Not Significant	Mann-Whitney	Not Significant	Mann-Whitney	<b>Significant</b>
Q2	Logistic Regression	Not Significant	ANOVA	Not Significant	Mann-Whitney	<b>Significant</b>

Table 4: Conducted tests for hypothesis 1

### 8.3.2 Hypothesis 2

To assert hypothesis 2, many dependent variables were checked. The questionnaire contained six questions, two for each of the three categories of understanding. Analyses were conducted on score, time and efficiency for the total set of questions, each type of understanding (control, data, and organization) and each separate question. Thus in total 30 different dependent variables were tested. To maintain some oversight, first the total set of questions was analyzed, then the three types of understanding and then each individual question. Each time, a table is provided with a summary of the conducted tests and their significance. First, tests were conducted for the total score, time and efficiency.

Shapiro-Wilks tests showed that the total score, the logarithmic transformation of total time and the square root of the efficiency were all normally distributed ( $P > 0.05$ ). This allowed for the use of parametric tests. An ANOVA showed no significant interaction effect between the application of the Plural method and the type of process on the total score for understanding the process models ( $F(1,70)=0.15$ ,  $P=0.7037$ ,  $\omega^2=-.007$ ). Similarly, another ANOVA showed no significant interaction effect between the application of the Plural method and the type of process on the total time required for answering the understanding questions of the process models ( $F(1,70)=0.84$ ,  $P=0.3636$ ,  $\omega^2=-.002$ ). Finally, an ANOVA showed there was no significant interaction effect between the application of the Plural method and the type of process on the efficiency of answering all six questions ( $F(1,70)=0.58$ ,  $P=0.4483$ ,  $\omega^2=-.004$ ). This implies no support for Hypothesis 2. The conducted tests are summarized in table 5:

Q1-Q6	Score		Time		Efficiency	
	Test	Result	Test	Results	Test	Results
	ANOVA	Not significant	ANOVA	Not Significant	ANOVA	Not Significant

Table 5: Conducted tests on total set of questions

Next, the analyses were performed for the three types of understanding. Shapiro-Wilks tests showed that the score for data flow and organizational aspects were normally distributed ( $P > 0.05$ ). Furthermore, the logarithms of times for the three aspects were all normally distributed, as were the square roots of the efficiency for those aspects. The one dependent variable which could not be transformed into a normal distribution was the score for the control flow questions. Therefore, Mann-Whitney tests were used to obtain the following results:

Group 1 (Mean; Median)	Group 2 (Mean; Median)	Test ( $W_s$ )	z-value	Sig.	Size (r)
Classical (1.5;1.5)	Plural (1.52;2)	445	-0.392	N.S.	(-)
ClassicalxFirst(1.4;1)	PluralxFirst(1.5;2)	114	-0.725	N.S.	(-)
ClassicalxSecond(1.6;2)	PluralxSecond(1.54;2)	113	0.246	N.S.	(-)

Table 6: Mann-Whitney results for understanding control flow of the Complaint Handling process (score)

Group 1 (Mean; Median)	Group 2 (Mean; Median)	Test ( $W_s$ )	z-value	Sig.	Size (r)
Classical (1.21;1)	Plural (1.52;2)	468.5	-1.268	N.S.	(-)
ClassicalxFirst(1.22;1)	PluralxFirst(1.67;2)	83	-1.290	N.S.	(-)
ClassicalxSecond(1.21;1)	PluralxSecond(1.33;1)	112.5	-0.307	N.S.	(-)

Table 7: Mann-Whitney results for understanding control flow of the CAPA process (score)

As tables 6 and 7 indicate, no significant effects were found for the Plural method on the score of understanding the control flow of the different process models. An ANOVA showed the main effect

of the application of the Plural method on the time to answer the control flow questions was not significant, ( $F(1,70)=2.07$ ,  $P=0.154$ ,  $\omega^2=.01$ ). The interaction effect between the type of process and the type of model on the time to answer the control flow questions was insignificant as well ( $F(1,70)=1.46$ ,  $P=0.232$ ,  $\omega^2=-.005$ ). No significant interaction effect was found between the between the type of process and the type of model on the efficiency to answer the control flow questions correct ( $F(1,70)=0.26$ ,  $P=0.6134$ ,  $\omega^2=-.007$ ).

After conducting ANOVA, a significant main effect was found between the scores of understanding data flow for the Complaint Handling process and the CAPA process ( $F(1,70)=8.92$ ,  $P<0.01$ ,  $\omega^2=-.007$ ). However, no significant interaction effect was found between the type of process and the type of model on the scores for the data flow questions ( $F(1,70)=0.70$ ,  $P=0.405$ ,  $\omega^2=-.002$ ). Furthermore, the main effect of the application of the Plural method on the time to answer the data flow questions was not significant, ( $F(1,70)=1.76$ ,  $P=0.189$ ,  $\omega^2=.007$ ). The interaction effect between the type of process and the type of model on the time to answer the data flow questions was insignificant as well ( $F(1,70)=0.47$ ,  $P=0.494$ ,  $\omega^2=-.005$ ). Finally, the interaction effect between type of process and type of model on the efficiency to answer data flow questions correct was insignificant as well ( $F(1,70)=1.34$ ,  $P=0.251$ ,  $\omega^2=-.005$ ).

ANOVA showed there was no significant main effect between the scores of understanding organizational aspects for the Complaint Handling process and the CAPA process ( $F(1,70)=0.01$ ,  $P=0.909$ ,  $\omega^2=-.01$ ). No significant interaction effect was found between the type of process and the type of model on the scores for the data flow questions ( $F(1,70)=2.15$ ,  $P=0.147$ ,  $\omega^2=.01$ ). The main effect of the application of the Plural method on the time to answer the questions on organizational aspects was not significant, ( $F(1,70)=2.07$ ,  $P=0.15$ ,  $\omega^2=.01$ ). The interaction effect between the type of process and the type of model on the time to answer the organizational questions was insignificant as well ( $F(1,70)=1.46$ ,  $P=0.232$ ,  $\omega^2=-.005$ ).

It was interesting to see however, that when the ANOVA model was used to assert the efficiency to answer the questions about organizational aspects, there was a significant interaction effect between model type (plural versus classical) and process (Complaint Handling versus CAPA),  $F(1,74)=15.28$ ,  $P<0.0005$ ,  $\omega^2=0.12$ ). This indicates that in either one of those processes, there was a difference between classical and Plural models. To further analyze this difference, a student t-test was used. It turned out that for the Complaint Handling process, the questions about organizational aspects were answered significantly worse for Plural than for the classical models ( $t(43)=3.0363$ ,  $P<0.005$ ). This result seems to partially reject hypothesis 2. All conducted tests for the three types of understanding are summarized in table 8.

	Score		Time		Efficiency	
	Test	Result	Test	Results	Test	Results
<b>Q1-Q2 (Control)</b>	Mann-Whitney	Not significant	ANOVA	Not Significant	ANOVA	Not Significant
<b>Q3-Q4 (Data)</b>	ANOVA	Not Significant	ANOVA	Not Significant	ANOVA	Not Significant
<b>Q5-Q6 (organization)</b>	ANOVA	Not Significant	ANOVA	Not Significant	<b><u>ANOVA &amp; t-tests</u></b>	<b><u>Significant</u></b>

Table 8: Conducted tests on three types of understanding

For further analysis of hypothesis 2, each individual question was asserted as well for differences among groups, based on scores, times and efficiency to obtain correct answers. This may provide additional insights in the effect of the Plural method on understanding process models. All scores of individual questions were asserted using a logistic regression due to their binary nature. It turns out for none of the six questions the hypothesized predictors resulted in a higher likelihood of correctly answering a question ( $P >> 0.05$  in all cases). Next, the difference among groups in times to answer those questions was asserted. Though none of these times were normally distributed (Shapiro-Wilk's test:  $P < 0.05$ ), the logarithmic transformations of these times were normally distributed ( $P > 0.05$ ), allowing for parametric testing.

For all six questions, an ANOVA was conducted to assert significant differences in time it took for answering these questions among groups. The only question for which a significant difference was found for the hypothesized groups was for question 5, one of the questions about organizational aspects. A significant interaction effect was found between the type of model (Plural versus classical) and process (Complaint Handling versus CAPA) on the time to complete this question,  $F(1,70)=6.20$ ,  $P=0.001$ ,  $\omega^2=0.09$ . The ANOVA only shows there was a significant difference, not in what direction. Therefore, a student t-test was conducted to assert these differences. This test showed that for the Complaint Handling process the time to answer question 5 was significantly shorter for a classical model than for a Plural model ( $t(43)=-3.28$ ,  $P=0.001$ ). This difference did not occur for the CAPA model. This result supports the partial rejection of hypothesis 2.

The final assertion of hypothesis 2 includes measuring the difference between efficiency of groups to answer the individual questions. For these analyses, Mann-Whitney tests were used due to the lack of normality in distribution. These tests showed no significant difference between the Plural and Classical models for efficiency in answering questions one and two. However, it was found that question three was answered more efficient for classical models rather than Plural models, for both CAPA ( $W_s=380$ ,  $z=2.191$ ,  $P=0.029$ ,  $r=0.33$ ) and Complaint Handling ( $W_s=487$ ,  $z=2.139$ ,  $P=0.032$ ,  $r=0.32$ ). Furthermore, when controlling for ordering effects, it was found that for the CAPA process, answering question three for the second model done significantly more efficient for classical models rather than Plural models ( $W_s=70.5$ ,  $z=2.370$ ,  $P=0.0178$ ,  $r=0.49$ ). This effect still stands after a strict Bonferoni correction for two consecutive tests. This did not occur when the model was the first one shown. This is in partial rejection of hypothesis 2.

Similar results occurred for assessing the difference in efficiency to answer question four between groups. Mann-Whitney tests showed that for the CAPA process, the questions were answered more efficient for the classical models than for the Plural models ( $W_s=314$ ,  $z=3.729$ ,  $P=0.0002$ ,  $r=0.56$ ). When controlling for the ordering effect, it was found that answering question four for the CAPA process on the first model was done (borderline) significantly more efficient for the classical model rather than the Plural model ( $W_s=90$ ,  $z=2.985$ ,  $P=0.028$ ,  $r=0.65$ ). Like in a previous case, Bonferoni correction tightens the significance level to  $\alpha=0.025$ , but due to the strict nature of Bonferoni it could be argued the effect does exist. Both the r-values of these effects are rather large, implying a very strong effect. No significant differences were found for other orders or in the Complaint Handling process.

For question five, similar analyses showed that for the Complaint Handling process answers for a classical model were given more efficient than for plural models ( $W_s=462.5$ ,  $z=2.581$ ,  $P=0.0098$ ,

r=0.38). When controlling for a potential ordering effect, it was found that when such a Complaint Handling model was the first model shown, the answer was given more efficient for a classical model rather than for a Plural model ( $W_s=123$ ,  $z=3.059$ ,  $P=0.002$ ,  $r=0.62$ ). Similar effects were not found for other orders, processes or model types. These effects are in partial rejection of hypothesis 2. All conducted tests for the six individual questions are summarized in table 9.

	Score		Time		Efficiency	
	Test	Result	Test	Results	Test	Results
Q1	Logistic Regression	Not Significant	ANOVA	Not Significant	Mann-Whitney	Not Significant
Q2	Logistic Regression	Not Significant	ANOVA	Not Significant	Mann-Whitney	Not Significant
Q3	Logistic Regression	Not Significant	ANOVA	Not Significant	<u>Mann-Whitney</u>	<u>Significant</u>
Q4	Logistic Regression	Not Significant	ANOVA	Not Significant	<u>Mann-Whitney</u>	<u>Significant</u>
Q5	Logistic Regression	Not Significant	<u>ANOVA &amp; t-tests</u>	<u>Significant</u>	<u>Mann-Whitney</u>	<u>Significant</u>
Q6	Logistic Regression	Not Significant	ANOVA	Not Significant	Mann-Whitney	Not Significant

Table 9: Conducted tests on all six individual questions

The results of the analysis on the efficiency of the individual questions, lead to the partial rejection of hypothesis 2. Understanding aspects regarding the flow of data was performed significantly worse for Plural models rather than for Classical models. Furthermore, one of the questions about organizational aspects was consistently done significantly worse, based on multiple criteria.

### 8.3.3 Hypothesis 3

Hypothesis 3 was asserted by performing analyses on items asserted on a likert scale. These scales were neither interval nor ratio scaled, meaning it could not be guaranteed that each set of two consecutive answer-categories cover constant difference. This meant that it did not make sense to conduct parametric tests like the student t-test or the ANOVA, as they assume an absolute zero point and constant intervals between answer categories. Therefore, all analyses to assess hypothesis 3 were Mann-Whitney tests. Whenever a certain effect was found through multiple test (let's say N tests), the usual significance interval of  $\alpha=0.05$  will be tightened through the Bonferoni correction of  $\frac{\alpha}{N}$ , to correct for the inflation in Type-I error.

For each observation, four types of questions were asked. Two items to assess the participant's intention to use the process model, three items about the perceived usefulness, three about the perceived ease of use and three items about the ability of the process model to provide the needed information. For each of these four sets of items, scales were created by calculating the average and assessing their reliability through Cronbach's Alpha. This resulted in the following scales: Intention to Use ( $\alpha=0.93$ ), Perceived Usefulness ( $\alpha=0.95$ ), Perceived Ease of Use ( $\alpha=0.83$ ) and Information Retrieval ( $\alpha=0.86$ ). Though some sources state a cutoff point of  $\alpha=0.70$  might be too strict, for these reliabilities it does not matter; they easily surpass that threshold (Cortina, 1993) (Kline, 1999).

For the assessment of this hypothesis, the averages of the scales among different groups were compared. This was done using the aforementioned Mann-Whitney test, while controlling for

ordering effects (First x second), process type (Complaint Handling x CAPA) and originating group (Student x random Philips employee x Plural agent). It was found that none of the hypothesized relationships were significant ( $P > 0.05$ ). In none of the cases, it turned out to be that either of the groups had a significantly different opinion about the models. To check as to why this was the case, all values for the different dependent variables are displayed in table 10.

Model type	Construct	# obs	Mean	Std. Dev.	Min	Max
CAPA	<i>Intention to use</i>	23	4.65	1.51	1.00	7.00
Classical model	<i>Perceived usefulness</i>	23	4.78	1.09	3.00	7.00
	<i>Perceived ease of use</i>	23	3.42	1.27	1.67	6.67
	<i>Information retrieval</i>	23	4.71	1.08	2.67	6.67
CAPA Plural model	<i>Intention to use</i>	21	4.93	1.57	1.00	7.00
	<i>Perceived usefulness</i>	21	4.25	1.30	2.00	6.67
	<i>Perceived ease of use</i>	21	3.92	1.66	1.00	6.00
	<i>Information retrieval</i>	21	4.90	1.33	2.00	7.00
Complaint Classical model	<i>Intention to use</i>	20	4.23	1.76	1.00	7.00
	<i>Perceived usefulness</i>	20	3.98	1.40	1.00	6.67
	<i>Perceived ease of use</i>	20	3.50	1.47	1.00	6.00
	<i>Information retrieval</i>	20	4.52	1.50	1.33	6.67
Complaint Plural Model	<i>Intention to use</i>	25	3.98	1.96	1.00	7.00
	<i>Perceived usefulness</i>	25	3.93	1.63	1.00	7.00
	<i>Perceived ease of use</i>	25	3.00	1.26	1.00	6.00
	<i>Information retrieval</i>	25	4.39	1.34	1.67	7.00

Table 10: Descriptives of TAM-related questions per model and process type

As can be seen from table 10, participants (on average) had some intention to use these process models, thought they were somewhat hard to use but quite useful. Furthermore, they indicated they thought the model did provide the information they needed to some extent. The crux, however, is for all these measurements *there was no significant difference among the hypothesized groups*. This leads to the conclusion that there is no support for hypothesis 3. Discussion about why this might be the case can be found in the next chapter.

## 9. Discussion

This section of the report is a critical discussion of the results from the previous sections. All hypotheses of this research concern relative advantages of the models created through the Plural method over models created through classical modeling techniques. Therefore, results about those hypotheses are discussed first. However, the Plural method has been applied in full and therefore more results have surfaced from this research. They are discussed afterwards.

### 9.1 Discussion of experiment

Though there were good reasons to hypothesize prevalence of the Plural method's models over classical models, the results of previous sections suggest otherwise. Out of the initial three hypotheses, two were partially rejected and one found no support. Table 11 recaps the different hypotheses and the reason for rejection or inconclusive results for the different conditions of process and ordering effects. Afterwards, an extensive discussion is written to explain the results as they were.

Hypothesis	Result	#	Process	Significant results (order of model)
<i>H1: A process model acquired through the Plural Method is easier to adapt to changes than a conventional process model.</i>	Partially rejected	1	Complaint	Score for updating favors classical (1 <sup>st</sup> )
		2	Complaint	Efficiency update 1 favors classical (1 <sup>st</sup> )
		3	CAPA	Efficiency update 2 favors classical (2 <sup>nd</sup> )
<i>H2: A process model acquired through the Plural Method is easier to understand than a conventional process model.</i>	Partially rejected	4	Complaint	Efficiency understanding organizational aspects favors classical (both)
		5	Complaint	Time Q5 favors classical (both)
		6	Both	Efficiency Q3 favors classical (both)
		7	CAPA	Efficiency Q3 favors classical (2 <sup>nd</sup> )
		8	CAPA	Efficiency Q4 favors classical (both)
		9	CAPA	Efficiency Q4 favors classical (1 <sup>st</sup> )
		10	Complaint	Efficiency Q5 favors classical (1 <sup>st</sup> )
<i>H3: A person is more willing to use the models created through the Plural Method, rather than using conventional business process models.</i>	No support found	11	Both	None of the hypothesized differences among groups were significant

Table 11: Results of asserting hypotheses

As the table indicates, the classical models outperformed the Plural models on multiple fronts. The rejection of two hypotheses was quite surprising, as there were good reasons to favor Plural models over classical models when understanding and updating them. The highly modular nature of Plural models would allow the user to quickly identify what block would be relevant to the question and then look inside it to see if the required update should occur at that place. In a classical model, a lot more information was present on the top level which would ‘overload’ the reader and thus make this update more difficult. Understanding would be easier to do for the Plural models through similar mechanics as well. The results showed more significant differences in the Complaint Handling process than in the CAPA process. To analyze inherent differences between these two models, a set of process model metrics was counted and displayed in table 12 (Reijers & Mendling, 2011).

		CAPA			Complaint Handling		
		Top	Bottom	Total	Top	Bottom	Total
<b>Classical model</b>	<i># nodes</i>	72	114	186	63	34	97
	<i># sequence arcs</i>	67	112	179	64	40	104
	<i># message arcs</i>	14	0	14	7	0	7
	<i># gateways</i>	18	27	45	22	11	33
	<i># sub-processes</i>	14	0	14	5	0	5
<b>Plural model</b>	<i># nodes</i>	37	143	180	42	141	183
	<i># sequence arcs</i>	30	134	164	37	144	181
	<i># message arcs</i>	29	47	76	17	24	41
	<i># gateways</i>	0	38	38	6	48	54
	<i># sub-processes</i>	19	0	19	17	0	17

Table 12: Metrics of researched process models

As table 12 shows, both the classical and Plural models of the CAPA process were rather modular. Where the Plural model had 19 sub-processes, the classical model also had 14 sub-processes.



However, in the Complaint Handling process there was a far larger structural difference between classical (5 sub-processes) and Plural (17 sub-processes). This indicates that if there were some kind of effect related to the usage of sub-processes, it was far more likely to surface in a process where this difference was more pronounced (i.e. the Complaint Handling process rather than the CAPA process). Furthermore, it turns out that the difference in number of nodes between classical and plural was far larger for the Complaint handling process (183 Plural vs 97 Classical) than for the CAPA process (180 Plural vs 186 Classical). So the difference between findings for CAPA and Complaint Handling may also be explained by the ratio of nodes between their respective classical and Plural models. Analysis of the rejected hypotheses combined with table 12 leads to a few questions:

- If there is such a strong difference in modularity of the classical and Plural model for the Complaint Handling process, why does the model with little to no modularization repeatedly outperform the highly modular one?
- If there is so little difference in modularity for the respective CAPA models, why were any significant differences found in the first place?

An explanation for the first point could be that the added benefit in understanding the Complaint Handling process through modularization is counteracted by the increase in size. So if there would be a significant increase in understanding through the usage of modularity, it may be completely turned over since the modular model also has almost twice the number of nodes of the non-modular model. This possibility cannot be ruled out by current experiments, as the other process (CAPA) provides too little difference in modularity between the classical and the Plural model. Follow-up research could show how modularity and size interact. More about this possibility is stated in section "Future Research".

Even though the questions have been validated by an expert of the processes (see section 7.2), another explanation for that first point might lie in the nature of the questions. Though initially the idea that modularity supports understanding and maintenance makes sense due to information hiding and encapsulation, there may be nuances to this phenomenon. If a participant is presented with a completely modular model and the question can be answered by looking at the behavior of a single sub-process, this modularity may be beneficial to understanding. If the exact same question should be answered for a model without any modularity, this may be harder to do. However, if the exact same two models should be used to answer a question which requires the participant to look into more parts (or sub-processes), modularity may hamper understanding. This notion of 'local' and 'global' questions was originally coined in previous research (Reijers, Mendling, & Dijkman, 2011). The authors concluded, though with some caution, that modularization fosters understanding, even more so the case of local questions.

As it turned out, most of the questions could be categorized as 'global' (i.e. they required to look at more than one sub-process). For the CAPA model, three questions were 'local' and five were 'global'. In the case of the Complaint Handling process, this distribution was two questions 'local' and six 'global'. Even though the paper by Reijers et. al. (2011) stated that modularity was more instrumental to understanding in the case of local questions, they doubted whether the opposite effect was true and that global questions could be answered more easily by non-modular models. Though there were cases where the classical (non-modular) model outperformed a Plural model for a global question (e.g. result #2, see table 11), there were also cases where the classical (non-modular)

model outperformed the Plural model for a local question (e.g. result #6). Furthermore, participants may not have recognized a local question when they saw one. A question which could very well be answered by looking at only a single block may still have lead the participant to look in other blocks to ensure a correct answer. More about that phenomenon is mentioned later on. Therefore, it seems to be that it cannot simply be stated that the global questions were easier for classical models and that the local questions were easier for the plural models. It was a general trend that classical models outperformed Plural models, regardless of the local or global nature of the question.

Thus, it seems that the strongest effect influencing the results so far is that the reported benefit of modularity in previous research may be dampened or even completely countered when the modular model is significantly larger than the non-modular model. No clear explanation for the significant results of the CAPA process has been provided so far. It was found that when it comes to the matter of understanding, classical models frequently performed better than Plural models, while there was very little difference in modularity between classical and plural for CAPA. So there must have been other forces working to foster understanding in the classical model. When looking at table 12, one can observe another striking difference between the Plural model and the classical model for the CAPA process.

With similar levels of modularity, the classical model displayed more control flow on the top level than the Plural model did on that level (67 sequence arcs versus 30, 18 gateways versus 0). With little difference in modularization on the top level, but with a strong difference in control flow elements, the explanation for the difference in performance may lie in the latter. However, when inspecting the questions which resulted in significant differences for the CAPA model, it was found that only questions 3 and 4 were answered more efficient for the classical model. This is surprising, as these were the questions asserting understanding of the data flow. Data flow was specifically the aspect which was thought to be most easy to understand in the Plural models. Furthermore, the questions asserting understanding of control flow found no significant differences between classical and Plural models.

The aspect on which questions 3 and 4 were answered better for the classical model was efficiency, obtained by dividing the score for the question by the time to obtain that score. This could mean that though questions 3 and 4 did not require control flow to answer the question itself, control flow may have guided the reader to reach the point where the answer can be found. This implies that though it has been established in previous research that modularity fosters understanding, 100% modularity is not the best way to apply modularity in a process model if the goal is to maximize efficiency in understanding. Current research has analyzed two highly modular models and showed that the fully modular model was not as efficient to understand as the model with both modularity and some additional control flow.

There is a nuance to the previous discussed results, however. It may have been the case that when a participant found the answer of a question, he/she would continue looking through the model to make sure no exceptions to their answer were found. If that was the strategy participants used, the results of the current study should be viewed in a different light. Such a strategy would favor smaller models and models which provide additional structure. Then that would further explain why the smaller model and the model with more control flow elements outperformed the larger model and the model with no control flow elements on the top level. Current research was not set up to control

for this phenomenon however. The only current test would be to compare global and local questions based on time. If participants did not take a significant different amount of time to answer these types of questions, the previously mentioned strategy may be the truth. However, as there are so many exceptions to such results, they would not be meaningful. Follow up research could control for this phenomenon and test if it even exists.

The final interesting result of the experiment was the insignificance of hypothesis 3. Regardless of groups (Plural agent x Student x Employee), type of process (CAPA x Complaint Handling), type of models (Classical x Plural), type of metric (Intention to use x Perceived ease of use x Perceived usefulness), one group did not respond significantly different from another. The questions were always stated very clearly *“for this type of process models...”*, which indicates people saw no difference between the first presented process model and the second presented process model. I.e. people did not see a difference between Plural and classical models. This is not surprising, as most participants were unaware of process modeling and BPMN in particular. Since people saw no difference between the two process models, it is not surprising they answered the questions identical for the second time. They answered the questions based on their opinion of process modeling in general. This leads to the conclusion that though there may be an actual difference in Technology Acceptance Model items between Plural and classical models, this will only surface if the difference is more pronounced than it has been in this study. Following chapters will build on this conclusion.

## 9.2 Discussion of modeling sessions

Though the research was about analyzing the process models which were a result of the Plural method, employing the Plural method itself gathered some interesting results as well. The first thing that immediately stood out at the very first session was the willingness of participants to talk about their troubles. Even though current research had little influence on the way work is being executed in Philips Healthcare, people were quick to point out where they felt their process could be improved. For instance, complaint handling experts immediately identified the troubles with classification of a complaint. One participant stated that *“a big problem with complaint handling right now is classifying complaints as either similar or duplicate. People recognize a word in the title of the complaint and, completely unjustified, classify it as a similar”*. One participant of another process was particularly harsh about his process, calling some crucial parts of the process *“executed in an abysmal way”* and stated that *“some potentially safety-related questions are steered towards preferred and easy answers”*. These remarks, stemming from the “blue-collars” could help paint a clear picture on the performance of the process. What was apparent, above all, was the complete honesty of participants. While they stated where improvements could be made, they also indicated where the work was of high quality.

Though the original papers about the Plural Method described only three phases, through this thesis a fourth phase is proposed to optimize value of the Plural Method. A ‘Management Feedback’ phase should incorporate creation of a report containing the process models, analysis on those models and a list of relevant managerial insights. As the Plural Method allows the coordinator to have a rather private session with the agents, it appears that people are feeling ‘free to vent’. Their troubles should definitely not be ignored. Even if their remarks were already familiar for management, this level of feedback could provide additional value of the Plural Method.

Another important finding concerns the suitability of the Plural Method for different processes. In three out of four processes the modeling sessions went fluent. This was the case for the complaint handling process, the CAPA process and the risk management process. The common factors between these processes include the fact that they are all rather stable processes. This means that the process has been going on for quite some time and no major changes have been made nor are they pending. Unlike the FCO process however, as that process would receive a major revision in a few months' time and was in the middle of a redesign effort. This was due to the highly complex and uncertain nature of the process. These attributes caused some confusion during the FCO plural sessions, as people started describing a hybrid between the 'old' situation and the 'new' situation. This would have been fine if everyone agreed on the same level of mixture between the two versions; it would have resulted in a single, clear picture of the process. This was not the case, however. Therefore it seems to be that a process in a redesign phase is not suitable for the Plural Method.

Maybe one of the most interesting findings of the modeling sessions was the difference between the obtained classical and Plural models. In all cases, it turned out to be that the Plural models contained more details, more feedback loops and more exceptions. In some cases additional roles were identified during the creation of the Plural models. This might be explained by the way the process models come about. The classical models were BPMN translation of textual procedures which operate on a rather high level. However, once meetings were set with Plural agents, more detail emerged in these models and thus creating the Plural models. The emergence of this previously uncharted behavior seems to support the qualitative value of the Plural method as a process discovery tool. Though it cannot be concluded that the Plural method is the *only* method that would accomplish this behavior, it *has* been shown that this specific method allows for detection of inconsistencies and deviations from top-down processes. Though similar results could have emerged from using other subject-oriented process modeling techniques, the notion of using an operation - which is exclusive to the Plural method as of yet- was referred to as "*intuitive*" and "*useful*" by plural agents. This result does not add to answering the main research question of this study, but it does add to the body of knowledge about the Plural method.

## 10. Limitations

Though the current research has been set up in the best possible way, limitations to the validity of the results have occurred. This chapter describes the identified limitations to the current setup. The first limitation lies with the inability to fully quantify the quality of the Plural models. The ability to understand and update the models was asserted through the questionnaire, but this is not the only way the performance of these process models can be asserted. The questionnaire asserted pragmatic quality of process models (i.e. can they be used?). The syntactic quality of the process models is fostered by verifying the process models with the BPMN syntax. However, the SIQ framework of process model quality as discussed in paragraph 4.5 includes one other type of process model quality: semantic quality (Reijers, Mendling, & Recker, 2010). The qualitative results of this research showed that the Plural models contained more information (i.e. were more 'complete') and sometimes showed where mistakes occurred in the classical models (i.e. plural models were more 'correct'). The inability of current research to quantify this difference in semantic quality made it harder to answer the original research question.

Another limitation to the current research is the dependence on the modeling skills of the researcher. Though the researcher has experience in process modeling, this mainly stems from practice in college lectures. It cannot be excluded that the experience of the researcher progressed through the phases of the project, which would add to the quality of the processes. This could be a threat to the validity of analysis techniques where models are compared. A related limitation lies in the order in which the types of modeling techniques were applied. In all cases the as-is model was created first and the Plural model second. This may lead to a bias as insights gained in the as-is model may have influenced the Plural models. Though the Plural models all have been made by the agents, the coordinator could unintentionally have influenced the results obtained. Performing the modeling sessions in reverse order (i.e. first the Plural model, second the as-is model) for half of the processes could allow to control for this issue.

A fourth identified threat to the validity of current research is the nature of the case study. In such a setting, it is not uncommon for the researcher to start 'believing' in the method. This level of commitment to the method may cause a bias by 'favoring' the researched method. Even though the experiment results showed that the classical method outperformed the Plural method, it may have been mitigated by unintended preference. A somewhat clumsy mistake by the researcher was the lack of measurement on time spent for the classical method. The times for all Plural sessions were recorded neatly, but this information cannot be analyzed in a meaningful way due to the lack of data regarding the classical method. Therefore, it cannot be assessed if there was disproportionate effort placed in the Plural models.

The final limitation to current research lies with the structure of the assessed models. As can be seen from table 12, there was a clear difference in modularity for the Complaint Handling process. However, this difference was significantly smaller for the CAPA process, which may have prevented additional significant results from surfacing. Therefore, it may have been better for the validity of this research to allow less modularity in the classical models. This would ensure a significant difference in modularity between the two models, which would allow for discovery of significant effects (should they exist).

## **11. Implications**

Current research has impact on several groups whom concern process modeling. Though research results may be interesting to managers and management results may be interesting to researchers, the two types of implications are split for clarity.

### **11.1 Implications for Research**

The execution of current research has led to an increase in understanding of the Plural Method and modularity in process modeling. Previous research has already shown that modularity does not foster understanding equally in all cases (Reijers, Mendling, & Dijkman, 2011). Current research has uncovered more factors to this field of research. The results of the experiment point towards a tradeoff in the usefulness of modularity, the size of the process model and the amount of control flow displayed on the top-level. This study has shown that 100% modularity is not the optimal way of using modularity to foster understanding of a process model.

Though the previous results are interesting for research to any kind of modularization in process modeling, this research has implications to research about the Plural method itself as well. As application of the Plural method entails discussion with employees, it is vital their input will be provided as feedback for management. Therefore, a fourth phase to the Plural method is proposed where management feedback is offered. This updated Plural framework is displayed in figure 20. This additional phase can be accessed by the Description and conflict resolution phase, as that phase provides most intimate contact with the Plural agents. This management feedback is very important to obtain most value out of a Plural effort, but does not have to influence the rest of the Plural project.

### 11.2 Implications for Managers

Performing this study has several interesting implications to Philips Healthcare management (and general management for that matter). The application of the Plural models showed directly where some potential inconsistencies in the current process models of the business units are present. These issues can now be discussed or resolved by the responsible process owners. Moreover, this research has shown the value of including process participants when modeling, or at least describing, business process models. The original reason for Philips Healthcare to have centralized, top-down models was to align the way different business units perform similar processes; Q&R processes in particular. Though there is merit to this way of working, it may hamper the quality of individual units. The most common cited problem with the current procedure among experts was the feeling that forces, more powerful than their business unit, decided on their current way of work. Jokingly, some people referred to this phenomenon like some holy commandment: *“Thou shalt follow this procedure”*. Current research has shown how plenty of problems can be identified and subsequently solved by granting process participants a high stake in the modeling efforts. This may be a powerful notion for upper management to reconsider the way they want the different business units to be aligned.

## 12. Future Research

To accommodate further research in the topic of subject-oriented process modeling, the Plural method or modularity in process modeling, lessons can be learned from current research. Based on the results of this thesis, along with the outlined limitations, future researchers could improve their work. Current research was about looking into the relative performance between Plural models and classical models. It has been shown that classical models may be easier to understand and update than Plural models, while the Plural models seem to identify more accurate behavior. For future research, it is recommended to attempt quantification of the semantic quality of these process models. Furthermore, the Plural method is about more than just the process models it creates. It has

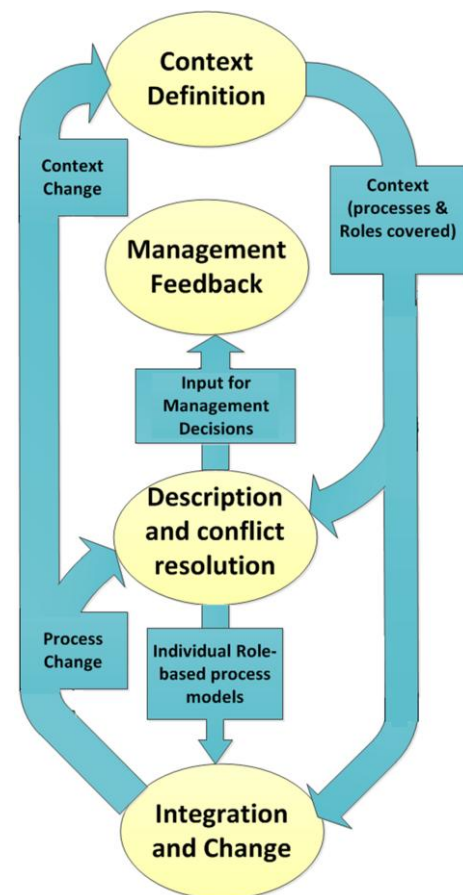


Figure 20: Proposed new Plural framework

been hypothesized in previous work that the Plural method will increase the feeling of process ownership and empowerment among participants (Türetken & Demirors, 2013). If future research could shed light on this hypothesis, the full value of the Plural method may be shown.

For research about the modular nature of the process models, it may be valuable to compare collapsed and expanded ('flattened') versions of the same model. This way, the semantic differences between these two models are completely neutralized; each model contains the exact same behavior as the other. The only difference between these two versions of the same model would be the amount of information displayed on the top level. By comparing these two views of the same process model, insights could be gathered about the phenomenon of modularity. When doing so, it would be advisable to ensure some of the questions of an experiment are global in nature, while others are local in nature. Having a well-found experiment to assert these notions may allow future research to verify or disprove the moderation of question nature on the relationship between modularity and understanding of process models.

Current research has indicated the presence of a preferable balance between several aspects of a process model. It may be that modularity fosters process model understanding, but the experiment has shown fully modular models are not as easy to understand as modular models with some additional control flow. Furthermore, it has been shown that the added benefit of modularity of process models may be canceled out if that implies the total model size is higher than that of a non-modular model. Future research about modularity could benefit from looking into the balance between model size, the extent of modularity and the presence of control flow elements on the top level. It may be the case that all three of these factors influence understanding of process models.

The discussion of the results mentioned the possibility that participants kept looking through the model, even after they already found the answer. This could be because participants want to make sure their answer is correct, or they may not even have recognized the correct answer. This could be checked by logging the html tooltip script used to display the sub-processes. If future research used the same setup as this research, the amount of times the html tooltip script is accessed could be counted and stored in a variable. If this is combined with asking questions to which the answer is available at the beginning of the model and questions at the end of the model, it can be asserted whether people keep looking through the model even though they have already found the answer.

Another recommendation to future research lies with the way hypothesis 3 in current research was asserted. If future researchers wish to compare two process models based on aspects of the Technology Acceptance Model, it is advisable to compare these two directly. Rather than asking TAM questions for model A first and then the same questions for model B, a setup could be designed where both models are shown and preference between the two options is asked based on the TAM questions. This way, there may be a more pronounced cue available that one of these models is different from the other. If they do not have a preference of one over the other, they can still indicate so.

Finally, the Plural method could really improve if researchers took an interest in developing a modeling tool for the Plural method. For this research, several workarounds had to be used in order for an available modeling tool to be suitable. However, if a repository-based tool was developed specifically for the Plural method, research about the method could really show its value.

### 13. Conclusion

This research has analyzed process models in an effort to validate the Plural method; a framework for having process participants model their own process. The main research question of this thesis was: *“To what extent do the models created through the Plural Method outperform those created through classical process modeling techniques?”* After the Plural method was applied on four different processes in the Quality Management System of Philips Healthcare, two out of these four processes were used in an online experiment. It turns out that though participants indicated to see no difference between the Plural models and the classical models, understanding-related questions were answered better for the classical models than for the Plural models. It also turned out that it is easier to update a classical model to a given change rather than a Plural model.

Furthermore, the results of the questionnaire showed that the complete modularity of the Plural models is not the best way to foster understanding in process models. The results indicate there exists some preferred balance between showing control flow and modularity on the top level of a process model as well as a trade-off between modularity and total process model size. To add some nuance to the lackluster performance of understanding and maintaining the process models: it did turn out to be the case that Plural models contained more information and found some faults not identified by the classical models. This result was not expected at the beginning of this research, which is why the effect was not hypothesized. Still, this unexpected result adds to the body of knowledge of the Plural method and any modeling tool incorporating process participants in general.

After having conducted all experiments and analyses for current research, the answer to the original research question is that the Plural models are more complete and correct than the classical models. However, it was found that the classical models were easier to understand and maintain than their Plural counterparts. This implies that the Plural method provides a powerful tool for process discovery, but that its models are not optimal. The way process models in the Plural method come about should be adapted in such a way that the resulting process models are not a 100% modularized anymore.



## References

- FDA - *What we do*. (2013). Retrieved Februari 17, 2014, from FDA: <http://www.fda.gov/AboutFDA/WhatWeDo/default.htm>
- Chinosi, M., & Trombetta, A. (2012). BPMN: An introduction to the standard. *Computer Standards & Interfaces, 34*, 124-134.
- Conger, J. A., & Kanungo, R. N. (1988). The empowerment process: Integrating theory and practice. *Academy of management review, 13*(3), 471-482.
- Coombs, C., Doherty, N., & Loan-Clarke, J. (2001). The Importance of User Ownership and Positive User Attitudes in the Successful Adoption of Community Information Systems. *Journal of End User Computing, 13*(4), 5-16.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology*(78), 98-104.
- Curtis, B., Kellner, M. L., & Over, J. (1992). Process Modeling. *Communications of the ACM, 35*(9), 75-90.
- Davenport, T. H., & Short, J. E. (1990). The New Industrial Engineering: Information Technology and Business Process Redesign. *Sloan Management Review, 13*(4), 11-27.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340.
- Davis, F. D. (1993). User Acceptance of information Technology: System Characteristics, User Perceptions and Behavioral Impacts. *International Journal of Man-Machine Studies, 38*(3), 475-487.
- Davis, F. D., & Venkatesh, V. (1995). Measuring User Acceptance of Emerging Information Technologies: an Assessment of Possible Method Biases. *System Sciences. Proceedings of the Twenty-Eighth Hawaii International Conference on IEEE*, (pp. 729-736).
- Devaraj, S., & Kohli, R. (2003). Performance Impacts of Information Technology: Is Actual Usage the Missing Link? *Management Science, 49*(3), 273-289.
- Dillman, D. A. (2000). *Mail and internet surveys: The tailored design method* (2 ed.). New York: Wiley.
- Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. (2013). *Fundamentals of Business Process Management*. Springer.
- Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaption to task constraints. *Annual review of psychology, 47*(1), 273-305.
- Field, A. (2009). *Discovering Statistics using SPSS*. Sage Publications.
- Field, A. (2011). *Discovering Statistics Using SPSS*. Thousand Oaks, California: SAGE Publications.

- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: an introduction to theory and research*. Addison-Wesley.
- Fleischman, A., Raß, S., & Singer, R. (2013). *S-BPM Illustrated*. Springer.
- Galivan, M. J., & Keil, M. (2003). The user–developer communication process: a critical case study. *Information Systems Journal*, 13(1), 37-68.
- Giaglis, G. M. (2001). A taxonomy of business process modeling and information systems modeling techniques. *International Journal of Flexible Manufacturing Systems*, 13(2), 209-228.
- Hammer, M. (2003). *The agenda: What every business must do to dominate the decade*. Random House LLC.
- Hammer, M., & Champy, J. (1993). *Reengineering the Corporation: a Manifesto for Business Revolution*. New York: Harper Collins.
- Hamscher, W. (1994). AI in Business Process Reengineering. *AI Magazine*, 15(4), 71-72.
- Henry, S., & Humphrey, M. (1988). *Comparison of an Object-Oriented Programming Language to a Procedural Programming Language for Effectiveness in Program Maintenance*. Virginia Polytechnic Institute and State University.
- Iden, J. (2012). Investigating Process Management in firms with quality systems: a multi-case study. *Business Process Management Journal*, 18(1), 104-121.
- Jeston, J., & Nelis, J. (2014). *Business Process Management*. Routledge.
- Kappelman, L. A., & Gunes, C. S. (1995). End-user training & empowerment. *Journal of Systems Management*, 46(5), 36-41.
- Karahanna, E., Straub, D., & Chervany, N. (1999). Information Technology Adoption Across Time: A Cross-Sectional Comparison of Pre-Adoption and Post-Adoption Beliefs. *MIS Quarterly*, 23(2), 183-213.
- King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management*, 43(6), 740-755.
- Kline, P. (1999). *The handbook of psychology testing* (2nd ed.). London: Routledge.
- Kohlbacher, M. (2010). The effects of process orientation: a literature review. *Business Process Management Journal*, 16(1), 135-152.
- Kumar, V., Smart, P. A., Maddern, H., & Maull, R. S. (2008). Alternative perspectives on service quality and customer satisfaction: the role of BPM. *International Journal of Service Industry Management*, 176-187.
- Lee, Y., Kenneth, K. A., & Larsen, K. R. (2003). The Technology Acceptance Model: Past, Present and Future. *Communications of the Association for Information Systems*, 1, 752-780.

- Legrís, P., Ingham, J., & Colletette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information & Management*, 40(3), 191-204.
- Lindland, O. I., Sindre, G., & Solvberg, A. (1994). Understanding quality in conceptual modeling. *IEEE*, 11(2), 42-49.
- Lindsay, A., Downs, D., & Lunn, K. (2003). Business processes - attempts to find a definition. *Information and software technology*, 45(15), 1015-1019.
- Mayer, B. (1988). *Object-oriented software construction* (2 ed.). New York: Prentice hall.
- McCormack, K. (2001). Business Process Orientation: do you have it? *Quality Progress*, 34(1), 51-60.
- Melenovsky, M. J. (2005, July 25). Business Process Management's Success Hinges on Business-Led Initiatives. *G00129411*.
- Mending, J., Reijers, H. A., & Recker, J. (2010). Activity labeling in process modeling: Empirical insights and recommendations. *Information Systems*, 35(4), 467-482.
- Mending, J., Reijers, H. A., & van der Aalst, W. M. (2010). Seven process modeling guidelines (7PMG). *Information and Software Technology*, 52(2), 127-136.
- Mending, J., Reijers, H., & Cardoso, J. (2007). What makes process models understandable? *Proceedings of the 5th International Conference of BPM* (pp. 48-63). Brisbane, Australia: Springer Berlin Heidelberg.
- Mending, J., Strembeck, M., & Recker, J. (2012). Factors of process model comprehension - Findings from a series of experiments. *Decision Support Systems*, 53(1), 195-206.
- Menon, S. T. (2001). Employee Empowerment: An Integrative Psychological Approach. *Applied Psychology: An International Review*, 50(1), 153-180.
- Mutschler, B., Reichert, M., & Bumiller, J. (2007). Unleashing the Effectiveness of Process-oriented Information Systems: Problem Analysis, Critical Success Factors, Implications. *Journal of Systems, Management and Cybernetics*, 38(3), 280-291.
- Mutschler, B., Reichert, M., & Rinderle, S. (2007). Analyzing the Dynamic Cost Factors of Process-Aware Information Systems: A Model-Based Approach. *Advanced Information Systems Engineering* (pp. 589-603). Springer Berlin Heidelberg.
- Nakagawa, S. (2004). A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioral Ecology*, 15(6), 1044-1045.
- OMG, B. a. (2006). *Business Process Modeling Notation Specification: Final Adopted Specification*. Needham: Object Management Group.
- OMG.org. (2011, Januari 3). *Business Process Model and Notation (BPMN)*. Retrieved from BPMN.org: <http://www.omg.org/spec/BPMN/2.0/PDF>

- Pan, C. S., Shell, R. L., & Schleifer, L. M. (1994). Performance variability as an indicator of fatigue and boredom effects in a VDT data-entry task. *International Journal of Human-Computer Interaction*, 6, 37-45.
- Parkes, A. (2002). Critical Success Factors in Workflow Implementation. *Proceedings of the 6th Pacific Asia Conference on Information Systems*.
- Parnas, D. L. (1972). On the criteria to be used in decomposing systems into modules. *Communications of the ACM*, 15(12), 1053-1058.
- Perneger, T. V. (1998). What's wrong with Bonferoni Adjustments. *BMJ*, 316(7139), 1236-1238.
- Philips Healthcare. (2013). *Healthcare in een notendop*. Retrieved Februari 17, 2014, from Philips: <http://www.philips.nl/about/company/businesses/healthcarehighlights/index.page>
- Ravesteyn, P., & Batenburg, R. (2010). Surveying the critical success factors of BPM-systems implementation. *Business Process Management Journal*, 16(3), 492-507.
- Recker, J. (2010). Opportunities and Constraints: The Current Struggle with BPMN. *Business Process Management Journal*, 16(1), 181-202.
- Recker, J. (2012). Modeling with tools is easier, believe me"—The effects of tool functionality on modeling grammar usage beliefs. *Information Systems*, 37(3), 213-226.
- Recker, J., & Dreiling, A. (2011). The effects of Content Presentation Format and User Characteristics on Novice Developers' Understanding of Process Models. *Communications of the Association for Information Systems*, 28(6), 65-84.
- Reijers, H. A., & Mendling, J. (2011). A study into the factors that influence the understandability of business process models. *Systems, Man and Cybernetics, Part A: systems and humans, IEEE Transactions on*, 41(3), 449-462.
- Reijers, H. A., Mendling, J., & Dijkman, R. M. (2011). Human and Automatic Modularizations of Process models to enhance their comprehension. *Information systems*, 36(5), 881-897.
- Reijers, H. A., Mendling, J., & Recker, J. (2010). Business Process Quality Management. In *Handbook on Business Process Management* (pp. 167-185). Berlin: Springer Berlin Heidelberg.
- Reijers, H., & Mendling, J. (2008). Modularity in process models: Review and effects. In *Business Process Management* (pp. 20-35). Springer Berlin Heidelberg.
- Reinhardt, W., Schmidt, B., Sloep, P., & Drachsler, H. (2011). Knowledge Worker Roles and Actions - Results of Two Empirical Studies. *Knowledge and Process Management*, 18(3), 150-174.
- Rito Silva, A., & Rosemann, M. (2012). Processpedia: an ecological environment for BPM stakeholders' collaboration. *Business Process Management Journal*, 18(1), 20-42.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury park, CA: Sage.
- Saidani, O., & Nurcan, S. (2007). Towards context aware business process modeling. *8th Workshop on Business Process Modeling, Development and Support*.

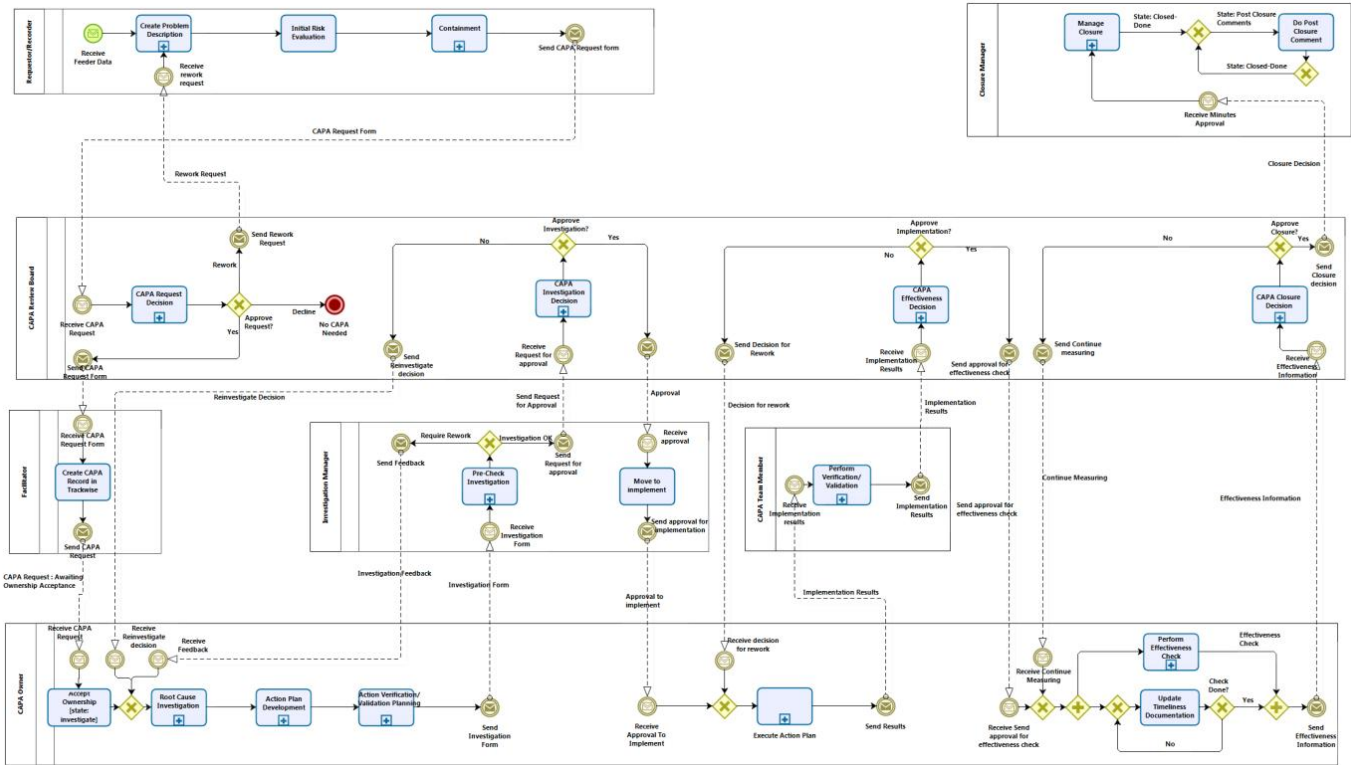
- Seddon, P., & Yip, S. (1992). An Empirical evaluation of user information satisfaction (UIS) measures for use with general ledger accounting software. *Journal of Information Systems*, 6(1), 75-92.
- Sichel, D. E. (1997). *The computer revolution: An economic perspective*. Brookings Institution Press.
- Stoddard, D. B., Jarvenpaa, S. L., & Littlejohn, M. (1996). The Reality of Business Reengineering: Pacific Bell's Centrex Provisioning Process. *California Management Review*, 38(3), 57 -76.
- Taylor, F. W. (1911). *The Principles of Scientific Management*. New York: Harper & Brothers.
- Thomas, K. W., & Velthouse, B. A. (1990). Cognitive elements of empowerment: An "interpretive" model of intrinsic task motivation. *Academy of management review*, 15(4), 666-681.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.
- Türetken, O., & Demirors, O. (2007). An approach for decentralized process modeling. *Software process dynamics and agility: proceedings of the International Conference on Software Process (ISCP 2007)* (pp. 195-207). Minneapolis, Minnesota: Berlin: Springer.
- Türetken, O., & Demirors, O. (2011). Plural: A decentralized business process modeling method. *Information & Management*(48), 235-247.
- Türetken, O., & Demirors, O. (2013). Business Process Modeling Pluralized. *S-BPM ONE 2013*, 34-51.
- van der Aalst, W. (2000). Workflow Verification: Finding control-flow errors using Petri-net-based techniques. In W. van der Aalst, & et. al., *Business Process Management* (pp. 161-183). Berlin Heidelberg: Springer Verlag.
- van der Werf, J. M., van Dongen, B. F., Hurkens, C. A., & Serebrenik, A. (2008). Process discovery using integer linear programming. In *Applications and Theory of Petri Nets* (pp. 368-387). Springer Berlin Heidelberg.
- van Dongen, B. F., De Medeiros, A. A., & Wen, L. (2009). Process mining: Overview and outlook of petri net discovery algorithms. *Transactions on Petri Nets and Other Models of Concurrency II*, 225-242.
- Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: four longitudinal field studies. *Management Science*, 46(2), 186-204.
- Wolf, C., & Harmon, P. (2006). *The State of Business Process Management*. Business Process Trends.
- Wolf, C., & Harmon, P. (2012). *The State of Business Process Management*. [www.bptrends.com](http://www.bptrends.com).
- Yan, Z., Reijers, H., & Dijkman, R. (2010). An Evaluation of BPMN Modeling Tools. *Proceedings of the Second International Workshop BPMN 2010* (pp. 121-128). Potsdam, Germany: Springer-Verlag Berlin Heidelberg.
- Yourdon, E. (1989). *Modern Structured Analysis*. Englewood Cliffs, New Jersey: Prentice Hall.
- zur Muehlen, M., & Ho, D. T.-Y. (2006). Risk Management in the BPM Lifecycle. *Business Process Management workshops* (pp. 454-466). Springer Berlin Heidelberg.

# Appendices

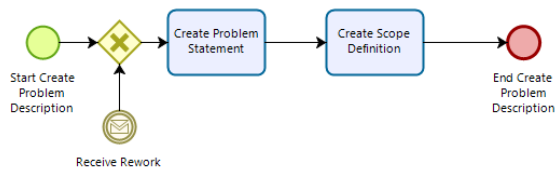
## Appendix A: the CAPA process

### A1: CAPA classical

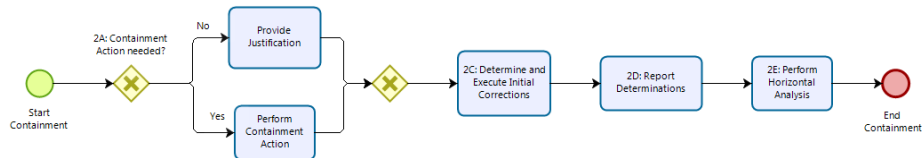
#### Main process



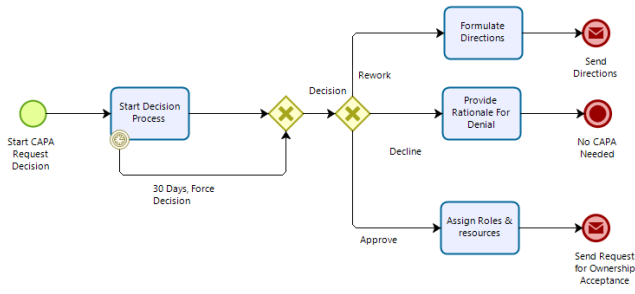
Create problem description:



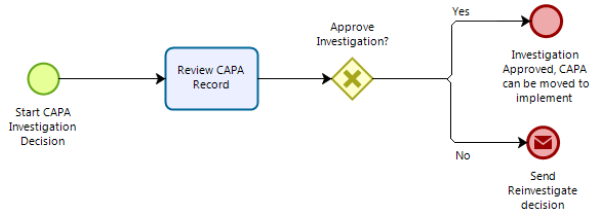
Containment



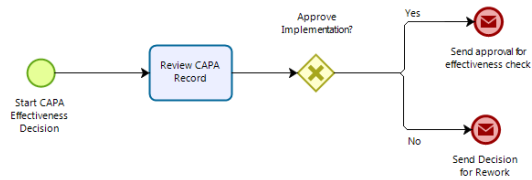
CAPA Request Decision



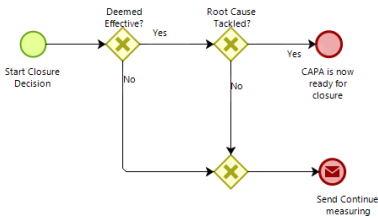
### CAPA Investigation Decision



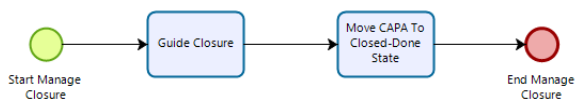
### CAPA Effectiveness Decision



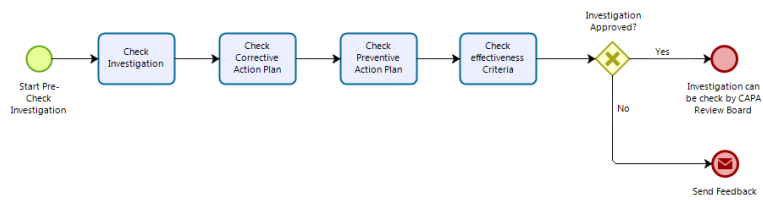
### CAPA Closure Decision



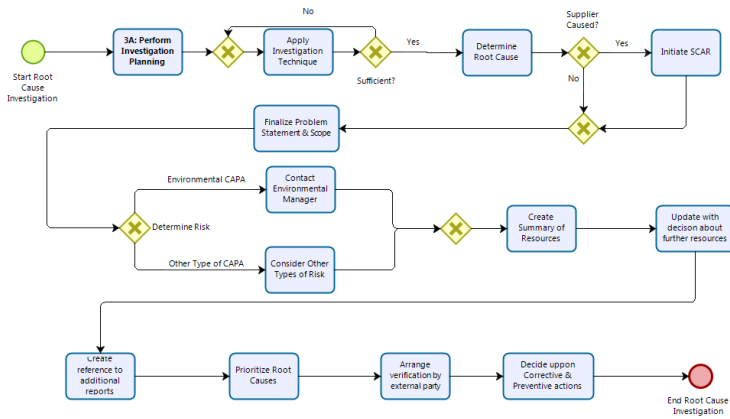
### Manage Closure



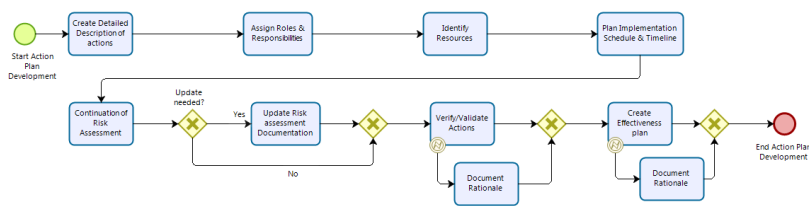
### Pre-Check investigation



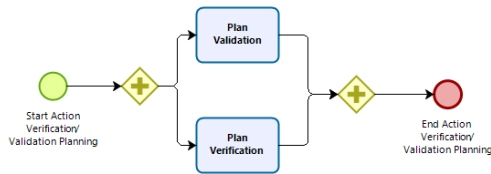
### Root Cause Investigation



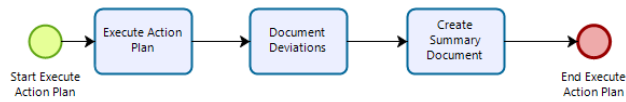
### Action Plan Development



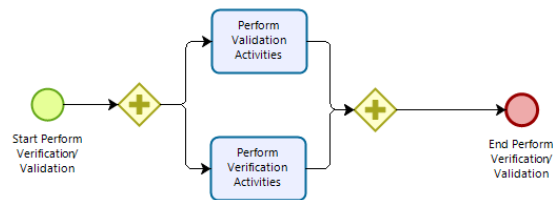
### Action Verification/Validation planning



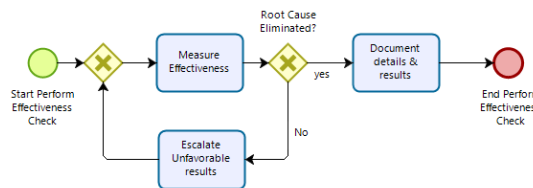
### Execute action plan



### Perform verification validation



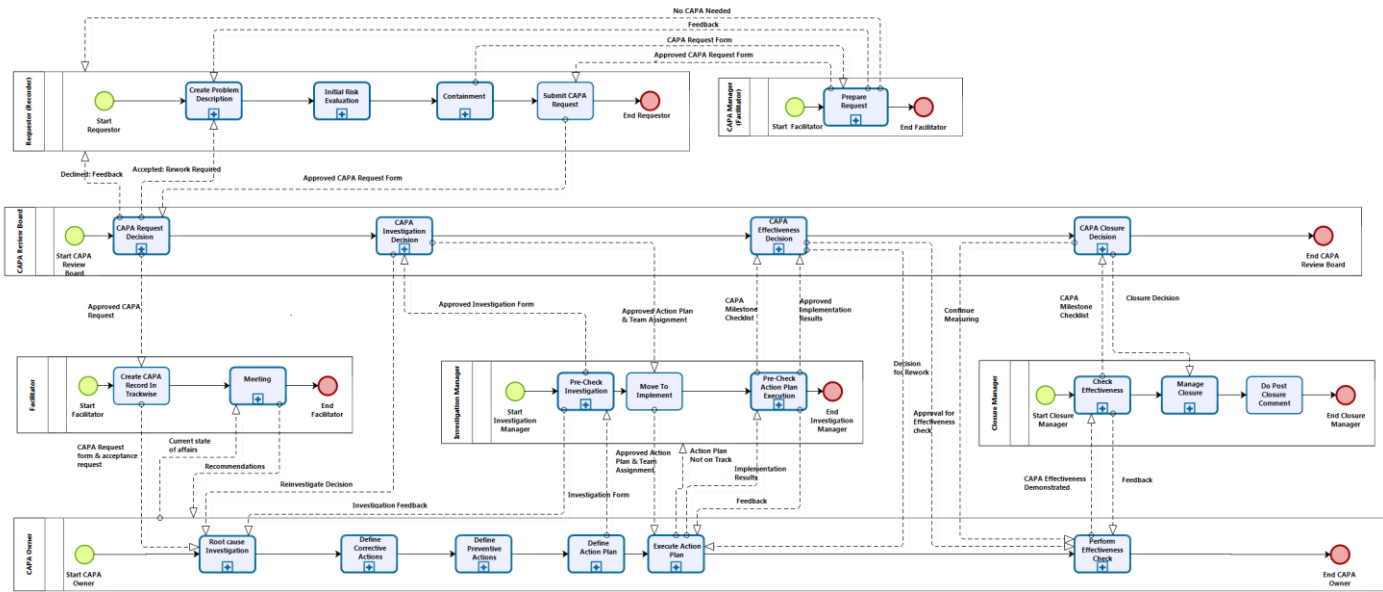
### Perform effectiveness check





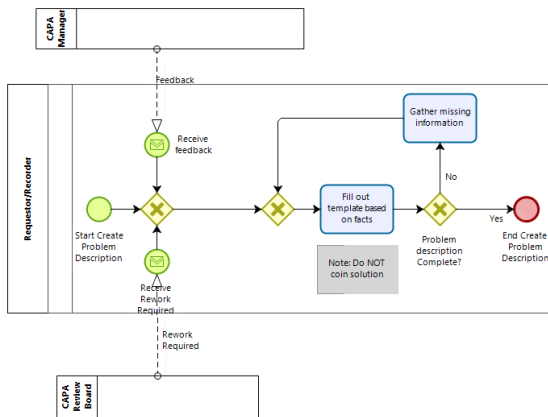
## A2: CAPA Plural

Main process:

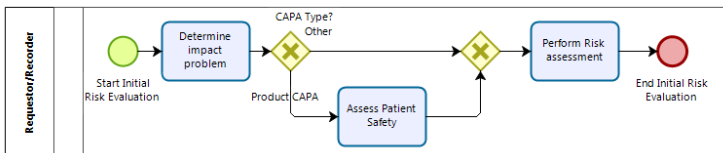


Operations:

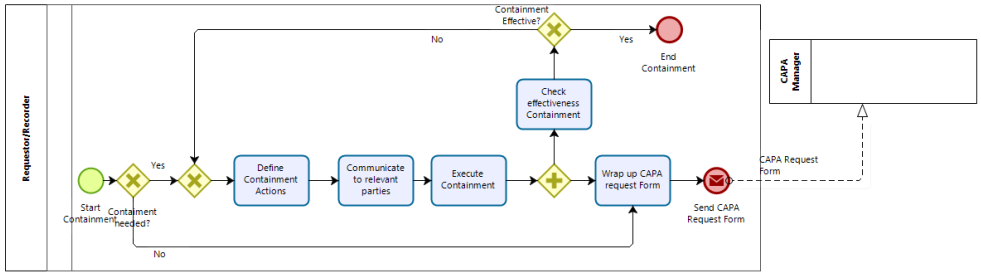
Create problem description:



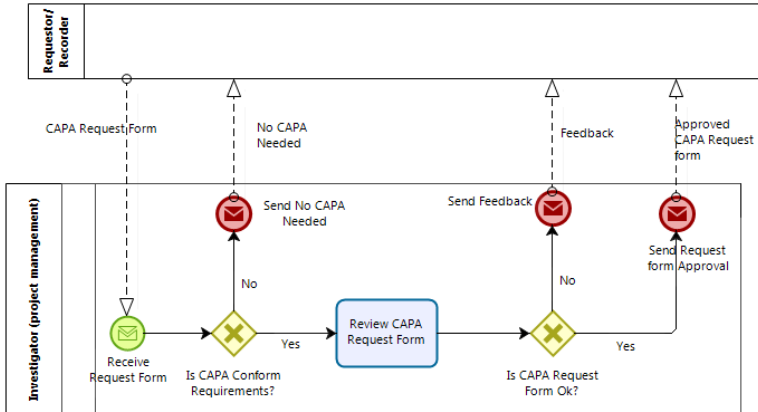
Initial Risk evaluation



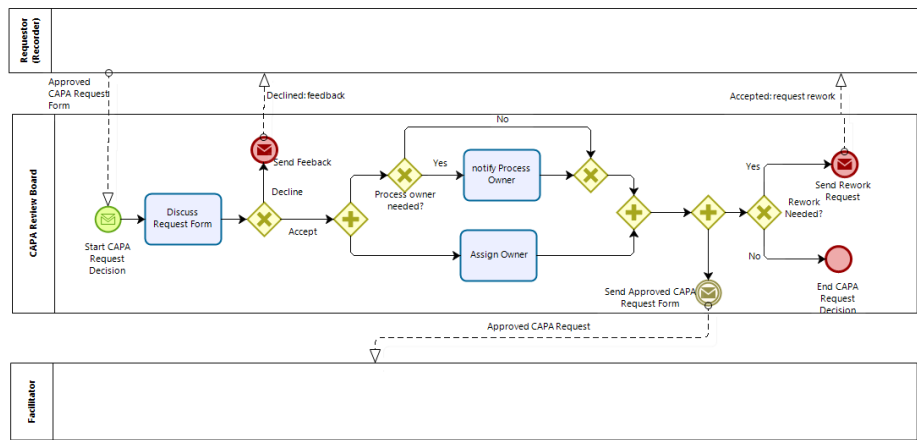
Containment



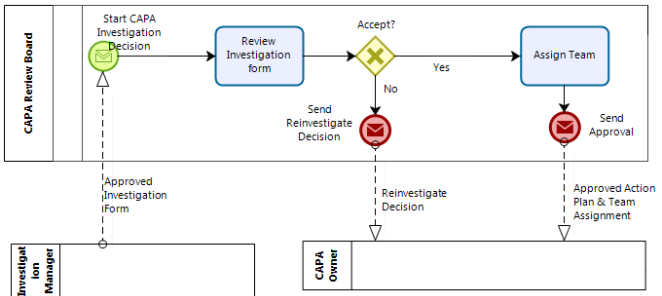
**Prepare Request**



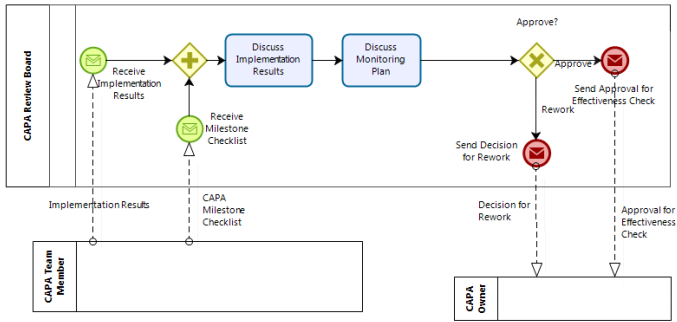
**CAPA Request Decision**



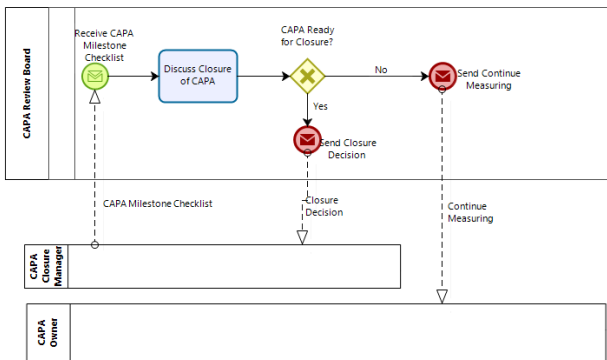
**CAPA Investigation Decision**



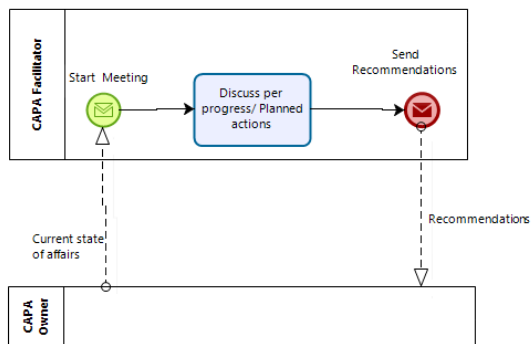
**CAPA Effectiveness Decision**



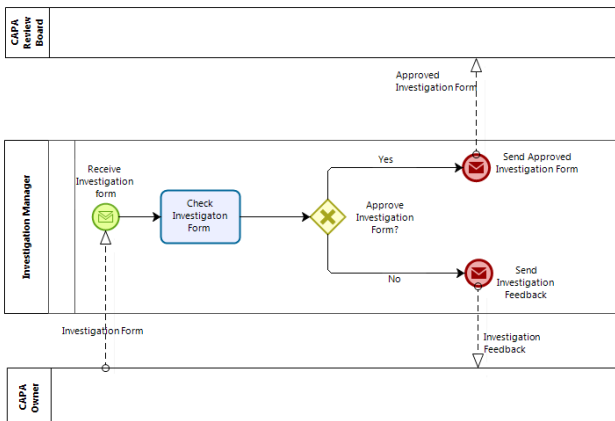
### CAPA Closure Decision



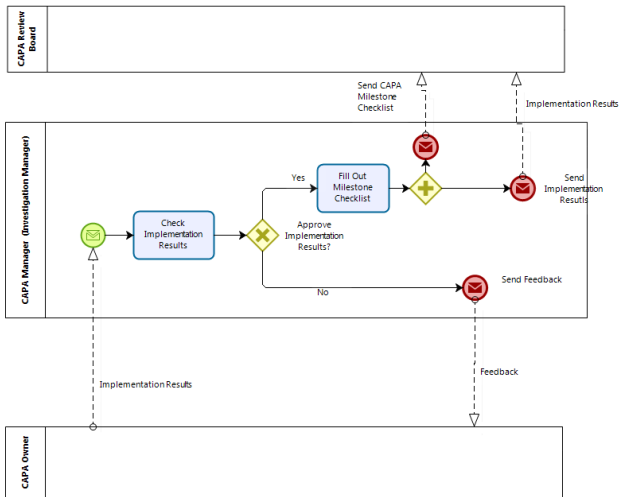
### Meeting



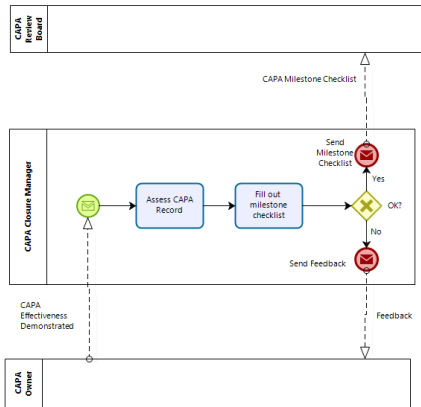
### Pre-Check Investigation



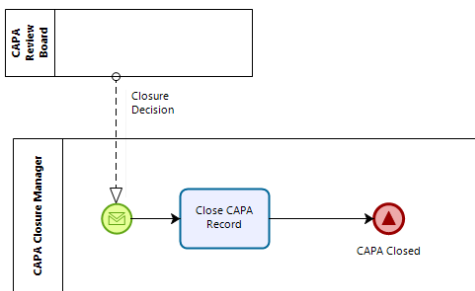
### Pre-Check Action Plan Execution



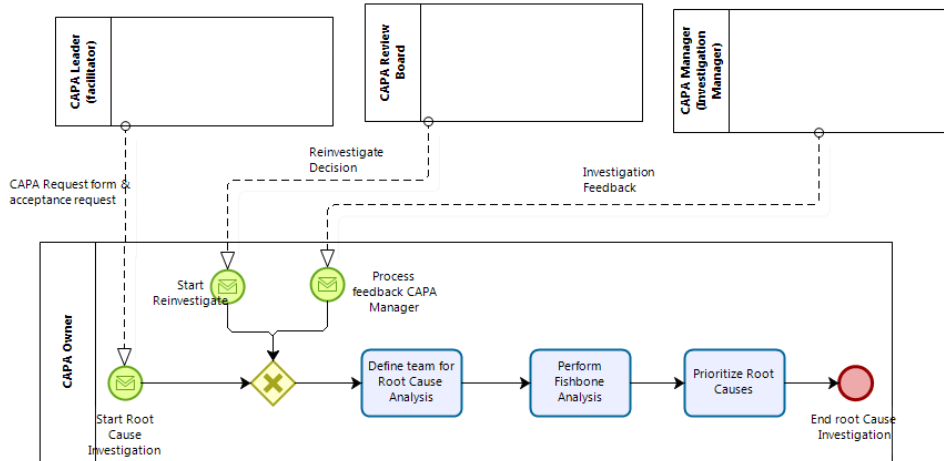
### Check Effectiveness



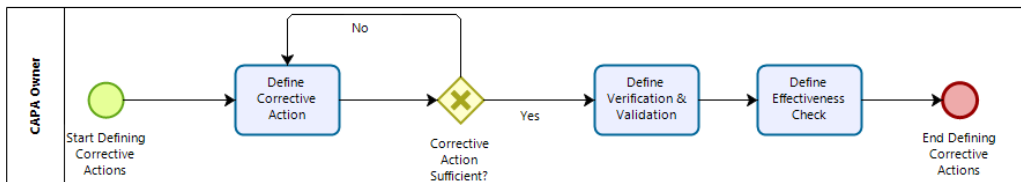
### Manage Closure



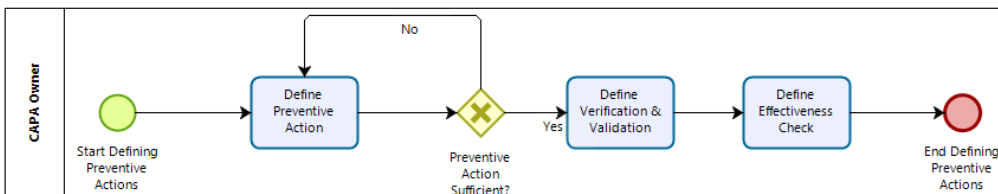
### Root Cause Investigation



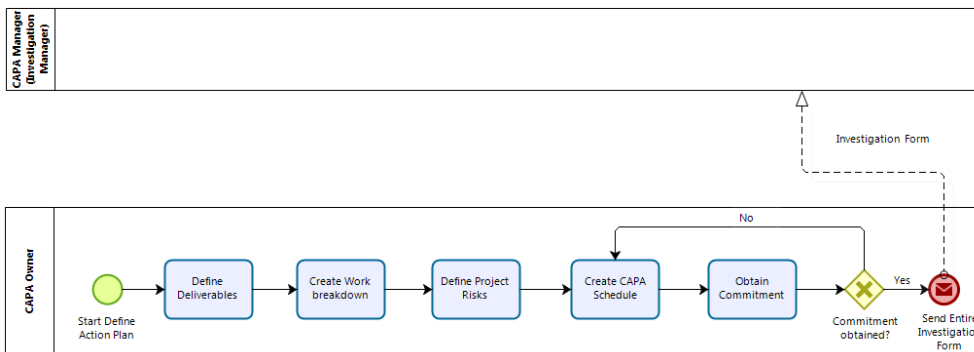
### Define corrective actions



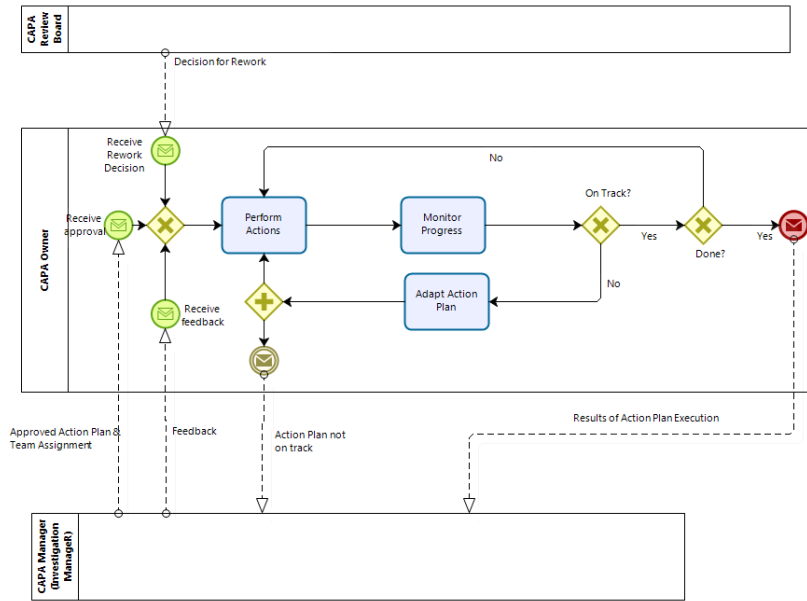
### Define Preventive actions



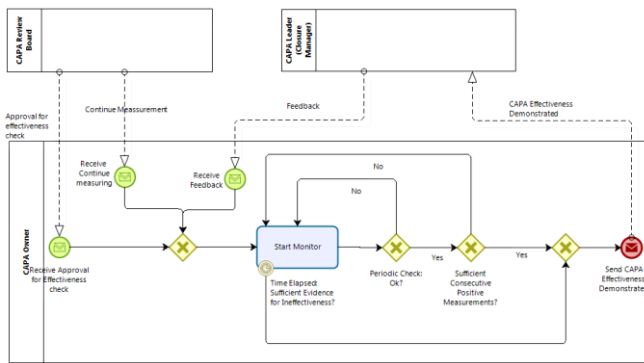
### Define Action Plan



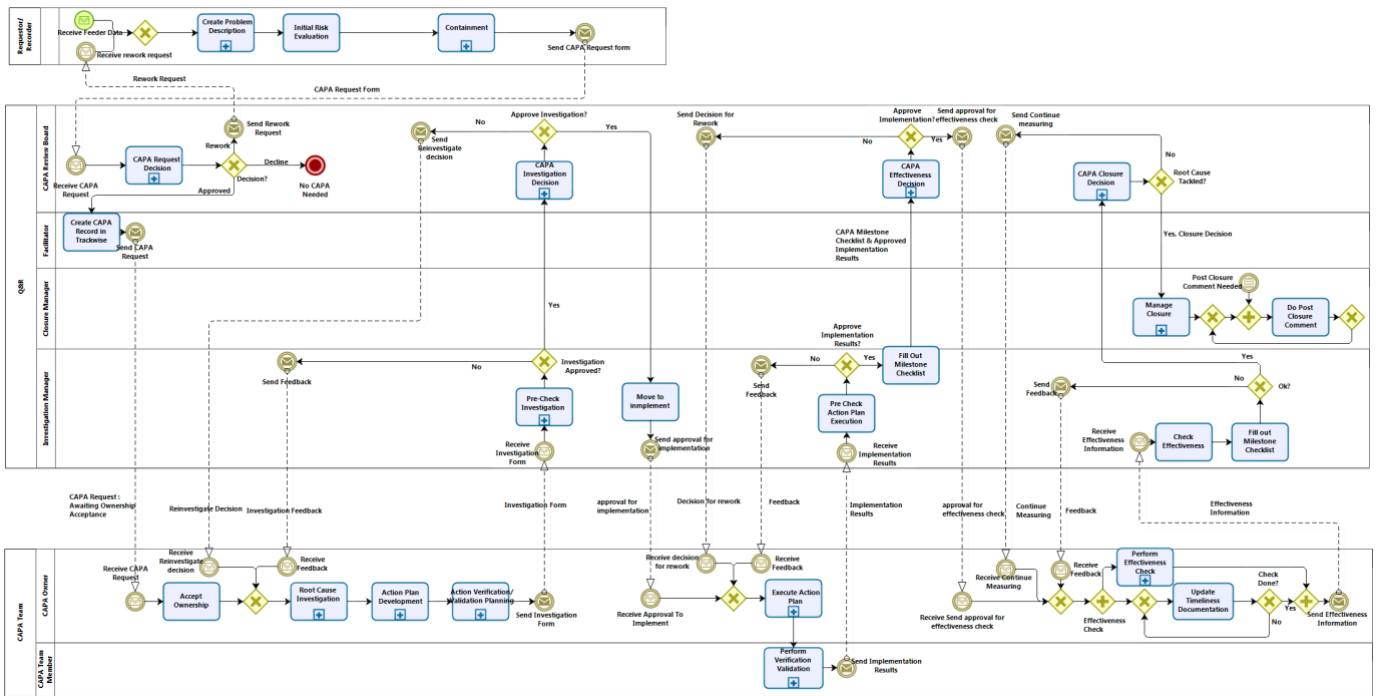
### Execute action Plan



### Perform Effectiveness check



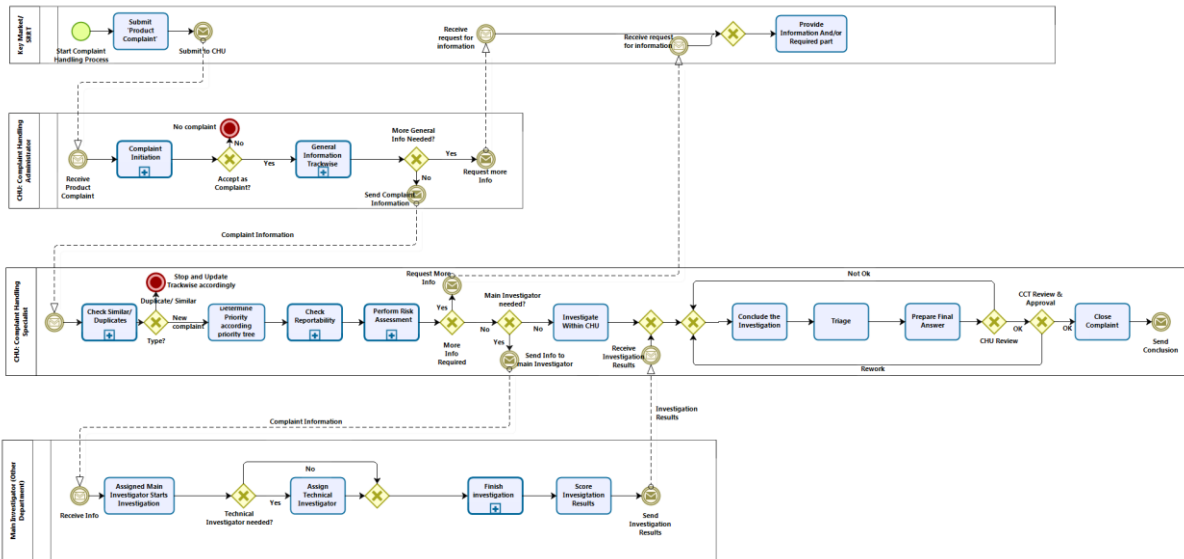
### A3: CAPA Classical adapted for experiment



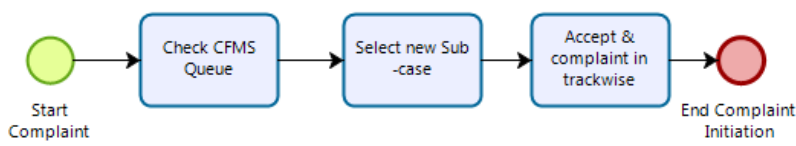
### Appendix B: the Complaint Handling process

#### B1: Complaint Handling Classical

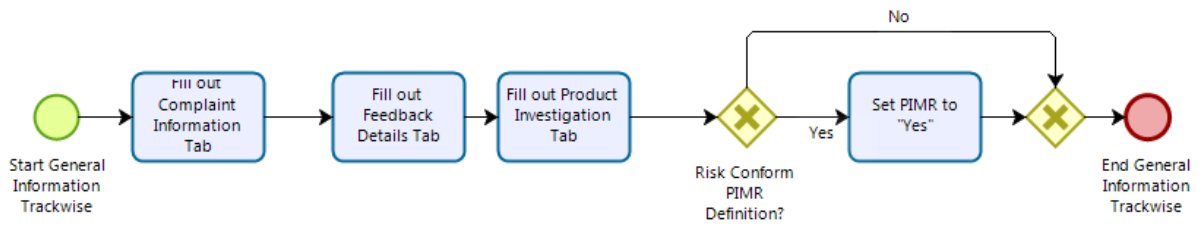
Main process:



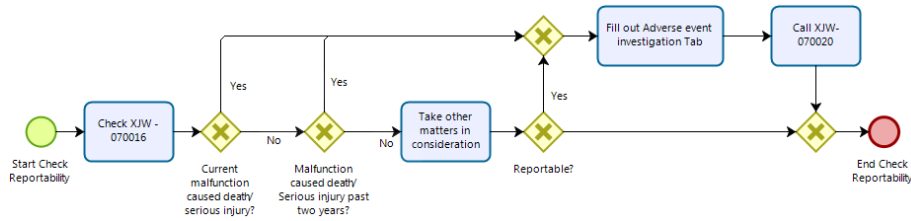
Complaint Initiation:



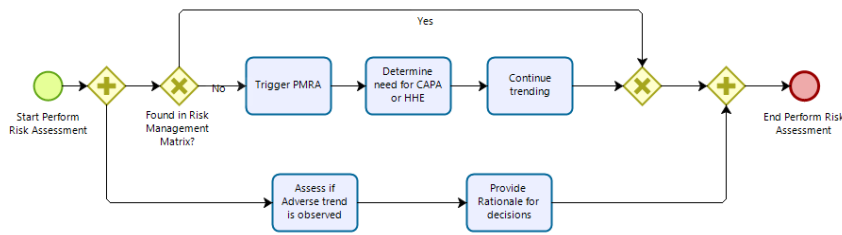
Fill out general information Trackwise:



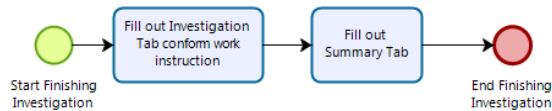
Check reportability



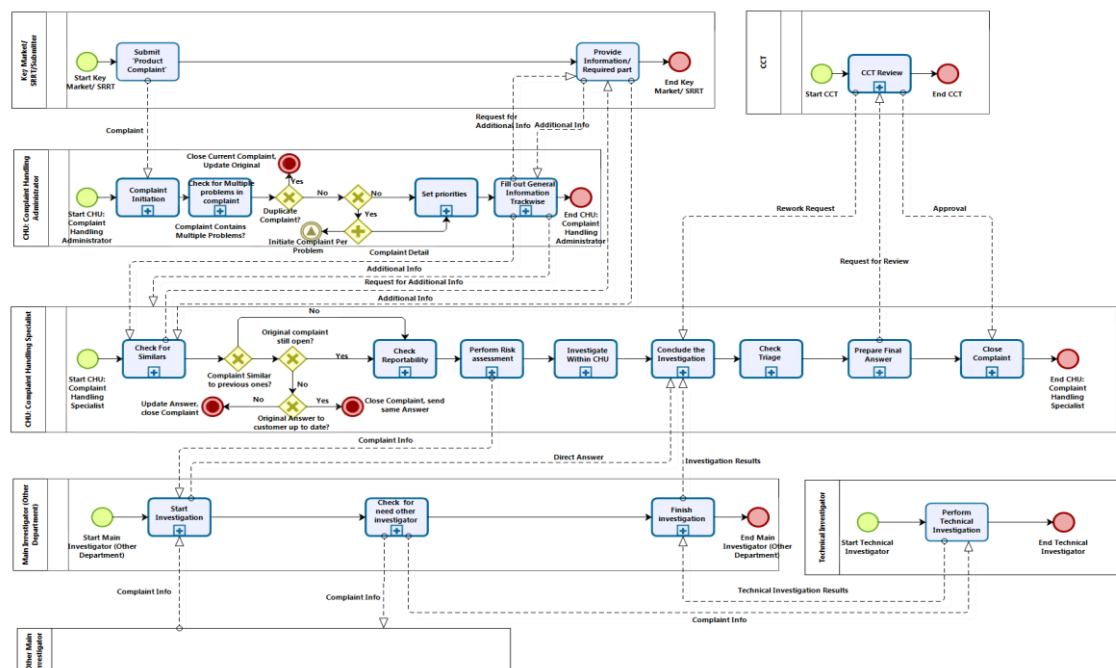
Perform risk assessment



Finish investigation

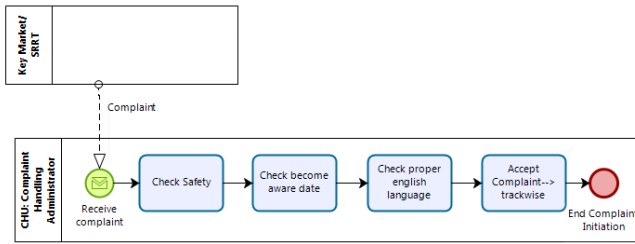


## B2: Complaint Handling Plural

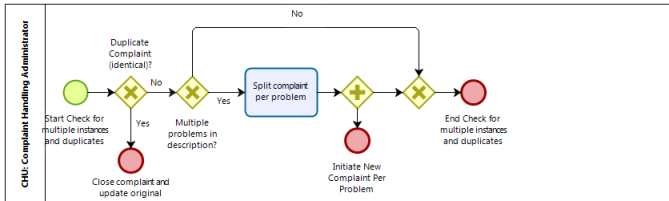




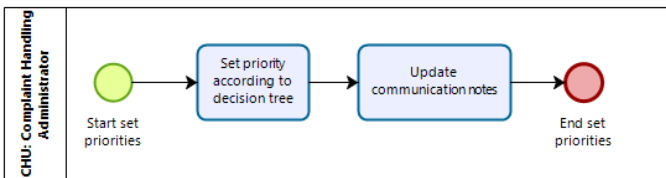
## Complaint Initiation



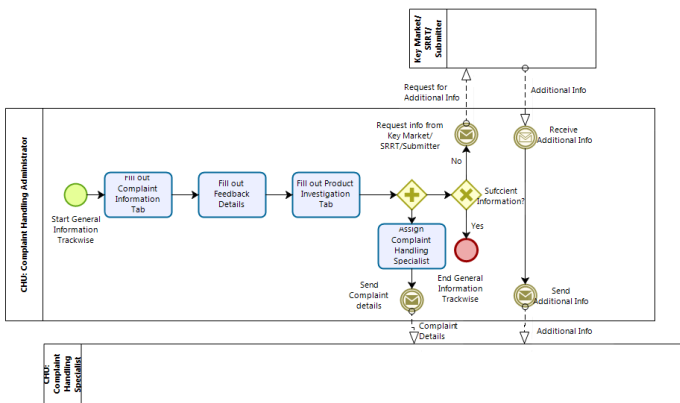
## Check for multiple problems in complaint



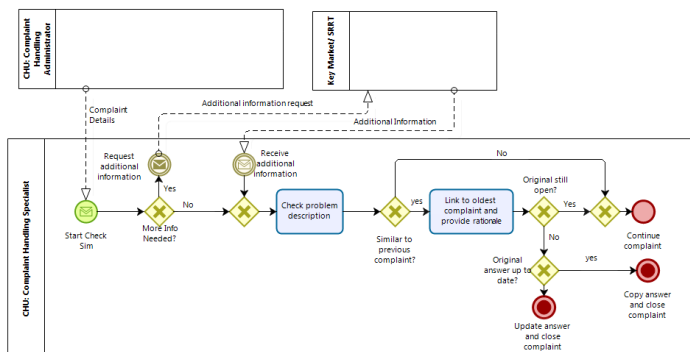
## Set priorities



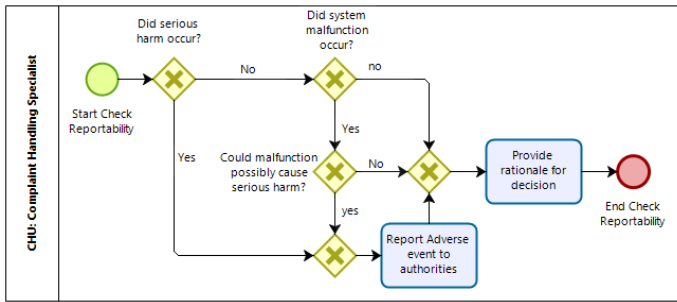
## Fill out General information Trackwise



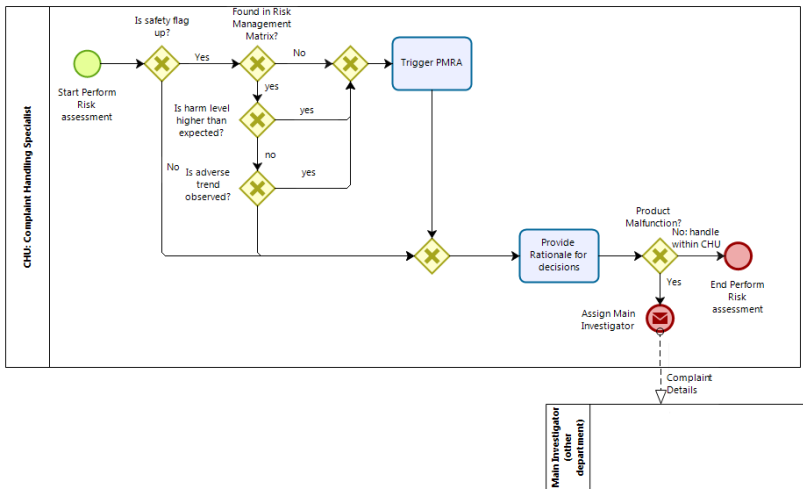
## Check for similars



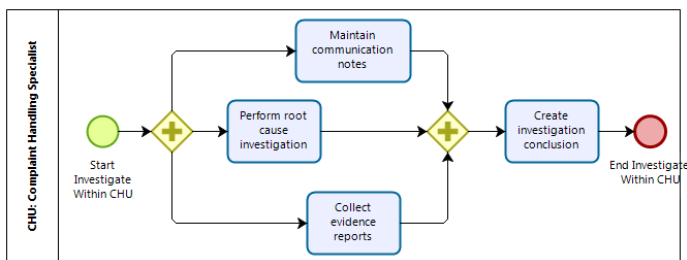
## Check reportability



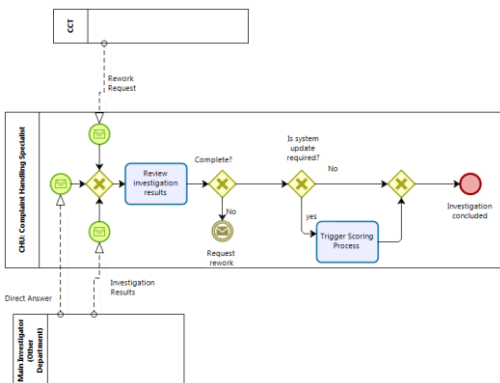
## Perform Risk Assessment



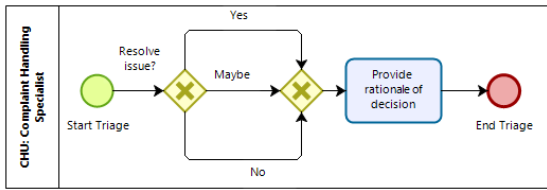
## Investigate within CHU



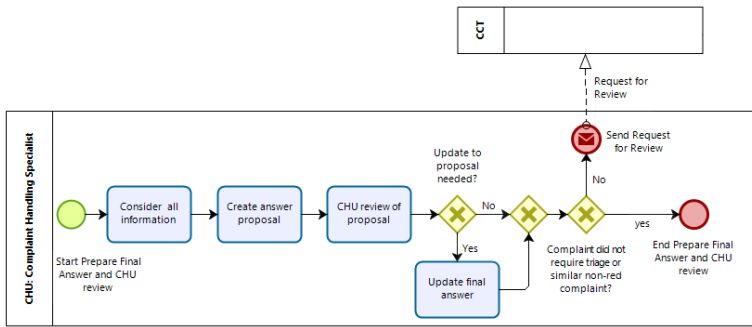
## Conclude the investigation



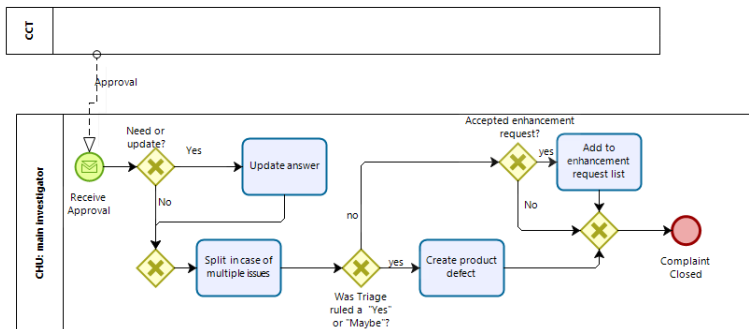
### Check Triage



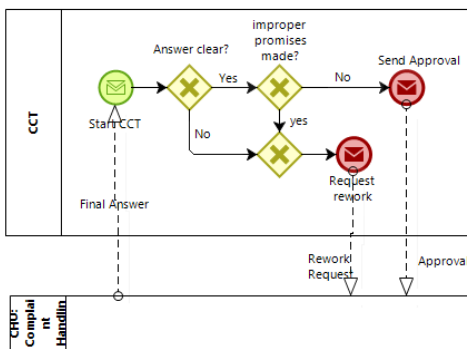
### Prepare final answer



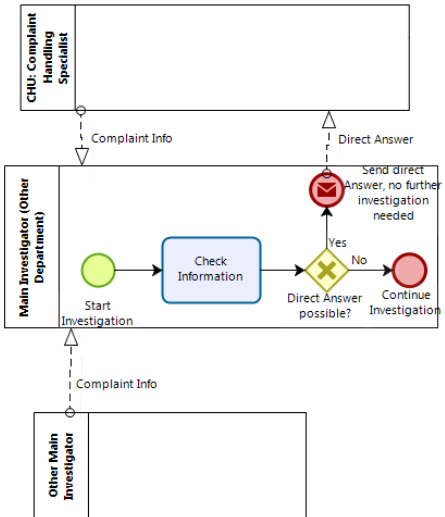
### Close Complaint



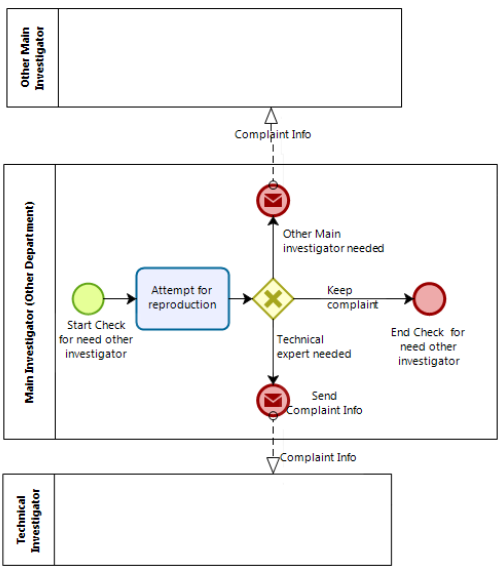
### CCT Review



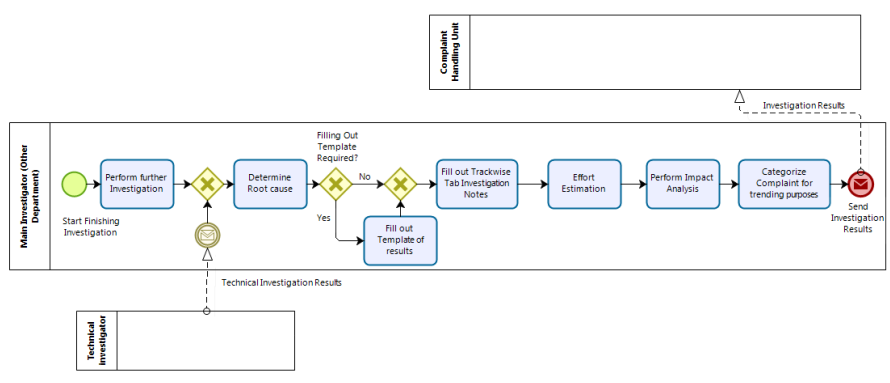
### Start investigation



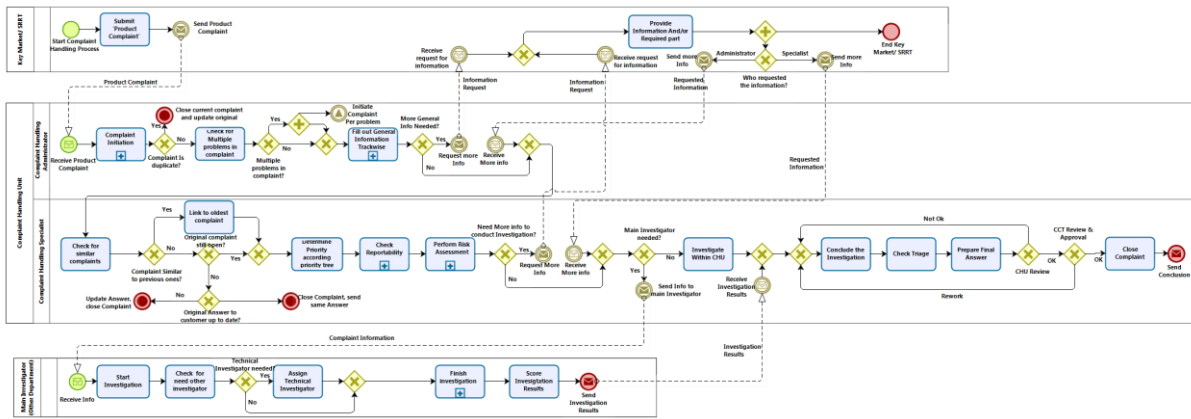
Check for need other investigator



Finish investigation



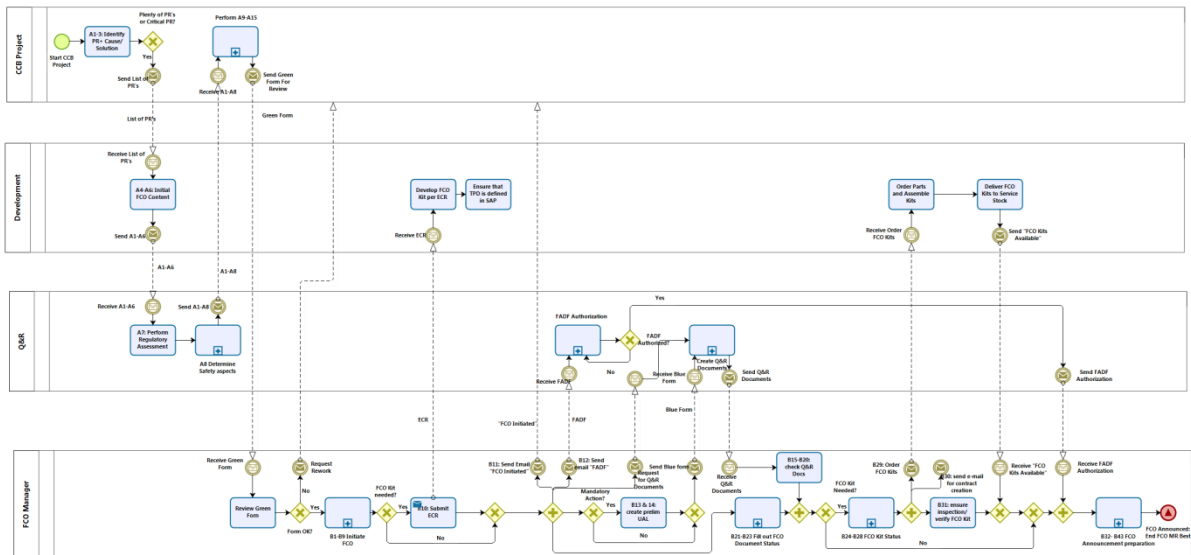
### B3: Complaint Handling Classical adapted for experiment



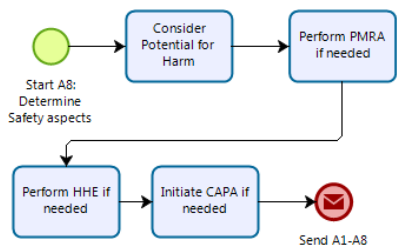
### Appendix C: the FCO process

#### C1: FCO Classical

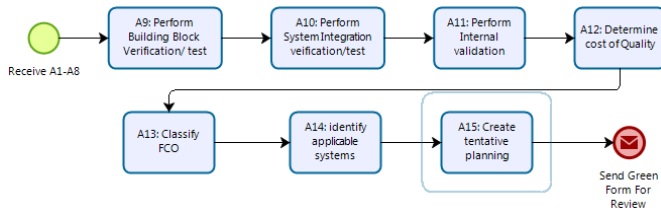
Main process:



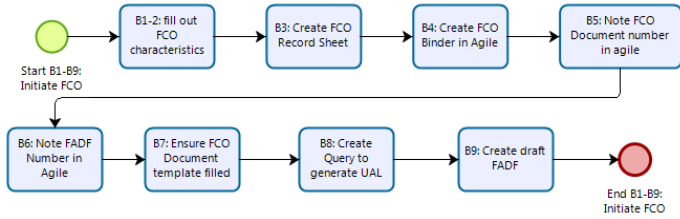
#### A8: determine safety aspects



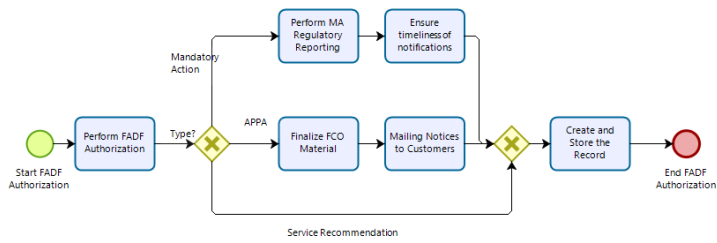
#### Perform A9-A15



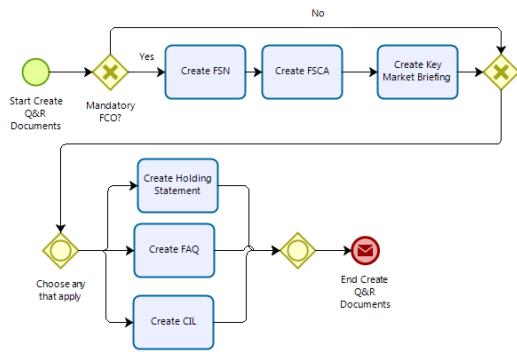
### B1-B9: initiate FCO



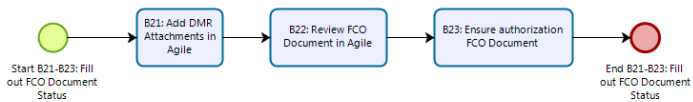
### FADF Authorization



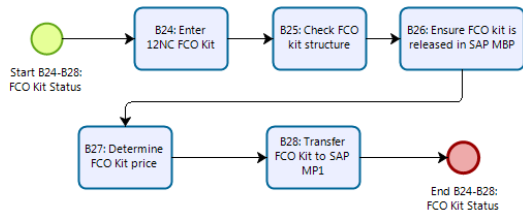
### Create Q&R Documents



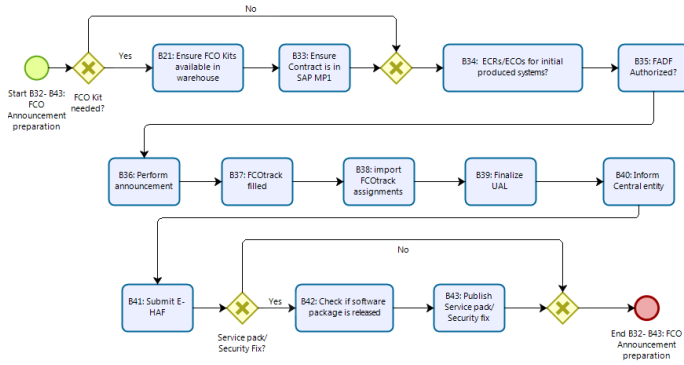
### B21-B23: Fill out FCO Document Status



### B24-B28: FCO kit status

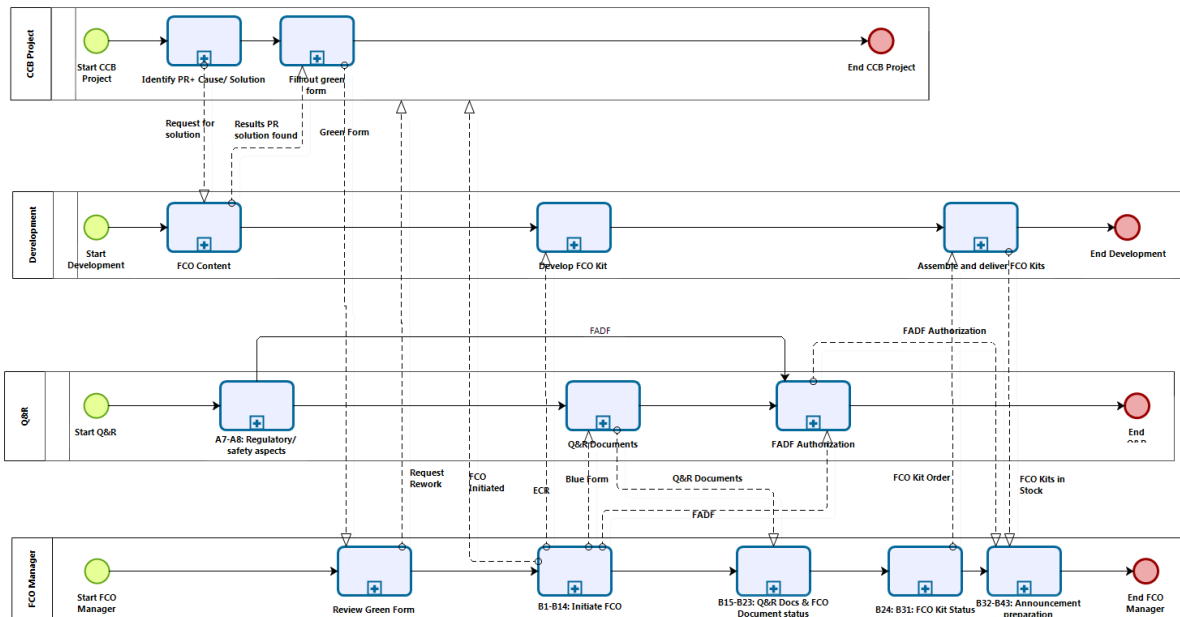


### B32-B43: FCO Announcement preparation

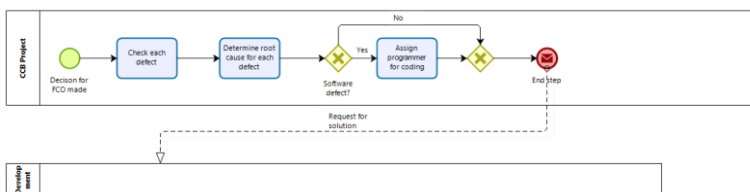


### C2: FCO Plural (incomplete)

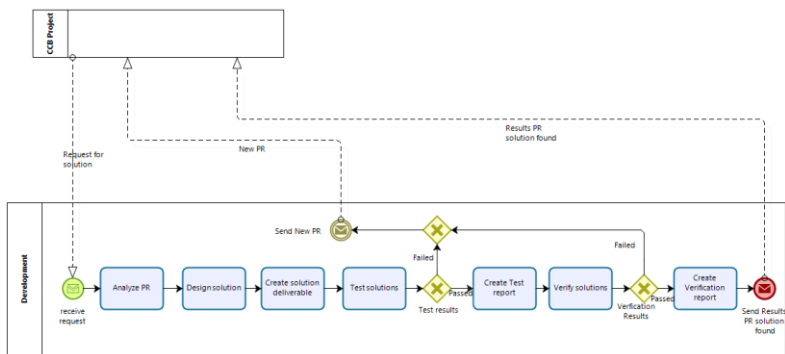
Main process:



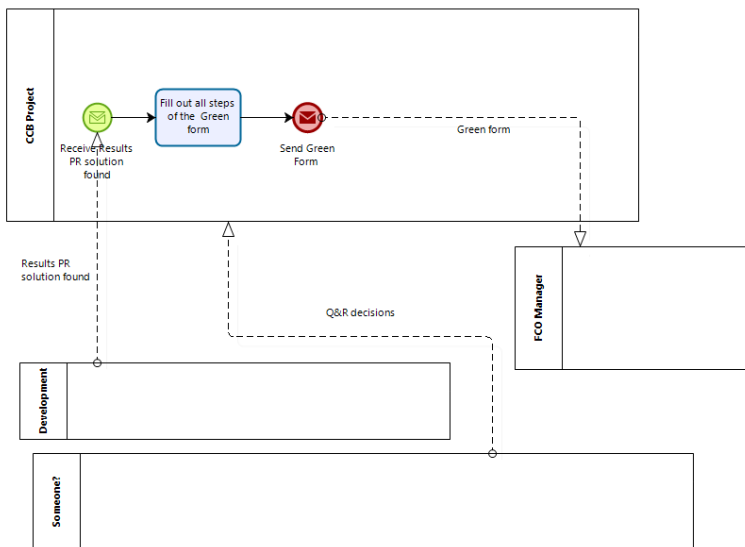
### Identify PR + Create solutions



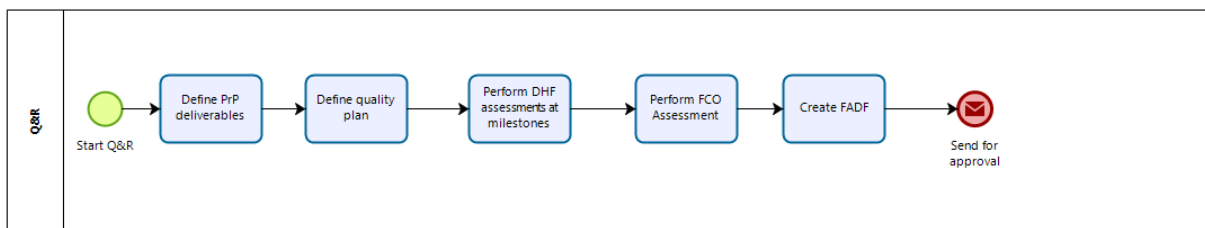
### FCO Content



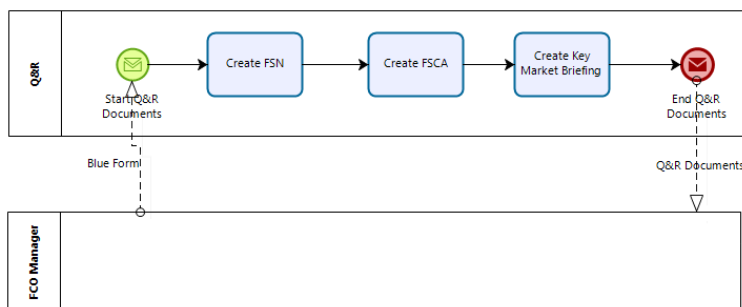
Fill out green form



A7-A8: Regulatory/Safety aspects

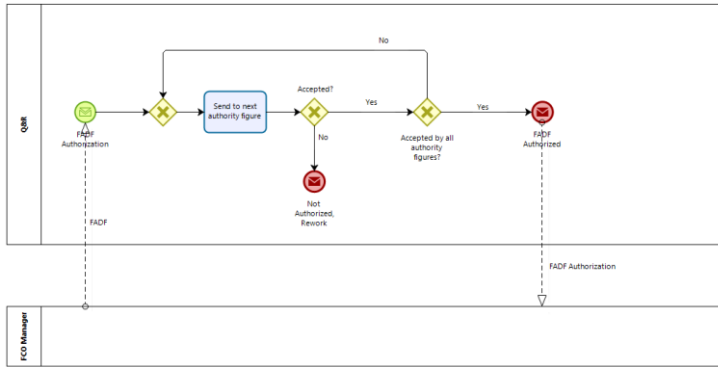


Q&R Documents



FADF Authorization

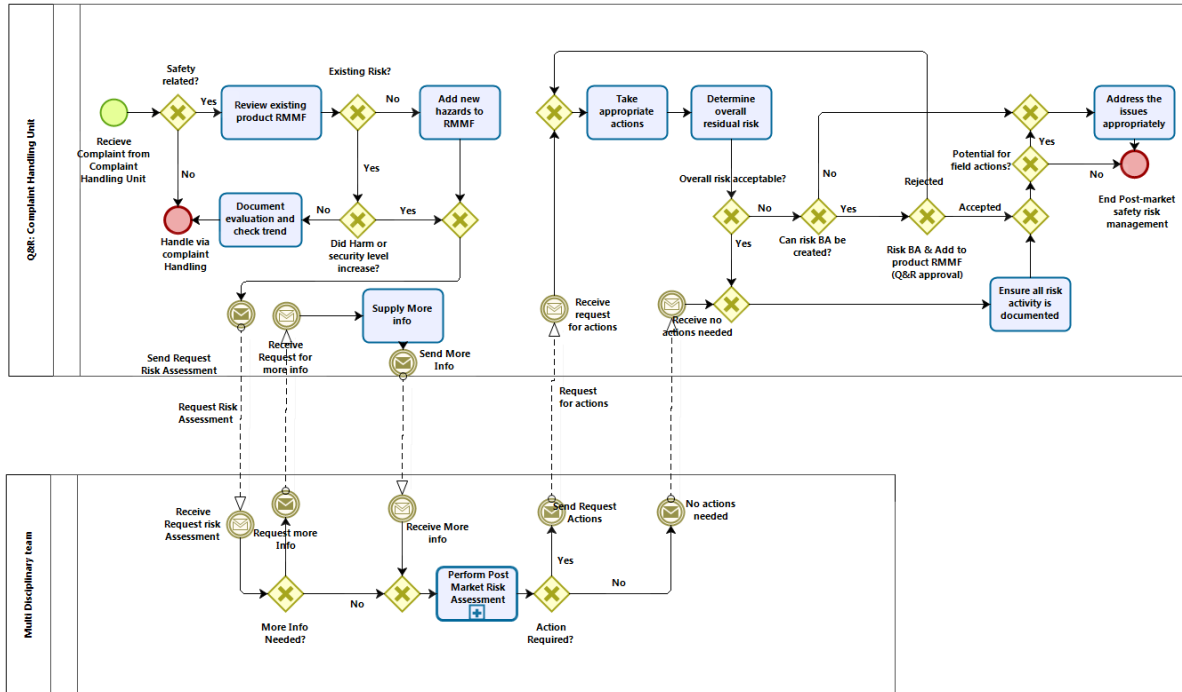




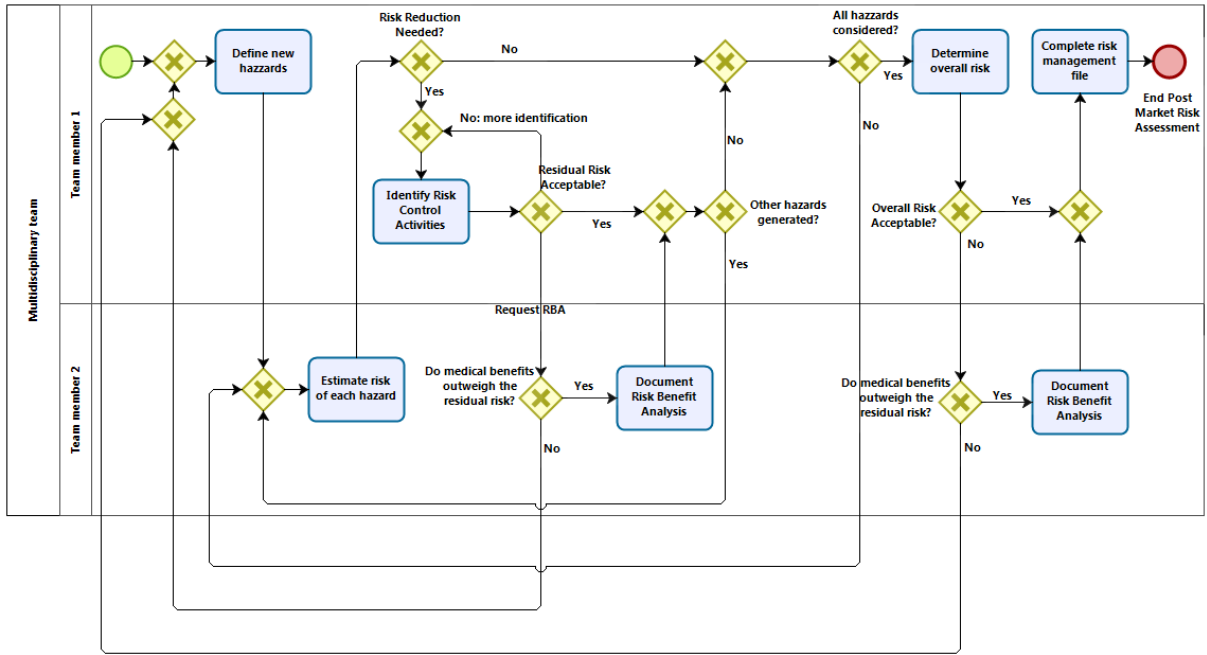
## Appendix D: Post Market Risk Management

### D1: Risk Management Classical

Main process

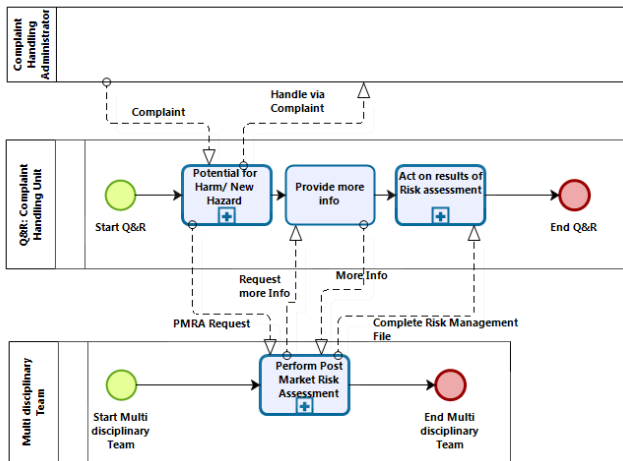


Perform Post Market Risk Assessment

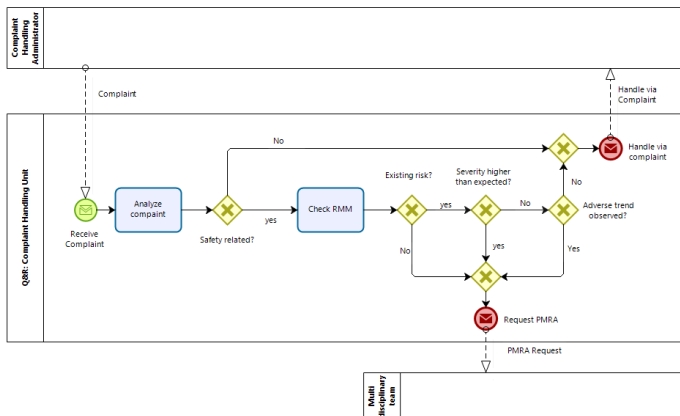


## D2: Risk Management Plural

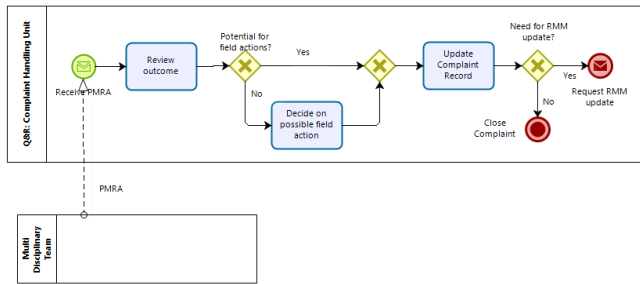
Main process:



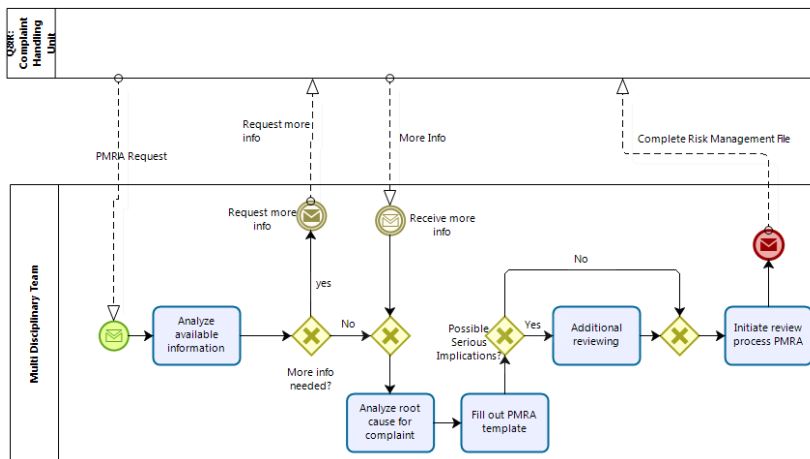
## Potential for Harm/ New Hazard



## Act on Results of Risk assessment



## Perform Post Market Risk Assessment



## Appendix E: Questions and Questionnaire

### Appendix E1: constructed questions

Three processes, 8 questions each: 2 about control flow, 2 about data flow, 2 about organizational aspects, 2 about maintenance

#### CAPA

- Control flow
  - o Is it possible to already start executing the action plan before the CAPA Review Board has approved it? (answer: No)
  - o Can the CAPA Owner start to perform the effectiveness check directly after CAPA Review Board has approved the Effectiveness Decision? (Answer: Yes)
- Data Flow
  - o Is the closure decision the only input needed for the "Manage closure" task? (answer: yes)
  - o Do the requestor/recorder and the CAPA Owner ever communicate directly? (answer: no)
- Organizational aspects
  - o Can the CAPA Owner still perform actions after the CAPA has been closed? (answer: no)
  - o Does the CAPA Review Board ever move a CAPA to implement? (answer: no)
- Maintenance

- Suppose that after an FDA inspection, it was decided that from now on all potential root causes should be considered. Will the removal of the step “Prioritize Root Causes” be sufficient to incorporate the change? (Answer: yes)
- Suppose that the manager of this process wants to improve the throughput time. This should be achieved by skipping the steps “Pre-Check Investigation” and “Move to Implement”. Would this change the deliverables which are required from the CAPA Owner? (answer: no)

### Complaint Handling

- Control flow
  - Is a rationale for the made decisions always provided when performing risk assessment? (answer: yes)
  - Is the Complaint ever split up into multiple complaints? (answer: yes)
- Data flow
  - Can the task 'conclude the investigation' be performed without receiving the investigation results from the Main Investigator? (answer: yes)
  - Is it possible that the key market gets more than two information requests? (answer: no)
- Organizational aspects
  - Can anyone other than the Main Investigator assign a technical investigator? (answer: no)
  - Does anyone other than the Complaint Handling Administrator check for duplicate complaints? (Answer: no)
- Maintenance
  - Suppose the business unit wants to improve the satisfaction of its key customers by communicating more often about what is being done about their complaint. This is done for high priority customers by sending an update to the Key Market/SRRT whenever the complaint investigation starts and ends. Will this update require more than one role to update their behavior? (Answer: yes)
  - Suppose that an FDA inspector reported that insufficient rationale is provided when a complaint is classified as ‘Similar’. From now on, every time a complaint is classified as ‘Similar’, evidence should be provided as to why that’s the case. Does this affect the way roles other than the Complaint Handling Specialist perform their tasks? (answer: no)

### Appendix E2: TAM (and related) questions

- Intention to use
  - Assuming I have access to this type of process models, I intend to use them.
  - Given that I have access to this type of process models, I predict that I would use them.
- Perceived usefulness
  - Using this type of process models improves my performance in my job.
  - Using this type of process models in my job increases my productivity.
  - Using this type of process models enhances my effectiveness in my job.
- Perceived Ease of Use



- Interacting with this type of process models does not require a lot of my mental effort.
- I find this type of process models to be easy to use.
- I find it easy to get this type of process models to do what I want to do.
- Information Retrieval
  - I believe this type of process models adequately meets the information I was asked to provide.
  - This type of process models is efficient for providing the information I need.
  - This type of process models is effective for providing the information I would need.

### Appendix E3: other questions

- Process modeling experience
  - How often do you encounter process models in your practice? (Never/Less than once a month/more than once a month/Daily)
  - When did you first encounter a process model in your practice? (Never encountered a process model/Less than once a month/less than a year ago/less than three years ago/more than three years ago)
- Descriptive questions
  - What is your age? (Younger than 20 years old/between 20 and 29 years old/between 30 and 39 years old/between 40 and 49 years old/between 50 and 59 years old/60 or older)
  - What is your gender? (Male/Female)
  - What is your highest form of completed education? (High school/MBO/Bachelor's degree (or HBO)/Master degree/PhD/Other, please specify)
  - What is your functional background? (Software engineering/Mechanical engineering/electrical engineering/systems engineering/manufacturing/logistics/Sales/Marketing/Finance/Other, Please specify)

### Appendix E4: Relevant screenshots of questionnaire

Start page:

**Master's Thesis Questionnaire Processes**

Dear colleague,

thank you very much for your participation in this study!

My research is about several processes within Philips and this questionnaire will be used to gather data. Together with several Philips colleagues, I have made models of these processes. With your help, we will try to determine the quality of these models.

In the invitation you received from me, a username and a password were given. You can enter this information in the fields below and then start the questionnaire. This is to facilitate online security for us all.

Username

Password

*Please note that you will remain anonymous at all times and no individual results will be reported to management, only averages. The results of this questionnaire will not be linked to your identity. The questionnaire will take about 30-40 minutes to complete.*

Please click "Next" to continue the study.

Coffee Explanation:

#### ☒ Cup of coffee?

It is rather important for this study to finish the questionnaire in one go. There is absolutely no problem with you taking your time to answer a question, but taking a coffee break might influence my results. Therefore, if you'd like to have a cup of coffee this would be the perfect time to do so.



*(disclaimer: individual experiences may differ. No guarantees are given that coffee served from the machines at PH MR resembles this picture)*

Please click "Next" to continue the study.

## Intro to process models

#### ☒ The process models

The questions in this questionnaire will be accompanied by a process model. A process model is a picture of a process to keep a certain overview. This is somewhat similar to a flowchart, but uses a richer notation. On each page you'll see a question and a process model below it, so you don't have to memorize it. The information needed to answer the question is always present in the process model.

If you have domain knowledge of the process, please try to answer the questions based on the model rather than your experience. The process model may not display all exceptions to the procedure like you expect them to.

The next page will show you a short legend about the process models to help understand it.

Please click "Next" to continue the questionnaire.

## BPMN tutorial

Process Modeling Legend

A process model is a picture of a process. It is very similar to a flowchart and can contain any of the following blocks:

Name	Description	Looks
Role	This long, black rectangle indicates a role within a process. All activities inside this role are considered his/her responsibilities.	
Activity	This blue square indicates some task is being performed. The description or activity of the task is mentioned on its label.	
Sub-Process	It's often useful to group some activities together and put them in a sub-process. This can be recognized by the '+' symbol on the blue square. To view the contents of a sub-process, you can simply place your cursor over the image. (You can try this out now).	
Sequence	This straight arrow indicates an order of event. The source of the arrow happened before the activity it points to can happen.	
Exclusive Choice	If, at some point, a choice is made, this yellow diamond with the 'X' is used. It means that either one of the outgoing arrows will be chosen. This element can also be used to merge paths. In that case, <u>only one of the incoming arcs has to be active.</u>	
Parallel activities	This element is used to indicate the following activities can happen at the same time. This allows for parallel activities. This element can also be used to merge paths. In that case, <u>all incoming arcs have to be active.</u>	
Sending information	If an activity results in sending information, this element is used. It means that a message is being sent.	
Receiving information	When someone receives information, this yellow circle is being used. The difference with the previous element is that with a receiving circle the 'envelope' is white rather than black.	
Message Flow	When a message is sent from one role to another, a message flow is used. This is a dashed line with a label.	

If, at any point in time, you feel the need to review this legend, you can hover your mouse over the Question mark icon to get this legend (you can test this on the top of the page).

Please click "Next" to continue the questionnaire.

What to expect

What to expect

You'll first receive a short explanation the way a process functions. You'll be shown the process model below it and you can have a look through it. Just hover your mouse over the different blocks to get a feeling what's going on in there. You don't have to memorize the model, you'll get a copy for each question. **So please don't spend too much time on the initial description.**

You'll be shown two process models, made through two different conventions. For each of these models 8 "Yes/No"-questions will be asked. After each set of 8 questions, you can indicate whether you thought this was difficult to do and your opinion about the model can be given.

Please click "Next" to continue the questionnaire.

Scrolling clarification

↳ The process models are big

Due to the size of the process models, you'll sometimes need to scroll across the image to navigate. You can use the arrow keys of your keyboard as well if you prefer to do so.

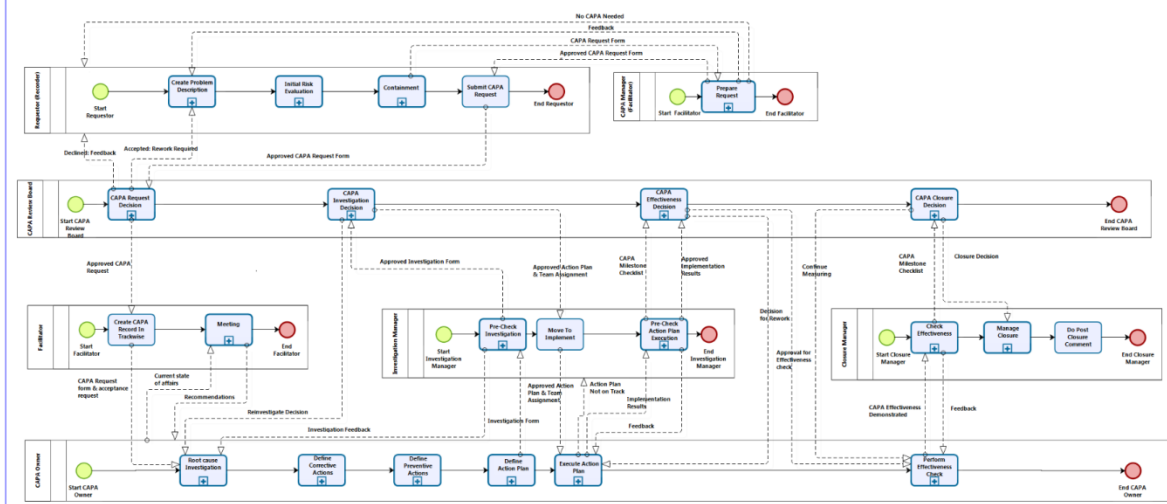


Please click "Next" to continue the questionnaire.

Example of introduction to a process:

↳ The CAPA Process

CAPA stands for Corrective Action & Preventive Action. It is an important process which can be triggered when certain Key Performance Indicators (KPI's) are over their allowed limits for some time. In the CAPA process, a dedicated CAPA owner will contain the current problem, perform a root cause analysis and ensure the problems do not occur anymore in the future. At four controlled points, a CAPA Review Board will give a formal go/no-go decision. Various other roles provide checks and feedback to improve the quality of the investigation. Please take a closer look at the process and then click "Next" to continue the process.

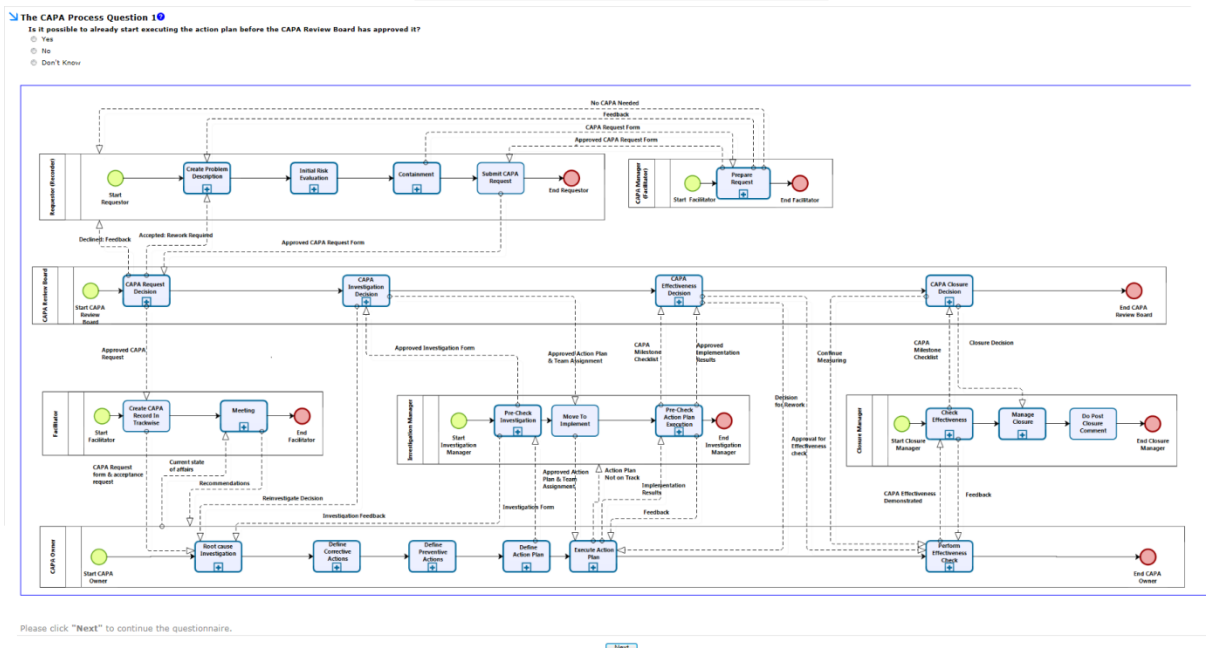


Please click "Next" to continue the questionnaire.

Next

Example of a question to that process (hovering over a sub-process shows its contents, just like in figure 11).





End of questionnaire (after all model questions and descriptive questions)

**That's it!**

Thank you very much for your participation in this study. You've helped TU/e, Philips (and me!) a great deal!

If you have any interest in the results of this study, you can indicate this by checking the box below. If you do so, you will receive an e-mail with the summarized results of my study. Please note that your e-mail address will not be used for other purposes, nor will it be shared with third parties.

Yes, please send me an e-mail with a summary of the results of this study.

Please click 'Next' to wrap up the questionnaire.

Closing of browser

You may now close the browser.

## Appendix F: Descriptive results experiment

STATA descriptive output for overall performance on Score, time and efficiency:

Variable	Obs	Mean	Std. Dev.	Min	Max
score_total	89	4.707865	1.560902	1	8
time_total	89	637.7416	334.0736	228	1971
eff_total	89	.009043	.004904	.0010147	.0232558
score_cont~1	89	1.438202	.6563511	0	2
time_control	89	189.0674	123.1377	36	734
eff_control	89	.0113499	.0098963	0	.0555556
question1	89	.7191011	.4519846	0	1
Q1_time	89	104.7865	75.90447	26	433
efficiency1	89	.0103445	.0097475	0	.0384615
question2	89	.7191011	.4519846	0	1
Q2_time	89	84.2809	74.25662	7	382
efficiency2	89	.0103781	.0095481	0	.0384615
score_data	89	1.337079	.7062034	0	2
time_data	89	146.618	95.97662	38	523
eff_data	89	.0127053	.0106357	0	.0526316
question3	89	.6292135	.4857521	0	1
Q3_time	89	74.73034	46.02093	14	246
efficiency3	89	.0079847	.0083689	0	.0384615
question4	89	.7078652	.4573204	0	1
Q4_time	89	71.88764	65.30826	11	342
efficiency4	89	.0099077	.009654	0	.0384615
score_org	89	1.269663	.7033046	0	2
time_org	89	126.8652	78.60433	24	446
eff_org	89	.01446	.0135014	0	.0833333
question5	89	.7191011	.4519846	0	1
Q5_time	89	63.85393	45.83408	12	233
efficiency5	89	.0102546	.0097739	0	.0384615
question6	89	.5505618	.5002553	0	1
Q6_time	89	63.01124	48.63816	12	326
efficiency6	89	.0068506	.0080758	0	.0333333
score_update	89	.6629213	.6561565	0	2
time_update	89	175.191	125.2789	42	640
eff_upd	89	.0047089	.0054979	0	.0238095
question7	89	.2359551	.4269999	0	1
Q7_time	89	96.73034	73.30893	10	439
efficiency7	89	.0027279	.0063409	0	.0384615
question8	89	.4269663	.4974398	0	1
Q8_time	89	78.46067	73.42189	18	420
efficiency8	89	.0062611	.0091752	0	.0384615

STATA descriptive output for CAPA performance on Score, time and efficiency:

Variable	Obs	Mean	Std. Dev.	Min	Max
score_total	44	4.886364	1.68738	2	8
time_total	44	671.1364	312.922	301	1790
eff_total	44	.0084793	.0046917	.0011173	.0232558
score_control	44	1.363636	.7182313	0	2
time_control	44	228.1818	132.9503	71	734
eff_control	44	.0077154	.0055967	0	.0222222
question1	44	.7954545	.4080325	0	1
Q1_time	44	107.2273	81.39964	28	420
efficiency1	44	.0110864	.009034	0	.0357143
question2	44	.5681818	.501056	0	1
Q2_time	44	120.9545	78.4323	26	382
efficiency2	44	.0082715	.0096321	0	.0357143
score_data	44	1.75	.4380188	1	2
time_data	44	123.2955	63.14862	38	298
eff_data	44	.0181542	.0106913	.0034483	.0526316
question3	44	.75	.4380188	0	1
Q3_time	44	68.75	46.24914	21	246
efficiency3	44	.0093986	.0077957	0	.027027
question4	44	1	0	1	1
Q4_time	44	54.54545	32.56939	13	155
efficiency4	44	.0137631	.0078501	.002381	.0357143
score_org	44	1.204545	.7649151	0	2
time_org	44	128.4773	63.11322	40	318
eff_org	44	.0118809	.011003	0	.0465116
question5	44	.6363636	.4866071	0	1
Q5_time	44	68.63636	41.29085	13	202
efficiency5	44	.0084079	.0091129	0	.0357143
question6	44	.5681818	.501056	0	1
Q6_time	44	59.84091	38.09532	17	188
efficiency6	44	.0066862	.0075739	0	.027027
score_update	44	.5681818	.6611381	0	2
time_update	44	191.1818	124.5939	49	584
eff_upd	44	.0034501	.0044298	0	.0153846
question7	44	.25	.4380188	0	1
Q7_time	44	108.5	83.12851	10	439
efficiency7	44	.0025941	.0056459	0	.0217391
question8	44	.3181818	.4711553	0	1
Q8_time	44	82.68182	66.43016	18	303
efficiency8	44	.0053297	.0093008	0	.0357143

STATA descriptive output for Complaint Handling performance on Score, time and efficiency:

Variable	Obs	Mean	Std. Dev.	Min	Max
score_total	45	4.533333	1.423823	1	7
time_total	45	605.0889	353.9916	228	1971
eff_total	45	.0095941	.0050949	.0010147	.0219298
score_cont~1	45	1.511111	.5886125	0	2
time_control	45	150.8222	100.1291	36	549
eff_control	45	.0149037	.0117957	0	.0555556
question1	45	.6444444	.4840903	0	1
Q1_time	45	102.4	70.96395	26	433
efficiency1	45	.009619	.010449	0	.0384615
question2	45	.8666667	.3437758	0	1
Q2_time	45	48.42222	48.54121	7	219
efficiency2	45	.0124379	.0091042	0	.0384615
score_data	45	.9333333	.6875517	0	2
time_data	45	169.4222	115.9746	38	523
eff_data	45	.0073775	.0074706	0	.0277778
question3	45	.5111111	.505525	0	1
Q3_time	45	80.57778	45.54942	14	244
efficiency3	45	.0066021	.0087597	0	.0384615
question4	45	.4222222	.4994947	0	1
Q4_time	45	88.84444	83.05939	11	342
efficiency4	45	.0061378	.0098375	0	.0384615
score_org	45	1.333333	.6396021	0	2
time_org	45	125.2889	91.97495	24	446
eff_org	45	.0169818	.0152678	0	.0833333
question5	45	.8	.4045199	0	1
Q5_time	45	59.17778	49.90094	12	233
efficiency5	45	.0120603	.0101566	0	.0384615
question6	45	.5333333	.504525	0	1
Q6_time	45	66.11111	57.38634	12	326
efficiency6	45	.0070114	.0086212	0	.0333333
score_update	45	.7555556	.6451059	0	2
time_update	45	159.5556	125.3516	42	640
eff_upd	45	.0059397	.006176	0	.0238095
question7	45	.2222222	.4204375	0	1
Q7_time	45	85.22222	61.00108	11	344
efficiency7	45	.0028587	.0070162	0	.0384615
question8	45	.5333333	.504525	0	1
Q8_time	45	74.33333	80.20995	18	420
efficiency8	45	.0071719	.0090617	0	.0384615

## Appendix G: data and data processing

To provide other researchers with the possibility to examine or replicate the data and analysis techniques used in this study, the following link is provided. It links to a Google Drive folder, where everyone with the link has access to. The folder contains three files: an annotated data analysis script and two data files; one with 23 cases and one with 28 cases. To read the data and execute the script, the statistical package STATA is needed. However, the data analysis script can be read in some text software, like Notepad++. The annotations allow the reader to understand what was done for this research. The faith is placed in all readers to use this link responsibly, without impeding other people from using it as well.

Link to Public Google drive folder:

<https://drive.google.com/folderview?id=0B201NrcK4yppY1U5T2FUVHpYWEE&usp=sharing>