Why companies cannot use carbon paper to develop products

- A multiple case study of experimental product development processes

Elin Grahn (50160) & Erik Gruvfors (50311)

Abstract:

To avoid developing products that customers do not want, many companies use iterative experiments in their product development processes. Today's existing frameworks regarding iterative experiments provide generic tools for how to conduct experiments within product development. However, limited research has been conducted on how individual company characteristics, i.e. contexts, affect the development process. This paper uses a qualitative multiple case study of eleven Swedish startups towards the aim to create further knowledge on how company characteristics, in terms of contextual factors, affect the way they conduct experiments in their product development process. The results show that companies with a higher level of maturity and previous innovation performance, experience, have a more structured experimental process and at the same time use an experiment type that conducts multiple experiments simultaneously whilst companies with a lower level of maturity have a more unstructured process and conduct sequential experiments. Smaller companies focus more on cost and criticality in their sequencing whilst larger companies have the capability to overlook the more critical tests in favor of testing features of a higher level of uncertainty. Startups with a physical product are more likely to conduct theoretical experiments using smaller sample sizes on blind test groups while companies with non-physical products create prototypes to test features on larger sample sizes that are aware of the experiments. Finally, companies with a high level of R&D intensity are more likely to test many alternatives within each experiment, set based, whilst companies with less intensive R&D investments focus on one sub-experiment and alternative at a time, thus point based.

Keywords: Product development, Innovation, Iterative experiments, Startup, Context

Stockholm School of Economics

Bachelor Thesis 2016

Tutor: Mattia Bianchi

Acknowledgements!

Mattia Bianchi, our faithful and patient tutor, for taking us under your wings when circumstances quickly changed and for all valuable tips and constructive feedback. Grazie mille.

Niklas Modig, for all your encouragement, invigorative texts and for truly being a reason for the completion of this thesis. You are the king.

Thanks to all who in some way have participated in this study and contributed to the fundamental data and findings, and it is true, we certainly could not have done it without you.

To our beloved **family and friends**, for your dedication and immortal support to us all the way through no matter what. We made it!

Stockholm School of Economics, May 23rd 2016

Gruvan & Grahnen

"Once upon a time, a frog and an octopus,
Met on a software project, that was deep in the bush.
The frog said, "you know, all these projects are the same;
Over the time we fill with our work the gap that we find
Between the burgeoning product, and our dreamed intent."
"Oh, no" objected the octopus, "they cannot be the same;
They come in all forms or shapes and sizes and colours,
And we cannot use the same tools and techniques
Like the cobbler shop, one size does not fit all."

(Kruchten 2011)

Table of Contents

1. Beginning

1.1. Why you should read this

According to the British tech and science journal Wired UK, Stockholm is the second largest producer in the world of unicorns[1] after Silicon Valley (Wired.co.uk 2015-07-30). When looking at innovation, Sweden is in third place of the most innovative countries in the world where Mojang, Skype, Spotify and Minecraft are only a few of the Swedish startups[2] that have now grown into multinational companies (Forbes 2015-11-11). A startup usually starts with an initial business idea or vision. For that idea to grow and to attract investors and venture capitalists, the business creator needs to prove that their business idea has traction (Ries 2011). When it comes to companies testing their products on the market there are many different frameworks and ideas (Wetter 2016). But the question still remains, does the process need adaptation to individual companies or can they use carbon paper to copy other companies' processes?

Galileo Galilei is thought to be one of the first to conduct experiments to test hypotheses (Drake 1999). Today, experiments have become a big part of product development especially software development (Bosch 2012). This has created a way of running businesses and driving innovation more efficiently to ensure the success of a company. By involving customers early in the development process and testing early versions of the product continuously, companies can constantly tweak and adapt the product to the customer needs before going live to the mass (Ries 2011).

Lean Product Development (LPD), The New New Product Development Game (NNPDG), SCRUM, Design Thinking (DT) and Lean Startup (LSU) are all examples of frameworks that have changed the way companies validate their innovation (Liker and Morgan 2006; Takeuchi and Nonaka 1986; Schwaber 1995; Brown 2008; Ries 2011). These frameworks only provide generic tools on how to develop products using an iterative experimental process (Poolton and Barclay 1998). Research highlights the importance of adapting the product development process after individual projects (Browning et al 2006). We have found limited research stressing the fact that the context of the companies could have an effect on the experiment process (Kruchten 2011).

The aim with this thesis is to move the knowledge forward and provide the readers with a more general understanding of if and how specific company characteristics such as size, maturity and product type can affect their individual innovation process when it comes to certain aspects of their experimentation.

1.2. From the laws of gravity to the lean startup

An experiment is a course of action to test a hypothesis which has been done for over hundreds of years to prove a particular effect and what is causing it (Griffith 1992). Some of the first experiments were by Galileo Galilei who studied the laws of gravity (Drake 1999).

[1] A label on companies that has succeeded in building a revenue of one billion dollars or more (Wired.co.uk 2015-07-30)

[2] "A startup is a human institution designed to create a new product or service under conditions of extreme uncertainty" (Ries 2011)

As Liker and Morgan (2006) state, Toyota created LPD by adapting their Toyota Production System and was one of the earliest versions of experimental product development. They understood the importance of front loading the development and listening to what the actual customers wanted very early in the process before they started the development. They also started to use what was called "set based concurrent engineering" which meant conducting multiple experiments parallel to achieve a more rapid development process.

A few years later, Takeuchi and Nonaka (1986) introduced the NNPDG which later evolved into SCRUM as a part of the agile innovation framework (Schwaber 1995). This approach was very different from the old sequential approach. A new holistic approach was developed to be able to cope with shorter product life cycles and increased demand for innovation. This new approach was different from the old one and the different parts of the development teams worked simultaneously using experiments to test if what they developed was something that worked and the customers wanted. These experiments were conducted as early as possible during the process.

Out of the NNPDG came then SCRUM which based all practices on iterations which. The traditional development processes used a linear approach whereas SCRUM introduced a development loop. This loop helped companies to adapt in an environment that is both complex and ever-changing (Schwaber 1995). DT was introduced with IDEO and has a very similar approach to product development and innovation. It uses a human centered approach where the needs and preferences of the customers are tested using experiments. The experiments often use prototypes and rapid prototyping that should not take more effort than to get the necessary feedback to develop the idea (Brown 2008).

A few years ago Eric Ries introduced a framework called the lean startup which also uses an experimental approach. One of the core concepts of LSU was the term minimum viable product, or MVP as most programmers know it. By using a MVP companies test different business model hypotheses to create validated learning and eventually a sustainable and growing business. When it comes to the actual concept of LSU and MVP people see it more as an "old wine in a new bottle". Today there are numerous frameworks around the topic of management and entrepreneurship regarding how to, in the most efficient way, experiment and test a product where LSU is one of the newest frameworks (Eisenmann et al 2011).

1.3. Problem area

Throughout the last decades, many new frameworks have been developed to drive innovation and make product development processes better. Frameworks such as LPD, The NNPDG, SCRUM, DT and LSU have changed the traditional way of developing products by involving the customer earlier and working with experiments to test the products many times throughout the development process (Liker and Morgan 2006; Takeuchi and Nonaka 1986; Schwaber 1995; Brown 2008; Ries 2011). The problem with frameworks like these ones which focus on startups' general needs, is that they only provide generic tools for how the product development process should look, as mentioned earlier (Poolton and Barclay 1998). Researchers have found problems when these are applied "out of the box" without any adaptation to the specific characteristics of the company itself and limited research has been done to try and find characteristics that affect agile innovation (Kruchten 2011). Given the fact that there is limited research that examines how the context of companies affect the way that

they conduct experiments within their product development, we see this as a theoretical gap that we aim to fill, or at least help to move the knowledge forward, with this study.

1.4. Purpose and research question

The purpose of this study is to examine and move the knowledge forward about how companies are using experiments in their product development processes given the literature gap further presented in the next chapter. Due to the limited research that exists today about how practices should be used in a given combination of company factors this thesis will examine the question on a company level (Kruchten 2011). Meaning that we study individual cases and look at their business as a firm rather than looking at an entire industry or individual projects within companies. We aim to create knowledge on how different companies are conducting experiments in their product development process based on their specific contexts. Our purpose is to empirically research if there are relationships between the contexts of which the companies operate in such as their size, age, geographical distribution and past innovativeness performance and their way of conducting experiments in terms of the dimensions such as cost of experiments, structure of experimental process and techniques to analyze experiments.

With this as background it leads us to asking the following research question: Are there firm specific contextual factors that affect the experiment dimensions within product development processes, and if so, how do they appear?

1.5. Expected knowledge contribution

By studying this topic in-depth regarding cause and effect, we hope that this thesis clarifies how different aspects of a company and their prerequisites affect the testing of their products and what is best suited for their particular situation when it comes to innovation processes within product development. However, we do not aim to set specific guidelines for each company but instead focus on finding patterns. There are differences in terms of context that have an effect and is something that companies should take into account when conducting experiments.

1.6. Disposition

This thesis is divided into seven chapters. Firstly, an introductory chapter to introduce the reader to the research topic of the thesis. Then a literature review to explain where the current research gap is in today's research which is followed by a theoretical framework serving as a basis for the discussion later on in the thesis. Chapter four is the research method and approach we have used in this study, where we present how we proceeded when conducting the data collection and analysis and introduce the data selection. The fourth chapter ends with a presentation of the reliability and validity of thesis before moving on to chapter five in which we present the findings. It starts with a brief presentation of all the cases in the study to give the reader a feeling of the empirics and afterwards the results we have found after filtering the empirics through our theoretical framework. Chapter six is our discussion where we analyze our results and connect them to the theories we present in chapter three. Finally, chapter seven which presents conclusions, implications and gives suggestions on further research.

2. Literature review

This chapter reviews current literature and findings about experiments in general to introduce the reader to the concept of experiments. It continues with a review of the literature about product development processes using iterative experiments and weave these together with focus on their similarities. Further on, the current research gap that is found in the literature today is discussed.

2.1. An overview of existing frameworks for experimentation within product development

Thirty years ago the NNPDG was developed in response to the traditional sequential product development (Takeuchi and Nonaka 1986). This evolved into SCRUM as a part of agile innovation which further emphasized the importance of iterations and experiments in the product development process (Schwaber 1995). DT continued with the importance of listening to customers in the development (Brown 2008) and five years ago LSU was introduced which framed this modern way-of-working with iterative experiments towards startups (Ries 2011). All of these frameworks provide generic tools for how companies should work with iterative experiments to drive innovation and develop new products (Poolton and Barclay 1998). However, we find limited research that take into account how differences between companies in terms of their contexts and characteristics and what effect it has on the iterative experimental process. This indicates that there is a theoretical gap to fill.

LPD, that evolved from Toyota, and its Toyota Production System could be seen as one of the first versions of experimental product development since it contains many of the framework elements such as defining customer value and avoiding the development of something that customers do not want through front loading of the development process (Morgan and Liker 2006). The customer focused product development was then developed further when the NNPDG was introduced and highlighted the importance of an iterative experimental process (Takeuchi and Nonaka 1986). Throughout the years more frameworks have been created, as mentioned above, such as DT, SCRUM and LSU. All of these present generic and flexible tools that do not take the company context into account (Poolton and Barclay 1998). Research have shown that even though the structure of each development phase is very similar, the specific activities and the dimensions of the activities and experiments differ a lot such as, duration and cost (Browning et al 2006). Therefore, we find that there is a demand for research that examines and identifies if there are contextual factors that affects the dimensions of the experimental process within the product development.

To further identify the existing research gap, the literature review below will examine the key features of the different frameworks regarding experimental product development. First comes an introduction to the definition of what an experiment is and following is experimentation within a business context. Thereafter, we explain the similarities between the existing frameworks and where the specific theoretical gap is identified to create a foundation for the research.

2.2. What is an experiment

An experiment is defined as a course of action to test a hypothesis; this is made to find what causes a particular effect on the test object. An experiment tries to show what outcome occurs when a specific factor is manipulated (Griffith 1992).

2.3. Experimentation within business

In today's competitive market innovation is important to develop a sustainable business (Ries 2011). In the traditional product development this was performed using a sequential approach where the product was first tested towards the end of the development (Takeushi and Nonaka 1986). Studies by Sosna et al (2010) have shown that implementing an experimental approach to a business product development can be effective and in an uncertain context, the trial and error approach can help companies to innovate their business model. At the same time not all projects and organizations will react similar to failure and therefore it is essential for entrepreneurs to understand the importance of learning from experiments that did not turn out as expected. Experiments within business is something that is a constant process of trial and error. To fine tune the product after the demands of customers (Sosna et al 2010).

2.4. The customer is the designer

When looking at the different frameworks for experimental product development, customer opinions and preferences are of great importance. Toyota with TPS were some of the first to realize this and made sure to listen to customers before starting any development processes (Likert and Morgan 2006). SCRUM, as a part of agile innovation, also emphasizes the importance of involving customers early in the process where each initiated development phase starts with an evaluation of the customer requirements and how the current system could be enhanced by looking at them or what a new feature should look like (Schwaber 1995). One of the major frameworks that really takes the customer preferences into account is DT. It uses a human-centered approach where user needs and preferences of thoughts on functions, packaging, marketing, etcetera are acquired by direct observations (Brown 2008). LSU builds its entire model around the product testing on customers through their build-measure-learn loop that is explained in the next section (Ries 2011).

2.5. The structure of a continuous process

The experimental product development process is presented by many frameworks as a continuous process often referred to as a loop. The NNPDG projects work with a continuous process using trial and error experimentation to solve problems. As a result, to close contact with information sources outside the firm and this way of working they can respond quickly to changes (Takeuchi and Nonaka 1995). Schwaber (1995) presents SCRUM as a framework divided into three phases; pre-game, game and post-game. The game-phase, also called the sprint phase, is a fairly structured experiment phase that takes place over one to four weeks. It is a loop consisting of development-wrap-review-adjust. This is something that really defines the experiments within SCRUM, a feature is developed under a short period of time to then be tested and reviewed by talking to customers or looking at sales and marketing. Features or items that do not lead to successful results are put in backlog for redevelopment in a new experiment until market fit is reached (Schwaber 1995).

Experiments in the DT framework use a loop in which projects will go through while the ideas are refined. So just like with SCRUM, the experiments are built upon a loop process. These phases of the loop are called inspiration-ideation-implementation where the product will loop through, especially, the first two phases many times as ideas are developed. By using rapid prototyping,

creating a prototype using only as much time as it takes to create a product generating the necessary feedback, the goal is to minimize the time during which experiments are conducted (Brown 2008).

Eric Ries (2011) describes the LSU framework to include a looped process as well. The work orbits around building a product in order to measure the outcome and learning from it, especially from failures (Ries 2011). Based on the feedback and learnings from previous experiences, the loop starts from the beginning again whilst the goal is to maximize the speed of the loop. What differentiates LSU from earlier frameworks is that it evaluates the entire business model instead of just the product itself (Eisenmann et al 2011). LSU uses something called a minimum viable product in their experiments which is just like rapid prototyping, a simple version of the product or service that the company is developing. This is done by only developing the minimum necessary for testing a specific feature on customers. When the MVP is developed it is tested and then the feedback is used to create a new version to test again (Ries 2011).

2.6. Driven by a vision

The frameworks reviewed in this study put a lot of emphasize on the fact that managers should have a vision and high goals for the development team. Looking at the NNPDG a broad goal is set from the top management positions, then the team uses iterative experiments to reach that goal. The managers set up extremely high goals to push their employees to test hypothesis through experiments to be able to develop new products (Takeuchi and Nonaka 1995). DT takes on the same focus where a vision is set in the beginning of a project but how to reach that vision is not decided but is instead something that is developed over time through the continuous process described in the section above (Brown 2008). Within SCRUM a manager defines the initial content and set the vision of the project. Thereafter the vision is used to control and steer the direction of the project but just like DT is the iterative experiments and continuous loops that drives the project forward towards that vision (Schwaber 1995). LSU has the same type of approach where the focus is on a long-term vision to find what their customers actually want and reach that vision. For this reason, the LSU process is provided, to help reach their visions (Ries 2011).

2.7. Teams

When reviewing literature about these frameworks they talk a lot about the teams behind the companies. According to Takeuchi and Nonaka (1986) the team uses experiments to develop their own agenda and own concept. The teams are given a direction by top management (the vision, as mentioned in the section above) from which they start conducting experiments to develop the product. The product development team works as a unit where multiple processes are ongoing at the same time, testing and experiments are being done simultaneously in different parts of the team. The rhythm of the individuals become the rhythm of the whole group and this pulse moves the development and experiments forward (Takeuchi and Nonaka 1986).

Takeuchi and Nonaka (1986) also states that as a result of the close contact with information sources outside the firm and this way-of-working the teams can respond to changes rapidly. This also helps them to acquire knowledge and skills making the team versatile. The learnings work both across multiple levels, from individuals to the whole group and corporation, and also across multiple functions where team members learn from other areas than their own.

The project teams are working mostly on their own without control from the management except from the initial directions and goals. The management makes sure to add enough checkpoints to keep the project stabile and prevent chaos. This helps the team to conduct experiments without getting off track from the initial goal without impairing their creativity (Takeuchi and Nonaka 1986).

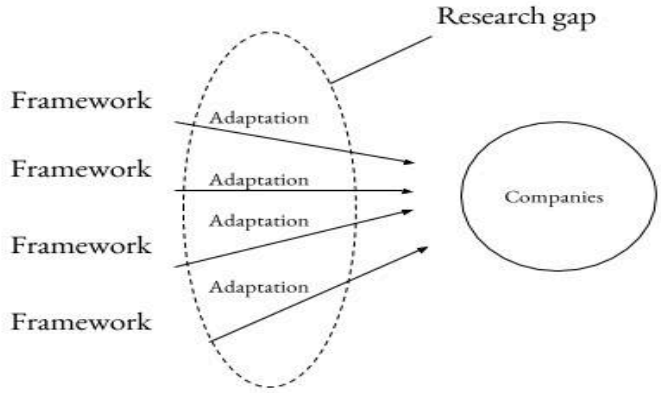
2.8. Carbon paper cannot be used to recreate experiments

Above we have reviewed relevant literature about different frameworks and tried to weave them together. The review shows that the current literature mainly focuses on providing generic frameworks of how companies should work with the customer as a designer and have a continuous development process where the products are tested early on. According to them, all of this should be driven by a vision that eventually aims to generate a product that the customers want.

However, some research show that the product development process is complex and needs to be adapted individually and cannot just be copied from one project to another (Browning et al 2006). Other researchers say that the use of frameworks like agile innovation that come from software development are likely to fail when they are used "out of the box". This means that they have not adapted to the context that they are used in, that is in some way far away from the context that it was originally created for (Kruchten 2011). As we have shown above many frameworks about the iterative experimental processes are very similar, and therefore we make the assumption that not only agile innovation, which is related to the software development, has to be adapted to its context but all iterative experiment processes. Researchers have found that to adapt an agile innovation process a company must look at what their context is to be able to adapt the dimensions of their process instead of forcing a project to use all practices from a framework (Kruchten 2011).

These findings stress the fact that there is a gap in current research about how the iterative experimental process could differ depending on different contexts of companies and is also the reason for why we aim to fill this gap with this study. To try to find out if there are underlying differences.

Figure 1 - The research gap



To be able to test and try to fill the research gap outlined above in figure 1, the following chapter will create a framework which categorizes different contextual factors we have found to be relevant and

look into if and how these can affect the experiments. The experiment cases will be defined by dimensions that separate them from to show their differences. More about this in the next chapter.

3. Theoretical Framework

This chapter aims to identify a framework based on today's knowledge about iterative experiments within product development to create a solid foundation to support the empirical analysis further on in the thesis. This will be our glasses through which we analyze our empirical data. It is divided into two parts: firstly, a number of contextual factors that supposedly affect the outcome of the experiment dimensions and secondly, a set of dimensions that are based on the literature review and what we have found to be in common for frameworks about iterative experiments. This is followed by a presentation of the method in the next chapter.

Figure 2 - The theoretical framework

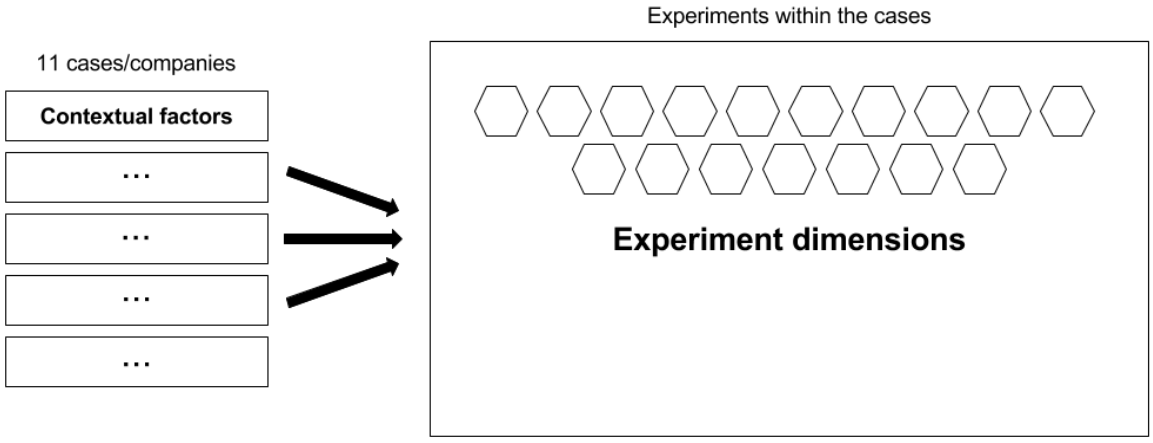


Figure 2 above illustrates the relationship between the eleven contextual factors and their potential effect on seventeen experiment dimensions. The arrows between are the link which this thesis aims to examine. This is the foundation of the theoretical framework of this thesis and will be further explained below.

3.1. Contextual factors

Part one of the framework focuses on the contextual factors of the actual companies studied in this thesis. Eleven different factors have been used in order to describe the context in which the experimental process takes place. As we showed in the literature review there is a current gap regarding how different contexts of companies could affect the experimental process (Kruchten 2011) which is also the reason for why we have chosen to examine it further. The contextual factors have been chosen by looking at two things; (1) what is presented in the literature review about what the frameworks emphasize and (2) on the basis of Kruchten's findings from his study where he creates a framework of contextual factors including these factors:

- Business domain
- Number of instances
- Maturity of the organization
- Level of innovation

- Culture
- Size
- Stable architecture
- Business Model
- Team distribution
- Rate of change
- Age of system
- Criticality
- Governance

Kruchten (2011) only applied this framework on a case of different software companies and found that there is an importance in identifying the context for software companies. However, limited research is found on how this is affecting other types of companies using similar experimental product development processes and boosts our drive to compare the experimental processes of companies with software products in relation to companies with other types of products.

Below is a presentation of the contextual factors we have chosen and the reason for why we find each of them relevant to look at when evaluating the company processes.

I. *Size of the company*: The number of employees within the company; small = ten or less, large = more than ten. Size of a company is usually measured in terms of revenues or market share (Rothaermel and Deeds 2004). However, startups usually have negative revenues during their early phases and to instead use the number of employees as measurement is a most relevant alternative (Shan et al 1994). Almeida et al (2003) further presents that this factor is relevant for experimentation since it can affect the overall company capacity, research availability, and limit what is possible to test and achieve. The size of the startup may play an important role in the design of the experiment dimensions. However, not every firm is utilizing the technological opportunities that can come from external learning and therefore, size may be the characteristic that explains why the external knowledge differs. A larger startup can have enhanced potential and abilities to exploit these technological opportunities but nonetheless, this might decrease the motivation to use informal learning mechanisms since larger firms have resources to exploit learning from formal mechanisms such as alliances (Almeida et al 2003). Size is also argued to be helping innovation in companies. Larger firms have more diverse capabilities and the ability to handle failures (Damanpour 1992). Nonetheless, some say that smaller firms have the advantage of being more flexible meaning they have a greater ability to adapt to changes (Hage and Mintzberg 1980).

II. *Maturity of the company:* The maturity level of the company; young = still exploring and experimenting with core concepts of the company/focusing on making it fly, mature = focusing on improving existing features of the product. Paulk et al (2002) describes it as when a company once is established the products continue to increase in both importance and size. The problems that are associated with this process can, according to the framework of the Capability Maturity Model more easily be managed through a sustained and focused effort that comes from the maturity of the overall firm (Paulk et al 2002).

III. *Product type:* Physical = the product of the company is a physical product such as an accessory or a gasket, non-physical = mainly software-based products such as smartphone applications. Dickson and Ginter (1987) highlight the ongoing discussion regarding physical and non-

physical products effect on areas such as market segmentation, product differentiation and marketing strategy; whether the outcome of virtually based products differs from the outcome of physical products.

IV. *Experiment innovation:* The level of innovation within the experiment under development; simple = low level of innovation, complex = high level of innovation. Based on the innovation level the risk of the company might be increased or decreased. For example, a completely new product that does not exist on the market has a high level of innovation and if this product does not sell well the company has spent a lot of money in vain. The goal of the experiment process is to drive innovation in many cases (Ahmed 1998). The different levels of innovation should hence be interesting to look at how it affects the actual experiment dimensions.

V. Geographical distribution: Where the workforce is located; co-located = in one place e.g. HQ, scattered = in more than one place. The distribution of the team is often correlated with the size of the project and company. Increase in distribution also increases the need of communication, coordination and control. Within software development scattered individuals is common when it comes to open-source development (Holmstrom et al 2006). The assumption that this factor is also relevant to look at when examining companies that are not software developers but using iterative experimentation processes just like software companies in Holmstrom et al's study.

VI. *Customer participation*: Low or high level of customer participation in the experimentation processes regarding workshops, feedback, customer test subjects, etcetera. According to many frameworks, the inclusion and participation of customers within innovation processes are crucial for the company to rely on (Ries 2011). "We must learn what customers really want, not what they say they want or what we think they should want" said Eric Ries and means that this only can be achieved by involving the customer in the process.

VII. *Newness to market*: Does relevant market players within the same industry already exist or is the company a market pioneer; low = competitors already exist, high = a pioneer. Research from Abbie Griffin (1997) shows that the cycle times of product development increases with the newness of the product. At the same time the use of cross functional teams reduces the cycle time and the use of a formal process decreases cycle time. The more complex a product is and the higher level of newness it has, thus less of the design can be transferred from a previous version the more advantages can be found in a formal process thanks to its shortening of cycle time (Griffin 1997).

VIII. *R&D intensity*: How much capacity in terms of time and cost that the company invests in R&D; low or high intensity. The iterative experiments aim to reduce the time to reach product market fit (Ries 2011). R&D generates innovation but it also helps the company to get external information from the environment (Cohen and Levinthal 1989) which according to Ries (2011) is the goal of experiments.

IX. *Past innovativeness performance*: How much innovation experience does the employees working with the experiment have from previous and similar projects; low = little experience and learning along the way, high = a lot of experience to implement on the current experiment. If a project has a large amount of assumptions and known inputs from an earlier version or previous project it tends to simplify the product development process. If a team have to create everything from scratch it requires a greater level of creativity to handle problems and issues due to lack of knowledge (Griffin 1997).

X. *Funding*: If the company is self-sufficient and funded on private equity, venture capital or a mix of both; divided = both venture capital and private equity, private equity = company equity only, venture capital = fully financed by external investors. To get funding, companies need to show that their business model is working (Sherman 1998). Since the aim of the experiments is to find traction and find something that the customers want (Ries 2011) this becomes an interesting factor to look at when studying the cause and effect of the experiment dimensions.

XI. *Type of industry:* What type of industry each specific case is active within. The product itself is described further in the results section under each case presentation. Audretsch and Thurik (1999) introduces the possibility of a potential influence regarding a young innovative company and the industry it is active within. They however mention the related complications based on the fact that preferable all other factors remain constant, which in this case is nearly impossible.

3.2. Experiment dimensions

The second part of the framework will look at the dimensions of the experiments. As our literature review shows, earlier studies have only looked at presenting generic tools for the experimental approach to product development. Since the purpose of our research is to look at how companies' different contexts affect their experiment processes, this section will explain what specific dimensions of the experiments we will look at to be able to find differences in the experiment process within our different cases and draw conclusions to how those differences are affected by the company differences in terms of their context.

The experiment dimensions were chosen with the literature review as a foundation and the frameworks presented there. By looking at them and the finding similarities within them we have developed twelve different dimensions that are specified below and explained according to what the different frameworks emphasize within each dimension.

I. *Type of unknown tested*: What part of the company that is tested in the experiment; technical = something related to the software such as application features or algorithms, market = testing how the market responds to certain actions of the company/segmentation/customer service/etcetera. LPD focuses on testing technical features which are very specific. Agile innovation focuses completely on software test object and their ideation comes from user stories and are needed when choosing which unknown to be tested. DT overlaps agile innovation by also focusing on user stories when defining what type of unknown should be tested. What differs between them is that DT have a broader perspective and makes this applicable for any company regardless of what industry. LSU takes this dimension to an even higher level of abstraction and evaluate the entire business model when conducting experiments but its foundation is still the market.

II. *Sequencing criteria*: What the reason is for conducting experiments in a sequence; criticality = some things need to be tested before other things/test objects are of different importance, uncertainty = experiments rely on the results from previous experiments (for example the engine of a car which is completed before testing what colour of the car the market prefers), cost = the company cannot afford multiple simultaneous experiment due to capacity of cost and time. These different criteria can either stand alone or be combined in pairs. All the

frameworks focus the sequencing on testing the most critical hypothesis first. Otherwise there are no specific criteria that the frameworks focus on and is something that the contextual factors could affect.

III. *Experiment type*

A. *(1)*: At what stage of the process that the experiment subject is tested; theoretical = only existing in mind/based on for example surveys and screenings, prototype = a virtual or physical product produced during the experimentation process with limited features. Ries (2011) emphasizes that experiments should be conducted as early as possible in the process, therefore we have chosen to include the dimension of in what stage the product is when it is tested.

B. *(2)*: Exploratory testing = unplanned and unscripted testing/no defined expected actual outcome, V&V = Verification and Validation/procedures for checking if the experiment meets requirements and specifications that fulfill the intended purpose. In the LSU framework a lot of emphasize is put into conducting hypothesis before the experiment is done do then validate that hypothesis depending on the feedback from the test (Ries 2011). The opposite strategy would be an unstructured test with an exploratory focus.

C. *(3)*: Parallel testing = a number of experiments are run simultaneously, sequential = the experiments are run after each other/not simultaneously. Takeuchi and Nonaka (1986) highlights the importance of overlapping development phases meaning that experiments are run simultaneously, therefore we have included this dimension.

D. *(4)*: Set based = the experiment tests many alternatives, point based = the experiment tests only one alternative. Experiments can be conducted to test multiple alternatives often referred to as A/B-testing (Ries 2011).

E. *(5)*: Blind test = the subjects of the experiments are not aware of it and cannot adjust their reactions, aware = the test subjects are aware of the experiment. When test subjects are aware of the experiments research have shown that they tend to modify their behavior (Thomke and Manzi 2014) meaning that the company risks to get feedback that is false.

IV. *Choice of environment*: Where the experiment takes place; in vitro = in a laboratory environment, in vivo = in a real environment/a field experiment. None of the frameworks specifically state when an experiment should be conducted in a specific environment. LPD focuses on creating learning from where the work is done, hence it focuses more on conducting experiments in a real environment rather than doing experiments in a laboratory environment. The other frameworks emphasize the involvement of real customers as early as possible in experiments but do not involve specific environments to be used for the experiments. This is something that could be explained by our contextual factors.

V. *Cost of experiment*: How much money the company has spent on a specific experiment in relation to the company's overall financial performance; low = no or little money spent, high = large amounts of money. All of the frameworks have the same focus regarding cost of the experiments, which is to minimize costs through the use of experiments and thereby minimize the costs of the experiments themselves. Their way of minimizing cost is however something that varies across the frameworks. To keep costs to a minimum the frameworks aims to find potential problems as early as possible. NNPDG and LPD do not explain a specific way of doing this compared to the other frameworks. Agile innovation wants

companies to work with small experiments throughout the game-phase to minimize costs and at the same time evaluate problems. DT involves another dimension to minimize the costs which is rapid prototyping that involves new techniques such as 3D-printing to get fast feedback from customers and keep costs down. LSU introduces a new way of testing to minimize experiment costs which is the MVP. It is used to evaluate the entire business model using an MVP that is focused on testing very specific features.

VI. *Experimentation process structure*: If the the experimentation process is clearly defined or not. The NNPDG is working with a dynamic process of experiment where the learning comes from trial and error and therefore it is less defined than for the other frameworks. The other frameworks define the experimentation process as a loop, however they choose to label it differently. LPD is working after the process of manage, improve and to continuously learn. DT calls it inspiration, ideation and implementation. LSU calls it build, measure and learn. SCRUM calls it develop, wrap, review and adjust. All of these loops are basically doing the same thing for the companies using the frameworks. They are the core of the actual experiment process for the development. Something is created to be tested and then that is tested through an experiment, after that reviewed and adapted or rebuilt to be tested again. The mutual and ultimate goal is, as mentioned above, to maximize the speed in the loop to eventually create a product that the customers want. Agile innovation differs slightly in terms of letting the loop being part of a larger framework. With one phase before and after the loop that are more structured and consists of defined process.

VII. *Experiment speed*: The amount of time that is spent on an experiment from the first idea until it is tested. All the frameworks are trying to reduce the speed to make the development as fast as possible but not in terms of a finished product but rather a product that could be used for an experiment to generate feedback. This appears different within the frameworks. SCRUM has the most defined timeframe of all the frameworks with development cycles of one to four weeks. DT and LSU have a very similar approach where both frameworks emphasize minimum time being spent on development to be able to generate the necessary feedback. This is being done with rapid prototyping in DT and MVP in LSU which basically are the same things with different names. The NNPDG and LPD do not explain specifically how companies should go from idea to experiment fast but they are focusing on overlapping development where different features are tested in overlapping phases rather than sequential.

VIII. *Fidelity*: Accuracy of reporting detail during the experiment in terms how real the experiment is. The higher fidelity the more circumstances are taken into account (Snyder 2003). This is subject to how far the development has come. Agile innovation, DT and LSU emphasize a low fidelity in the beginning where specific features are tested leaving out all other variables after that the fidelity increases as the product gets closer to a finished product. The NNPDG and LPD state that it is important to test the most critical features early in the process however it does not specifically state the level of fidelity within these tests, but the assumption is made that it follows the same state as the other frameworks with a low fidelity in the beginning which increases along with the development.

IX. *Requirements*: If the requirements of the experiments are affected by fixed or variable factors; constant = the requirements are the same all the experiment through, changing = the requirements change one or more times during an experiment. Audretsch and Thurik (1999) connect the requirements from customers and the market with the flexibility of a company and its processes could therefore result in an essential pattern in this study.

X. *Sample size*: The size of the sample group that the experiment is tested on is not something that the frameworks explain or determines. LSU emphasizes the importance of small test groups of early adopters for the first experiments and can therefore be assumed that the other frameworks also focus on smaller sample sizes in the beginning to then scale up as the fidelity level gets higher. In the corporate world, on which important decisions are based, the size of the data sample is of crucial importance (Morse 2000) and a factor that market players in these industries can use to reduce their risk by taking on more and more customers.

XI. *Techniques to analyze experimental data*: How the experiment performers evaluate the outcome/results; subjective assessment = more of a gut feeling based on different variables such as previous experience and time limit, metrics = one or more variables that are assessed and measured in all experiments. Agile innovation focuses more on software development and therefore uses metrics rather than subjective assessments because the most important thing to validate is if the feature is being used or not. LSU also highlights the use of metrics to be able to describe cause and effect. Other frameworks do not specifically state what type of metric to use as it seems to be subject to what is tested.

XII. *Test subjects*: If the test subjects are a group of people or a virtual assessment basis; real = people or other living organisms, unreal = virtual or theoretical. This is also not stated by the frameworks of what kind of test subject to use. LSU talk about using early adopters when testing because they can ignore the lack of features when testing MVPs.

With the help of this framework we can see that all of the reviewed frameworks are in fact very similar, and general, in the dimensions that they focus on. This framework will be used as a filter to look at our cases individually and compare them to understand if there are contextual factors that in the outcome can be evident to affect the dimensions of the experimental process.

4. Method

4.1. Research approach

According to Andersen (1998) a research can be either inductive or deductive where inductive means that you use the gathered empirical data to make general theoretical conclusions. On the opposite side is deductive which is when you generate hypotheses from existing literature to draw conclusions. In our research we chose to use a combination of the inductive and deductive approach due to the following reasons. A combination where both and inductive and deductive approach is used is called abductive approach and was chosen for this study since it best describes our research methodology. We chose this approach since the aim of our study is to make an exploratory study to create further knowledge about how the contexts can affect the outcome of a certain phenomenon (Andersen 1998).

4.2. Research method

The questions we aim to answer in this thesis require thorough formulations and complete in-depth answers and are reasons for why the study method is focusing on quality rather than quantity (Bryman and Bell 2011). Hence the detailed qualitative data collection. We do also, through the various elements of the thesis, wish to describe, explain and interpret data in an exploratory manner and at a deeper level since it today exists limited knowledge about whether the context of companies

have an effect on their development process (Ahrne and Svensson 2011; Stige et al 2009). Therefore, a multiple case study was chosen since it enables an in depth analysis of each case which helps to find contexts that separates them from each other (Yin 2003). A quantitative method was not chosen since it often has a deductive focus meaning that it tests theories (Bryman and Bell 2011). Our research is exploratory and does not aim to test theories but rather explore how companies uses the frameworks that we have reviewed. Therefore, a quantitative method is not relevant.

4.3. Sampling strategy

These eleven interviews of startups included in this study are within eleven different industries. Stockholm was chosen as the geographical sampling area mainly due to convenience, but also the fact that Stockholm is one of the most innovative cities in the world (Wired.co.uk 2015-07-30) makes it an even more interesting market to observe and examine. Our study began with contacting three business incubators: SSE Business Lab, STING (Stockholm Innovation & Growth) and SUP46. Through these three we got in contact with twenty-nine different companies that were contacted by email of which eleven companies were willing to participate in the study and contribute with empirical data.

We have chosen to limit our study to startups only, given the contextual factors included in the theoretical framework. This since the decision to focus on a single industry allows us to assume startup-level homogeneity and thereby increases the possibilities to find and interpret heterogeneity in the cases' experimental dimensions (Bianchi et al 2011). Another reason for why startups were chosen as the sample group is because they operate in an environment of extreme uncertainty with very limited resources which forces them to quickly find traction to be able to succeed (Ries 2011). The unique context of startups makes it interesting to study, as well as the sampling selection that is ranging from completely new companies that have just launched to startups that has begun their international expansion. These also differ widely in terms of their implementation of experiments in their development process. All cases included in this study have in-house R&D processes based on the fact that if they used a third party actor for innovation and development, it would require this study to interview the partners of each specific case as well.

In order to cover and include as many aspects as possible in our results, the selection process has been made based on two fundamental criteria:

> Variety: to be able to generalize study results to some extent we need a relatively large and varied selection to start from. For that reason, the interviewees are a combination of people with different experiences and backgrounds but all with some sort of connection to product testing. We have however come to the understanding that this thesis will not be able to reach statistical generalizability due to the limits that will be presented in the end of this chapter.
>
> Detailed access: from these interviewees we wished to get as much detailed information as possible, especially when it comes to soft variables that can be hard to measure numerically, what in this thesis are called the experiment dimensions. As a consequence, the potential interviewees were chosen foremost based on access and also their role with regards to the others. This is also one of the reasons for all cases being based in Stockholm.

4.4. Data collection process

All companies are based in Stockholm and were interviewed using a semi-structured method in order to enable flexibility to follow up on questions while it at the same time maintains an overall configuration and direction to analyze and identify similarities and differences (Bryman and Bell 2011). In the appendix we present the key characteristics of the companies that have been studied in a result table.

When it comes to the number of cases included in this study, a multiple case study format was chosen based on the lack of well-formulated theories regarding this research area, as presented in the literature review above, and since the aim is to examine how the concept of the phenomenon of interest affects the outcomes (Ellram 1996). Each of the eleven case studies function as a self-contained experiment itself with its unique contribution to the overall experiment as a whole, enabling in-depth analysis both within-case and cross-case (Ellram 1996; Yin 2003). Based on the first step in this contingent approach, the study further on aims to find patterns of cause and effect.

Yin (2003) further presents a discussion of whether the case identities of a study should be real or anonymous depending on the context. It can be raised at two levels; either of the entire cases or of individual persons within the cases. He argues that the most desirable option for multiple case studies is to exclude all identities for both cases and interviewed individuals within the cases for two main reasons. Firstly, the averted risk that the reader could recollect any other information he or she has gained of the specific cases from another source prior to reading this study and thereby could influence the interpretation of this study. Secondly, Yin explains that the total disguise of names also simplifies the overall study review. However, he also says that the use of anonymity needs caution since valuable and relevant background information of each case can be lost and also complicates the composing of the case. Based on the case study research performed by Yin we have therefore chosen to exclude all names of both individuals and cases to prevent subjective assessments from readers and chosen to retain and present background information regarding each case in order not to lose any relevant contributions to the results. Since this study aims to examine if and how different characteristics of startups impinge their way of conducting experiments within the innovation processes, the background information of each is of crucial importance and therefore presented.

4.4.1. Pilot study

During the first weeks of this thesis our aim was to study the phenomenon of LSU and its entry on the Swedish market. After the pilot study the aim was rewritten due to realization that the research focus on LSU was too wide together with similar content findings in each of the four cases presented during the first interviews. Even though these were conducted on the basis of studying LSU, we still consider these cases relevant to the updated research question regarding innovation processes within startups since the interview method was open-ended and the answers cover the same areas as covered in the main study. For this reason, those four opening cases are included in the main study.

The aim with the pilot study was to gain a general understanding of the phenomenon of a different framework process than the one we ended up studying but what we found was that there are many similar frameworks within innovation through experiments, as shown in the literature review. They have different names and partly varied process formulations, but many of the key concepts can be

found in several of them. As a first step towards this general understanding we met with Erik Wetter, managing director at SSE Business Lab. The discussion with Erik focused on general topics regarding his view on LSU, what it is and where it comes from, as well as the development processes of startups within the accelerator.

The key takeaways from this initial interview is that (1) the LSU framework is only a repackaging of previous and existing frameworks; "just like CrossFit, the best parts of existing alternatives are put together in a new single format", (2) this repackaging of frameworks is a sort of management fad and finally (3) that it is a huge positive advantage that these reformulated ones exist because of the too linear and conservative course literature education uses today which for example can say that a pitch needs to be perfect but it in reality does not, it needs to grow.

4.4.2. Main study

Part two of the study was to contact a wide range of startups with different characteristics in order to cover as many different aspects as possible (presented in appendix). Even though the selection of companies and their variation cannot be considered to meet the generalizability (Bryman and Bell 2011) it provides a deeper insight from different contexts. All interviews were conducted with a founder of the company or the CEO at their offices which were situated at SSE Business Lab and SUP46 except from the bowtie and housing companies who had their own offices. An interview script was written before meeting with any of the interviewees (presented in the appendix) and is based on Robert Yins strategy to use both open-ended and direct interview methods in order to cover all case aspects (Yin 2003). All of the interviews were conducted verbally and was transcribed from a recorded audio file to a digital word processing program. The interviews were transcribed by the ourselves to avoid errors and misinterpretations (Ahrne and Svensson 2011). The analysis of the empirics was done using techniques from Miles and Huberman 1984, following these four steps. 1. Data Categorization; 2. Data Contextualization; 3. Preliminary within-case analysis; and 4. Cross-Case analysis. This structure for analysis combined with the use of semi-structured interviews enhance the reliability of this study (Yin 2003).

Something that all of these cases have in common is that they all have their experimentation and innovation processes in-house (a demand in this study), include their customers in their processes and are focusing on building the business through experiments. Otherwise they would not be of value to this thesis because of its aim to examine the specific causes and effects of these processes.

4.5. The data

During the first part of the data collection, qualitative indications and guidelines were found for where we should direct our focus regarding both startups and the way they work in general, together with a fundamental understanding as to how different frameworks work and their similar characteristics. The key takeaways were that the phenomenon of LSU would be too diffuse to examine on the Swedish market today due to the its relatively recent entrance, the processes within startups today are based on logical thinking and a number of different frameworks and that all these are basically the same. Furthermore, these frameworks are very metaphysical and serve as guidelines for startups in general, thereby not providing any specific guidelines for a specific kind of startup.

During the second part of the data collection, the eleven case studies provided information regarding both company characteristics and their experimental processes as well as stories from each case regarding the development processes from an idea until now. The interviews were conducted in the same manner with a set base of essential questions leading up to an information bank covering the same information from each case. This information was categorized into contextual factors and experiment dimensions.

4.6. Coding and analysis of the data

Coding of data is one of the most important parts of a qualitative study. The aim of the coding is to make sense of the textual and unstructured data to be able to analyze it (Basit 2003). We chose to code our data after the two parts of our research, company context and experiment dimensions. The context factors and experiment factors were picked from the findings of the literature review and the theoretical framework that was created. After the data was transcribed we coded each case according to their performance on each factor and dimension. For simplicity, the variables have been color coded and presented in table 1 and 2 below.

Table 1 - A guidance to how to interpret the tables 3-13 representing each case's contextual factors

Small	Size of the company	*Large*
Young	Maturity of the company	*Mature*
Physical	Product type	*Non-physical*
Simple	Experiment innovation	*Complex*
Co-located	Geographical distribution	*Scattered*
Low	Customer participation	*High*
Low	Newness to market	*High*
Low	R&D intensity	*High*
Low	Past innovativeness performance	*High*
Private equity	Funding	*Venture capital*

Table 2 - A guidance to how to interpret the result tables 14-18 presenting patterns in experiment dimensions

Market	Type of unknown tested	*Technical*
Cost/Critical	Criteria used for sequencing experiments	*Uncertainty*
Theoretical	Experiment type (1)	*Prototype*
Exploratory testing	Experiment type (2)	*Validation & Verification*
Sequential testing	Experiment type (3)	*Parallel testing*
Point based	Experiment type (4)	*Set based*
Blind test	Experiment type (5)	*Aware*
In vitro	Choice of environment	*In vivo*
Low	Cost of experiment	*High*
Undefined	Experiment process structure	*Defined*
Slow	Experiment speed	*Fast*
Low	Fidelity	*High*
Low	Importance of evaluation	*High*
Constant	Requirements	*Changing*
Small	Sample size	*Large*
Subjective assessments	Techniques to analyze experimental data	*Metrics*
Real	Test subjects	*Unreal*

4.7. Data reduction

Data reduction is needed to be able to establish a meaning of the data and works like a preliminary analysis (Miles 1979). We created a coding scheme where the dimensions and factors were combined which we filled with our coded data from the interviews. This made it easier to find patterns in the

data and see how the company contexts interacted with the way the companies conducted their experiments.

4.8. Research quality

The nature qualitative and explorative study design indirectly makes it difficult to include only objective assessments of collected data since interviews are our main source of information at the same time as it brings out more trustworthy results. To minimize the effects of subjective assessments which could make the credibility lower due to the risk of misinterpretation of the social reality during the interviews we have both participated in all parts of the data collection (Bryman and Bell 2011).

The *external reliability,* that is to what degree this study could be replicated, is relatively low. This is because it is impossible to "freeze" a social situation and all of the circumstances around it to try and recreate it. However, we have tried to minimize this by choosing a broad variety of companies to get a more nuanced picture of reality (Bryman and Bell 2011).

The *internal reliability* that is if there are more than one observer of the research that agree on what is gathered from the research object. This is something that we believe is quite high in this study since we have both participated in all parts of the data collection (Bryman and Bell 2011).

The *external validity,* that is how the findings can be generalized in other contexts. This is a problem with the qualitative research method and something the reader should keep in mind when reading this. We provide a snapshot of eleven cases from the Swedish market, but to be able to generalize our findings more research is needed in other types of contexts (Bryman and Bell 2011).

The *internal validity,* that is how well the observations lead to finding contingent relationships. This is one of the strengths of the qualitative study since it allows in depth information of these relationships that we later have analyzed. The internal validity is therefore seen as acceptable given the timeframe that was available for conducting this research. A longer period to observe the cases would have given the study even higher validity (Bryman and Bell 2011).

5. Results

The first part of the results presents the cases that make up the empirics to give the reader a background about each case and what they do and what their processes look like. In the second part we present the results we have found when comparing the different cases.

5.1. Case summaries

5.1.1. The young bowtie company

This company was founded during the last year of the founders' high school education as a part of an entrepreneurship course in collaboration with Ung Företagsamhet. The course ends with all student startups competing in the Swedish Championship in Young Entrepreneurship where this company won awards in two categories at the Stockholm award ceremony. Their product is handmade bow-ties made of recycled clothes donated by Swedish celebrities such as Stefano

Catenacci, Lena Endre and Thomas Ledin. The focus of the four founders is to change the attitude of the society to make recycling cool. This company is also one of the youngest cases included in this study with less than one active year since launch and very little previous experience within experimentation. Their journey began after a seminar with a young entrepreneur who encouraged them to use what they had to obtain, including networks, and that is how they ended up in the fashion industry as a young Stockholm-based company with very little experience. Even though the founders are close friends since before becoming business partners they have no previous knowledge from each other in this situation, hence a very floating innovation process and theoretical approach to new ideas before applying it on the market. However, they made sure before starting production, by including customers in the development process through surveys and interviews, that they had a customer base that could ensure sales to cover costs and repayment to investors and thereby reduce the risk in advance. The funding was split between the founders and venture capitalists 10/90. Their product is relatively classic with a constant demand, fast experiment speed and a set degree of innovation depending on garment donations and flexibility that comes with them working on a small scale. Their initial experiments were conducted on friends and families without them knowing what the test actually was about, to later shift towards more high-fidelity experiments focusing on one experiment at a time in order to reduce costs and evaluate their progress on the market.

Table 3 – The contextual factors of the bowtie company

Size	Maturity	Product type	Experiment innovation	Geographical distribution	Customer participation	Newness to market	R&D intensity	Past Performance	Funding

5.1.2. The fragrance subscription

A subscription service that simplifies finding new fragrances. The service sends you one seven milliliter bottle of perfume every month with different fragrances according to a chosen theme. Their team consists of three people. Their process started during the summer 2015 with extremely simple experiments without a physical product or prototype. Tests was being done without mentioning the product at all just to see what people's attitudes were towards perfumes. After that they created the first MVP, due to the fact that they did not have the funding to build a stock of hundreds of perfumes they had to test this using an MVP. This MVP was tested in a closed environment for a focus group to gather information about their attitude to the subscription service. By now they have realized that their initial idea of a library of fragrances was too expensive. So what they launched was an MVP of that library with four different themes that the customer could choose and every month the theme contains a different fragrance according to the theme. This was launched just before Christmas and turned out to be something that the customers liked due to the fact that it eliminated the option of having to choose a fragrance. Today, six months later, their MVP has become their actual product that they sell and opens up for regular dialogues, which are crucial for the fundamental and future product demand, between the company and their customers regarding the service and specific fragrances included in the different themes. Based on the need to keep costs down and the low level of past innovativeness performances, experiments are focused, sequential and point based. Their innovation processes are flexible and fast.

Table 4 – The contextual factors of the fragrance subscription

Size	Maturity	Product type	Experiment innovation	Geographical distribution	Customer participation	Newness to market	R&D intensity	Past Performance	Funding

5.1.3. The proactive housing company

This company serves as a quick, secure and smooth way to deal with the secondary housing market. It was founded in the spring 2014 with a first product version launched the same year during fall with backing from Sweden's largest housing site Hemnet which makes the funding consist of both private and venture capital. Four full-time employees and a number of part-time employees work daily with a mature and continuous development of the product through innovation experiments in order to reach their goals by the end of this year. Some of the employees have a lot of experience from either the housing market or entrepreneurship. The processes are made up of many small cycles building up to these overall goals and their unstructured process stages can be generalized to: a flood of ideas that pop up from everywhere all the time, testing, replanning, development and thereafter iterations. Between testing and replanning customers contribute with evaluation metrics. The many high-fidelity experiments are usually conducted by their test/growth team, take the longest about a month and are parallel and set based. According to them it is not possible to have longer and expensive processes in the quickly changing housing market, and especially not as a startup due to the high burn rate of capital, therefore they work with multiple experiments and projects simultaneously. Some results might fail in some aspects but can be relevant in other experiments so parallel testing has proven to work very well in their case. One advantage of being a software startup is that they quickly can go back to how something was before an experiment if it does not work. They have also found it extremely important to early ensure a demand or need to minimize potential risks in the future.

Table 5 – The contextual factors of the housing company

Size	Maturity	Product type	Experiment innovation	Geographical distribution	Customer participation	Newness to market	R&D intensity	Past Performance	Funding

5.1.4. The social stock market service

This is a small, young and co-located company with a financial platform for stock market services, free for users, tailored to a younger audience who have different preferences for how they consume information, in this case financial information. It has a social network connected where the companies will be a part of the community. It will be a place for communication and interaction for equity investors and companies. The company has four employees and was started two years ago. Their process started with a lack of programming skills which forced them to create a mockup to conduct their first experiments. The first prototype was not even a functioning prototype but just design mockup of how the platform was supposed to look. This was tested in a closed environment to show the traction for investors which was hard. When they finally found a person who wanted to invest that person was someone with a lot of experience in terms of users, ease of design and his technical knowledge of how to put this together. So they made a first version and thought they were ready, just had to launch it and pick up a few million. However, the more experiments they conducted the more feedback they gained about new features to build. These iterations were very

unstructured with a short term focus. They use metrics in terms of if and how much a new feature is being used and releases a new feature once every month. Experiments are usually simple but processes fast and set based parallel testing practiced on a large test group who is aware of the experiments.

Table 6 – The contextual factors of the stock market service

Size	Maturity	Product type	Experiment innovation	Geographical distribution	Customer participation	Newness to market	R&D intensity	Past Performance	Funding

5.1.5. The simplified website tool

This company is a free-form web design tool currently within the incubator, SSE Business Lab. Enabling entrepreneurs to build websites and apps without coding. They are a team of three people and launched their first version in January 2015. They were aware that they could not develop the final product in one step and therefore divided the development into phases and evaluated these using a set of metrics. In the first stage the only experiment subjects involved were interaction designers and not entrepreneurs. They asked them what they were missing in today's tools and from that created a prototype and as it evolved they let the designers test and interact with the products. Eventually they started to test the product on their potential customers to get feedback and today perform undefined and high-fidelity experiments with focus on parallel, set based verification and validation testing. A lot of the venture capital is put into R&D which they back with years of experience. The experiment costs are low, sample sizes large and test group aware of the experiment.

Table 7 – The contextual factors of the website tool

Size	Maturity	Product type	Experiment innovation	Geographical distribution	Customer participation	Newness to market	R&D intensity	Past Performance	Funding

5.1.6. The composing hunting application

This company develops a user-friendly system working for all hunters via SMS, smartphone applications or computers and helps solving tasks that take away time from the hunting itself such as reporting, compilation and debriefing for shooting teams and fields of maintenance. The main difference from other products in this thesis is that this one is season-based to just before hunting season during fall and for same time afterwards. This means that when other companies can measure and evaluate their progress on a daily basis and have a fast experiment speed, the hunting application company can measure after each season and year. Today five part-time employees work together during the year before, during and after the season to activate the users (hunters). This application is not as complex but more ad-hoc compared to other applications in this study due to its first use as an administration tool only which later has been developed to what it is today. The idea came from a farmer who paid the founders in advance which ensured capital and usage as well as a solid foundation for data collection. They launched their first prototype in a classical and non-scalable way but worked in the way that they had something to measure and develop. The customers really liked the product and are today included in the process through a twelve-hour customer support all year around which plays an important role in the company's business. This also works as security for the farmers who quite often are unaccustomed to technology. There cannot be any changes to the

product without an open line to the users since it revolves around hunting. Both service and feedback wise. This hunting application usually tests technical features in low-cost, sequential and set based tests in a high-intensity, undefined and communicated experimentation environment.

Table 8 – The contextual factors of the hunting application

Size	Maturity	Product type	Experiment innovation	Geographical distribution	Customer participation	Newness to market	R&D intensity	Past Performance	Funding

5.1.7. The connecting influence marketers

This company acts on the market of influence marketing by connecting brands with influencers in social media and helping them become the hottest brands in social media. It has two stakeholders: customers and influencers. Social media influencer marketing is the practice of building relationships with the people who can build relationships for you. The company has a sort of influencer bank of bloggers and artists, amongst others, which they present a relevant segment from to a specific customer and case based on target group and campaign purpose. Today most of the processes are manual due to a complexity and uncertainty regarding the platform with the goal to digitalize more over time and eventually become an online platform complemented with a manual initial customer contact to build confidence and trust. The beta version of the platform was built after signing their first customer and is the reason for why they use mostly manual communication via personal meetings, phone calls, mail and Google Docs today in all projects whilst the platform is further developed with each customer and campaign. The seven employees have different responsibilities and previous experiences resulting in a relatively undefined way-of-working and enables set based parallel testing in correlation to the pressure of being a part of an accelerator for a limited period of time. These experiments and campaign outputs are evaluated by using a number of metrics such as the most important one, *commitment*, which is based on likes and comments. These metrics make it possible for them to present expected values to the customer before launching a campaign together. This has not always worked and is the reason for why the business plan has changed over time to focus on commitment, credibility and relevance rather than range. Due to the early maturity of the company and manual A/B testing changes happen fast and give direct results, such as a change in influencer preferences in a Google Doc; the next time the customer is on the phone and gives feedback, they will now if the experiment gave better results or not. Based on what kind of product the customer is selling a campaign generally runs for three to four weeks, sometimes even just a few days, but with multiple experiments conducted simultaneously as the campaign during this period of time. Prototypes are A/B-tested during real customer campaigns together with the customers to constantly improve the product which makes the experiments easy and quick to implement and evaluate. Even though the influence marketers invest a lot of capacity in their innovation progress and sample sizes are large, the experiment costs are low. Their financials are made up of both soft funding from Almi, convertible loan from when they entered STING and private equity capital.

Table 9 – The contextual factors of the influence marketers

Size	Maturity	Product type	Experiment innovation	Geographical distribution	Customer participation	Newness to market	R&D intensity	Past Performance	Funding

5.1.8. The self-learning photo sharing company

The photo sharing was founded in 2014 and is currently a part of the accelerator SSE Business Lab. The product is a smartphone application for sharing photos with any person or group of people just by using hashtags. It all began with the founder having a problem herself and started searching for other people encountering the same issue. She quickly realized that she cannot program and therefore built the first MVP in a couple of hours using another smartphone application to come up with a design mockup to use as a guide for external programmers at hackathons. Fifteen workshops, deep interviews and surveys with different people were conducted and came to the same conclusions. All through the journey of this company hypotheses have been used and reformulated. These processes up to this stage barely cost a thing. The next step was to build a prototype and test it on forty persons in her surrounding who came with feedback which resulted in a second prototype that took two weeks to build and cost SEK 10.000. After the second prototype and its launch in Apple's TestFlight Beta Testing tool the founder reached out to friends and family, even Google and universities, to find someone who could build the product. Eventually she ended up doing a very simple version herself that impressed investors and programmers and resulted in her being able to pressure prices and time limits. The theme in this company's processes is to always write hypotheses, keep things as simple as possible, test, collect feedback, measure and evaluate, rewrite hypotheses and reformulate the process. In this case feedback was of great importance since the founder claims it to never be enough to work in theory when it comes to startups, customers really need and will use their products. The founder says that they do not work as lean today as they did in the beginning due to the importance of feedback. Today they have a long list of to-dos that they need to catch up with. The innovation process decreases in speed and increases in complexity as the company and user group grow. To manage such a high intensity of R&D on large test subject groups in a small company with undefined processes, the experiments are conducted sequentially with a set based focus. Due to private investments only the main reason for their way of testing new innovations is to keep costs low.

Table 10 – The contextual factors of the photo sharing company

Size	Maturity	Product type	Experiment innovation	Geographical distribution	Customer participation	Newness to market	R&D intensity	Past Performance	Funding

5.1.9. The flexible food delivery company

This Finnish company began one and a half year ago with its first development process. Their goal was then to develop a mobile solution for eating at restaurants and take-out and started out with a Finnish entrepreneur picking out a team of six people. These six have heavy programming backgrounds and said that they together were going to create the best app for food since people will always want to have food. Six months later, they released their first app which then had attractive features but the problem was that people wanted to have their food delivered, not just ready for take-out. When the app instead was launched with a partner and now included home delivery they began to grow rapidly. The partner did not have enough structure which forced them to create their own operations. Recently, they have launched their application in Sweden with six people at the Swedish head office and around thirty people in total including drivers. Unfortunately, their initial research regarding laws and restrictions was not enough. This have forced them to adapt their business model

to start using MVPs of their way of hiring drivers and the supply of vehicles. They know the goals with their MVPs, which is to apply the Finnish business model, but to get there they have to use MVPs to make it fit on the Swedish market. With their solid venture capitalist backing they can afford to fail in the short run. Therefore, they perform multiple sub-experiments within a number of experiments exploring new features and services and implement these on real users. The costs of these experiments are however low due to the competences within the company that quickly can respond to certain demands and feedback from customers with just a few clicks. This even though the overall level of experiment innovation is relatively complex as a result of the new market requirements and the uncertainty.

Table 11 – The contextual factors of the food delivery company

Size	Maturity	Product type	Experiment innovation	Geographical distribution	Customer participation	Newness to market	R&D intensity	Past Performance	Funding

5.1.10. The disappointed delivery service

When one of the founders tried to transport a treadmill during a move he was disappointed in the existing alternatives so he started his own delivery service together with his friend. Their aim is to become the world's best delivery experience by enabling local companies to deliver to the customer's door. To overcome the issue with them lacking programming skills they brought a developer from an earlier project of theirs with them and had a first platform up and running in two weeks. Since launch they have won many awards and received venture capital from many sources. One of the first learnings was that one cannot do everything for everyone which they say is a classical startup mistake. In the beginning they used the drivers' own phones but realized that they could save up to three weeks by using company phones to send out new updates digitally. To counteract stress and delays they moved two additional developers to Stockholm and are now active in seventeen Swedish cities with their eyes on the international market. By using algorithms and route optimization in an otherwise conservative and non-digital industry their drivers can cut time by 50 % per hour. Due to huge successes the company is growing fast and putting pressure on themselves and constantly changes and updates processes. This has also been proven to work very well since they are growing with the organization as well as in complexity. According to the founders, startups need to understand that everything does not go right the first time but also that a startup needs the quest to always grow and improve even though the reality sometimes might state otherwise. Startups need to dare; build an MVP as quick as possible and test it. Today, this delivering company realizes that they have a fully operational system and that it cannot break down and therefore always include new customers in the process. Their innovation process is built up by two week sprints originating from the SCRUM framework from after they hired a SCRUM coach who taught them to better work as a growing team with clarified processes. An experiment usually takes two weeks which ends with a demo for the whole Stockholm office, they believe that transparency is the only way to succeed. "The whole point of employees is to employ people who are smarter than yourself and thereby develop." After the meeting the new version is sent out to all drivers during Friday night when no one is using the application for a trial period until Tuesday the week after. If the operations team gives thumbs up after those first days, the new version is evaluated and implemented permanently. All solutions must be general solutions since the company is aiming high but product changes are rarely big. The same goes for the customers, they are encouraged to chat with their quickly

responding customer support and give feedback when they have received the delivery. A challenge that comes from being a fast-growing organization is to always assure quality in all parts and to earn money in an industry with low margins. This startup never does something without a profit, because it is for them necessary. Their cost structure is relatively flat with set costs per hour and car. Daily metrics are used and stored in a data base for future use and development. Due to their structured way of working they can, even as a large business partner, be flexible in their complex and highly innovative experiments through parallel and set based processes. This, together with good relations between the office and drivers, also enables them to keep costs low.

Table 12 – The contextual factors of the delivery service

Size	Maturity	Product type	Experiment innovation	Geographical distribution	Customer participation	Newness to market	R&D intensity	Past Performance	Funding

5.1.11. The advanced portfolio company

The oldest company of the cases included in this study and has been around for over ten years. It is also the largest company with over 900 employees in roughly seventy companies. This case is slightly different compared to the other ten, but was nonetheless chosen to be included since this study is focusing on companies' way-of-working and therefore contributes with relevant and complementing data. Their products could be everything within bio-, clean-, medtech and advanced materials such as material solutions for industrial applications and pet food. They start with an idea from an external researcher, bring business competences to the project and find a customer to develop the product for using the idea from the researcher. The processes that follow are very similar to a startup and is another reason for why this case should be included. The customers are introduced to the project early on during the sketch phase to minimize risks by avoiding the development of something that the end-customers do not want. Thereafter they use iterative experiment processes, which they have named "the creative dance", to test different hypotheses and theoretical approaches before implementing their ideas on the physical products. Since they develop these kinds of products it is even more important for them to make sure that the product is truly working. Due to their complex experiment innovation and high-risk physical products in terms of biotech products that needs rigorous testing, experiment processes are slow and conducted in blind laboratory environments before field tested on large sample sizes. However, their venture funding and high R&D intensity enable parallel and set based testing that are thoroughly evaluated using metrics. As a result of the product types, costs are high even in early stages of the experiments and hence increasing the potential risk.

Table 13 – The contextual factors of the portfolio company

Size	Maturity	Product type	Experiment innovation	Geographical distribution	Customer participation	Newness to market	R&D intensity	Past Performance	Funding

5.2. Case findings

5.2.1. Maturity + experience = structure + parallel testing

"So we sit with our last money from the agency who invested hours working with us and try to scoop up what the next project is but have no idea what it will be yet" - The fragrance subscription, low level of maturity and innovation experience.

"I knew it would take years to build the entire product and therefore tried to break down the product in several steps. Thus came the idea to break down the development into three different parts" - The Simplified website tool, high level of maturity and innovation experience.

Table 14 below shows the findings of evident relationships in certain contextual factors' effect on experiment dimensions. The first result is the maturity of the company and its influence on the structure of experimental processes. The higher level of maturity of a company, the more structured and defined does experimental processes seem to be. This becomes clear when comparing the bowtie and fragrance companies with the website and portfolio companies. The first two companies mentioned have the lowest level of maturity, young companies, and are working with a very loose and unstructured process in their experiments. We can also see if we look at these similar cases that they evolve and reach a higher level of maturity as their experimentation processes gets more structured. However, the website tool and the portfolio company have a higher level of maturity in their current stage and present a well-defined and structured process by letting the experiments become the core of the entire company's innovation process.

The case findings also show that there is another contextual factor that affects these two dimensions of experiment process structure and testing type; past innovativeness performance. We find that a higher level of structure is more likely to be found in companies with more experience from conducting experiments in other previous contexts. The companies of the website tool and the delivery service have a high level of past innovativeness performance from earlier projects which influences the structure of their experimentation process and parallel testing. The fragrance company which has a low level of past performance within innovation lets their experiment process become more unstructured and sporadic. For some parts of their development they exclude experiments completely. In our sample of companies, we can see that the companies with a lower level of structured experiment only focus on the current experiment and do not look further beyond the potential results of that experiment. Companies with a higher level of structure do not only structure the experiment itself more but also the process of multiple experiments throughout the innovation process. The housing company, the stock market service and the website tool stand out with all variables alike the other cases except that they have undefined experimentation processes even though they are mature and experienced teams using parallel testing. This result shows that there is an effect from contextual factors on the dimensions which will be discussed further in the next chapter.

Table 14 – First result (also presented in the appendix as the result table)

CONTEXTUAL FACTORS	The Bowtie Company	The Fragrance Subscription	The Housing Company	The Stock Market Service	The Website Tool	The Hunting Application	The Influence Marketers	The Photo Sharing Company	The Food Delivery Company	The Delivery Service	The Portfolio Company
Maturity of the company											
Past innovativeness performance											
EXPERIMENT DIMENSIONS											
Experiment type (3)											
Experimentation process structure											

5.2.2. Size + innovation + distribution = sequencing

"What can we do to test it? We cannot buy in 350 scents and offer it. We do not have that money. We cracked the idea of creating fragrance themes or playlists just like Spotify." - The fragrance subscription, small company with limited resources, had to focus on cost and criticality

The empirical study, as presented in table 15 below, shows that there are three contextual factors and one experiment dimension that are associated by the same pattern; the size of the company, level of experiment innovation and geographical distribution as company factors and sequencing criteria regarding the dimension.

We find that smaller, simpler and co-located companies are more likely to have cost as a sequencing criteria while larger, more complex and scattered companies focus on uncertainty and criticality. When comparing the different cases, we can see that the bowtie and fragrance companies are very small in size, especially in their early stages, which forces them to focus on sequential testing based on cost, meaning they could only test one thing at a time due to lack of capacity in terms of time and money. Some of the smaller companies also, or instead, focus on criticality in their sequencing of experiments. We find that smaller companies have to prioritize their experiment sequence more in consideration of how critical the feature is that is being tested. If we compare this to the food delivery, and the portfolio companies, which have a much larger company size and a much more complex innovation level, we can see that their sequencing focus shifts towards uncertainty rather than cost and criticality. This mean that they can focus their experimenting on testing the features which have the highest uncertainty and rely the experiments afterwards based on the results. Empirics show that a company that is larger in size have the capability to overlook the more critical test in favor of testing features which have a higher level of uncertainty.

Table 15 – Second result (also presented in the appendix as the result table)

CONTEXTUAL FACTORS	The Bowtie Company	The Fragrance Subscription	The Housing Company	The Stock Market Service	The Website Tool	The Hunting Application	The Influence Marketers	The Photo Sharing Company	The Food Delivery Company	The Delivery Service	The Portfolio Company
Size of the company											
Experiment innovation											
Geographical distribution											
EXPERIMENT DIMENSIONS											
Sequencing criteria											

5.2.3. Physical = theoretical + blind tests + small samples

"I started with phone interviews in the summer, and tried to ask open-ended questions without asking straight out if you want to subscribe to perfume"- The fragrance company

"The process was that we had an idea, we had a framework that we wanted to do so we made a first version tested it on customers and thought well now, now we are ready, we just launch it here. The more people we met, the more we have expanded our product because of their feedback" - The social stock market service

The third finding presented in the empirics and in table 16 is that the dimension of experiment type, in terms of theoretical or prototype, is influenced by if the company is creating a physical or non-physical product. The results show that companies with a physical product are more likely to conduct theoretical experiments and companies with non-physical products tend to conduct experiments using prototypes. These findings have a relation with the dimension of whether the test group of an experiment is exposed to it using blind tests or if the test groups are aware of it. The findings show that companies with a physical product focus on blind tests and that the test subjects of virtual product companies are aware of the experiment when participating. The bowtie and fragrance companies are both producing physical products as the main focus of their business and use theoretical experiments and blind tests. The portfolio company uses the theoretical experimentation approach to all their experiments and uses blind tests in the beginning of their projects when conducting initial research but switches to aware test groups further on in the process. All other cases include non-physical and software-based products which are all using prototypes and test subjects who are aware of the experiment. We have also found empirical evidence that companies with a physical product test their experiments on smaller sample sizes, whilst non-physical product companies use larger sample sizes.

Table 16 – Third result (also presented in the appendix as the result table)

CONTEXTUAL FACTORS	The Bowtie Company	The Fragrance Subscription	The Housing Company	The Stock Market Service	The Website Tool	The Hunting Application	The Influence Marketers	The Photo Sharing Company	The Food Delivery Company	The Delivery Service	The Portfolio Company
Product type											
EXPERIMENT DIMENSIONS											
Type of unknown tested											
Experiment type (5)											
Sample size											

5.2.4. R&D intensive = set based + metrics

Patterns in table 17 also show that the level of R&D have a relationship with what technique the company is using to analyze the results from their experiments and whether the experiment type is point or set based. The first relationship is between a low level of R&D intensity and point based experiments, meaning it only tests one alternative. These are evaluated using subjective assessments. The other relationship is between a high level of R&D intensity, set based experiments meaning that they test multiple alternatives simultaneously. These are evaluated made using metrics.

Table 17 – Fourth result (also presented in the appendix as the result table)

CONTEXTUAL FACTORS	The Bowtie Company	The Fragrance Subscription	The Housing Company	The Stock Market Service	The Website Tool	The Hunting Application	The Influence Marketers	The Photo Sharing Company	The Food Delivery Company	The Delivery Service	The Portfolio Company
R&D intensity											
EXPERIMENT DIMENSIONS											
Experiment type (4)											
Technique to evaluate											

5.2.5. Customers | Evaluation | Test subjects | In vivo | Costs | Fast speed

These values presented in table 18 are unchanged between the different cases. Some of the contextual factors and experiment dimensions have been found to be constant for all cases and are therefore assumed to not have an effect on or to not be affected by other variables. These are the factor customer participation and the two dimensions regarding the importance of evaluation and test groups. This means that all eleven cases find it important to involve customers in the experimentation process using feedback/workshops/customer support/etcetera. In their development process they continuously evaluate the output of implemented experiments and use real test subjects to collect intrinsic and substantial data to evaluate.

Three additional dimensions follow the same pattern except in one or two cases which categorizes them with the unchanged and constant variables. These are choice of environment, cost of experiment and experiment speed. The reason for this variation will be discussed in the next chapter.

Table 18 – Fifth result (also presented in the appendix as the result table)

CONTEXTUAL FACTORS	The Bowtie Company	The Fragrance Subscription	The Housing Company	The Stock Market Service	The Website Tool	The Hunting Application	The Influence Marketers	The Photo Sharing Company	The Food Delivery Company	The Delivery Service	The Portfolio Company
Customer participation											
EXPERIMENT DIMENSIONS											
Choice of environment											
Cost of experiment											
Experiment speed											
Importance of evaluation											
Test subjects											

5.2.6. No evidences of patterns

For three contextual factors we cannot find empirical evidence that they affect the dimensions of the experiment process. These are the contextual factors newness to the market, funding and type of industry together with what type of unknown that is tested. We also cannot find empirical evidence for the variations in four of the experiment dimensions. The experiment dimensions that we cannot find any belonging patterns to experiment type regarding exploratory or verification and validation, fidelity and requirements.

6. Discussion

In this section we will analyze our results that we have presented in the chapter above. The results will be analyzed using our framework and the theories we have presented within the factors and dimensions. We will show that some of the factors have a relationship in our result model but we cannot draw a conclusion that there is a contingent relationship.

6.1. Maturity and innovation experience create structure

As our results show, there is a clearly defined influence from the maturity of the company and their past innovation experience on how defined their experiment process is. We find that companies with a lower level of past innovation experience focus on solving problems with a short term focus which may cause new problems in the future. On the opposite side, the firms with a higher level of experience and a much more structured process with long term focus. This is consistent with the findings from Griffin (Griffin 1997), who show that a project or a firm with lots of known inputs from an earlier version or project tends to simplify the product development process compared to when a company has to create everything from scratch which requires a greater level of creativity to handle problems and issues due to lack of knowledge. If one takes this further it means that mature companies, with their defined way of working, can handle multiple experiments conducted at the same time whilst young companies need to focus on one experiment at a time. This results in mature companies using parallel testing and young companies sequential testing within their experiments.

The fact that startups operates under conditions of extreme uncertainty and the fact that they have very limited resources may be the reason to why they have a very unstructured process. They have many unknowns that has to be answered but due to the extreme uncertainty, a result from one experiment can change the environment for the other planned experiments completely which makes

it hard for them to structure the process over time. The unique situation of startups can also be the reason for why we find some variation in the housing, stock market companies. They have undefined processes even though they have experience but we can see, looking at their process over time, that as they grow they get more and more defined processes. This further show that even though experience have an influence on the structure level of the experiment processes, the startup environment involves many uncertainties and therefore a larger amount of experience is required to have a more structured process. This is also supported by the maturity of the company from Paulk, Weber, Garcia, Chrissis and Bush (Paulk et al 2003) as we find just like them that a company with a higher level of maturity has a higher level of structure in their processes. Looking at this result through the glasses of our research question we can clearly see that there are differences between companies experiment dimension and that maturity and experience is something that have an impact on the experiment process in terms of structure.

6.2. For startups, size matters

The second result presented was that the sequencing criteria for the firms is associated with their company size and the innovation level within their experiments. Firms that have a smaller size uses cost and criticality when sequencing their experiments while larger and more innovative firms focuses on uncertainty instead. Looking at Damanpour's findings (Damanpour 1992) that a company's size affects innovation capability, we can see that our results are consistent with their findings by implying bigger firm has a greater ability to drive innovation. We can see that the smaller companies do not have the abilities to use uncertainty as sequencing criteria. This can be linked to the Damanpour's findings that larger firms have the ability to take more risks by a higher level of experiment innovation and handle unsuccessful experiments better financially. Smaller firms are completely dependent of validating experiments focusing on cost and criticality. They do not have the same abilities as larger companies to handle failures and therefore minimize risks by a lower and more confident innovation level. The extreme uncertainty of startups makes this even more difficult and adds more pressure on the companies to test the most critical features at the same time as they do not have the resources to conduct multiple experiments. A smaller company size can nonetheless be an advantage in the ever-changing environment of startups. Smaller firms can be more flexible and adapt to changes faster, a big advantage for startups. However, in this research most companies are still so small in terms of size that even the larger firms included in the study still have the ability to rapidly adapt to changes. We instead find that larger firms have the financial abilities to fail and according to Eric Ries (2011), failure is what drives learning when conducting experiments.

Looking at table 15, we can see that there is one additional contextual factor that seems to have a connection to the size of the company and guts to support innovativeness; geographical distribution. One example that contradicts this finding, that implies larger companies to be more geographically scattered whilst smaller companies sit together, is the fact that the large delivery company actively chose to move two programmers up to their head office in Stockholm from Malmö in order to keep the experiment process more efficient. Their product innovation is complex but is built up of an administrative and progressive foundation at their head office with drivers all around the country which makes their business scattered. The core of the company is still co-located which objects with the research of Ågerfalk and Fitzgerald that larger projects and companies need to grow geographically in order to handle new challenges such as communication, coordination and control.

This might be something that these companies will evolve into in the future but currently not a result that can be ensured based on the findings in this study.

This finding show that the context of a company is something that is complex and makes it difficult to answer our research question. We see that size have an effect on the sequencing factor which shows that the differences are affected by the context. However, at the same time there are similar patterns in the geographical location factor but we cannot find empirical evidence to support that that factor is affecting a specific dimension.

6.3. Differences between physical and virtual products

The third finding based on the case interviews is the contextual factor of product type and its effect on the three dimensions; sample size, experiment type (1) and (2). According to Dickson and Ginter (1987), the physical or non-physical feature of product type should affect the innovation process through segmentation strategy. By that they mean customers find other product characteristics to be of varied importance and value them differently, but little previous literature clarifies exactly how the perceptions should and do differ. This relationship is nonetheless evident when looking at the test subject awareness which for companies with physical products has been blind tests and aware test groups for companies with non-physical products. When a customer tests a physical product, he or she subconsciously adjusts their perception of it. Therefore, these companies benefit from conducting experiments on test subjects that are not aware of the specific experiment characteristics. The eight cases with non-physical products included in this study unambiguously show the relationship between non-physical products and aware test subjects, test groups who are aware of the experiment. On the one hand, no correlation in previous research can be found that supports this pattern of physical products initially being theoretically tested on smaller sample sizes and prototypes of non-physical products are tested on larger sample sizes. On the other hand, Symanowitz (2013) says that as a sample size gets bigger and bigger during a company's development, they are likely to get closer and closer to the true underlying value, with less and less variability. The CMO of the housing company also said that software companies have the luxury to in an extremely fast and smooth way go back to an earlier version if a new feature fails whilst the same do-it-right process for physical product companies can take weeks. This might make test processes longer and more expensive which the majority of startups cannot afford. It could also be the reason for why companies, such as the bowtie and fragrance companies, chose to make a theoretical first version of the product instead of making a prototype straight away to reduce the potential risk that follows after the experiment implementation. Symanowitz's words together with the ones from the housing CMO therefore give support to the result that the contextual factor of product type affects the experiment dimensions of theoretical vs. prototype, blind test vs. aware test subjects and the sample size. Looking at this finding in connection to our research question this provides further evidence that there are contextual factors that have an effect on the certain experiment dimensions.

6.4. High R&D intensity equals parallel testing

The fourth result finds that there is an influence from the contextual factor R&D intensity on the dimension experiment type (4) which is divided between point or set based testing. When looking at theories this relationship becomes quite obvious. Like Wesley, Cohen and Levinthal (1989) presents, R&D does not only exist to generate a new product but also to gather external information from the

environment it operates. Information that could be extremely valuable in the changing environment startups operate in. Since startups wants to generate as much information as possible through their experiments to make sure they develop a product the customers actually want we can see that the influence from the intensity of R&D on point or set based testing is strong. Companies that focus on having a high intensity of R&D are more likely to use parallel testing rather than sequential testing. They can maximize the output of their innovation processes by enabling capacity to concentrate on multiple sub-experiments researching one hypothesis. Less R&D intensive companies cannot divide the capacity available between more than one sub-experiment at a time, hence the point based testing.

The other dimension that table 16 shows to be affected by the level of R&D is the techniques companies use to analyze data generated from experiments. Some companies mainly use subjective assessments whilst other companies mostly use metrics which is what is shown in table 17. However, based on case findings all companies use metrics to some extent. Some to see their long-term progress according to a few numerical values and some specific figures of merit to analyze their short-term results from certain experiments. The connection is therefore too diffuse to be validated and as a result this dimension cannot be explained by the level of R&D intensity only.

When using these results to try and answer our research question we once again find that it is a complex question to answer. These findings show that there are contextual factors that have an effect on the experiment dimension but once again we find patterns that seems to be affecting a specific dimension but we cannot find enough empirical evidence that supports that the effect on the experiment dimension, techniques to analyze data is affected by the contextual factor R&D intensity as the pattern in table 16 suggests.

6.5. Constant patterns

In our results we find some contextual factors that are constant for all cases but still does not affect any particular experiment dimension. These are customer participation, importance of evaluation and test subjects. Looking at the literature review we find that all of the reviewed frameworks highlights the importance of involving the customers in the experiment process and that is something we find that all cases that have been studied also do however we do not find that this influences any experiment dimension but is rather something that the companies have to do to gain validated learning which Ries (Ries 2011) also emphasizes. To gained the validated learning they have to know exactly what the customers think and that is why all the cases have the same level of importance of evaluation, without the evaluation there is no point in including the customer. This also explains why all cases have the same type of test subjects as real persons, since they conduct their tests on their customers.

There are also three experiment dimensions that are the same for all cases except one or two. Firstly, the choice of environment is constant for all cases where everyone is using field experiments except for the portfolio company which conducts more laboratory experiments. This is mainly because of them working with high risk bio-, clean- and medtech products that has to be validated before they can be tested on customers. Secondly, the cost of experiments is low for all the companies except for the portfolio company which has relatively high experimental costs due to the same reason as for why they conduct laboratory experiments; a great complexity of product innovation together with

high R&D intensity which results in high risks. This means they have to conduct more experiments of a higher complexity before they can launch their products. Thirdly, the experiment speed is constant and fast for all cases except for the hunting application and the portfolio company which both have a slower experiment speed. Based on the two relatively constant dimensions mentioned above together with this third one of experiment speed make it clear that the portfolio company is an exception to some patterns due to its product characteristics. The reason for the hunting application's slow experiment processes is explained by their season-based product which hinders them to test their product all-year-round by only activating customers during the hunting season. Another reason for the varied experiment speed was presented by the CMO of the housing company who said that startups do not have enough capacity, in terms of time and capital, to conduct longer experiments due to their industry. Even though these two companies stand out due to their product characteristics, there are different

Some markets change rapidly which would make experiments conducted over a longer period of time insignificant by the time they are implemented and evaluated. However, the correlation between experiment speed and what industry a company conducts business within is not visible in the majority of the cases in this study and cannot be concluded as an evident result. The same goes for the cases of the hunting application and the portfolio company where different product characteristics are affecting the protruding relationships and therefore cannot be substantiated as an evident pattern.

7. Conclusions

7.1. One size does not fit all

In this research we study the dimensions of experiments within product development. We focus on investigating if and how companies are influenced by contextual factors that affect how their dimensions differ from each other. This has been done by reviewing the current literature about innovation and product development through experiments which have shown that today's literature provides generic tools that are not adapted to differences in companies' contextual factors. Limited research has been found that the context of the company is an important factor that affects the product development process (Kruchten 2011; Browning et al 2006). After that a theoretical framework was developed as a tool for analyzing the companies' contexts and how they have conducted experiments in their product development process.

We find empirical evidence that there are influences from some of the contextual factors on specific dimension of the experiments. The results show that companies with a higher level of maturity and previous innovation experience have a more structured experiment process and at the same time uses an experiment type that conducts multiple experiments simultaneously rather than sequential experiments whilst companies that have a lower level of maturity have a more unstructured process and conducts sequential experiments. We also find that size of the company has an effect on what type of sequencing criteria that is used for the experiments. Smaller companies focus more on cost and criticality in their sequencing whilst larger companies have the capability to overlook the more critical test in favor of testing features which have a higher level of uncertainty. Additionally, we find that physical products are more likely to conduct theoretical experiments using smaller sample sizes

using blind tests while companies with non-physical products create prototypes to test features on larger sample sizes that are aware of the experiments. Our last finding is that companies with a high level of R&D intensity are more likely to conduct multiple sub-experiments within bigger experiments, set based, whilst companies with less intensive R&D investments focus on one sub-experiment at a time, thus point based.

We also find some contextual factors that are constant for all cases without affecting any specific dimension as well as the test subjects. Customer participation, the importance of evaluation and real test subjects are all three the same for all cases and can be seen as the foundation for startups that use experiments for innovation and product development. This is also emphasized by the NNPDG, LPD, DT, agile innovation and LSU which we assume is the reason for it being found in all cases. Startups also generally seem to keep their costs of experiments low, conduct them in the field and analyzing the results using both metrics and their gut feeling but cannot be completely assured on the basis of this study.

To sum up our findings and reconnect to our research question, we want to highlight that our findings show that there are differences in the way companies conduct experiments that are affected by their context. Just as the frog and the octopus let us know in the introductory story; one size does not fit all. However, we also find patterns in some contextual factors that seem to affect specific experiment dimension but our empirical evidence cannot prove that there is a contingent relationship for these. This shows the great complexity of evaluating company contexts and that this is something that need to be researched further.

However, the differences we find show that the generic tools that are presented by the different frameworks we have reviewed should be used with caution if they are used "out of the box" and companies should try and adapt them according to their context and not use carbon paper to copy the processes.

Table 19 - Evident relationships between contextual factors and experiment dimensions

	Sequencing criteria	Experiment type (1)	Experiment type (3)	Experiment type (4)	Experiment type (5)	Structure of process	Sample size
Size							
Maturity							
Product type							
R&D intensity							
Past performance							

Table 20 - Definition of colour coding in table 19

	Yes, the experiment dimension is influenced by the contextual factor
	No, the experiment dimension is not influenced by the contextual factor

7.2. Implications

This exploratory study has contributed with empirical result that in accordance with the purpose of this research have moved the knowledge forward to try and fill the identified knowledge gap. The findings stress the fact that it is important for companies to be aware of what context they are in when they conduct experiments in their product development processes. This could be particularly valuable knowledge for managers in startups that are using this type of product development strategy since they are operating under conditions of extreme uncertainty.

The framework provided in this study could be useful for managers to look at when deciding what type of strategy they should use for their experiments. But also to use to determine what context they operate in.

7.3. Future research and recommendations

For future research on this topic we recommend a longer time period for data gathering, both qualitative and numerical data. By doing this a greater number of variables can be measured and function as a metric. The goal should be to; in a clearer and more detailed manner paint a picture of what the actual results say and mean and only then ensure how newly started companies should test their products.

Future research should also try and test this research question on companies that have grown larger and are not classified as startups any more. When doing that it would also be suitable to research these across different industries.

Finally, we suggest that future research should investigate how the financing of the company is affecting the process: "A scientist is driven by reaching equity to develop his or her knowledge, an entrepreneur is driven by the knowledge to reach equity." - Amin Omrani, The Advanced Portfolio Company.

8. References

Ahmed, Pervaiz K. "Culture And Climate For Innovation". European Journal of Innovation Management 1.1 (1998): 30-43. Web.

Almeida, Paul, Gina Dokko, and Lori Rosenkopf. "Startup Size And The Mechanisms Of External Learning: Increasing Opportunity And Decreasing Ability?". *Research Policy* 32.2 (2003): 301-315. Web.

Ahrne, Göran and Peter Svensson. *Handbok I Kvalitativa Metoder*. Malmö: Liber AB, 2011. Print.

Andersen, Ib. *Den Uppenbara Verkligheten*. Lund: Studentlitteratur, 1998.

Audretsch, David B and A. R Thurik. *Innovation, Industry Evolution, And Employment*. Cambridge, U.K.: Cambridge University Press, 1999.

Basit, Tehmina. "Manual Or Electronic? The Role Of Coding In Qualitative Data Analysis". *Educational Research* 45.2 (2003): 143-154. Web.

Bianchi, Mattia et al. "Organizing For External Technology Commercialization: Evidence From A Multiple Case Study In The Pharmaceutical Industry". *R&D Management* 41.2 (2011): 120-137. Web.

Bosch, Jan. "Building Products As Innovation Experiment Systems". *Software Business*. 1st ed. 2016.

Brown, Tim. "Design Thinking". *Harvard Business Review* 6 (2008): n. pag.

Browning, Tyson R., Ernst Fricke, and Herbert Negele. "Key Concepts In Modeling Product Development Processes". *Systems Engineering* 9.2 (2006): 104-128. Web.

Bryman, Alan and Emma Bell. *Business Research Methods 3rd ed.* Oxford: Oxford University Press, 2011. Print.

Cohen, Wesley M. and Daniel A. Levinthal. "Innovation And Learning: The Two Faces Of R&D". *The Economic Journal* 99.397 (1989): 569. Web.

Damanpour, F. "Organizational Size And Innovation". *Organization Studies* 13.3 (1992): 375-402. Web.

Dickson, Peter R. and James L. Ginter. "Market Segmentation, Product Differentiation, And Marketing Strategy". *Journal of Marketing* 51.2 (1987): 1. Web.

Drake, Stillman. *Essays On Galileo And The History And Philosophy Of Science, Volym 3*. Toronto: University of Toronto Incorporated, 1999.

Eisenmann, Thomas, Eric Ries, and Sarah Dillard. "Hypothesis-Driven Entrepreneurship: The Lean Startup". *Harvard Business Review* (2011): n. pag.

Ellram, Lisa. "The Use Of The Case Study Method In Logistics Research". *Journal of Business Logistics* 17.2 (1996): n. pag.

Griffin, Abbie. "The Effect Of Project And Process Characteristics On Product Development Cycle Time". *Journal of Marketing Research* 34.1 (1997): 24. Web.

Griffith, W. Thomas. *The Physics Of Everyday Phenomena*. Dubuque, IA: Wm. C. Brown Publishers, 1992.

Hage, Jerald and Henry Mintzberg. "The Structuring Of Organizations: A Synthesis Of Research.". *Administrative Science Quarterly* 25.3 (1980): 547. Web.

Holmstrom Olsson, Helena et al. "Global Software Development Challenges: A Case Study On Temporal, Geographical And Socio-Cultural Distance". University of Limerick Institutional Repository, 2016.

Kruchten, Philippe. *The Frog And The Octopus - Experience Teaching Software Project Management*. 1st ed. Vancouver: University of British Columbia, 2011.

Kruchten, Philippe. "Contextualizing Agile Software Development". *Journal of Software: Evolution and Process* 25.4 (2011): 351-361. Web.

Miles, Matthew B. "Qualitative Data As An Attractive Nuisance: The Problem Of Analysis". *Administrative Science Quarterly* 24.4 (1979): 590. Web.

Morgan, James M and Jeffrey K Liker. *The Toyota Product Development System*. New York: Productivity Press, 2006.

Morse, J. M. "Determining Sample Size". *Qualitative Health Research* 10.1 (2000): 3-5. Web.

Paulk, M.C. et al. "Capability Maturity Model, Version 1.1". *IEEE Software*. 10.4 (1993): 18-27. Web.

Poolton, Jenny and Ian Barclay. "New Product Development From Past Research To Future Applications". *Industrial Marketing Management* 27.3 (1998): 197-212. Web.

Ries, Eric. *The Lean Startup*. New York: Crown Business, 2011.

Rothaermel, Frank T. and David L. Deeds. "Exploration And Exploitation Alliances In Biotechnology: A System Of New Product Development". *Strat. Mgmt. J.* 25.3 (2004): 201-221. Web.

Schwaber, Ken. "SCRUM Development Process". *Business Object Design And Implementation*. J.V. Sutherland. 1st ed. Texas: N.p., 1995. 117-134.

Shan, Weijan, Gordon Walker, and Bruce Kogut. "Interfirm Cooperation And Startup Innovation In The Biotechnology Industry". *Strat. Mgmt. J.* 15.5 (1994): 387-394. Web.

Sherman, H. and D. S. Chappell. "Methodological Challenges In Evaluating Business Incubator Outcomes". *Economic Development Quarterly* 12.4 (1998): 313-321. Web.

Snyder, Carolyn. *Paper Prototyping*. San Francisco, Calif.: Morgan Kaufmann, Elsevier Science, 2003.

Software Engineering Institute. *Key Practices Of The Capability Maturity Model Version 1.1*. Pittsburgh: Carnegie Mellon University, 1993.

Sosna, Marc, Rosa Nelly Trevinyo-Rodríguez, and S. Ramakrishna Velamuri. "Business Model Innovation Through Trial-And-Error Learning". *Long Range Planning* 43.2-3 (2010): 383-407. Web.

Stige, B., K. Malterud, and T. Midtgarden. "Toward An Agenda For Evaluation Of Qualitative Research". *Qualitative Health Research* 19.10 (2009): 1504-1516. Web.

Takeuchi, Hirotaka and Ikujiro Nonaka. "The New New Product Development Game". *Harvard Business Review* 1 (1986): n. pag.

Thomke, Stefan and Jim Manzi. "The Discipline Of Business Experimentation". *Harvard Business Review* 92.12 (2014): 70-79.

Venkataramanan, M. 2015, "*Europe's hottest startups 2015: Stockholm*", Wired UK, no. 9, pp. 26-02-2016.

Wharton School, 2015, *"How Stockholm Became A 'Unicorn Factory'"*. Forbes, N.p. 23-02-2016.

Yin, Robert K. *Case Study Research*. Thousand Oaks, Calif.: Sage Publications, 2003.

9. Interviews

Bengtsdahl, Tobias. *Main study*. 2016-04-25. Interviewed in person by Elin Grahn & Erik Gruvfors. Company office, Stockholm.

Bernadotte, Fredrika. *Pilot/Main Study*. 2016-03-16. Interviewed in person by Elin Grahn & Erik Gruvfors. SSE Business Lab, Stockholm.

Gradin, Alexander. *Main study*. 2016-04-18. Interviewed in person by Elin Grahn & Erik Gruvfors. SUP46, Stockholm.

Grahn, Ebba. *Main study*. 2016-04-20. Interviewed in person by Elin Grahn & Erik Gruvfors. Company office, Stockholm.

Hamilton, Fredrik. *Main study*. 2016-04-26. Interviewed in person by Elin Grahn & Erik Gruvfors. Company office, Stockholm.

Hoffman, Henrik. *Pilot/Main study*. 2016-03-24. Interviewed in person by Elin Grahn & Erik Gruvfors. SUP46, Stockholm.

Höglund, Joakim. *Main study*. 2016-04-21. Interviewed in person by Elin Grahn & Erik Gruvfors. SUP46, Stockholm.

Kjellqvist, Elisabet. *Main study*. 2016-04-20. Interviewed in person by Elin Grahn & Erik Gruvfors. SUP46, Stockholm.

Omrani, Amin. *Pilot/Main study*. 2016-03-18. Interviewed in person by Elin Grahn & Erik Gruvfors. Company office, Stockholm.

Söderlund, Tom. *Pilot/Main study*. 2016-03-16. Interviewed in person by Elin Grahn & Erik Gruvfors. SSE Business Lab, Stockholm.

Wetter, Erik. *Pilot study*. 2016-03-16. Interviewed in person by Elin Grahn & Erik Gruvfors. SSE Business Lab, Stockholm.

10. Appendix

10.1. Table of results

	Type of Unknown Tested	Sequencing Criteria	Experiment Type (1)	Experiment Type (2)	Experiment Type (3)	Experiment Type (4)	Experiment Type (5)	Choice of Environment	Cost of Experiment	Experimentation Process Structu	Experiment Speed	Fidelity	Importance of Evaluation	Requirements	Sample Size	Technique to Analyze Experimental	Test Subjects
The Bowtie Company	*Market*	*Cost*	*Theoretical*	*Exploratory testing*	*Sequential*	*Point based*	*Blind tests*	*In vivo*	*Low*	*Undefined*	*Fast*	*High-fidelity*	*High*	*Constant*	*Small*	*Subjective assessments*	*Real*
The Fragrance Subscription	*Market*	*Cost*	*Theoretical*	*Exploratory testing*	*Sequential*	*Point based*	*Blind tests*	*In vivo*	*Low*	*Undefined*	*Fast*	*Low-fidelity*	*High*	*Constant*	*Small*	*Subjective assessments*	*Real*
The Housing Company	*Technical*	*Cost/Criticality*	*Prototype*	*Verification & Validation*	*Parallel*	*Set based*	*Aware*	*In vivo*	*Low*	*Undefined*	*Fast*	*High-fidelity*	*High*	*Changing*	*Large*	*Metrics*	*Real*
The Stock Market Service	*Technical*	*Cost/Criticality*	*Prototype*	*Verification & Validation*	*Parallel*	*Set based*	*Aware*	*In vivo*	*Low*	*Undefined*	*Fast*	*Low-fidelity*	*High*	*Constant*	*Large*	*Metrics*	*Real*
The Website Tool	*Technical*	*Criticality*	*Prototype*	*Verification & Validation*	*Parallel*	*Set based*	*Aware*	*In vivo*	*Low*	*Undefined*	*Fast*	*High-fidelity*	*High*	*Constant*	*Large*	*Metrics*	*Real*
The Hunting Application	*Technical*	*Cost/criticality*	*Prototype*	*Verification & Validation*	*Sequential*	*Set based*	*Aware*	*In vivo*	*Low*	*Undefined*	*Slow*	*High-fidelity*	*High*	*Constant*	*Large*	*Metrics*	*Real*
The Influencer Markets	*Technical*	*Cost/criticality*	*Prototype*	*Verification & Validation*	*Sequential*	*Set based*	*Aware*	*In vivo*	*Low*	*Undefined*	*Fast*	*High-fidelity*	*High*	*Changing*	*Large*	*Metrics*	*Real*
The Photo Sharing Company	*Technical*	*Cost*	*Prototype*	*Verification & Validation*	*Sequential*	*Set based*	*Aware*	*In vivo*	*Low*	*Undefined*	*Fast*	*High-fidelity*	*High*	*Constant*	*Large*	*Metrics*	*Real*
The Food Delivery Company	*Market*	*Uncertainty*	*Prototype*	*Exploratory testing*	*Parallel*	*Set based*	*Aware*	*In vivo*	*Low*	*Defined*	*Fast*	*High-fidelity*	*High*	*Changing*	*Large*	*Metrics*	*Real*
The Delivery Service	*Technical*	*Uncertainty*	*Prototype*	*Verification & Validation*	*Parallel*	*Set based*	*Aware*	*In vivo*	*Low*	*Defined*	*Fast*	*High-fidelity*	*High*	*Constant*	*Large*	*Metrics*	*Real*
The Portfolio Company	*Technical*	*Uncertainty*	*Theoretical*	*Exploratory testing*	*Parallel*	*Set based*	*Blind tests*	*In vitro*	*High*	*Defined*	*Slow*	*Low-fidelity*	*High*	*Changing*	*Large*	*Metrics*	*Real*

10.2. Pilot study – fundamental interview script

- What does experimentation mean to you?
- What are, according to you, the most important parts?
- In what way do you implement this in the processes of your company?
- What are your thoughts on the way-of-working and outcome?
- Are you, and if so why, using specific frameworks? Have you tried any other framework?
- Do you have any general comments regarding this topic?

10.3. Main study – fundamental interview script

- What does your innovation process look like today? Has it always looked the same?
- Who are involved in this process?
- How do you usually test your product and new experiments?
- How do you evaluate your experiments? What do you use to evaluate the experimental outcome?
- What is the time frame for a typical experiment from idea to launch?
- Has this process changed over time? E.g. fewer iterations, bigger loops.
- If yes to the question above, what has been the reason for change?
- What has worked out well/not worked out?
- What is the next step in the business plan/experimentation process?
- Within the company, do you have any previous experiences and if yes, in what area?
- How often do you release a new version of the product to your customers?
- How often do you evaluate and reflect upon the output?
- Do you document all changes and experiments?
- What does the financials look like? Venture capital, private equity, divided between both.
- Basic company characteristics and the company story.