Refining Big Data

The impact of data analytics in the retail industry

Abstract

We live in the most dynamic and expansive age of all time with immediate and virtually unrestricted access to information. This is changing the business landscape in the way companies interact, operate and make decisions. In this golden age of data, successful companies are able to refine data into valuable insights to reveal the most profitable path to every customer. Despite Big Data's endless opportunities, the research within the area is scarce. This study has its focus on the retail industry with its access to large volumes of rich consumer data creating untapped potential. The purpose of this study is to provide a deeper understanding of how Swedish retailers are exploiting Big Data, especially within marketing. A qualitative case study consisting of eight in-depth interviews with consumer insight experts from some of Sweden's largest retailers was conducted. The results show that Swedish retailers are only capturing a fraction of the potential value of Big Data. Out of the four components in the marketing mix, promotion was shown to receive far more attention than the others. Furthermore, all retailers were found to possess dynamic capabilities and only a few have the prerequisites to develop adaptive capabilities. Whether these two findings are connected will require further research.

Keywords

Big Data, marketing capabilities, retail industry, resource-based theory

Authors

Emma Wikefeldt, 23376
Stina Itzel Westerberg, 23301

Presented

May 23, 2017

Tutor

Per Andersson

Examiner

Gustav Almqvist

Acknowledgements

Table of Contents

Definitions

Data mining: The process of refining collected data into insights.

Omnichannel: A cross-channel business model used by companies to improve the customer experience.

Transactional data: Data that is derived directly from business transactions.

Customer loyalty program: A company community where customers can gain offers and discounts.

Legacy systems: Outdated computer systems that may still be in use because its data cannot be changed to newer or standard formats, or since its application programs cannot be upgraded.

Behavioral data: Data generated by or in response to a customer's engagement with a business.

Weblog: A website that in chronological order displays the postings by one or more individuals and usually has links to comments on specific postings.

1. Introduction

"Data is the new oil" - Shivon Zilis, Bloomberg Beta

Data is said to be the new oil of the 21st century. In its raw form, it is quite worthless but for the ones who understand how to refine it, there are tremendous value to be captured. The difference between the two is however that data is no longer a scarce resource. It has been said that the era of Big Data began at the point at which the cost of storing data dropped below the cost of deleting it (Hofacker et al., 2016). Today 20% of the world's information is stored on the internet, while the other 80% is privately stored within companies and organizations (Fortune, 2016). For decades, companies have been making business decisions based on transactional data stored in relational databases. However, beyond this structured data lies a potential gold mine of unstructured data: weblogs, social media, email, sensors and photographs that can be used to retrieve useful information.

When Big Data is processed and analyzed in combination with structured data, enterprises can develop a more thorough and insightful understanding of their business, which can contribute to enhanced productivity, a greater competitive advantage and increased innovation. Foremost, it can provide fresh insights about customer behavior and identify market trends early on, making big implications for the marketing departments within organizations. Raconteur (2013) is writing that "Big Data is turning marketing inside out and smart companies are making sure their marketing campaigns are founded on figure". Marketing executives with access to rich, insightful and accurate data are able to make better decisions based on evidence at a given time, rather than on intuition or laboratory-based consumer research.

But the availability of massive volumes of new data also creates challenges. The retail industry has been using structured data in business decisions for years, and is according to McKinsey & Company one of the industries that have progressed the most within data analytics during recent years. In the private sector it is further estimated that a retailer using Big Data to its fullest capacity has the potential to increase its operating margin by more than 60 percent (Manyika et al. 2011). With the access to high volumes of consumer Big Data, retailers have unlimited opportunities to make informed marketing decisions based on insights and offer more customer oriented solutions. However, how retailers are exploiting the opportunities that arise with Big Data is today unknown, especially in Sweden. We therefore find it interesting to study to what degree Swedish retailers are exploiting the potential of Big Data in their marketing decisions. Further, it becomes interesting to investigate if any barriers in particular are restraining the retailers and what capabilities will be required for further development.

1.1 Purpose and research questions

The purpose of this study is to provide a deeper understanding of how Swedish retailers are exploiting Big Data, especially within marketing, and to reduce the current research gap within the area. This will be examined by answering the following research questions:

R1: How is Big Data influencing marketing and value creation in retail?

R2: What are the major hindrances for Big Data exploitation amongst retailers today?

R3: What marketing capabilities are required of retailers for further development?

In order to answer these research questions, we will investigate how some of the largest retailers in Sweden are using Big Data in their marketing practices today.

1.2 Delimitations

Due to practical reasons Sweden was selected as the geographical area to study. And since we believe the interviews would provide the most substance when performed face-to-face, the study was limited to retailers with headquarters in Stockholm, when possible. Since all retailers distribute their products and collect data nationwide, we assume that this will not significantly affect the results.

Since the study is based on interviews from only eight companies, the sample is not representative for all Swedish retailers. Also, since only one employee per company has been interviewed, the results are based on this employee's point of view and cannot be generalized to represent the entire company. This is partly due to time limitations, both from our and the companies side, but in most cases also because the retailers only have one employee working within the area of our research. The study is therefore limited in terms of how other departments within the companies are using Big Data. Had the study been expanded for a longer period of time, it would have been of interest to include several more interview objects to get a broader perspective of our research area. However, since the interviews were conducted with the employee most knowledgeable within the area we believe to have gained valuable insights despite the limited time.

1.3 Expected contribution

The main expected contribution with this study is to develop a deeper knowledge and understanding of how Swedish retailers are exploiting Big Data. We will also give recommendations on how retailers can create value by overcoming the barriers identified in the empirical study.

1.4 Defining consumer Big Data

The characteristics of Big Data were first described by Laney (2001) who identified three key attributes of large data amounts: variety, volume, and velocity, commonly known as the three Vs (IBM, 2012; Lycett, 2013; Oracle, 2012; Erevelles et al., 2015). These are defined as:

- *Volume* refers to the "bigness" property (Hofacker et al., 2016). Enterprises are overwhelmed with ever-growing data of various types, reaching terabytes, even petabytes, of information.

- *Velocity* refers to the rate at which the digital processes makes Big Data even bigger (Hofacker et al., 2016) or the continuous rapidity of data creation (Erevelles et al., 2015).

- *Variety* refers to new formats and types of data. This new data consists of words, images, videos or other non-numeric consumer output, which is not suitable for traditional statistical analysis (Hofacker et al., 2016).

Although the three Vs are used to define and differentiate consumer Big Data from large-scale data sets, two additional Vs are important when collecting, analyzing and extracting insights from Big Data: veracity and value (Ebner, Bühnen, & Urbach, 2014; Lycett, 2013). Veracity emphasizes the need to be aware of data quality (IBM, 2012). Thus, the veracity of Big Data is a major issue at a time where the volume, velocity, and variety of data are constantly increasing (IBM, 2012; Oracle, 2012). The ever-increasing amounts of Big Data lead to the question of value. The task is to eliminate unimportant and irrelevant data, so that the remaining data is useful. The remaining applicable data need to be valuable for obtaining insight and area-specific interpretation (Lycett, 2013). The challenge is to identify what is relevant and then rapidly extract that data for timely analysis (Oracle, 2012). Taking these different extensions of the original definition into account, (Ebner, Bühnen, & Urbach, 2014) define Big Data as a:

"phenomenon characterized by an ongoing increase in volume, variety, velocity, and veracity of data that requires advanced techniques and technologies to capture, store, distribute, manage, and analyze these data."

2. Theoretical framework

In the following section we will present the framework used to analyze the empirical findings and answer our research questions. Part 2.1 presents the framework chosen and part 2.2 explains how we have applied it in our study.

2.1 Resource-based theory (RBT)

Resource-based theory provides an explanation of Big Data's impact on marketing. RBT explains how a company's resources combined with various data structures can generate a sustainable competitive advantage. (Erevelles et al., 2015). The framework introduces three different types of data; structured data such as scanner or sensor data, records, files, and databases which have been available to marketers for a long time. Unstructured data include textual data e.g. blogs and text messages and non-textual data e.g. videos, images and audio. The majority of unstructured data is collected from social media. Semi-structured data consists of various types of software capable of organizing the unstructured data.
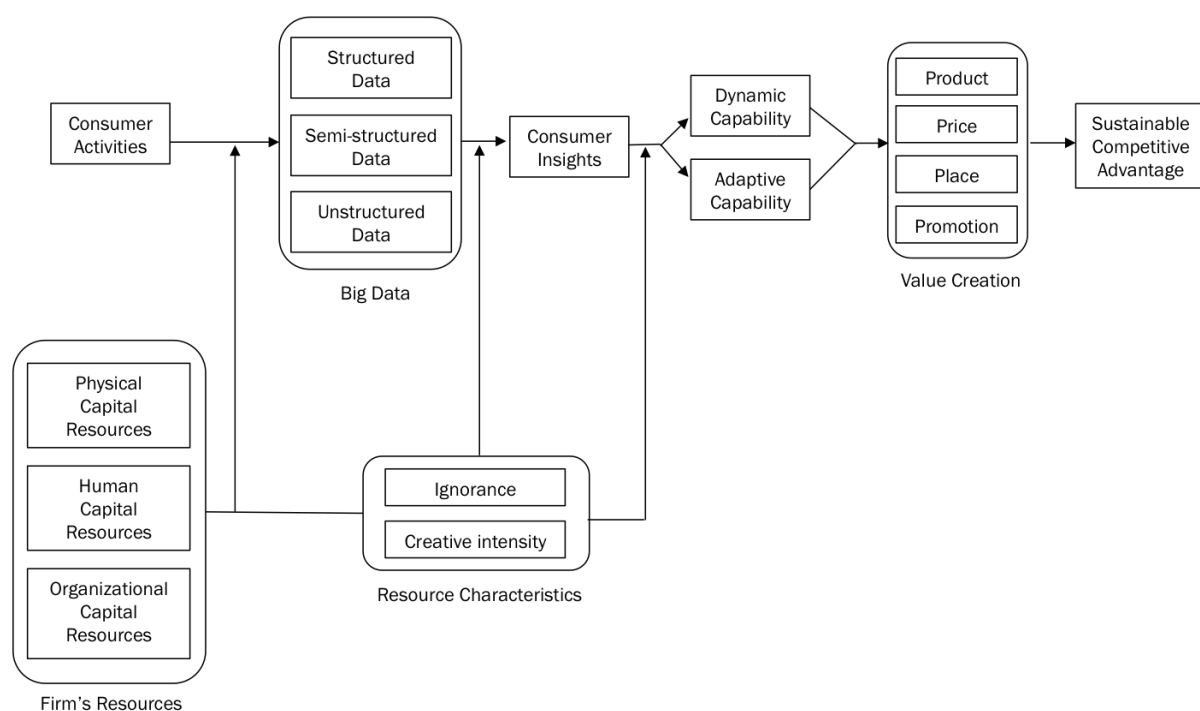
Figure 1. Resource-based theory

The framework further suggests that both tangible and intangible resources are importing for storing and analyzing various kinds of data. Following, it is important for firms to be aware of what they know and what they do not know which Erevelles et al. (2015) refers to as ignorance. The framework also state that creativity is crucial when firms are handling Big Data as they must find new ways of creating and implementing the insights into marketing activities. Thereafter, a firm's potential to analyze and implement the insights drawn from the data depends on whether the firm has dynamic or adaptive capabilities, which will be explained further in section 2.2.3. Enhanced by insights from Big Data, adaptive and

9

dynamic capabilities lead to value creation through the different components of the marketing mix. Finally, the value creation may result in a sustainable or temporary competitive advantage.

2.2 Focus areas of the framework

The framework has its major focus on what research approach (deductive or inductive) that should be applied when approaching Big Data and does not focus on how Big Data actually can create value or how to measure it. Hence, we have chosen to mainly focus on the later part of the model and look into which of the four components in the marketing mix that creates the most value when using Big Data analytics. We will also investigate if the studied companies are possessing dynamic and/or adaptive capabilities based on how the firms exploit their resources.

However, to be able to analyze whether the retailers have adaptive and/or dynamic capabilities, it is necessary to first investigate their business processes when analyzing Big Data. Therefore, we will first analyze our empirical findings using the marketing mix to be able to draw conclusions of which capabilities the retailers' posses. Finally, these findings combined will form the basis for discussion regarding what capabilities and resources that are required to gain a competitive advantage through data exploitation in the retail industry.

2.2.1 Value creation through the marketing mix

The marketing mix consists of 'four key decision areas that marketers must manage so that they satisfy or exceed customer needs better than the competition'. These four elements are product, price, promotion and place. The benefits gained from the four key decisions will lead to increased value for the firm (Jobber, 2012). In this study we refer to value creation as when insights from Big Data create value to customers that competitors cannot achieve or when a firm's bottom line is improved.

The product decision refers to the physical goods or services produced by the company. New product development is an important element as technology and demand changes. It is therefore important to stay up to date to be competitive and deliver customer value. By analyzing collected data, companies can gain insight on what kind of product attributes the customers prefers and can then tailor new products to their target group. One example of this is Netflix. They investigated which shows that had previously been popular and based on those insights they created the series *House of Cards*. The show became a real success and is a clear example of how successful product development can be made through data insights (New York Times, 2013).

The price refers to the actual price that customers pay for a product. As the price affects the value perceived by the customer, it is an important factor in the purchasing decision. By

implementing a price-optimization system based on both internal and external data sources, companies can set their prices after the exact customer demand.

Promotion is all activities made to create awareness of the product or service to the target audience (Jobber, 2012). In terms of Big Data, it can involve measuring efficient communication channels and what types of messages that gains the highest reactions. The big amount of information available about customers can also simplify the process of customizing offers and hence make the targeting more accurate.

The place decision concerns factors such as distribution channels, location of stores, methods of transportation and inventory levels. The overall goal is to locate the product where the customer is most likely to buy it (Jobber, 2012). One example of a company who really has applied their data resources on place is Amazon. By analyzing order history, product search history and shopping chart activity, they can predict when a customer will purchase a specific item and then begin to ship the product to the closest hub before the customer has submitted the actual order (Erevelles et al., 2015).

2.2.2 Firm's resources

Resources include physical capital resources, human capital resources, and organizational capital resources (Barney, 1991). When it comes to Big Data, physical capital resources include software or platforms used by firms to gather, store or analyze Big Data. Traditional software might not be capable of analyzing Big Data (Bharadwaj, El Sawy, Pavlou, & Venkatraman, 2013). Firms therefore need to build a platform capable of storing and analyzing large amounts (volume) of data continuously flowing in real time (velocity) from many different sources (variety) (Davenport, Barth, & Bean, 2012). Human capital resources include data analysts and strategists with knowledge of how to capture information from consumer activities, as well as manage and extract insights from Big Data. Organizational capital resources include an organizational structure that enables the firm to transform insights into action.

2.2.3 Dynamic and adaptive capabilities

Dynamic capabilities are about reacting to a rapidly changing business environment and being able to adapt organizational skills, resources and functional competences after these changes. Adaptive capabilities are about foreseeing and being proactive to changes in the external environment and does not, as dynamic capabilities, come from a specific change in the organizational structure (Erevelles et al., 2015).

When analyzing Big Data, a firm has dynamic capabilities when the organization is able to react and adapt to the consumer insights (Erevelles et al., 2015). When analyzing transactional data, firms might for example see that customers tend to buy a certain

combination of products together. The company can then react to that information by promoting this specific product combination in stores or in offers.

A company that has adaptive capabilities can through data predict what kind of products the customers are likely to demand. Instead of reacting on data of what customers have demanded in the past, the firm is now one step ahead and predicts what they will demand in the future. Weiss et al. (1998) argues that predictive outcomes of Big Data exploitation have the highest potential payoffs and the most precise results. The goal of data mining should therefore be to achieve predictive outcomes.

2.2.4 Connecting capabilities to competitive advantage

A competitive advantage is when a company is performing better than its competition in creating customer value or by achieving low cost advantages, and can be derived from the marketing decisions (Jobber, 2012). By analyzing insights from Big Data, companies can gain knowledge about their customers that the competition does not posses. The ability to collect and capitalize on Big Data resources can therefore generate big competitive advantages for firms.

Erevelles et al. (2015) suggests that the value creation from Big Data may result in a sustainable or temporary competitive advantage. However, since the firms in our study have not worked extensively with Big Data for a period long enough to measure whether a competitive advantage has been created, nor whether it is sustainable or temporary. We have therefore chosen to limit our investigation to discuss whether the companies have the capabilities required to achieve a competitive advantage in the future.

2.3 Theoretical connection to research questions

These focus areas from the Resource-based theoretical framework by Erevelles et al. (2015) will lay the foundation for the analysis of our empirical data and help us answer our research questions:

R1: How is Big Data influencing marketing and value creation in retail?

R2: What are the major hindrances for Big Data exploitation amongst retailers today?

R3: What marketing capabilities are required of retailers for further development?

The marketing mix variable will be used to investigate and answer the first question (R1), the firm's resources variable will assist when answering (R2) and the dynamic and adaptive capability variables will help to answer (R3).

3. Methodology

3.1 Scientific approach

Even though Big Data is not a new phenomenon there is not much literature written on it's impact on value creation. We have therefore chosen to conduct an exploratory study with the aim to gain knowledge and a deeper understanding of how Swedish retail companies are using Big Data to create value and what future impact it will have for these companies when it comes to creating a competitive advantage. In order to get a deeper understanding of how the studied companies are exploiting Big Data, a qualitative research method was chosen over a quantitative (Malhotra, 2004). The choice to conduct a qualitative study is further motivated to contribute to the knowledge gap in this research area (Merriam and Nilsson, 1994).

3.2 Study design

In order to gain an understanding of the relationship between Big Data and value creation, the study follows an inductive approach as it aims to generate new theory rather than testing existing ones (Bell, 2006). An inductive approach was also chosen to avoid the complexity and constraints that comes with making assumptions from the beginning, and instead allowed us to shift between theory and data while knowledge was gained during the process (Bryman and Bell, 2015).

Due to the exploratory nature of the study, a case study methodology was used. According to Yin (2013) this is a suitable research methodology when the research focuses on contemporary events in real situations and can answer how and why certain events occur. The choice of method is further supported by Eisenhardt (1989) who recommends case studies when the study aims to provide a description of an event and when the study concerns a relatively new topic, which holds true for Big Data. The case study methodology has received criticism for example by Abercrombie, Hill and Turner (1984) arguing that it only provides hypothesizes and is therefore not scientifically generalizable. However, since the aim of the study is not to provide generalizable conclusions, but rather to describe the current Big Data landscape within Swedish retail organizations and provide future recommendation for value creation we find the method suitable for our study. Other scholars later support that the method should be seen as a strength (Dubois and Gadde, 2002) and a valid research method (Flyvbjerg, 2006).

3.3 Case selection

The empirical area of this study focuses on Swedish retailers, this selection was made because we found it interesting to study companies which are in direct contact with the customers and have direct access to customer data. In the private sector, Manyika et al.

(2011) estimate that a retailer using Big Data to its fullest capacity has the potential to increase its operating margin by more than 60 percent. Because of these reasons, we found the retail industry to be an interesting area to study.

Based on the criterion stated above a list of relevant cases to include in the study was created, however, not all of these companies were available for interviews. Due to the scope of the research, we were not able to secure organizations to be part of the study in advance and hence, the selection of cases ultimately studied is a convenience sample (Onwuegbuzie and Leech, 2007). In total the study is based on eight interviews with IKEA, ICA, Willys, Hemtex and four companies that will be handled anonymously throughout the study based on their request.

3.3.1 Selection of interviewees

As a starting point for our selection of interviewees we contacted the Market research company GfK who suggested additional suitable companies to study and provided us with contacts to these, along with the companies earlier proposed by us. This allowed us to contact the most relevant employees and likely contributed to the high response rate. During the initial contact with the companies, via email or telephone, we introduced the focus area of our study. In most cases we could book a meeting right away, and in a few cases we were redirected to another employee within the company who was more knowledgeable within the specific area. As our study focuses on how the companies are creating value by exploiting Big Data most of our interviewees are employees in charge of consumer insights with hands on knowledge within the area. The sampling principle and biased selection process of interviewees is supported by Ritchie et al. (2013) as a qualitative study does not require to achieve statistical representation. Hence, it could be motivated to only interview one employee per company to achieve a good understanding of the research area, since the employees interviewed are the ones most knowledgeable within the studied area.

Table 1. Overview of interviewees

No	Company	Employee	Position	Located in	Interview type	Interview date
1	IKEA	Anna Höjman	Market intelligence consultant	Helsingborg	Phone	22.02.2017
2	Company A	Employee A	Head of CRM	Stockholm	Face-to-face	28.02.2017
3	Company B	Employee B	Customer Analyst	Stockholm	Face-to-face	02.03.2017
4	Company C	Employee C	Head of business development and omnichannel	Stockholm	Face-to-face	16.03.2017
5	Company D	Employee D	Senior business analyst	Stockholm	Face-to-face	22.03.2017
6	Hemtex	Mona Adawi	Head of CRM	Borås	Phone	29.03.2017
7	ICA	Charlotta Halfwordsson	Head of Consumer Insights	Stockholm	Face-to-face	11.04.2017
8	Willys	Patrik Eriksson Ta	CRM Analyst	Göteborg	Phone	12.04.2017

Table 1 gives a short summary of all respondents represented in the study. The companies who wishes to be anonymous in the study has been given a pseudonym ranging from Company A to D. Similarly, the respondents from these companies will be referred to as Employee A to D.

3.4 Data collection

After initial gathering and study of relevant theory, research questions and interview questions aiming to capture the essence of the research area were formulated. The interviews were constructed and conducted in a semi-structured manner, supported by for example Quader (2007), who argues that interviews should be semi-structured when the data is characterized by general information within a specific area. Semi-structured interviews are also recommended by Bryman and Bell (2015) when conducting a qualitative study. Each interview originated from the same structure and set of questions, but the interviewees were encouraged to elaborate their answers on relevant topics, and follow-up questions were asked when deemed necessary. During the data collection process, the set of interview questions were revised and improved in order to increase relevance and get the most valuable information out of each interview. However, the interview questions were not changed to an extent that would compromise the consistency of the data gathered from the different interviews.

To create a sense of security, five out of eight interviews were held face-to-face at the interviewee's office, based on their choice of location (Trost, 2010). Three interviews were held over telephone due to the geographic distance. Both of the authors participated in all the interviews to capitalize on insights in the data (Trost, 2010). This also allowed us to take on different roles during the interviews in accordance with recommendations of Eisenhardt (1989). One of us asked the questions and was responsible for guiding the interview in the right direction, while the other took notes and had a more reflective role. Fortunately, all respondents gave us approval to record the interviews, which enabled us to be more focused during the interview, to come up with relevant follow-up questions and to reflect on the interviewee's behavior (Bryman and Bell, 2015). The interviews lasted between 35 and 60 minutes and were held in Swedish.

Finally, in order to decrease potential bias, the interviews were summarized within 24 hours individually, and thereafter discussed together. In order to facilitate the analysis, all interviews were transcribed word by word into written documentation.

3.5 Research quality

While there are several developed methods of varying effectiveness to ensure the quality of quantitative research, there is no fully developed method for evaluating qualitative studies (Flick, 2009). There are currently two central alternatives discussed in the literature. First is the option to apply the classic measurements of validity and reliability while the second option is to develop a new method of appropriate criteria to evaluate qualitative research. However, according to Flick (2009), none of these alternatives offers a satisfactory solution to the problem of grounding qualitative research. Bryman and Bell (2015) also suggests that the meanings of the terms validity and reliability needs to be altered, or alternative criteria for evaluating qualitative research needs to be applied. Some writers have suggested that qualitative studies should be judged or evaluated according to criteria quite different from those used by quantitative researchers. Lincoln and Guba (1985) and Guba and Lincoln (1994) propose that it is necessary to specify terms and ways of establishing and assessing the quality of qualitative research that provide an alternative to reliability and validity. They propose two primary criteria for assessing a qualitative study: *trustworthiness* and *authenticity*. Trustworthiness is made up of four criteria: *credibility, transferability, dependability* and *confirmability.*

3.5.1 Credibility

Establishing credibility of findings requires both ensuring that the research is carried out according to good practice and submitting research findings to the people being studied to confirm that the researchers have understood correctly. The latter technique is often referred to as *respondent validation*. In order to ensure a good correspondence between our findings and the perspectives and experiences of the research participants all interviewees were provided with the quotations for approval as well as the final thesis.

3.5.2 Transferability

As our research is focusing on studying several individuals within organizations sharing certain characteristics, the qualitative findings can be generalized for the Swedish retail industry but may not hold true in other contexts. As Guba and Lincoln put it, whether or not findings 'hold in some other context, or even in the same context at some other time, is an empirical issue' (Lincoln and Guba 1985: 316). Instead, the findings should be viewed as what Geertz (1973a) calls 'thick description', which is a general description of a company

culture. Guba and Lincoln argue that a thick description provides others with a database for making judgments about the possible transferability of findings to other environments.

3.5.3 Dependability

As a parallel to reliability in quantitative research, Guba and Lincoln propose the idea of dependability and argue that to establish the quality of research in terms of this criterion of trustworthiness, researchers should adopt an 'auditing' approach. To ensure the dependability of our study and that proper procedures were being followed, complete records were kept during all phases of the research process including problem formulation, selection of research participants, fieldwork notes, interview transcripts, and data analysis decisions.

3.5.4 Confirmability

Confirmability is about ensuring and recognizing that even though complete objectivity is impossible in business research, the researchers can be shown to have acted in good faith. In other words, we have not allowed personal values or theories affect the research and the findings derived from it.

3.5.5 Authenticity

In addition to these four trustworthiness criteria, Guba and Lincoln suggest the criteria of authenticity which raise a wider set of issues concerning the political impact of research.

• *Fairness:* since all interviews were conducted with employees at their respective departments, the data represents their perceptions and exclude the perceptions of other stakeholders.

• *Ontological authenticity:* the research can hopefully help retailers gain at a better understanding of the potential value creation from Big Data.

• *Educative authenticity:* the research can hopefully also help retailers to better appreciate the perspectives of other stakeholders.

• *Catalytic authenticity:* the research could potentially act as a momentum for retailers to come up with ideas for how to change their current situation.

• *Tactical authenticity:* the research might empower stakeholders to take the steps necessary for engaging in action.

4. Empirical findings

Our empirical findings are structured around three focal themes which will be presented in the following section. These themes emerged during our analysis, where we identified converging topics describing major areas that have been affected by Big Data. In the first part we will focus on marketing capabilities and how Big Data has affected the different components in the marketing mix. Next, we explain the major hindrances for retailers today and finally we investigate what capabilities that are required for further development.

4.1 Big Data's impact on marketing decisions

4.1.1 Product

Several of the studied companies state that Big Data has had a considerable impact on their product decisions. Willys, for example, says that the start of their loyalty club four years ago has had a radical effect on their product strategy. Company C also states that they have found substantial value in product range optimization.

> *"We have certainly found value in product range optimization through Big Data. And we're talking significant values."*

Company A claims that they today know a lot more about who their customers are and that this permeates the whole organization. They have increased the number of customer surveys performed in order to understand the driving forces behind the customer's purchase intentions. This enables them to develop a more relevant range of own-brand products, of which they can control the product development themselves. Company D are also performing data analysis on their own-brand products. The data tells them which consumer groups that are buying what products, and from that information they can form hypotheses which they later investigate in consumer panels. Thereafter they can modify the product attributes, for example the packaging design, to appeal to the target group. ICA is also using data analysis in their product decisions to gain a deeper understanding of what product ranges that attracts different customers. For Company B, the product decision is more related to the gross profit:

> *"What products that we purchase is not data driven in that sense, it's rather a question of what is selling and what is not."*

But Company B are on the other hand using so called market basket analysis, where they investigate which products that are purchased together and what products are substitutes and competitors to one another. Hemtex are also performing market basket analysis which

helps them to be more accurate when creating new collections and in the placement of products in store. Hemtex says that they wished that the data insights had a higher impact on product decisions than what it has today, and that they are continuously working on becoming more accurate in their targeting.

> *"We strive to do that everyday. Today it's more that I mediate an insight that the supply department can bear in mind when they purchase their product range."*

4.1.2 Price

When asking the interviewees if Big Data has had any implications on their pricing strategy, Company D tells us that they are working with data insights to perform price-optimization and Company C says that it is something they are currently developing:

> *"Yes, we are actually working on that right now. I hope it will lead to significant business value for us." (Employee C)*

> *"We hope it will lead to increasing marginal revenues by getting a certain percentage of customers to switch from one product to another that is slightly more profitable. Each swap can generate quite a lot of money in the long run." (Employee D)*

Company D tells us that their price-optimization system is quite comprehensive and allows them to test several different scenarios and run simulations. These are then analyzed by analysts and category managers together to understand what implications a specific price-scenario will lead to:

> *"/.../ it's a combination of pure data driven information and category expertise."*

ICA is using data to follow up on how customers are reacting to price increases. In that way they are trying to understand what implications a price change will have on their customers. Company A says that they do not want to compete on price because their biggest competitor has a low cost advantage. But on the other hand they offer customer loyalty deals and targeted discounts which they base on purchase history. In that way transactional data is affecting the final price to the customer, rather than the starting price. Company B are also using purchase history and targeted discounts in a similar way:

> *"Approximately 40% of goods sold in total are not sold at full price, and that makes it quite important to keep track of where you are losing gross profits. /.../ that is how we control the price rather than the price setting."*

Hemtex tells us that the data does not affect their pricing decisions today. It is however something they would like to be able to use the data for, but product development is of higher priority when it comes to implementing data insights.

4.1.3 Place

> *"I would definitely say that which stores we choose to keep or not are based on how profitable they are." (Employee A)*

None of the retailers state that they use Big Data to any larger extent when deciding where to place their physical stores or in other decisions concerning distribution. The same goes for decisions regarding assortment online versus offline. Some state that they simply have not looked into these parts yet while others claim that there are other aspects that determine where new stores are opened and how e-commerce is expanded. For existing stores, the bottom line will always dominate the decision whether to keep a store or not. However, Big Data can be used to identify a gap in the retailer's store network and give information about where the customers are coming from and how far they have travelled to shop at that specific retailer.

Company D explains that during the establishment phase, decisions are mainly built on more traditional market analysis rather than Big Data. These analyses are often complemented by purchased external data with information such as purchasing power, number of inhabitants and economic outlook in the area. Information from municipalities regarding future construction plans are also taken into account.

4.1.4 Promotion

The vast majority of retailers interviewed agree that promotion is the part of the marketing mix where Big Data has had the greatest impact to date. The most significant effects can be seen in how retailers control their communication by targeting offers to their customers. Data has made it easier to gain insights of how much value different communication strategies offer.

> *"/.../ it's mainly in our promotion where we can see that the large amounts of data on purchase history has influenced what and how we communicate." (Employee A)*

Direct marketing has been greatly impacted by the increased access to consumer data. All interviewed retailers send personalized offers to their customers. These offers are to a large extent based on the customer's purchase history and are complemented by offers on similar products that the customer ought to be interested in. The large amount of consumer data

available today allows retailers to make informed communication decisions instead of guessing what might work:

"/.../ there's no reason to guess when we have so much data on our customers." (Hemtex)

A large advantage is to be able to proactively communicate with customers based on the insights derived from consumer data. These insights allow retailers to adapt their communication and customize it to their different customer segments. However, several respondents state that they are just in the starting-blocks of working with this type of customer segmentation.

Another interesting analysis performed by retailers in order to find potential customers is the twin customer analysis. Based on an analysis of the retailer's existing customers, they find so-called twins to these to target. Twin customers are prospects that are equivalent to their existing customers and have great potential to become future customers. This process can streamline the retailer's recruitment of new customers and allow them to reach many more customers. This type of analysis is trending among retailers today.

All retailers interviewed have access to structured data in form of their own customer purchase data, which is rich in information and available in massive volumes. This data can be split between different product categories and provides the retailers with information regarding what products their customers are buying, when and where the purchase is made, how often they buy certain products and how much they spend on each category. This information is then used to segment customers based on purchase behavior, which in turn becomes the underlying determinator when directing the communication to customers. Company A reveals that this information is what determines the whole campaign year in terms of purchase and campaign periods:

"We use all the data we can, so to speak, in order to make sure we make the right decision."

According to the retailers interviewed, communication is one of the most important usage areas for Big Data. Both in the aspect of making the communication relevant for the customers as well as for the retailers not to waste money on communication in inefficient channels.

The challenge facing many of the retailers is how to measure the actual effect of their marketing efforts. What they would like to know is if it has any effect on sales or customer loyalty in the long run and how to identify it. These are issues that the retailers are trying to grasp:

"It's very difficult to measure ROI on these particular type of activities, a campaign can be visible, but what does it really mean?" (Employee D)

New sources of data from web and online are also available to retailers, however most retailers claim that they are not using this data to any larger extent. A lot more could be done with this data, however, some retailers admit to being too traditional and somewhat behind in the progress of exploiting Big Data in general, and digital data in particular.

When asked if the access to Big Data has affected which communication channels the retailers are using, most agree that the media mix has definitely been affected. However, the greatest impact can be noted in how they exploit the different channels rather than affecting the actual choice of which channels to use. This change comes from the fact that the retailers, with help from Big Data, now have the possibility to measure the effect of communication through different channels in a much more reliable way. With these new measurement tools the retailers can adapt their communication until they reach the desired outcome. IKEA are analyzing multiple data sources when deciding which media to invest in:

"We are working with data mining and analyzing many different data sources in a database to be able to do more advanced analysis of purchasing patterns, media investment analysis, target group analysis etc."

Traditional communication channels such as text messages and email are still the ones used most frequently by the companies studied. One explanation given by Company B is that it becomes clear when the communication ceases to be relevant, as the customers simply stops to open the messages. These traditional channels provide very accurate tracking to see what customers open and click on, and can also be easily linked to purchase history to make the information richer. With the amount of data available today, retailers can personify customers in ways not viable earlier, which enables them to send out more accurate offers. With regards to communication, digital data provides information about how their customers are acting online. For example, retailers can analyze which channels feeds the most customers to their website. By connecting this with other data sources understand which customers are most likely to make an actual purchase and which customers are that are returning to the website. This is used to a large extent in order to prioritize which channels to use by understanding which ones that generate the highest returns. Even though new opportunities are opening up with the increasing amount of data, the more traditional retailers are still looking at the costs as the primary basis for which channels to use:

"We see that different messages works in different channels, but I would still say that it's a matter of cost, organizations look at and think of costs first, you can not see past the costs, it is always primary." (Employee B)

4.1.5 Value creation

All interviewees agree that Big Data has created value for their companies, however the difficulty lies in how to measure it. Willys says that their biggest value gain is that they now have proof to back up their decisions and the fact that they can clearly determine whether they are doing the right things:

"We can put it in relation to something"

Hemtex also agree that data helps them in determining whether something is working and that the use of data insights show positive results in their campaigns. They are continuously analyzing the effects of their activities:

"/.../ by adapting the communication based on the data we collect and the analyzes we do; we can see an uplift on our campaigns."

Company A says that they run control groups on all their communication, and that their test methodology helps them extract as much value as possible from their activities. They say that they can especially see an uplift in their targeted communication, and how much it has actually generated in monetary terms:

"/.../ we can also get a better control over the communication, along with insights of how much it's worth in crowns."

Company B on the other hand reveal that they do not measure the value that data analysis has created for the company. They say that it is unfortunate that they do not, because that would motivate the management to allocate more money to the consumer analytics department, allowing them to develop even further:

"I would say that Company B haven't been that good in gathering the facts or evidence to show that it has improved"

However, Employee B states that it is obvious that the various analytical tools used on large amounts of both structured and unstructured data, has generated value and that companies will not be able to manage without it in the future.

4.1.6 Perceived progress in Big Data exploitation

When asked to rank their company on a scale of 1-10 based on how far they perceive that their company have come with their Big Data analytics, all studied retailers placed

themselves on the lower range of the scale with a mean of approximately 4.4. Most of them are aware that there is much more potential to be captured. In section 4.2 we will discuss the major hindrances that are keeping the retailers from progressing in their Big Data analytics.
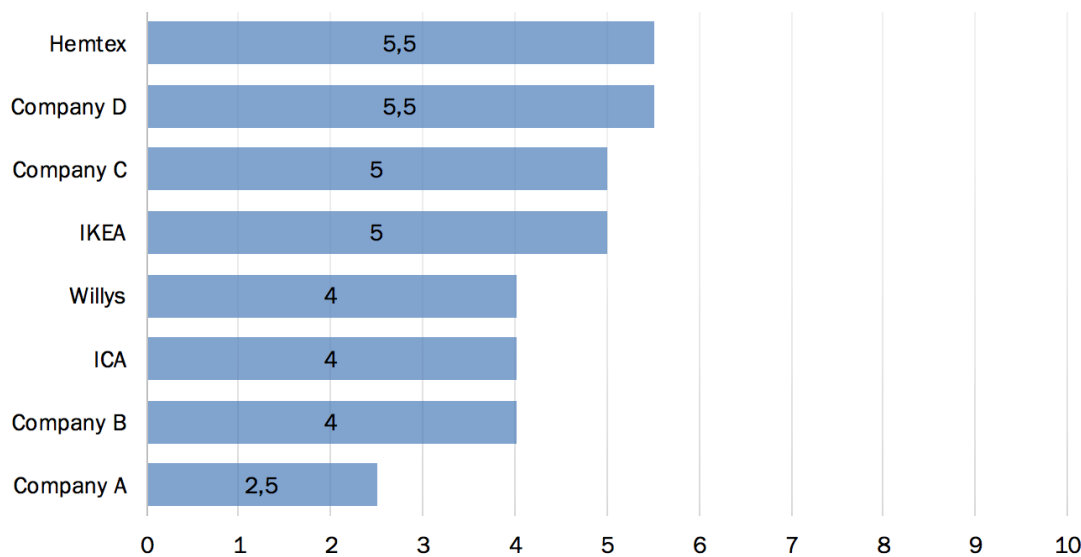

Figure 2. Retailers perceived progress in Big Data exploitation

4.1.7 Concluding remarks

After investigating which of the four key decisions in the marketing mix that has gained the highest value from Big Data analysis, it is clear that the promotion decision has been significantly affected by Big Data. All interviewed companies agree that Big Data has changed how they communicate with their customers and how they exploit their communication channels. Some also experience that it facilitates how they track the performance of their communication. Some have also found considerable value in the product decision through product range optimization on their own-brand products and others through market basket analysis. But overall, the product decision has not received as much attention as the promotion decision. The same refers to the price decision where only two out of the eight retailers has realized the value of price-optimization. Many use targeted discounts that affects the final price, but it does not provide the same insights as price-optimization. None of the companies has involved Big Data in the place decision, and their main motivation is that profitability will always be the main aspect here. All retailers agree that they have already gained substantial value through Big Data, but that it is hard to measure how big that value actually is. In summary, Big Data available to retailers today affects customer segmentation, which in turn influences the customer strategy and thereby affects the overall marketing decisions.

4.2 Hindrances for value creation

"I think that everyone in the industry have just scratched the surface, so far" (Willys)

All retailers interviewed agree that Big Data holds a lot of potential for the future. Many say that they hope it will help them become more relevant for their customers through targeted communication enabled by better segmentation. They believe that further automatization will take over the manual tasks and help the organizations to be more effective in their decisions.

"I believe that we will be able to become a company with more products and bundled services – a more customer driven company, and hopefully we will automatize more, which will facilitate gathering and analyzing data and becoming more effective." (IKEA)

"It might imply that we can provide the right product's so that we don't waste time, money and energy on developing a product that will not be well-received by the customers. It might lead to customers getting a better relationship with ICA because they feel that ICA is more relevant." (ICA)

Company C are hoping that it will help them improve their customer satisfaction and that it will lead to an overall more profitable organization. They say that the next step to take is to go into machine learning. This is something Company D also believes involve a lot of opportunities for the future, but is progressing slowly due to the extensive investments required. Employee B believes that insights derived from Big Data will lead to a more customized and precise targeting. They also believe that e-commerce will take over and that having an omnichannel strategy will become more and more important to satisfy customer needs.

Access to data is no longer a problem, in contrary retailers are overwhelmed by the massive volume of data available and does not have a clear strategy for how to exploit it in a way that can create value for the company. According to Company B, many retailers are somewhat lost when it comes to what possibilities that lie ahead and what their actual purpose is with data analysis. Willys and the other retailers agree that the challenge today is figuring out what data that is valuable and be able to refine it to develop relevant offers to customers, and at the same time be profitable:

"There is an ocean of information available, you just have to find what to use it for and how we can use it in a good way." (Willys)

When discussing the barriers that prohibit the retailers from exploiting Big Data and reaching their full potential two main themes becomes obvious; IT-architecture and competence. Within these time and organizational support are also factors that hinder development.

4.2.1 Physical capital resources

"The IT-architecture needs to change and allow the different systems to talk to each other"
(IKEA)

One of the main problem retailers are facing today is the company's current IT structure and the fact that different data sources are not connected, which makes it difficult and time consuming to analyze and derive insights from different data sets. According to ICA retailers does not have the technical conditions to analyze and gain valuable insights from the accessible data. The constrained resources are of course a matter of time and money, but also knowledge as stated by Company D:

"The challenge here is how to synchronize all data sources available - that's a big question."

A typical problem for traditional retailers is the fact that the different IT systems currently in use have been implemented during a long period of time. Company B explains that some IT systems might have been around for 20 years, then another system was introduced 10 years ago to cover another area and then an additional system was implemented three years ago for another purpose, and on top of these they have a text messaging system. It has been a real struggle to link these different systems together, and can often feel like a hopeless task for traditional companies. Because of this, Company B has just recently started calculating the gross profits on their products. Allocating more storage capacity and more machines is primarily a problem of funding according to Company C. Company D agrees:

"Everything that requires an initial investment from the start is often a barrier in itself"

Traditional companies such as Company B are lagging behind when it comes to exploiting digital data:

"I think it's much easier for new companies that don't have any legacy, old systems are standing in the way to allow for new ones."

4.2.2 Human capital resources

Another barrier for value creation mentioned by all interviewed retailers is the lack of competence. According to Company C, management does not understand the value that connecting different databases can create. Both Company D and Willys also raise the issue that the organizations need more than just data analysts but also people with technical background, who are able to perform the necessary connections of the IT systems. Connecting systems and competence will be key. With more analysts they could also do a lot more. Company C tells us that retail companies who lack Big Data competence have large IT costs due to the fact that they have to hire external competence to solve functional problems. In order to obtain an overall solution companies, have to recruit and understand the value of competence with programing skills:

"In New York or Silicon Valley companies understand that it is very cheap for what they create. They have the ability to solve tomorrow's problems. Risk Capitalists understand, but not a retailer. And that's a challenge." (Employee C)

4.2.3 Time

Time constraints is another issue identified. Company A, who announces time as the primary barrier, explains that if they had the time they would be able to do a lot more sophisticated analysis, which in turn would free up time for product development and communication. The massive volumes of transactional data available also comes with natural limitations in terms of processing the data. Not only does it require a lot of computing power but primarily it is extremely time consuming to figure out what to look for and be able to make something useful out of it.

4.2.4 Organizational capital resources

Company C mentions that there is an organizational resistance and a lack of support to implement Big Data on a large scale. One big problem is that each department have invested in their own database and therefore do not want other departments to take advantage of it. This results in time consuming discussions regarding who should be responsible for what instead of capitalizing on the value. According to Company D, there are a lot of low hanging fruits that could be solved, but these require organizational support and receptiveness which takes time to build. Several companies believe that the lack of organizational support might be due to the fact that there are not enough proof showing that investing in Big Data will generate positive results. Company C says that it is because people want to work with something that they feel comfortable with, all humans have a need of security. This

expresses itself by people putting all their time and effort in finding reasons of why the new solutions would not work, as a way to protect your own position and competence. It is a great challenge that the receiving organization, who are responsible for their own development does not have an invested interest at the individual level to create Big Data:

"This is a human problem, and therefore, it takes very good leadership to guide the company through such a transformation." (Employee C)

Another issue is old patterns that needs to be broken, which is not something that happens over a day, adaptation takes time. The analytic department at Company D has during the past five years fought to become a prominent part of the organization and to have a greater saying in the strategy process. As Hemtex describes it, the organizational structure can also be a barrier:

"Even though we are all working in the marketing department, there are still silos within the department in a way."

The customers' needs are constantly changing, which is why Hemtex believe that they will never be done with their analysis. They also express a concern regarding the technical resources that are not developing in the same pace as the customers needs. Retailers are somewhat stuck in their old ways of thinking; they have the sales perspective rather that the customer perspective.

4.2.5 Concluding remarks

In order to create value from Big Data companies will have to find ways to refine the available data. Time, money and knowledge are the main hindrances for further development which are all related to a lack of organizational support. Connecting systems and investing in the right competence will be key. Organizations will have to trust that Big Data can generate significant value and start investing for the future.

4.3 Dynamic and Adaptive capabilities

Many of the interviewees explain that they are working more reactive than proactive with Big Data today. Company D, for example, tells us that if sales are decreasing in a category, they put more emphasis on handling that current problem rather than working with questions concerning what could be done to prevent it in the future. According to Company D, the latter would be a better and a more analytical way of using Big Data. It is a problem that a lot of the data is based on current purchase patterns and that they are working towards becoming

better at focusing on predicting the future instead. This is something also mentioned by Willys:

> *"We can see that it is changing and how people are acting, but we don't see what they actually would like to buy. Instead it becomes kind of a guessing-game."*

Company A also argues that the transactional data analyzed today is rather poor in terms of providing the whole perspective around customer behavior. It does not say anything regarding the driving forces behind a purchase. However, this is something they are currently working on and are in the process of implementing in their own-brand product range.

> *"The transactional data is quite reactive; we can see what the customers have bought but it doesn't say much about what they want in the future or how they actually perceive Company A or their shopping experience." (Employee A)*

Employee B says that they are trying to work from the customers' perspective, while the other business areas in the organization are more focused on the product development and afterwards try to reflect on what type of customer that could potentially be interested in the product. Employee B receives many questions regarding why sales are suddenly going down on a product that sold very good last year. This indicates that they are working more descriptive, explaining how things relate. But he/she also mentions that they are working predictive, trying to understand and segment to be able to make more precise in their future targeting.

Company C state that it is very important to combine insights from both the present and the past when working with product range development and that it is essential to have a proactive approach to gain customer satisfaction:

> *"/.../ being proactive is vital. That is how we get the customers satisfied in the end. You solve the problem before the customer even realizes that the problem exists."*

IKEA and Hemtex also mentions the importance of truly understanding the customers. IKEA says that customer oriented products sell better, because you already know what the customer wants. Hemtex says that as a retail company, it is crucial to adapt and understand the customer in order to survive within the industry. With the fierce competition, focus will have to be on the customers needs.

> *"From us telling the customer what to buy, it will become the customer telling us what products to offer." (Hemtex)*

4.3.1 Concluding remarks

Concluding, the majority of our interviewees agree that the future success for retailers requires more emphasis on Big Data analysis, and that it becomes more and more important to be proactive when trying to understand customer behavior.

"It is essential for the customer experience to become more relevant, I believe that the more we know about our customers, the more our customers will require from us." (Willys)

5. Analysis

In this section we aim to answer our previously formulated research questions with aid from the chosen theoretical framework and findings from the empirical study. The analysis is divided into three main sections based on each research question. To ensure a better understanding of the presented findings, our analysis will be enhanced using complementary cases of companies who have succeeded in their exploitation of Big Data.

5.1 How Big Data is influencing marketing and value creation in retail

When looking at figure 2 in section 4.1, it can be said that in most cases the individual rankings correspond to reality. However, there is a notable difference when it comes to the interviewees awareness of what their companies actually could achieve with Big Data. This might be the reason why some retailers placed themselves higher on the scale in relation to the other studied firms, whom we perceive to have come further in their exploitation of Big Data. Further, our empirical findings show that far more emphasis have been put into promotion compared to the other components of the marketing mix. However, our empirical findings complemented with research from external sources tells us that the possibility to gain value through the other decisions are endless. So, how come promotion has received the most attention? Could it be due to lack of knowledge of how to exploit Big Data in the other marketing decisions? Or are the other decisions just not as highly prioritized?

One reason may be that promotion is the easiest to both extract and capitalize on, and therefore it has been higher prioritized within the organizations. The majority of the tools and resources used by the studied companies in promotion today does not require any massive investments or advanced competences, which facilitates implementation. For Swedish retailers to move forward in the other decision areas, they would need to break the barriers discussed in our empirical findings. We will discuss how this can be done when answering our third research question in section 5.3.

None of the studied companies are using data analytics in their place decisions. However, there are multiple innovative options to increase efficiency and profitability by making the shopping experience more convenient for customers. Two companies who have successfully capitalized on the place decision, are the former online furniture retailers Oak Furniture Land and sofa.com. Both retailers decided to open physical stores, which are now generating as much as 60% of sales. Reshaping the store network is another opportunity, and one approach is to keep a couple of flagship stores. These stores will act as marketing and service channels for the online business, and should be supported by a number of smaller outlets that offer convenient and customized product offerings. When it comes to distribution, the British retailer Argos is testing out with a new system in London, where

products are being delivered from large stores to smaller stores with new digital solutions. This enables all Argos stores in the area to offer delivery the same or next day for 20,000 products. These examples indicate that the physical store is not dead, it just needs to shoulder a different role and can even provide a competitive advantage. (McKinsey, 2016)

Our empirical study shows that only two of the interviewed companies are using price-optimization, even though there are substantial value to be captured here. There are today numerous companies offering price-optimization solutions. Based on various data sources as accumulated historical behavior and competitor's prices, price-optimization systems can even provide predictive analyzes. Researchers at MIT developed a price optimization system for the online fashion retailer "Rue La La", which increased their revenue by 9,7%. The opportunity to capture unmet demand through data driven pricing systems is enormous and there is a lot of profit to gain, especially in the retail industry. (Ferreira et al. 2015)

Out of the four components, product is probably the decision where there lies the most potential. Several of the studied companies say that they have gained substantial value from using Big Data in their product decisions. However, they are today working more with strategic decisions concerning existing products, such as product range optimization, rather than using the data in product development. Being able to offer products demanded by customers, will save a lot of time and money. According to McKinsey, Big Data analytics can help manufacturers decrease their product development and assembly costs by 50%. Using insights of customer demands derived from Big Data in product development will simply make the retailers deliver more relevant products for their customers. This should be the overall goal for any organization.

5.2 Major hindrances for Big Data exploitation

Two major barriers found in our empirical study that are restraining the retailers to capture value from the data available are lack of analytical talent and data being siloed within departments. These barriers are in line with the findings from McKinsey Global Institute's report *Big data: The next frontier for innovation, competition, and productivity,* 2011. In the report they predicted a 60% increase in net margin and a 0.5-1.0% annual productivity growth for US retailers. However, in their follow up study *The age of analytics: Competing in a data-driven world,* 2016*,* it becomes apparent that only 30-40% of this value opportunities had been captured. Similar observations hold true for the EU retail sector. Even though many of the opportunities described in the 2011 report remains uncaptured for US and EU retailers, our empirical findings suggest an even slower progress for Swedish retailers. In the meantime, the potential for value creation has grown even bigger.

Even though research is scarce, there are business cases showing that investing in data and analytical capabilities generates high returns. However, these high returns are mainly driven by only a few successful companies. Early adopters of Big Data are showing faster growth in operating profits. This in turn enables them to continue investing in data resources and analytics capabilities, which further increases their advantage. The best example is Facebook who created a platform capable of collecting outstandingly detailed data on billions of individual users. But not all winners in the data transformation are digital natives. Walmart, GE, Ferrari F1, and Union Pacific are examples of companies in traditional industries whose investments in data and analytics have resulted in outstanding returns (McKinsey, 2016). However, many Swedish retailers fail to recognize the size of the opportunities and hesitate to act on it due to the fact that returns are still unproven. This causes them to lag behind in multiple dimensions of data and analytical transformation, and the barriers are mostly organizational. The first challenge is incorporating data and analytics into a core strategic vision. The next step is developing the right business processes and building capabilities including both physical capital resources and human capital resources.

Data analytics has tremendous relevance for retailers, since they can gain valuable insights about their customers from transaction-based and behavioral data. However, organizational barriers including the difficulty of finding data analysts and breaking down information silos across large organizations are keeping the companies from realizing the full potential reward. The perception of the studied companies seems to be that the complexity and cost of analytics could outweigh the potential gains, particularly for those companies who have difficulties identifying the right technology and talent. The problem of data being segregated in silos across the organization continues to be unsolved, and installing replacement systems is difficult, time consuming and expensive.

5.3 Marketing capabilities required for further development

In order to overcome the above discussed barriers, the companies need to develop and enhance a handful of marketing capabilities. To start with, companies will have to take a step back and ask themselves what the data and analytics will be used for, how the insights can create value and how these values can be measured. Once the strategic vision is set the companies will have to develop the required IT-architecture as well as data collection and generation capabilities. Many traditional companies struggle with switching from legacy data systems to a more flexible architecture to store and analyze Big Data. Acquiring the analytic capabilities needed to derive valuable insights from data will be another vital step, the retailers may choose to add in-house capabilities or to outsource. They will also need to change their business processes to incorporate data insights into their daily workflow. This requires getting the right data insights into the hands of the right personnel within the organization. Finally, organizations need to build the capabilities of executives and managers

to understand how to use data-driven insights, and to begin to rely on them as the basis for making decisions. Organizations can take the first step by establishing an organizational culture and environment that is open to exploring new ideas and challenging established ways of thought. Getting people to be open to new ideas is only one step in the process.

Human capital was found to be one of the biggest barriers standing in the way of realizing the full potential of data and analytics. Translating data into business insights is the most important and most valuable part of the process. Great analytical capabilities can provide insights even from low-quality data, while poor analysis can destroy the potential value of high-quality data. Providing data analytics requires industry and functional expertise, and there is a limited pool of talent and organizations combining these skills. Even as tools and platforms improve, the need to combine analytical and industry expertise will continue to present a bottleneck, driving up the value of analytics from external service providers. The most successful companies in our study are combining analytic capabilities with industry or functional expertise. Nevertheless, the most value generated by analytics will be captured by the companies who has the knowledge to implement the data insights.

The first step in creating value from data analytics is ensuring access to relevant data. This may be straightforward in theory, but has been proven to be much more difficult for the companies in practice. Many of the organizations have a department and business unit structure that tends to create silos. As a result, it is difficult to share information seamlessly across those internal boundaries. Different departments may have unique data systems that are not integrated or may even be incompatible with one another. This issue is particularly evident in the larger companies with long histories and legacy systems, but was shown even in the younger companies. Managing data across these organizational barriers can be costly, and few companies have existing data systems that are capable of running analytics on a large scale. The organizations could also benefit from integrating new types of data from a range of external sources with their internal databases.

It becomes clear that all the interviewed companies are working extensively with responding to changes that have occurred in the past, and hence are possessing dynamic capabilities. They are not performing as well when it comes to predicting customer needs and we can therefore see a general lack of adaptive capabilities amongst the companies. However, among the studied retailers we found Company D to have the greatest adaptive capabilities due to their predictive price-optimization and general predictive mind-set. Others, like Company A, are currently working on implementing a more predictive product development strategy for their own brand products which will probably lead to developed adaptive capabilities in the near future.

All companies are however well aware of their deficiency in the field and highlights the importance of developing adaptive capabilities to move forward. Some interviewees even

mean that it is crucial for their survival in the retail industry. None of the companies have a clear motivation to why they have not yet developed more adaptive capabilities. The earlier barriers discussed in the empirical findings have probably all played their part in the matter, but it is evident that the major issue is a lack of knowledge of how to refine the data into insights of a more predictive nature.

To obtain a competitive advantage the companies will need more than Big Data, they will have to fully map the customer journey to truly understand customer motivations. This will require the development of adaptive capabilities to collect the right data and refine it into useful insights. Getting to that level of proactive insights will require observing customer behavior by analyzing the massive data sets provided. It will also require listening and monitoring, for example, how customers are interacting on social media. Finally, and most importantly, companies will have to *ask customers why they do the things they do* through traditional channels like surveys and focus groups. The companies need to observe customer behavior but also listen to what customers are saying and, crucially, ask for their insights. If the retailers can manage to implement these steps, they stand a good chance of improving their marketing capabilities and attain a competitive advantage. Company D concludes:

"We work with analysis of transactional data to see how it looks today, customer panels to understand what our customers wants and with qualitative research to get a deeper understanding - you need to have all three perspectives."

6. Concluding discussion

In the following section we will present and discuss our key findings based on the research questions posed and the results from our empirical findings and analysis.

6.1 Conclusions

6.1.1 How is Big Data influencing marketing and value creation in retail?

The results from our study indicate that Swedish retailers are only in the starting-blocks of capturing the potential value that data analytics can bring. Today they mainly focus on customizing promotional offers based on behavioral customer data. However, we found that several other marketing areas could profit from data insights. The reason for this prominent focus on promotional activities can be explained by the fact that the required data is accessible and can be refined using existing resources and does therefore not require any major investments. Even though the studied retailers are only scratching the surface, they have gained substantial value through data analysis. This indicates that there is considerable value still uncaptured for Swedish retailers and should be seen as an encouragement to advance further within Big Data analytics.

6.1.2 What are the major hindrances for Big Data exploitation amongst retailers today?

During our study it became evident that the major hindrance constraining Swedish retailers today is the lack of organizational support. This seems to be caused by an uncertainty within the management whether the large investments required to implement Big Data will generate sufficient returns. This results in insufficient financial means for investments in the right competence and analytical tools. Even though some companies are possessing the right competence they lack the data systems or the time required to perform the desired analysis. Another barrier is the inability to connect various data sources, either due to legacy systems or lack of knowledge. Organizational structures hindering data and insights to be shared between departments are also a problem amongst many of the retailers. These barriers together form a massive resistance for Swedish retailers to develop, and overcoming these hindrances will require large-scale investments.

6.1.3 What marketing capabilities are required of retailers for further development?

In order to win in the data transformation, the retailers must be able to collect valuable data, but most importantly be able to refine it and process it for an actual purpose. To advance,

organizations need to prioritize the consumer insights departments and provide them with the financial means to invest in the right tools and acquire competence with the right analytical knowledge.

The retailers must look for new possibilities instead of focusing on previous purchase behavior. They must learn how to predict the demand and always be one step ahead of the customer, and this requires the development of adaptive capabilities. These capabilities will be essential to create customer loyalty in today's competitive retail industry where the customers are posing higher and higher demands. To be able to predict customer needs companies need to combine behavioral data and insights from qualitative market research. Combining data sources will be essential in order to derive valuable insights and companies must therefore understand the importance of investing in compatible IT systems and competence with knowledge of how to use it.

6.1.4 Final conclusion and discussion

Our study concludes that most Swedish retailers are capturing only a fraction of the potential value of Big Data. Developing adaptive capabilities and investing in a compatible IT structure and analytical competence will be key for the retailers' survival. In the highly competitive retail industry companies need to fully adapt and understand their customers, they have to be customer oriented. In the future it will most likely be the customer telling the retailers what to offer instead of the retailer telling the customer what to buy. This highlights the importance of deriving insights from Big Data.

When looking at Big Data's impact on the marketing mix, we have seen a lack of focus in all other decisions except promotion. We have also noted a general lack of adaptive capabilities amongst the studied companies. One possible explanation might be that the majority of insights used in promotion decisions are based on transactional data. Since this data is rather reactive the companies' dynamic capabilities could explain why they are able to progress within this decision. This leads us to the question whether there might be a correlation between the lack of adaptive capabilities and the lack of focus on the price, product and place decisions. Proceeding within these marketing decisions will require combining several data sources, both internal and external, implying that the analytical approach required in these areas need to be of a more predictive nature.

6.2 Strategic and managerial implications

Our findings suggest several managerial implications in order for retailers to improve their marketing capabilities. Since acquiring the right analytical competence will be crucial for companies to develop further, they first need to ask themselves *what competence do we*

actually need? When the right competence is identified, managers will have to realize the value that skilled analysts can provide and consider it as an investment that can generate high returns. Secondly, the right competence has to be matched with the right technological tools. This implies further investments in the IT structure, either by connecting old legacy systems or building new ones able to combine various data sources. Finally, once the analytical foundation is in place it will be the task of the managers to make sure that the new insights are shared throughout the organization in order to support more informed business decisions.

6.3 Critique and limitations

Due to the exploratory nature and limited scope of our study the findings can not be generalized for the whole retail industry. The reached conclusion should not be seen as definite but rather as suggestions serving as a basis for future research. This is due to the chosen methodology and limitations discussed in section 1.2. By adding more cases, our findings generalizability could have been approved and provide a better reflection of the industry.

6.4 Future research

Due to the exploratory approach of the study, additional areas for further research has been identified. To broaden the research, it would be interesting to investigate other industries and countries in order to compare results and identify differences and similarities as well as possible learnings. It would be especially interesting to study how pure online-retailers are exploiting the potential of Big Data, since they operate entirely in the digital arena, and have access to new formats of data. Since our study is focused on companies selling consumer goods, it would also be interesting to perform a similar study with companies in the service industry to investigate how they are using Big Data analytics.

Due to the research gap within the area, the purpose of our study has been to explore the Big Data landscape within retail. To take our research further and to get a more throughout understanding, one could look deeper into how retailers can apply data insights in the marketing mix focusing on one of the four components. Further research should also investigate whether our hypotheses that adaptive capabilities are required to progress in the other marketing decisions holds true.

7. References

Abercrombie, N., Hill, S. & Turner, B.S. 1984, *Dictionary of sociology,* Penguin Books.

Barney, J. 1991, "Firm resources and sustained competitive advantage", *Journal of management,* vol. 17, no. 1, pp. 99-120.

Bell, J. 2006, "4. uppl", *Introduktion till forskningsmetodik,*

Bharadwaj, A., El Sawy, O.A., Pavlou, P.A. & Venkatraman, N.V. 2013, "Digital business strategy: toward a next generation of insights",.

Bryman, A. & Bell, E. 2015, *Business research methods,* Oxford University Press, USA.

Carr, D. 2013, "Giving viewers what they want", *The New York Times,* vol. 24, pp. 2013.

Davenport, T.H., Barth, P. & Bean, R. 2012, "How big data is different", *MIT Sloan Management Review,* vol. 54, no. 1, pp. 43.

Centri consulting (2017). *2017: The Year of Big Data?* (Retrieved Apr 15, 2017 from https://centricconsulting.com/2017-the-year-of-big-data_boston/).

Cisco (2014). *Visual networking index IP traffic chart.* Cisco Corporate Website (Retrieved Feb 8, 2017 from http://www.cisco.com/cdc_content_elements/networking_ solutions/service_provider/visual_networking_ip_traffic_chart.html).

Dijcks, J. 2012, "Oracle: Big data for the enterprise", *Oracle White Paper,*.

Dubois, A. & Gadde, L. 2002, "Systematic combining: an abductive approach to case research", *Journal of business research,* vol. 55, no. 7, pp. 553-560.

Ebner, K., Buhnen, T. & Urbach, N. 2014, "Think big with Big Data: Identifying suitable Big Data strategies in corporate environments", *System Sciences (HICSS), 2014 47th Hawaii International Conference on*IEEE,, pp. 3748.

Eisenhardt, K.M. 1989, "Building theories from case study research", *Academy of management review,* vol. 14, no. 4, pp. 532-550.

Erevelles, S., Fukawa, N. & Swayne, L. 2016, "Big Data consumer analytics and the transformation of marketing", *Journal of Business Research,* vol. 69, no. 2, pp. 897-904.

EY (2014) *Big data Changing the way businesses compete and operate* (Retrieved May 6, 107 from http://www.ey.com/Publication/vwLUAssets/EY_-_Big_data:_changing_the_way_businesses_operate/%24FILE/EY-Insights-on-GRC-Big-data.pdf).

Ferreira, K.J., Lee, B.H.A. & Simchi-Levi, D. 2015, "Analytics for an online retailer: Demand forecasting and price optimization", *Manufacturing & Service Operations Management,* vol. 18, no. 1, pp. 69-88.

Flick, U. 2009, *An introduction to qualitative research,* Sage.

Fortune (2016). *Why data is the new oil.* Fortune (Retrieved May 7, 2017 from http://fortune.com/2016/07/11/data-oil-brainstorm-tech/).

Flyvbjerg, B. 2006, "Five misunderstandings about case-study research", *Qualitative inquiry,* vol. 12, no. 2, pp. 219-245.

Hofacker, C.F., Malthouse, E.C. & Sultan, F. 2016, "Big data and consumer behavior: Imminent opportunities", *Journal of Consumer Marketing,* vol. 33, no. 2, pp. 89-97.

IBM (2012). *What is big data?* IBM Corporate Website (Retrieved Feb 8, 2017 from http://www-01.ibm.com/software/data/bigdata/).

IDC (2014). *The digital universe of opportunities: Rich data & the increasing value of the internet of things.* EMCompany A (Retrieved Feb 8, 2017 from http://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf).

Jobber, D. & Ellis-Chadwick, F. 2012, *Principles and practice of marketing,* McGraw-Hill Higher Education.

Laney, D. 2001, "3D data management: Controlling data volume, velocity and variety", *META Group Research Note,* vol. 6, pp. 70.

Lycett, M. 2013, "'Datafication': Making sense of (big) data in a complex world", *European Journal of Information Systems,* vol. 22, no. 4, pp. 381.

Malhotra, N.K. 2008, *Marketing research: An applied orientation, 5/e,* Pearson Education India.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. & Byers, A.H. 2011, "Big data: The next frontier for innovation, competition, and productivity",.

Merriam, S.B. & Nilsson, B. 1994, *Fallstudien som forskningsmetod,* Studentlitteratur.

Onwuegbuzie, A.J. & Leech, N.L. 2007, "A call for qualitative power analyses", *Quality & Quantity,* vol. 41, no. 1, pp. 105-121.

Quader, M.S. 2007, "Human resource management issues as growth barriers in professional service firm SMEs", *Journal of Services Research,* vol. 7, no. 2, pp. 115.

Raconteur (2013). *How big data is changing marketing.* Raconteur (Retrieved May 2, 2017 from https://www.raconteur.net/business/how-big-data-is-changing-marketing)

Ritchie, J., Lewis, J., Nicholls, C.M. & Ormston, R. 2013, *Qualitative research practice: A guide for social science students and researchers,* Sage.

Trost, J. 2010, "Kvalitativa intervjuer. 4: e upplagan. Lund: Studentlitteratur. Trost, Jan (2007)", *Kvalitativa intervjuer,* vol. 3.

Weiss, S.M. & Indurkhya, N. 1998, *Predictive data mining: a practical guide,* Morgan Kaufmann.

Yin, R.K. 2013, *Case study research: Design and methods,* Sage publications.

8. Appendix

Appendix 1: Interview guide

Background
- Current position
- Primary duties
- Educational background
- Previous employments and positions
- Time of employment at company X

Company information
- What does the organizational structure look like at company X?
- How many employees are working with consumer insights?

Big Data usage
- How is company X working with Big Data?
  - Which departments are using Big Data?
  - To what extent is company X implementing the insights from Big Data?
  - Are these insights shared between the different departments?
- Has the access to Big Data changed company X business processes when it comes to:
  - Product development and product assortment decisions?
  - Price setting?
  - Place decisions?
  - Promotion and communication?

Value creation
- Has the use of Big Data created measurable values for the company X?
  - Has Big Data given company X a better understanding of your customers needs?
  - On a scale 1-10 how far do you perceive that company X has come in Big Data analysis?
  - What opportunities do you believe that Big Data can generate for company X in the future?
  - What barriers are currently restraining company X?

Innovation
- Has the access to Big Data affected the innovation in company X?

Appendix 2: Translation of Citations

4. Empirical findings

4.1 Big Data's impact on marketing decisions

No quotes.

4.1.1 Product

Citation 1:

"Absolut så har vi funnit värde i sortimentsoptimering genom Big Data. Och då pratar vi betydande värden." (Company C)

"We have certainly found value in product range optimization through Big Data. And we're talking significant values." (Company C)

Citation 2:

"Vilka produkter som köps in är inte datadrivet på det sättet utan det är alltid snarare vad som säljer och inte säljer." (Company B)

What products that we purchase is not data driven in that sense, it's rather a question of what is selling and what is not". (Company B)

Citation 3:

"Vi jobbar varje dag för att sträva efter det. Idag är det mer att jag förmedlar en insikt som dom kan använda för att ha lite i bakhuvudet när de köper sitt sortiment." (Hemtex)

"We strive to do that everyday. Today it's more that I mediate an insight that the supply department can bear in mind when they purchase their product range. " (Hemtex)

4.1.2 Price

Citation 4:

"Ja, vi jobbar med det nu faktiskt. Jag hoppas det kommer att få betydande affärsvärden för oss." (Company C)

"Yes, we are actually working on that right now. I hope it will lead to significant business value for us." (Company C)

Citation 5:

"Vi hoppas att det ska leda till att man ökar sin marginalintäkt genom att få en viss procent av kunderna att byta från en vara till en annan som är lite mer lönsam. Så varje sånt byte kan ge ganska många kronor i slutändan." (Company D)

"We hope it will lead to increasing marginal revenues by getting a certain percentage of customers to switch from one product to another that is slightly more profitable. Each switch can generate quite a lot of money in the end. " (Company D)

Citation 6:

"/.../ så en blandning av just rent datadrivet med kategorikunskapen också." (Company D)

"/.../ it's a combination of pure data driven information and category expertise." (Company D)

Citation 7:

"Ca 40% av det som säljs totalt går inte till fullpris, och då är det ju ganska viktigt att ha koll på var man tappar bruttovinst någonstans"….. "det är så vi styr priset då, snarare än kanske själva prissättningen i sig." (Company B)

"Approximately 40% of goods sold in total are not sold at full price, and that makes it quite important to keep track of where you are losing gross profits"/.../"that is how we control the price rather than the price setting." (Company B)

4.1.3 Place

Citation 8:

"Jag skulle absolut säga att vilka butiker vi väljer att ha kvar eller inte baseras ju på hur lönsamma de är." (Company A)

"I would definitely say that which stores we choose to keep or not are based on how profitable they are." (Company A)

4.1.4 Promotion

Citation 9:

"/.../ det är mest promotion där vi kan se de större datamängderna på köphistorik har påverkat vad vi kommunicerar och hur vi kommunicerar." (Company A)

"/.../ it's mainly in our promotion where we can see that the large amounts of data on purchase history has influenced what and how we communicate." (Company A)

Citation 10:

"/.../ det finns ingen anledning att göra det när vi har så mycket statistik på våra kunder." (Hemtex)

"/.../ there is no reason to guess when we have so much data on our customers." (Hemtex)

Citation 11:

"Vi använder ju all den data vi kan så att säga, för att då se till att vi tar rätt beslut egentligen." (Company A)

"We use all the data we can, so to speak, in order to make sure we make the right decision." (Company A)

Citation 12:

"Det är väldigt svårt att mäta ROI på just den typen av aktiviteter, en kampanj kan ju lätt synas men vad betyder det egentligen?" (Company D)

"It's very difficult to measure ROI on these particular type of activities, a campaign can be visible, but what does it really mean?" (Company D)

Citation 11:

"Vi arbetar med data mining och analys av många olika datakällor i en databas för att kunna göra mera avancerade analyzer av köpbeteende, mediainvesteringsanalyser, målgruppsanalyser med mera." (IKEA)

"We are working with data mining and analyzing many different data sources in a database to be able to do more advanced analysis of purchasing patterns, media investment analysis, target group analysis etc." (IKEA)

Citation 12:

"Vi ser ju att olika budskap funkar i olika kanaler, men jag skulle ändå säga att det är en kostnadsfråga, organisationer ser och tänker kostnader först, man kan inte se bort från kostnader utan det är alltid primärt." (Company B)

"We see that different messages works in different channels, but I would still say that it's a matter of cost, organizations look at and think of costs first, you can not see past the costs, it is always primary." (Company B)

4.1.5 Value creation

Citation 13:

"Vi kan sätta det i relation till något." (Willys)

"We can put it in relation to something." (Willys)

Citation 14:

"/.../ genom att vi anpassar kommunikationen efter den datan vi får in och analyserna vi gör, så ser vi ju ett uplift på våra kampanjer." (Hemtex)

"/.../ by adapting the communication based on the data we collect and the analyzes we do; we can see an uplift on our campaigns." (Hemtex)

Citation 15:

"/.../ vi kan också få en helt annan kontroll över kommunikationen, och insikt över hur mycket värde ner till hur många kronor det ger." (Company A)

"/.../ we can also get a better control over the communication, along with insights of how much it's worth in crowns." (Company A)

Citation 16:

"Jag skulle säga att Company B har varit lite dåliga på att samla ihop den faktan eller bevisen på att det har blivit bättre." (Company B)

"I would say that Company B haven't been that good in gathering the facts or evidence to show that is has improved." (Company B)

4.1.6 Retailers perceived progress in Big Data exploitation

No quotes.

4.1.7 Concluding remarks

No quotes.

4.2 Hindrances for value creation

Citation 17:

"Jag tror bara vi har skrapat på ytan alla i branschen än så länge." (Willys)

"I think that everyone in the industry have just scratched the surface, so far." (Willys)

Citation 18:

"Jag tror att man kommer kunna leverera ett företag med fler produkter, tjänster och serviceerbjudande - ett mer kunddrivet företag och förhoppningsvis kommer vi automatisera och hjälpa med att samla ihop och analysera data och bli mer effektiva." (IKEA)

"I believe that we will be able to become a company with more products and bundled services – a more customer driven company, and hopefully we will automatize which will facilitate gathering and analyzing data, and becoming more effective." (IKEA)

Citation 19:

"Det innebär kanske att man tar fram rätt produkter så att man inte slösar tid och pengar och energi på att ta fram en produkt som sen inte mottas av konsumenterna. Det kanske innebär att kunderna får en bättre relation med ICA för att man känner att ICA är mer relevant." (ICA)

"It might imply that we can provide the right product's so that we don't waste time, money and energy on developing a product that will not be well-received by the customers. It might lead to customers getting a better relationship with ICA because they feel that ICA is more relevant." (ICA)

Citation 20:

"Det är ju ett hav av information som finns, det gäller bara att hitta vad vi ska använda den till och hur vi kan använda den på ett bra sätt." (Willys)

"There is an ocean of information available, you just have to find what to use it for and how we can use it in a good way." (Willys)

4.2.1 Physical capital resources

Citation 21:

"IT arkitekturen behöver förändras och göra att dem kan prata med varandra." (IKEA)

"The IT-architecture needs to change and allow the different systems to talk to each other." (IKEA)

Citation 22:

"Utmaningen där är ju just hur synkar man ihop alla dem datakällorna som finns - det är en stor fråga." (Company D)

"The challenge here is how to synchronize all data sources available – that's a big question." (Company D)

Citation 23:

"Och det där kräver en viss investering från början är ju ofta en barriär." (Company D)

"Everything that requires an initial investment from the start is often a barrier in itself." (Company D)

Citation 24:

"Jag tror att det är mycket lättare för nya företag som inte har någon legacy, allt det gamla som är i vägen för att kunna möjliggöra det nya." (Company B)

"I think it's much easier for new companies that don't have any legacy, old systems are standing in the way to allow for new ones." (Company B)

4.2.2 Human capital resources

Citation 25:

"I New York/ Silicon Valley så förstår bolagen att det är väldigt billigt för vad de skapar. De har möjligheten att lösa morgondagens problem. Riskkapitalisterna förstår, men inte en retailer. Och det är en utmaning." (Company C)

"In New York or Silicon Valley companies understand that it is very cheap for what they create. They have the ability to solve tomorrow's problems. Risk Capitalists understand, but not a retailer. And that's a challenge." (Company C)

4.2.4 Organizational capital resources

Citation 26:

"Det här är ett mänskligt problem och därför så krävs det mycket bra ledarskap för att lotsa ett företag igenom en sån transformation." (Company C)

"This is a human problem, and therefore, it takes very good leadership to guide the company through such a transformation." (Company C)

Citation 27:

"Nu sitter vi alla på marknadsavdelingen men det är fortfarande lite silos ändå." (Hemtex)

"Even though we are all working in the marketing department, there are still silos within the department in a way." (Hemtex)

4.2.5 Concluding remarks

No quotes.

4.3 Dynamic and Adaptive capabilities

Citation 28:

"Så vi kan ju se att det förändras och hur man beter sig men vi ser ju inte vad man kanske egentligen hade velat köpa. Utan det blir ju lite gissningslek." (Willys)

"We can see that it is changing and how people are acting, but we don't see what they actually would like to buy. Instead it becomes kind of a guessing-game." (Willys)

Citation 29:

"Köpdatan i sig blir ju ganska reaktiv, vi ser ju vad kunderna har köpt men den säger ju inte så mycket om vad de vill ha i framtiden eller vad de tycker om Company A eller sin köpupplevelse." (Company A)

"The transaction data is quite reactive; we can see what the customers have bought but it doesn't say much about what they want in the future or how they actually perceive Company A or their shopping experience." (Company A)

Citation 30:

"/.../ det är ju att vara proaktiv som är A och O. Det är så man får kunden nöjd i slutändan. Man löser problemet före kunden själv upplever att det var ett problem." (Company C)

"/.../ being proactive is vital. That is how we get the customers satisfied in the end. You solve the problem before the customer even realizes that the problem exists." (Company C)

Citation 31:

"Från att vi talar om vad kunden skulle köpa, så kommer det bli kunden om talar om för oss vad vi ska köpa in." (Hemtex)

"From us telling the customer what to buy, it will become the customer telling us what products to offer." (Hemtex)

4.3.1 Concluding remarks

Citation 32:

"Ja det krävs för att kundupplevelsen ska bli mycket mer relevant, jag tror att ju mer vi vet om våra kunder desto mer kräver våra kunder av oss också." (Willys)

"It is essential so that the customer experience will be much more relevant, I believe that the more we know about our customers, the more our customers will require from us." (Willys)

5. Analysis

5.3 Marketing capabilities required for further development

Citation 33:

"Vi jobbar med analys av transaktionsdata för att se hur det ser ut nu, med medlemspanelen för att se vad vill folk ha och med kvalitativa undersökningar för att gå in på djupet - man behöver liksom alla tre perspektiven." (Company D)

"We work with analysis of transactional data to see how it looks today, customer panels to understand what our customers wants and with qualitative research to get a deeper understanding - you need to have all three perspectives." (Company D)