Customer Metrics and How Store Performance is Related to Them

Bjorn P* and Matteo A

1Department of Int. Management and Economics, Copenhagen Business School, Denmark
2School of Business Society and Engineering, Malardalen University, Sweden

Abstract

This paper aims to give the reader an overview of the latest available metrics to track customer behavior in physical retail stores. This presentation is mainly based on a literature review and leads in the next step to the question about how useful are such metrics and how can they be used. Therefore, a representative study in Sweden was conducted to measure which factors have an impact on customer performance and which not. Based on statistical analysis the paper finds out that there is a significant relation between the usage of metrics and the profitability of the store.

Keywords: Marketing; Tracking; Customer metrics; Customer behavior; Physical retail stores

Introduction

Business and marketing management studies show that maintaining transactions with customers is essential for a sustainable growth [1]. However, the more the business grows, the more data and information need to be collected and interpreted. This fact can be critical for physical stores, often referred as “offline store” or “brick and mortar store” [2]. Brick and mortar means in this context that the customer can has a physical interaction with goods in the store (Ibid). The thesis aims to give the reader an overview of different metrics (e.g. retention rate, turnover per customer) that can be used to track customers in physical stores. The term metrics is defined in this paper that metrics are any type of measurement used to quantify data in a corporate environment [3]. This study will focus only on marketing metrics that are related to physical retail stores. In addition, this study shows which metrics are used by profitable stores.

In the 1980s, the science of shopping and with that the tracking of customers in stores became popular. Underhill describes that these studies were mostly carried out in the form of research projects and due to that the measurement of the customer was not constant and the metrics were not standardized [4]. At the time, data was collected only once a year, allowing stores to have less than once a year insights of how the customer acts. Advisory firms who also created the metrics later did the interpretation of the data.

Later on, graphical driven user interfaces were improved and led to the development of the internet and with that the possibility of online retailing [5], changing the way of how retailing can be carried out [6]. The risks and potentials derived from this development forced the stores to use metrics in order to measure and control their performance more effectively (Ibid). The metrics that were developed for this purpose gave companies the opportunity to track their activities and to improve their knowledge about the customer constantly [7].

Afterwards, a large number of online stores arose, which used modern technology to track and measure consumers [5]. The metrics used nowadays by online shops are highly standardized and contain measures like visits, queries, click-through and conversion rates [7]. In online stores the metrics are well connected, which allows the store manager to analyze how changes in the store influence the way customers act (Ibid). This can also be done frequently and in real time (Ibid). However, traditional physical stores might suffer from the inadequacy of their personnel capabilities and their information system (e.g. customer relationship system) to develop and use similar powerful metrics as online stores [5]. Accordingly physical stores lack in the possibility to carry out experiments in order to find ways to become more profitable (Ibid).

Problem statement

The background shows that physical stores have a deficit when it comes to the application of certain metrics. Shankar et al. states that when it comes to shopper marketing, there is a lack in standardized metrics for the physical store [7]. However, in the past years there was an increase in marketing metrics, which aims to give the user more insight into the performance of the marketing activity and the way customers act [8]. The usage of metrics in marketing helps to increase the accountability of the firm, justifying spending and the selection of marketing actions in general [9]. When the store is able to use metrics in order to identify drivers of customer behavior, the managers can use that knowledge to maximize profits of the store [10].

According to Shankar et al., traditional metrics to track customers and measure their performance are insufficient [7]. The performance definition is defined in this paper as how a customer contributes to the profitability of the company, this means that a customer with high performance buys more products on a more regular basis from the same store than one with a lower performance. Therefore, they need to be augmented and new metrics have to be tested [7]. Especially metrics that measure attention and reflection of customers could be interesting and would provide useful information for the shop manager for example cross buying metrics (Ibid). It might be possible that the data collection can be automated in order to calculate this metrics. The metrics can then be constantly calculated and analyzed on a day-to-day base. The resulting insights that the manager get from the metrics need to be implemented into applications in order to make the interpretation easy to understand for people at different management levels [7].

Academic research explored in the past years a wide range of metrics that are able to track the customer and their shopping cycle in more detail [7]. Most of these metrics that appear to be efficient...
in theory tend to be hard or too complex to track within a physical store [8]. In the physical world, store managers and personnel might not understand how to use such metrics. More commonly, they do not have access to proper tools to measure the metrics that can lead to a wrong usage and interpretation of the metrics [7]. Furthermore, it could be the case that the interpretation of some of the metrics is too complicated and would require too much time [11]. This would lead to inefficiency because of the complexity of the data or the lack of proper tools, especially for small businesses (Ibid). Based on that it can be necessary to define specific metrics related to the level of management (e.g. store manager, worker) where they are used [7]. Doing so, would lead to the fact that each level would get access to the required metric to carry out the job.

Research Questions and Hypotheses

The purpose of this study is to investigate to which extend metrics, that are discussed in current research publications, are used in physical retail stores. Through these questions, the reader gains more understanding of metrics that can be used in order to track customers effectively and efficiently. Furthermore, the study examines which metric groups are used by profitable stores and whether there are differences between physical retail stores which can be clustered by using attributes for example size, profitability and type of business. This leads to the following research questions:

RQ1: What ways are there to measure customers’ performance in physical retail stores?

RQ2: How do the profitability of the store and the usage of specific metrics correlate?

RQ3: Which role plays the complexity in how store managers see the usability and usefulness of the metric groups?

Based on these research questions and the prior problem statement, the four hypotheses are stated and further tested in this study.

The variety of metrics can range from those that are easy to use and implement like the sales figure [8], to more complex such as cross buying metrics which are forecasting and probability metrics that are based on statistical calculations and requires constant data collection in order to make reliable predictions [12]. Current research has shown that companies tend to use only single metrics although it states that some metrics cannot be used alone [13]. The complexity of some of the metrics requires knowledge and special training for the operating personnel as well as a complex IT system [14]. Furthermore, the operationalization is for various metrics not easy and requires knowledge about this and the resources for testing the metrics [8]. For this study we assume that knowledge for dealing with the metrics is linked to the education. Based on this the following hypothesis can be stated:

H1: The usage of metrics depends on managers’ education.

According to Shankar et al. [7], big retail chains have the possibility to use technology and knowledge to analyze and influence customer decisions within the store. They grow according to Reinartz et al. and dominate the market place [7,13]. From a management point of view, this can be seen as complex which requires a standardized metric system [7]. Furthermore, bigger companies have more IT capabilities that can be used in order to collect and analyze data [10].

H2: Stores that belong to a retail chain use more metrics.

Kumar et al. stated that specific metrics require special knowledge that can lead to a lower usability of the metrics [14]. The way of how easy the metric is seen, influence the way how useful participants of the survey see metrics due to a lack of understanding them [8]. This leads to the following hypotheses H3a and H3b.

H3a: The individual preferences of the usefulness of metrics are related with the usability.

H3b: Profitable stores rank the usability and the usefulness higher.

Getting to know the customer and understand its behavior is crucial for the success and profitability of the company [1]. Stores get through specific metrics more insights about how customers act in the store [9]. Based on this the following hypothesis can be drawn:

H4: Stores that use more metrics are more profitable.

In order to answer the research questions and to test the hypotheses, a review of customer metrics in current research will be shown to give an overview of common metrics that are stated by the recent literature.

Review of Customer Metrics in Current Research

In the past, metrics were used by marketers and managers to justify spending in marketing [8]. Latest development, however, has shown that there is an increase in marketing metrics, which increment the accountability of the firm [8,9]. Nowadays, metrics can measure various facts about how a customer acts in a store, from retention of customers to clicks on specific websites e.g. see Farris et al. The next part will present four groups in which the metrics for customer performance can be divided [15].

Customer value

Current research has developed multiple metrics to measure customer value, both on individual (e.g. customer lifetime value) and aggregate level (e.g. customer equity) [8]. Customer Value is seen in this case as the perceived value that a customer has towards a product and how he contributes with that to the sales of the store [16]. This group of metrics, according to current research, can be used by marketers to measure customer selections and the outcome of marketing campaigns [8]. By using this metrics, the manager can lead marketing efforts to those customers who can appreciate and collect the message that the firm develops and delivers [9]. The result of these marketing campaigns can be analyzed and measured with customer value metrics, which can help to draw future patterns for the return on marketing (Ibid). According to Venkatesan and Kumar, there are specific demographic drivers of customer value [17]. Based on these findings, it is possible to generate strategies to allocate resources more efficiently. The calculation is also from an overall firm perspective useful since there are linkages between the value of customers and the overall value of the firm [18]. However, many metrics of this group such as Customer’s Lifetime Value (CLV) are difficult to implement in a physical retail store according to Petersen et al. [8]. The importance of customer value as a system of measurement, lead to the development of specific metrics: for this study, 12 customer value metrics have been collected from current researches.

The first important metric that is possible to analyze in this group is the Customer’s Lifetime Value (CLV); this metric helps to predict possible future customer patterns by analyzing previous purchasing data [17]. In order to draw such pattern through the usage of sophisticated statistical models, the marketer needs to investigate the present value of the customer such as his or her contribution margin and the cost of marketing the firm is spending in order to reach the
customer (Ibid). The data collected by the marketer comes from each transaction the customer makes in a period of time which is usually not more than one month (Ibid).

A related metric to CLV is the Past Customer Value (PCV), which is used to measure the past profit of each customer [8]. This metric is calculated based on the same information as the CLV through the customer transactions within the last month [17]. However this metric focuses on the value that the customer had in the past instead of the value over the whole lifetime which include assumptions about the future value [8,11].

More complex to measure is the Customer’s Referral Value (CRV) [8]. This metric considers the number of attempts that a marketing campaign needs to be run before the customer reacts and addresses his or her behavior to the company [11]. The complexity of this metric derives from the amount of data necessary to build the statistical model and the range of time, usually between three to six months, depending on the industry (Ibid).

Based on databases, marketers can measure the Share of Wallet, focusing on how the customer spends his or her budget within the same business area. This metric can be used to calculate the share of the budget a customer spent in the store, compared to the spending’s in stores of competitors. For this metric external data about the size of the wallet and competitors is needed [8]. The results can direct the firm’s marketing efforts towards certain customers in order to increase their designated Share of Wallet and profitability (Ibid). Furthermore, it is stated that the share of wallet and the CLV are linked in the way that the share of wallet influences the CLV (Ibid).

A common metric, easy to recognize, is the Basket Size, which represents the quantity of products a customer buys in a single shopping session [19]. According to current research, customers with a smaller Basket Size tend to pay a higher price because they are pushed by an unplanned consumption need (Ibid). At the opposite, customers with larger basket size planned the shopping session and will stick with certain brands (Ibid). Both this opposite sizes indicate a relationship development between the customers and firm (Ibid). Customers with intermediate Basket Size seek for the lowest price which limits the development of relationships with the firm [20].

Analyzing this metric allows marketers and store managers to make certain adjustments in order to lead intermediate Basket Size customer to switch to a larger or smaller Basket Size (Ibid).

Another important metric is the Level of Price Discounts. Discounts are perceived by customers as a gain because the price they pay is less than the normal price [21]. Because of this benefit, the customer is more prone to develop a relationship and be loyal to the firm (Ibid). The level of price discount leads the customer to make a decision, whether to buy an item or more from the same firm faster because of a motivation by discounts [19].

Store managers often deal with issues like the return of items purchased by customers. This can be measured with a metric called Proportion of Returns, which shows how often a customer brings items back to the store [19]. If this number is high, it means that the quality of the product or the specifics are below standards and the level of customer satisfaction is low (Ibid). Store managers must put great efforts in order to deal with these situations because there is a threshold beneath where customers cut down the relationship with the firm (Ibid). The faster and more effective the responses are, the higher is the customer happiness which can lead to retain of loyalty and processes like positive word of mouth (Ibid).

Part of the customer value metrics analyze the level of interaction between customers and a specific firm [19]. Such metrics measure the frequency of these interactions whereby a high ratio creates a feeling of familiarity and trust (Ibid), which contributes to increase the firm’s profit. The most common frequency related metric is the Purchase Frequency, used by marketers to measure how many times a customer purchases from the same store (Ibid). The higher the Purchase Frequency is, the higher is the familiarity and the possibility that the customer returns in the future to buy again (Ibid). According to Morgan and Hunt [22], a high purchase frequency produces satisfaction and trust, while other authors observed that a higher ratio of this metric leads customers to develop positive impressions and engages benefits derived from the interaction itself [23].

Communications allow firms to inform and establish interaction with the customers, for this reason the Frequency of Marketing Communication is a critical issue to measure because it is able to influence customers’ choice [24]. According to Fournier [25], this metric is important in order to calculate the real range and effectiveness of the communication and the different thresholds that the message sent by the firm can reach. Upon a certain threshold, the customer perceives the communication as a way to develop a mutual beneficial relationship with the firm but when the frequency of marketing communication is too high and exceeds this threshold, the effect is the opposite (Ibid). Beyond this threshold in fact, the customer feels that the firm does not understand his or her needs but only pushing its products for profit’s sake (Ibid). The right balance for this metric should be in between, where the frequency is not too low nor too high (Ibid).

When the customer purchases, he or she can decide to use different stores (offline or online), services or firms [26]. Each one of these have different characteristics that can ease the purchase and usage of a specific channel instead of others (Ibid). A critical factor when talking about mortar and bricks stores is the distance between the stores and the customer; such issue can be measured through the Travel Cost Proportion. This metric calculates the sum of all travels the customer does in order to reach a specific store (Ibid). According to Bell et al. [27], customers tend to buy in stores near their residence as a result of low travel costs. Physical distance influences the development of relationships with the customer; if a customer establish the same level of relationship with two different stores or firms, he or she will be inclined to purchase in the closest store [19]. Travel Cost Proportion is a metric that can be only used for physical stores, in fact it is not applicable for online stores which do not have to deal with this issue [19].

Another discriminating factor between offline and online stores is the Immediate Product Availability (IPA). When the purchase is made in offline stores, the customer has the possibility to consume the product right after the product was bought [27]. Such experience and interaction with the product cannot be provided by online stores because of the range of time between the process of purchasing and the collection of the item (Ibid). The proper metric to measure this factor is the IPA. This metric measures the cost related to the customer when he or she can experience the immediate consumption of the product or has to wait for a specific period of time (Ibid). The lower this proportion the shorter will be the process of adoption of a store or firm for the customer (Ibid).

The Net Promoter Score as the last metric in this group describes the ratio of promoters (someone who promotes the shop by giving positive comments to third parties) to detractors (someone who spreads critiques over a certain store) of the store [8]. According to Reichheld [28], this
metric can be used to calculate and analyze customer loyalty. The data can be collected by using surveys. Furthermore, it will be possible to use this metric in order to analyze competitors’ customers and use them as a benchmark, if external data is available (Ibid).

**Cross buying**

According to Venkatesan et al. [19], cross buying is the process of purchasing products that belong to different categories during a single shopping session. Within this behavior, the customer, instead of buying different categories of products in multiple stores, prefers to make the whole purchase in the same round choosing only one store. The first important aspect of such behavior is the direct association between cross buying and customer satisfaction (Ibid). The reason for this is that the individual feels more comfortable making the purchase in one specific store (Ibid). Petersen et al. showed that cross buying consumers increase the revenue and with that the profit of the store [8]. One reason that can make the cross buying a challenge for physical stores is that it requires a more complex marketing strategy (Ibid). The combination of customer, product, and message must be chosen wisely in order to get the right effects (Ibid). For the process of identifying the right customers, there are different drivers available (Ibid). Most important is, when it comes to cross buying, to target the right customers with the right products at the right time [8]. According to Kumar et al. [11], there is a difficulty when it comes to the usability and usefulness of this metrics. This is especially the case when drawing conclusions for strategies in order to increase cross buying behavior [Ibid].

The first metric in this group is the Next Product to Buy Model (NPTB). The model is based on a forecasting equation that predicts which product a particular customer is most likely to buy at a specific time [12]. In order to perform this process, different statistical methods can be used. The most common and from various scholars advised is the logistic regression (Ibid). The reason for using this method above others is easy implementation which makes the NPTB model more easy to use [12]. Furthermore Stevens states that logistic regression is the most common used model when it comes to a violation of the normality assumption which is in consumer data often the case. In his research Knott et al. showed that the application of NPTB models help stores to generate more cross selling profits then other common heuristic methods which are not based on statistics [12]. Also if the logistic regression method is the best one, there can be other statistical models be used in order to create NPTB models (Ibid). However, according to Knott et al. [12], the statistical method that is used by the model has not a high influence on the possibility of predictions made by the model about the buying behavior of customers.

Likelihood of Purchase is another metric that can be used in order to get knowledge about the cross buying behavior of the customer [14]. In order to market a product better, it is important to be able to predict what products a customer will buy in the future (Ibid). This is often done in a probability cube where different customers or groups of customers will be shown on the y-axis (Ibid). The products or product groups will be displayed on the x-axis and the time on the z-axis (Ibid). The probabilities are estimated by the marketers (Ibid). First, which probability is there to purchase a certain product? After that the marketer has to assume a probability of when the products will be bought [14]. In order to overcome the inaccuracy of estimations of the marketer, another method was developed (Ibid). It uses the so called Bayesian estimation (Ibid). This is a mathematical and statistical method to specify the range of a weighting factor that could have produced the actual data which was collected during operating the store (Ibid). The improvement compared to the classical way is that this method uses not a single weighting factor like a traditional regression analysis (Ibid). A major problem most companies might face with this metric is, beside the complexity, that they might have a lack of enough data on all their customers to estimate relationships between drivers of purchases that are meaningful (Ibid). In addition, another problem could be that the number of data sets is just too big for computers (Ibid).

The third metric that this study wants to present in this category is the Past Purchase Frequency. With this metric the frequency of the past purchase of a customer is measured [8]. From the measurement of the frequency of the purchases of customers, the metric predicts future purchases of the customer (Ibid). The problem that current research predicts is that customers might not follow the same frequencies over time (Ibid). This holds especially for a non-contractual setting which can be found in a retail business to consumer market (Ibid).

The Inter Purchase Time can be seen as an improvement of the Past Purchase Frequency [17]. This metric measures the time between purchases of a customer. Base for this are data sets from the past which were collected while operating a particular store. To analyze the data, the model uses a generalized gamma distribution [8]. The assumption on which this method is based is that a customer is likely to reduce the frequency of purchase before ending a relationship [29].

The last metric of this group is the Likelihood of Staying Active (LSA). As input variables it uses the time of purchases and the total number of purchases of a particular customer [30].

\[
LSA = \left( \frac{\text{Average Inter Purchase Time}}{\Delta t} \right)^{\text{Number of Total Purchases}}
\]

\(\Delta t\) is the period of time between the first purchase of a customer and the current purchase.

The higher the number, the higher is the likelihood of staying active. This measure was introduced by Schmittlein et al. and used later by Kumar et al. [30,31]. Past research has shown that the likelihood of staying active increases with the channels that a customer uses for transactions [30].

**Customer acquisition**

It is crucial for the success of marketing that there is a high Customer Retention and a constant acquisition of new customers [8]. This leads to the need of measuring Customer Acquisition and retention (Ibid). It is important to not decide about acquiring customers in an isolated way [32]. According to Thomas [32], retention and acquiring are linked so that a focus on only one of them would lead to not long lasting customers. The goal should also in this case be to increase the CLV (Ibid). Reinartz et al. shows that the likelihood of acquiring new customers, relationship duration and profitability of the customer are linked [13]. However, it is therefore important to have trade-offs between acquisition, retention and profitability in order to maximize the profitability of the firm (Ibid). Verhoef mentions that direct marketing campaigns and/or loyalty programs can increase the purchase behavior of customers [33].

The first metric is the Cost of Acquisition. This metric shows the cost for acquiring a new customer [8]. The measure gives the manager or marketer a guidance which customers should be acquired and which not (Ibid). Previous research has shown that it is not the best to just focus on acquiring the least expensive customers [13]. Profitable customers can be expensive to acquire (Ibid). Therefore this metric should not be used alone since it would not lead to more value creation. It leads just to a reduction of the cost of acquisition (Ibid).
In the same context the metric Cost of Retention can be recalled. This metric describes the cost related to customer retention [8]. Also, this metric needs to be combined with others. Otherwise it leads to the same problems, as stated before, for the cost of acquisition (Ibid).

The next two metrics focus, instead of the cost, on the profit. The first metric is the Acquisition Profit. This metric shows the profit gained through the acquisition of a new customer [8]. Also this metric should be not maximized in isolation. Managers need to focus both on acquisition and retention [13].

The Retention Profit describes the profits gained through the retention of a customer [8]. The same points as for the acquisition profit are for this metric important (Ibid).

The last measure in this group is the Product Return Rate. This metric calculates the number of purchased products that are returned by the customer [8]. Product returns are according to Petersen et al. also key drivers of the CLV. The return of products gets often interpreted in a negative way and has an impact on customer value [34]. However, current research has shown that customers who return 5-10 % of their total purchases, are willing to purchase more in the future [8].

**Transaction with customers**

This group of transaction-related metrics is the most applied on store level [8]. All metrics of this group are combinable in a model that describes how the metrics influence each other (Ibid). An investigation from Schlesinger showed that conversion percentage and units purchased per transaction are important to monitor [35]. The end result of Schlesinger’s study was that these two metrics could be tied directly to future total store sales (Ibid) [35].

The Sales figure is a traditional number from accounting which can also be used to classify customers [35]. This figure fits also to the model developed by Atkinson where he stated that stores need to look at the connections between orders, revenue and visits and should not just focus on one of them [36]. However, sales are a highly aggregated figure so that the user does not get so much detailed insights.

Number of Transactions counts the numbers of transactions in a store [8]. Which can also give an overview of how successful the store is (Ibid). This can directly linked to sales according to the model presented by Schlesinger [35].

Traffic describes the number of visitors of the store in a particular time frame [8]. The next metric that needs to be mentioned is the Conversion Rate. This metric shows in percentage how many of the shopping people actually buy in the store [8]. A problem that occurs when just focusing on this figure is that managers could miss to look on average order size [36]. Just focusing on the Conversion Rate can also mean that the score rises and implies a good development (Ibid). This could be a deception because the value of the metric will also rise, if more shoppers buy but just for a low value [8].

The Average Order Size describes the average value of an order in a store [36]. The metric is the sum of the value of all products that were bought at an average transaction [8,36]. It is hereby important to see the metric in combination with the conversion rate [36].

The metric Average Units per Transaction shows the average number of purchased goods per transaction in the store. Current research has found out that this metric and also the conversion rate correlate strongly with future store sales [8]. According to Schlesinger, average units per transaction should be also a key metric in brick and mortar stores [35].

The Average Unit Price is defined as the price that a purchased good in a store costs in average [8].

To the model of Schlesinger [35], the metric Revenue per Visit can be added. This metric is, according to Petersen et al. [9], the average revenue per visit of a shopper in the store. The metric is also linked to the framework of Atkinson and represents a mixture of conversion rate, number of transactions and sales. However, it is not included in the original framework [36].

**Conceptual framework**

The following section presents the established framework based on the literature presented in the literature review. This framework contains the four groups of metrics that were used to structure the literature review. Each group contains various metrics that deliver specific information to the manager in order to track and measure the performance of the customer in the store. When looking at customers in a physical retail stores, different metrics can be measured in order to understand how they act in the store [8]. The business needs to understand all four groups of metrics in order to have a complete picture of the customer (Ibid). A full perception of these groups allows managers to gather the appropriate information combining metrics from different groups (Ibid). Because of this, stores have to perform in all metrics to be profitable (Ibid).

Stores need to acquire customers and retain those [13]. This leads to a stable customer base that is necessary for the future success of the store. Transactions with the customers are important since they lead to sales which are the basis for profit. Based on the transactions, the store can detect which customers contribute more to the overall sales and to the profit. Cross buying customers are, according to current research, more profitable than those who do not. Therefore, stores are advised to make the customer to cross buy and/or attract customers who are willing to cross buy. Finally, these actions of the customer and the company have to result in high customer value. This previous metrics have also an influence on the value that a customer has for the firm. With customer value metrics the store manager can measure the value and is able to direct its marketing activity better to the right customers. The following table presents the summary of the metric groups, these contain as well as the related literature and the core concepts of the group (Table 1).

This conceptual framework summarizes the main concepts stated by current research. It will further be used together with the method in order to analyze the results of the survey.

**Method**

**Selection of method and data collection**

As Shankar et al. state, it is important to find standardized metrics that measure customer performance in physical stores in the same manner as online stores [7]. Thus, this study aims to explore through current research which metrics, from a theoretical point of view, can be considered as useful and further to understand how such metrics are applied in physical retail stores.

The theoretical concepts about consumer metrics used in this study were collected and have been connected with each other within the literature review. The goal hereby was to see what current research has studied in connection to the research topic. The chosen concepts have been selected by reading scientific articles found in the data bases like Google Scholar, Web of Science, Elsevier. The keywords used to search such articles were: metric, marketing, customer, consumer,
measurement, tracking, KPI, performance, store, retail. All these keywords were used in a separate and conjoint way in order to get more results from the search.

A quantitative survey has been designed in order to collect the data (see appendix 2). In order to answer the research questions and to test the hypotheses in this study, the method of a quantitative survey is seen as more fitting then a case study-since the chosen method allows it to generalize and gives an overview on how methods are used in a variety of physical retail stores [37]. The possibility of getting deep into a specific case contributed not to the answering of the research questions.

For a more efficient collection, the survey had two versions: an online version sent via email or through LinkedIn and an offline variant, where the authors met the managers inside the stores which contributes to 83% of the collected data and was the main channel of data collection [37]. All in all 61 data sets were collected from the on- and offline survey. Thus, through these tools the data was collected within two weeks. This section presents the method behind this process of data collection.

The survey had the goal to collect primary data about how companies deal with metrics for measuring customer performance in physical retail stores. Primary data is according to Stevens et al. data that was collected for the first time [38]. This data was collected specifically for the project in order to have a strong matching between the purpose, theoretical framework and the data of the study.

**Choice of scale:** When deciding which scale suited the research best, it was important to make sure that both the respondents and the researchers have no difficulties in interpreting the scale and the result of the survey. Zander stated further that a scale with the options "sometimes", "often", etc. can lead to misunderstandings when asking people with different cultural- or knowledge backgrounds [39]. Therefore this study used the Likert Scale, when it comes to a differentiated statement for example "How useful are transaction metrics?" [37]. For other questions like "Do you know transaction metrics?" the options for the answers were given in a binary form with "Yes" and "No" (Ibid). This made the statements clear and with only these two options the respondent needed to decide (Ibid). In order to keep the answers of the study more structured the answers for the fix measurements were also given (e.g. industry, profit margin).

**Development of the Questionnaire:** The survey consisted of 27 questions. Seven of these were fixed questions to classify the store and the owner/manager. The survey had in total 61 respondents. Table 2 shows the questions used in the survey and the connection to the theories as well as the answer options.

<table>
<thead>
<tr>
<th>Group</th>
<th>Contained metrics</th>
<th>Core concept</th>
<th>Complexity</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer Value</td>
<td>CLV, PVC, CRV, Share of Wallet, Basket Size, Level of Price Discounts, Proportion of Returns, Purchase Frequency, Frequency of Marketing Communication, Travel Cost, Immediate Product Availability, Net Promoter Score</td>
<td>Measuring of Customer Value on individual and aggregate level. Future patterns for the return of marketing can be drawn.</td>
<td>Complex</td>
<td>(Petersen et al., [8]) Venkatesan and Kumar, [17]) Grupeta et al., 2004)</td>
</tr>
<tr>
<td>Cross Buying</td>
<td>NPTB model, Likelihood of Purchase, Past Purchase Frequency, Inter Purchase Time, LSA</td>
<td>Process of purchasing of products from different categories in one shopping session. Cross buying customers are more profitable, however it requires a complex marketing strategy.</td>
<td>Complex</td>
<td>(Venkatesan et al., [19]) Petersen et al., [8]) Knott et al., [12]) Kumar et al., [14])</td>
</tr>
<tr>
<td>Acquisition</td>
<td>Cost of Acquisition, Cost of Retention, Retention Profit, Acquisition Profit, Product Return Rate, Customer Retention and Acquisition is important to have stable profits. Trade-offs between acquisition and retention are needed in order to increase profit of the firm.</td>
<td>Not complex</td>
<td></td>
<td>(Petersen et al., [8]) Tomas, 2001) Reinharz et al., 2005) Verhoef, [33])</td>
</tr>
<tr>
<td>Transaction</td>
<td>Conversion Rate, Average Purchase per Transaction, Traffic, Average Unit Price, Number of Transactions, Average Order Size, Sales, Revenue per Visit</td>
<td>These metrics are mostly at store level applied. Schlesinger's research (2008) shows that these metrics have a direct connection to future sales.</td>
<td>Not complex</td>
<td>(Schlesinger, [35]) Petersen et al., [8]) Atkinson, [36])</td>
</tr>
</tbody>
</table>

Table 1: Framework based on the literature review (own illustration).

Simplicity and clearness are the factors that were chosen for the development of the questionnaire in order to make it easy for the respondents to understand the questions. For this reason, the survey was divided in sections for each group of metrics, providing a short description on what kind of metrics are included. This partition in groups is important in order to limit confusion or possible misunderstandings. This led to multiple "yes" and "no" questions instead of one question that asks for the used metric group. Based on this structure, it was easy for the respondent to recognize and focus on different groups for which they might have knowledge under different terminology.

For a more efficient collection, the survey had two versions: an online version sent via email or through LinkedIn and an offline variant, where the authors met the managers inside the stores which contributes to 83% of the collected data and was the main channel of data collection [37]. All in all 61 data sets were collected from the on- and offline survey. Thus, through these tools the data was collected within two weeks. This section presents the method behind this process of data collection.

The survey had the goal to collect primary data about how companies deal with metrics for measuring customer performance in physical retail stores. Primary data is according to Stevens et al. data that was collected for the first time [38]. This data was collected specifically for the project in order to have a strong matching between the purpose, theoretical framework and the data of the study.

**Choice of scale:** When deciding which scale suited the research best, it was important to make sure that both the respondents and the researchers have no difficulties in interpreting the scale and the result of the survey. Zander stated further that a scale with the options "sometimes", "often", etc. can lead to misunderstandings when asking people with different cultural- or knowledge backgrounds [39]. Therefore this study used the Likert Scale, when it comes to a differentiated statement for example "How useful are transaction metrics?" [37]. For other questions like "Do you know transaction metrics?" the options for the answers were given in a binary form with "Yes" and "No" (Ibid). This made the statements clear and with only these two options the respondent needed to decide (Ibid). In order to keep the answers of the study more structured the answers for the fix measurements were also given (e.g. industry, profit margin).

**Development of the Questionnaire:** The survey consisted of 27 questions. Seven of these were fixed questions to classify the store and the owner/manager. The survey had in total 61 respondents. Table 2 shows the questions used in the survey and the connection to the theories as well as the answer options.

Simplicity and clearness are the factors that were chosen for the development of the questionnaire in order to make it easy for the respondents to understand the questions. For this reason, the survey was divided in sections for each group of metrics, providing a short description on what kind of metrics are included. This partition in groups is important in order to limit confusion or possible misunderstandings. This led to multiple "yes" and "no" questions instead of one question that asks for the used metric group. Based on this structure, it was easy for the respondent to recognize and focus on different groups for which they might have knowledge under different terminology.

The questions were developed after constructing the theoretical framework for this thesis. Hereby it was important to collect information that helped to answer the research questions. The questions asked for information regarding the classification of the company (fix measurement), how store managers use metrics and how they see it in the future, as well as to the profitability of the store, this is illustrated in Figure 1.

At the end of the questionnaire, seven questions for fix measurement about the store and the manager were asked. Asking personal questions at the end is according to Fisher the best way to get the respondent answering all asked questions [40]. This is especially crucial for getting as much complete responses as possible when conduct one questionnaire survey. The questions can with that be divided into three groups that are presented in the following figure and that structure the analysis. The data for the group "Using metrics" was collected for each of the metric groups (e.g. Customer Value, Cross Buying, etc.)

**Operationalization:** The operationalization shows the asked questions and the connection of these to the theory to get an understanding why this questions were asked and how the answers will be connected to the theory when it comes to the analysis. Furthermore the options for answers are shown in the table below (Table 2).

**Distribution of the Questionnaire:** In order to embrace a wide share of brick and mortar stores, the questionnaires were delivered in two ways. The most efficient way, in terms of number of answers, was to visit the stores and ask for the store manager to answer the
questionnaire. In some stores, where the manager was not present, a letter was sent to the manager in order to inform him about the research and to ask if there is the possibility to get the questions answered. (See the information letter and flyer in appendix 1.) When asking the questions, the groups of metrics were presented by using examples, this was also explained in the online version. If questions regarding the understandability arose during the questioning process, these questions were clarified by explaining which metrics are in the groups or not.
Aside from personal handling, the collection of data was made through an online survey, designed with the same questions [37]. This design is shown in appendix 2. The link to the online form was provided to the managers by email or within a letter using a QR code. Furthermore, the link was shared through private store managers groups at the platforms LinkedIn and Google+, with permission of the group managers.

The main part of the responses (83%) came from the personal visiting of the stores. Only 17% of the responses came from the online survey tool. The stores in this survey were chosen randomly due to the time limit. However, all stores that participated in this survey fulfilled the criteria that they were physical and retail stores. The criteria to choose the respondents were limited in order to get a variety of different retail stores and with that be able to see patterns in the data that could state something about what influences the way metrics are used. This variety of physical retail stores helped to explore the reasons for using metrics and hence the development of concepts and constructs in this study.

Analysis method

Statistical techniques: The correlation analysis aims to measure the correlation between variables by using the Spearman’s rank correlation, for which, normality cannot be assumed [41]. Furthermore, this method focuses on the relative difference between specific metrics and business groups. In this study, the correlation analysis was used to check whether there is a correlation between the profitability and the usage of specific metric groups. In addition, the respondents with no knowledge about the metrics were not included into this part of the analysis.

In order to get more insights in how stores use metrics, the technique of regression analysis was applied and profitability became the dependent variable [41]. This method was chosen to gain a deeper knowledge for whether there is significance between an independent variable and the profitability. Thus, the beta value out of this analysis can give then additional information on how high is the impact: the higher the number—the higher the impact. Furthermore, the regression analysis in SPSS created an analysis of variance (ANOVA) table that gave information regarding the significance by presenting the p-value for each regression test. With this information it was possible to evaluate whether the regression model is useful in order to predict the relationship between the independent variables and the dependent variable.

The third major analysis in this paper was done by using the Wilcoxon-Signed-Rank-Test [41]. It was chosen in order to analyze the answers about the usefulness and usability. The data in this case is ordinal, which is not possible to analyze with higher statistical methods like regression, thus, this non-parametric test was chosen. The test itself compares the respondent’s answers, contrasting each time two variables and with that analyzing two metric groups at the same time. This resulted in a total of six comparisons for the four metric groups, comparing each variable. Once the analysis was done it was possible to rank the metric groups based on the individual usefulness and values of usability of the respondents.

Furthermore, the results of the survey were also graphically presented, using descriptive statistical analysis (see for a list of descriptive statistics Argyrous) [41]. These methods were used mainly to show relationships between the fix measures and the way stores evaluate the different metric groups. An overview of the factors that influence the implementation of the metrics can be seen in the analysis.

Reliability, Validity and Generalization: Reliability of the questionnaire depends on whether the study can be reproduced and repeated several times [37]. The method used for this study is quantitative; therefore it is crucial to think about reliability. Thus, in order to get high reliability it is necessary that tools and techniques used to gather the data ensure the possibility of repeating the research. According to Bryman and Bell, a study can be considered reliable when future researches result into similar outcome using a comparable process [37].

To get accurate information, only specific stores were targeted for this survey. The stores needed to be physical and needed to have retail operations. Furthermore, the survey was addressed only to owners and/or managers of the store in order to get reliable information about the metrics used by the store. Alongside, an online version of the survey (Google forms) was delivered, allowing to answer the questions freely and accurately. Through this process it was possible to ensure that the gathered data was reliable. Afterwards, the collected data was elaborated in excel files and further with SPSS system in order to apply statistical analysis. The statistical tools used to analyze this study are presented in the heading before.

Validity is also important for the sake of the study and it can be achieved by using theories and concepts in the questionnaire [37]. This was achieved by designing the survey along the literature review. Terms like customer value, cross-buying or transaction metrics were mentioned in the questionnaire and the theories were incorporated to answer the research question. The cause and effect relation, on which this study is based, was tested by using the statistical software SPSS from IBM. This cause and effect relationships are according to Fisher a way to reflect internal validity [40].

The range for which the result of this study is included can be generalized because such range depends on the sample size. In the correlation analysis the result had a meaningful difference to support the hypothesis. Also, the significance in the regression analysis was high enough with the given sample size to draw a meaningful conclusions out of the collected data [42-45].

Limitations

The first point that limits the following analysis is that the data was collected by a survey. The researchers were willing to keep the influence as low as possible to get answers that are representing the truth and are not influenced by third parties. However, it might be the case that for example the answers regarding the profitability might vary a bit from the true numbers because the managers wanted to make their store look more profitable.

Furthermore, due to a limited time frame that was given to conducting this research, the number of surveyed store is 61 respondents is limited which might result in slightly different results in future research. Also the type of industry that was surveyed might have an impact on the outcome of the study. In this case the fashion industry makes 41% of the responses followed by other with 26% and on third place electronic stores with ca. 8% of responses.

Results and Analysis

The analysis shows the empirical data on customer performance metrics based on the 61 collected data sets. The analysis gives a descriptive overview of the results and factors that influence the usage of metrics. After that it presents the connection between the usefulness and usability of metrics and the profitability. The second in depth...
look will be on the fact whether there is an individual preference of metrics. The last in-depth analysis will focus on the question whether there is a correlation between the profitability of the store and the usage of metrics. The analysis will focus at relations between the data of the three groups presented in Figure 1 above. First it will look at the relations between the fix measures and whether they have an effect on how metrics are used and implemented (the second group). After that the analysis looks on whether there is a relationship between the usage of metrics in stores and the performance of the store which was measured by using the profit margin.

**Circumstances that effect the implementation of metrics**

The first factor that influences whether a store uses a metric group or not is the type of business which can be an independent store or a store that belongs to a retail chain (Figure 2).

In combination with the following table it can be seen that stores that belong to a retail chain use much more often metrics than independent stores. Circa 75% of the stores that belong to retail chains use metrics but only ca. 50% of the independent stores use metrics. This fact verifies the hypothesis H2 that stated that stores belonging to a retail chain use much more often metrics than store that belongs to a retail chain (Figure 2).

This fact can be due to the complexity of metrics and with-the requirements in special educated personal and IT systems. This can be seen in the next figure where the usage of metrics divided by the metric group and the implementation of an IT system, for example an Enterprise Resource Planning (ERP) system or a Customer Relationship Management (CRM) system, is shown. The y-axis shows the number of stores that use the metric group. The figure presents only the respondents that use the specific metric and splits the answers up in using an IT system and not using an IT system (Figure 1).

The stores with an IT system are over all metrics groups more likely to use metrics. Often the data collection of metrics and the later analysis requires information technology to handle the big amount of data as it is especially for the cross buying metrics the case (e.g. Likelihood of Purchase) were only stores with a complex IT system use the metrics. Stores that are member of a retail chain also rank higher in the usefulness of metrics.

Furthermore, analysis shows that an education with a more theoretical base (High School and University) has an influence in the way how useful store managers see metrics. The results are in average 34% higher as with a non-academic (e.g. training on the job) background. This is backed up by the fact that ca. 80% of the managers form stores that belong to a retail chain have a university degree or at least a high school diploma. At the independent stores only slightly above 60% of the asked managers have a high school diploma or a university education. Connected to this analysis part is the first hypothesis that stated:

H1: There might be a difference in the usage of metrics depending on the education

Analysis has shown that there is a difference in the usage of metrics depending (on) the variables education and the usage or IT systems like ERP or CRM systems. It turned out that stores with more educated personnel (high school or university degree) used more metric groups then stores where the personal only had a non-academic education like training on the job. The following figure shows the average scores for the usefulness by independent stores and retail chains (Figure 3).

They gave on average a score of 4.35 and this is ca. 17% higher as the average value of 3.73 that were given by the independent stores over all metric groups. The reason for this is that stores that belong to a retail chain have better IT systems which can be seen in the following figure about the implementation of ERP and similar systems. This allows the stores to calculate complex metrics which increases the view on the usefulness of the metrics by the store manager. Furthermore, this requires more educated personal to run the system as well as to interpret the results (Figure 4).

The second factor that influences this result is the way that education of the personal is required in order to gain knowledge out of more complex metrics. The usage of metrics, moreover, is affected by the degree of implementation of the IT system. As a result, stores with such a system tend to use significantly more metric groups then those who did not. This influence between the usage of metric groups and the implementation of IT systems leads to the next hypothesis:

H2: Stores that belong to a retail chain use more metrics.

This hypothesis was confirmed by the analysis as shown above in Figure 3. The data shows that stores belonging to a retail chain tend to
use more metrics rather than independent stores. A possible reason to explain such trend can be connected to the previous hypothesis (H1), where education and data complexity are possible discriminant factors that limit the usage of metrics. Observing the degree of usability and usefulness reported within the survey, it is possible to draw a regular pattern where store managers tend to understand and use more simple metrics above those which are considered to be more complex and require a deeper knowledge.

Before analyzing the relationship between the usage of metrics and the profitability of the store, it can be useful to see whether there is an individual preference of metrics so that a ranking of metrics can be created. When looking at the ranking of the metrics depending on the evaluation of the usefulness of the metrics, the analysis shows these results (Table 4).

The result of this analysis shows that Acquisition metrics and Transaction metrics score always above the Cross Buying and Customer Value metrics. A ranking of the metrics can be created based on whether the group scores above all other with the positive ranks or not. For the usefulness this results in the following ranking of the metrics were Cross Buying and Customer Value are at the same rank (Table 5).

This shows that Transaction and Acquisition metrics are the once who get according to the data on individual basis the highest scores in usefulness. One reason for this can be that, as current research showed, these are the metrics that are easy to understand and also most widely used in practice. Thus, the results support prior research.

When looking at the usability, the results of the Wilcoxon test looked like follows (Table 6).

When comparing the pairs of metric groups presented in Table 6 and how they rank against each other, this results in the following ranking of the metric groups. The metric group that scores above all other in the pair comparisons is on rank one (in this case the Transaction metrics) (Table 7).

Also, in this ranking the metric groups that are based on current research are easy to understand rank higher in usefulness then the more difficult groups [8]. This can be explained by the fact that the metrics from these groups are easier to measure. Transaction metrics are for example to a high grade measurable based on the information collected by the cashier register. Interesting is that cross buying scores are at the last position which can be due to the reason that the metrics out of this group are the most complex once. The view on this individual preference shows that the hypothesis H3a can be verified:

H3a: The individual preferences of the usefulness of metrics go in hand with the usability.

This hypothesis H3a can be verified since data showed that for both the usefulness and the usability, Transaction and Acquisition metrics ranked first and second based on an individual preference.

**Contribution of metric groups to the profitability of the store**

The graphical analysis has shown that the profitability is in average higher for those stores who valued the usefulness and the usability higher compared to those stores who saw not so much usefulness. In the two following figures the colors indicate the usability and usefulness score given by the managers. A darker color was given for a low score (1) of usability and usefulness, while a lighter colors expresses a high score (5). The x-axis describes the average profitability of the stores that gave the scores (Figures 5 and 6).

This correlation shown in the last two figures, which describes the average profitability for each metric group divided by the score given for usefulness and usability, lead to the estimation that the profitability of the store can be explained by the view of usefulness of the metrics. Based on that the data was used to prove how the observed data looks in relation to the linear model of regression between the usefulness as independent variable and the profitability as dependent variable. This was done for all metrics and is presented in the following Table 8.

The plots presents that the average of the observed value lies close to the linear function. However the scattering of the data is huge. The correlation analysis shows that there is a non-random association between the variables profitability and the value of usefulness of the different metric groups. In order to quantify how strong the relationship between the variables is, the R Square needs to be taken into consideration. In this case the R Square indicates that 17% (Transaction) up to 42% (Cross Buying) of its variation is accountable for the relationship with one another. With looking on the ANOVA (F, df1, df2, and Significance) it has to be evaluated whether the regression model is better in predicting value then using the mean of the values.

**Table 4: Result Wilcoxon signed rank test usefulness.**
getting the F value. In this case it is lower than 0.05 so it indicates that the regression line is significant better than the prediction of values by using the mean. The value of the beta ($b_1$) describes the slope of the regression line. It turns out that an increase of 1 value in how the usefulness is seen, relates to an increase of 1.8 (Customer Value) to 3.1 (Cross Buying) percent points in profitability (Figure 7).

The reason for this might be that also stores, who do not use the metrics, still see the usefulness as high. The results of this analysis come close to the answer of the first research question. Beside the usefulness, the usability is a critical factor. This determines how stores evaluate the possibility of using metrics to track consumers. A lack in knowledge and or in technology might be factors that lead to low ranks in the usability of metric groups. In the usability it can be seen that there are the same patterns like with the answers for the usefulness (Table 9).

The profitability is in average higher at those stores who saw the usability as more easy (higher rank). By looking at the R square value for the usability of the metric groups as independent variable, it can be seen that 17% (Acquisition) to 23% (Customer Value) are accountable for the shown relationship between the usability of specific metric groups and the profitability of the store. Looking at the ANOVA, the

<table>
<thead>
<tr>
<th>Pairs</th>
<th>Negative Ranks</th>
<th>Positive Ranks</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usability_Acquisition-Usability_Cross_Buying</td>
<td>13</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>Usability_Acquisition-Usability_Customer_Value</td>
<td>13</td>
<td>21</td>
<td>8</td>
</tr>
<tr>
<td>Usability_Cross_Buying-Usability_Customer_Value</td>
<td>16</td>
<td>14</td>
<td>-2</td>
</tr>
<tr>
<td>Usability_Transaction-Usability_Acquisition</td>
<td>4</td>
<td>21</td>
<td>17</td>
</tr>
<tr>
<td>Usability_Transaction-Usability_Cross_Buying</td>
<td>4</td>
<td>27</td>
<td>23</td>
</tr>
<tr>
<td>Usability_Transaction-Usability_Customer_Value</td>
<td>3</td>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>Grand Total</td>
<td>53</td>
<td>125</td>
<td>72</td>
</tr>
</tbody>
</table>

Table 6: Result Wilcoxon signed rank test usability.

<table>
<thead>
<tr>
<th>Usability ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transaction</td>
</tr>
<tr>
<td>Acquisition</td>
</tr>
<tr>
<td>Customer Value</td>
</tr>
<tr>
<td>Cross Buying</td>
</tr>
</tbody>
</table>

Table 7: Ranking of metric groups based on usability.

Figure 5: Average profit by usability of metric groups.
significance is between <0.001 (Customer Value and Cross Buying) up to 0.001 (Acquisition and Transaction) with that it can be said that also in this case, the regression model is significant better in predicting values than the mean. The amount of increase in percentage points of the profitability that is explained by the increase of valuing the usability higher or lower ranges from 1.6 (Acquisition) to 2.8 (Transaction) (Figure 8).

To conclude, this regression analysis part it is important to clarify that these kind of relations represent a piece of the complete picture and that there is a difference in correlation and causation also if the significance is high (p<0.05). This leads to the verification of the hypothesis H3b:

H3b: Profitable stores rank the usability and the usefulness higher. For all metric groups there was a correlation between the usability/usefulness and the profitability. In all cases the linear regression model was a better fit to explain the relation then the calculation of the mean that already indicated a correlation. A critical point could be that the fact that more profitable stores have a more complex structure and defined reporting if they are part of a retail chain so that the managers have to collect more data so that the reason for the increase in profitability is not only the different view on the usefulness and usability. This could lead to future research.

With the view on the usage of metrics based on profit classes, the Hypothesis H4 can be verified. Stores who are more profitable use in average more metrics then those who are not so profitable (Table 10).

This leads already to the answer of the research question RQ2, where it was questioned if stores which use metrics are more profitable.
And hypothesis H4 can be verified:

H4: Stores that use more metrics are more profitable.

The analysis has shown that stores in the lowest profit class (0-5% profit margin) tend to use 1.75 metric groups in average compared to stores from the highest group (16+% profit margin) who used 3.21 metrics in average.

In order to answer the research question RQ2 that asked whether stores that use metrics to track customers are more profitable, a correlation analysis was performed. The result is shown in the next figure (Figure 8).

The result shows clearly that there is a correlation between the usage of metrics and the profitability of the store the analysis showed an increase in the profitability of stores who use metrics compared to the once who did not used the metrics of specific groups. This holds for all of the four metric groups. Furthermore, it can be seen that there is the biggest difference between using Cross Buying metrics or not. The increase in profitability is above 5%, whereby the usage of transaction metrics does not affect the profitability of the store so much. This indicated that Cross Buying metrics have a bigger impact on the profitability and with that gives that store managers more insights in order to run the store more profitable.

Conclusions

This study aimed to investigate the usage of metrics to track customer performance in physical retail stores. The initial literature review provides an overview of metrics that are, based on current research, the main metrics that can be used to track customers in physical retail stores. The metrics were divided into four groups: Customer Value,
Cross Buying, Acquisition and Transaction. Extant literature showed that the metrics from the Customer Value and Cross Buying groups were more complex when it comes to collecting, storing and analyzing the data. Therefore they require, more IT resources, while the other two groups were less complex and thus easier to implement and use [11].

The fact that the metric groups differed in the outcome as well as in the complexity, led to the research question number two which asked whether there is a correlation between the profitability of the store and the usage of specific metrics. The study’s result and analysis shows that there is a correlation between the usage of specific metrics and the store profitability. Basically stores, who use more metrics, are more profitable. This correlation is the highest in the group of Cross Buying metrics which are, according to current research as well as the findings, considered to be more complex to understand, use and to implement. However, it is important to state that correlation is not equal to causation. A resulting correlation does not necessarily means that this factor is the only reason for higher profitability. There might be more factors that play a role and might be linked to the usage of metrics e.g. how much control a store manager wants to have or whether the store has also an online store.

The fact that metric groups differ in complexity and that elaborated metric groups have a higher positive correlation with the profitability of the store is followed by the question, which role the complexity plays when it comes to how a store manager sees the usability and usefulness of specific metric groups. Research has shown that metrics that are easier to use, rank in the individual preferences for usability and usefulness higher than those groups with more complex metrics (e.g. Cross Buying). The requirement of IT systems as well high educated personnel makes, according to the analysis, the use of this metrics more difficult. However, companies who used this kind of metrics showed a significant increase in profitability which indicated that the investment in educated personnel as well as IT infrastructure is a success factor.

Managerial implications

The research in this study showed that manager’s get through the usage of metrics from specific metric groups more insights into their business. Groups of metrics that are more complex, according to the analysis, showed a higher correlation with the store profitability. This leads to the advice that managers should implement more elaborated metrics in order to get useful insights into their business. Thus, that holds also true for the store level in retail business. This helps the management of the store to get more detailed information and supports decision making right on time.

The research showed that the number of used metrics is crucial: more profitable stores use more metric groups. The increased number of metrics give store managers a more complete overview of the business and reduce the risk of focusing only on the increase of single measurements. However, it is necessary to mention that, according to current research, too many metrics would make decision making more difficult. Therefore, managers of physical retail stores should select metrics in a way that they add value to each other.

As stated before, the complexity of metrics leads to more valuable insights that influence the profitability of the stores. Managers are advised to also use the more complex metric groups and not just rely on the metrics that are easier to implement and to analyze. This fact leads to the implication that the IT system in the store should be enough developed to handle these metrics and deals with the amount of data used by calculating such metrics. In particular, this is the case when it
comes to Cross Buying metrics, where investments in a modern scalable IT system improve the way metrics can be calculated and analyzed.

**Future Research**

The study of customer performance becomes more and more elaborate. This complexity derives from the necessity of collecting new kind of data in order to improve the knowledge about the customer behavior.

The term "brick and mortar stores" seems to be not contemporary. Therefore, there could be a spill-over of knowledge from online operations to offline operations this leads to the development of the term "click and mortar stores" which are a combination of retail and online stores. In connection to this, future research should observe how online operations have an influence on how stores use offline metrics. It can be assumed that the online operations have a much higher impact on the knowledge about metrics then what the level of education has.

Nowadays, the market offers innovative technologies and software (e.g. 3D camera scanning) so that companies are able to track the customer’s pattern and analyze it, combining a multitude of factors and variables. The goal for these technologies is to reduce the gap between offline and online store in terms of customer tracking systems. For this reason such innovations are leading to an implementation of ways to track the customer inside the brick and mortar store, where more and more the customer is observed and analyzed in real time in order to allow the managers to have a quick response. This new technology and the effect on the tracking of customers can be a topic for future research.

Major technological companies like Google and IBM are leaders in this field, offering business solutions and improved software that allow other companies and managers to collect and analyze such data. However, the demand for such technology is growing fast which will lead soon to a possible overall cost reduction, allowing more businesses to use and implement their data collection process. A possible trend for this fast growing market is the opportunity to simplify the user experience through, for example, an improved usage of graphic tools, in order to make the data interpretation easier and accessible also for those managers who have lower education. These developments and which factor determine a better usability of such software for less educated managers could be answered by future research.

Furthermore, in the future it is plausible to believe that more companies will develop a deeper integration between online and offline store, leading to combination and mutation of already existing metric groups and the development of new hybrid metrics, able to track customer performance for both channels at the same time. This could lead, according to the authors of the thesis, to an advancement in Cross customer performance for both channels at the same time. This could lead to medium-sized enterprises. International Journal of Business Information Systems 3: 374-390.


