

Factors influencing the customer satisfaction rate in Dutch SME fitness centers

Master Thesis Business Administration

Abstract:

Currently there is a clear problem with obesity and with the business continuity of Dutch SME fitness centers. Literature indicates that increasing the customer satisfaction rate will improve the business continuity of fitness centers and will consequently help to decrease or stabilize the obesity problem. This research focusses on SME fitness centers since the increasing presence of rivaling low-budget fitness centers has resulted in the business continuity problems, while these rivaling low-budget fitness centers do not offer much in terms of service and guidance in comparison to many SME's. Therefore, SME's are contributing more towards solving or stabilizing the obesity problem in comparison to the larger low-budget fitness chains. Consequently, the aim of this study is to determine which factors influence customer satisfaction in Dutch SME fitness centers. In order to find which factors influence customer satisfaction at Dutch SME fitness centers, this study has combined results and opinions from previous relevant studies in order to create a new model that indicates which factors influence customer satisfaction. In order to test the hypothesized model, data has been collected from three Dutch SME fitness centers using a questionnaire. A principle component analysis was conducted in order to test the validity and to uncover an underlying structure based on the relatively large set of used variables. The resulting factors have been used in a multiple regression analysis in order to test the impact that they have on the dependent factor, customer satisfaction. The result of this study indicates that: Personnel, Physical environment, Program, Social environment, Supporting services and Outcome have a significant and positive effect on customer satisfaction in Dutch SME fitness centers.

Student: Mario Klasens
Student number: S1662686
Date: 21-04-2020
Study: Business Administration
Faculty: Behavioural, Management and Social sciences
Supervisor: Dr. A. Leszkiewicz

Acknowledgements

I would like to thank my thesis advisor Dr. A. Leszkiewicz for her guidance and advice, Dr. Leszkiewicz always helped me to the best of her abilities whenever I had questions regarding this research. I would also like to thank the cooperating fitness centers for giving me the ability to collect data, this was essential for the completion of this research. While many people experience studying as the best time of their life, I certainly did not have this feeling when I was completing my Master's degree. Due to certain developments I was not sure if I would be able to finish my study. However, due to the support of my parents, friends and partner I continued. Therefore, I am very grateful for their support throughout my study. Furthermore, I would like to thank Rich Piana for his inspiring words and motivating me to give it my all.

Enschede, April 2020

Mario Klasens

Table of contents

1. Introduction.....	4
2. Literature review	5
2.1. Customer satisfaction.....	5
2.2. Factors influencing customer satisfaction.....	6
2.2.1. Service quality and satisfaction models	7
2.2.2. Service quality and satisfaction models in the fitness industry	7
3. Model Development and Hypotheses.....	11
3.1. Hypothesizes	11
4. Research design & Methodology	12
4.1. Sample & data collection.....	12
4.2. Instrument development	12
4.3. Measurements	13
5. Results	13
5.1. Descriptive statistics.....	13
5.2. Missing data	14
5.3. Validity.....	15
5.4. Factor Analysis independent variables.....	16
5.4.1. Factor Analysis dependent variables.....	20
5.5. Reliability	21
5.6. Pearson Correlation Coefficient Analysis	22
5.7. Multiple Regression Analysis.....	22
6. Discussion & Conclusion.....	28
6.1. Theoretical contribution.....	28
6.2. Practical contribution	29
6.3. Research limitations	29
6.4. Suggestions for future research	29
7. References.....	30
8. Appendices	36
8.1 Questionnaire in English.....	36
8.2 Questionnaire in Dutch	39
8.3 Questionnaire measurement / question sources	42
8.4 Pearson Correlation Coefficient Analysis	44
8.5 Factor analysis with 7 factors	45

1. Introduction

Both worldwide and in the Netherlands the population is increasing significantly. Besides the increasing population, it has also been found that the percentage of people with obesity has been increasing significantly. In 1990, circa one in three adult Dutch citizens (35.1%) were moderately or severely overweight. Since then, the number of overweight Dutch people has risen to nearly half of all adults (48.8%) in 2017. The percentage of people which are extremely overweight (obese) has doubled from 6.2% to 13.9% over the same period of time (Volksgezondheidenzorg, 2017). It can be concluded that the total amount of people with obesity is not only increasing due to the increasing population size, but also due to the increasing percentage of people which are overweight. The purpose of this study is to measure what factors influence the customer satisfaction rate in Dutch SME fitness centers in order to improve business continuity and therefore help to decrease or stabilize the obesity problem.

Even though the obesity rate in the Netherland is increasing, it was found that people are becoming more conscious and proactive regarding their health (Hanspal & Devasagayam, 2017). This is reflected by the increasing amount of people in the Netherlands who participate weekly in a sports activity (NOC-NSF, 2017). Fitness is currently the most practiced sports activity in the Netherlands and is practiced by more than 3.076.000 people, which comes down to ~19% of the total population in the Netherlands (NOC-NSF, 2017). Estimates suggest that about two-thirds of the fitness practitioners are members of a fitness center (Rabobank, 2016). It has also been found that the percentage of the population who has a gym membership in the Netherlands has been steadily increasing the past 10 years (CBS, 2017; NOC-NSF, 2017). This is most likely the result of the increasing consciousness and proactiveness regarding health and the increasing obesity rate.

Besides the increasing number of practitioners, the total number of gyms in the Netherlands has also been increasing (CBS, 2017). This is most likely the result of the increasing population and increasing percentage of gym members, which increases market potential for fitness centers. The number of fitness centers in the Netherlands has grown by 78% between 2007 and 2018, which has resulted in a total of 1.790 fitness centers in 2018 (KSB, 2018). This growth is characterized by larger fitness chains which have increased their number of fitness locations significantly. The significant increase of fitness centers has also led to challenges in the fitness industry since the number of fitness centers is increasing at a faster pace than the people who have a gym membership (Rabobank, 2019). Due to this trend, competition between fitness centers is high and profit margins in the fitness industry are under pressure. The recent growth of large budget chains in the Dutch fitness industry has further contributed to the pressure on profit margins and an increased competition between fitness centers based on price. It is expected that these conditions will remain or increase due to the arrival of foreign fitness chains (Rabobank, 2019).

The decreasing market share of SME's and the increasing competition in the Dutch fitness industry may not only affect the future business continuity of the SME's, but may also have a negative impact on the overweight problem that has been described before. The large fitness chains which are often low-budget, have acquired a large market share in the Dutch fitness industry in recent years, partly due to the increasingly price-conscious consumer (Rabobank, 2019). However, the low-budget fitness centers focus on price and do not offer much in terms of service and guidance in comparison to many SME's (Rabobank, 2019). Therefore, it is likely that SME's are contributing more towards solving or stabilizing the obesity problem in comparison to the larger low-budget fitness chains. Difference in the effect of small and large fitness centers on the lowering of obesity has not been scientifically studied. However, there are studies which indicate that good guidance, which is more

present in smaller fitness centers (Rabobank, 2019), helps with the weight-loss process and the maintenance of lost weight (Painter et al., 2018; Hall & Kahan, 2018).

The problem regarding business continuity in SME Fitness centers is further increased due to the trend that consumers are becoming less loyal, and due to the relative low marketing budget of SME's in comparison to large fitness chains (Rabobank, 2019). Therefore, smaller fitness centers cannot fully rely on expensive means to acquire new members, while bigger fitness chains may. This makes it important for SME's in the fitness industry to focus on retaining customers. Furthermore, research indicated that attracting a new customer requires five times more costs than retaining a current one, and that loyal clients will bring in more new clients due to word-of-mouth (Reichheld and Sasser, 1990). Moreover, McCarthy (2007) describes that deregistering members is a major issue in the fitness industry. McCarthy (2007) describes that the quitting of members can be seen as triple loss due to the financial loss, the opportunity loss and the experiential failure for club members. Therefore, the creation of loyalty and the binding of customers should have a priority for fitness centers.

Generally, customer loyalty is often being conceptualized as the intention to purchase a product or service again from the same organization. Oliver (1997) has created a more comprehensive definition of loyalty. Oliver (1997) defines loyalty as a commitment to re-purchase a preferred service or product in the future, resulting in the purchasing of the same brand despite possible marketing efforts or situational influences that may result in switching behavior. Many studies have focused on customer loyalty and what variables positively or negatively influence customer loyalty. Gupta and Basumatary (2017) found that customer satisfaction is an important indicator of the likeliness that a client will make another purchase in the future. These findings have also been confirmed by other researchers (Helgesen, 2006; Hallowell, 1996). Suwono and Oktaria (2016) have also studied the effect of customer satisfaction on loyalty, but unlike the previously mentioned authors they have focused on fitness centers. It was found that there is a significant relationship between customer satisfaction and customer loyalty in fitness centers (Suwono and Oktaria, 2016). Therefore, it seems interesting to investigate what factors influence satisfaction rates in Dutch SME's fitness centers in order to improve customer loyalty. The subsequent aim of this study is to determine what factors influence customer satisfaction in Dutch SME fitness centers, in order to improve business continuity and therefore help to decrease or stabilize the obesity problem. This has resulted in the following research question:

"What factors influence the customer satisfaction rate in Dutch SME fitness centers?"

2. Literature review

In order to investigate the relationship between customer satisfaction as a dependent variable and the independent predictor variables, this study reviews literature regarding satisfaction and service quality in order to create a model. Knowledge regarding these topics was gathered from general studies on these items and from studies which have focused on the service industry.

2.1. Customer satisfaction

Customer satisfaction is a construct that has been extensively studied and discussed in literature. Customer satisfaction has been the topic of many studies due to its important role in the competitive environment, since it allows organizations to retain current customers and to increase the amount of new customers (Jiradilok, Malisuwam, Madan, & Sivaraks, 2014). Customer satisfaction can be defined as the situation where the provided product or service meets the expectations of the customer (Kotler, Cunningham, & Turner 2001). In a more recent study, Kotler and Keller (2012, p.128) identified satisfaction as: "a person's feeling of pleasure or disappointment which occurs as a result of the

comparison between product or service performance and expectation". Moreover, Kotler indicates that the degree of customer satisfaction is the key to the success of marketing strategy (Kotler & Keller, 2006). Oliver (2015) has provided another point of view regarding customer satisfaction, according to Oliver (2015) customer satisfaction is a judgement regarding the product or service, where the product or service should have provided the customer with a pleasurable level of consumption-related fulfillment. Due to the large amount of literature regarding satisfaction, there are many different conceptualization of the customer satisfaction. However, most conceptualizations are in line with the expectancy disconfirmation theory (McQuitty, Finn and Wiley, 2000). The expectancy disconfirmation theory proposes that the satisfaction level of a customer is the result of the difference between expected and perceived performance. Therefore, satisfaction (positive disconfirmation) occurs when the delivered product or service is better than anticipated by the customer. Dissatisfaction (negative disconfirmation) may also appear when the performance was worse than expected by the customer. If dissatisfaction occurs due to members not getting the services they expect, or even not exceeding their expectation it probable that the fitness center will be unable to retain this member (Talley, 2008). According to Robinson (2006), who studied customer expectations of sport organizations, expectations of a service are based on word of mouth, desires, needs, previous experiences, price, advertising and image. In this study, customer satisfaction is considered to be based on the expectancy disconfirmation theory, since most conceptualizations are in line with the expectancy disconfirmation theory (McQuitty, Finn and Wiley, 2000).

The large amount of conceptualizations indicate that there is a lot of literature regarding customer satisfaction. It has been a topic of interest for organizations and researchers since the research by Cardozo (1965) regarding customer effort, expectations and satisfaction. This is not surprising since researchers have found that an increased customer satisfaction rate can result in an increase in business performance. Reibstein (2002) confirmed this by explaining that it is important for organizations to keep consumers content with the delivered product or service, since satisfied customers are more likely to be loyal and therefore increase profitability of the organization. A research by Fornell (1992) adds to this by explaining that customer satisfaction is an indicator of organizational health and future prospects. Fornell (1992) also claims that customer satisfaction may benefit organizations due to reduced marketing costs, prevention of customer churn, an enhanced business reputation, and an increase in customer loyalty. Other authors which have studied the relationship between customer satisfaction and business performance confirmed these results and also found that customer satisfaction results in a positive word-of-mouth (Dimitriades, 2006; Olorunniwo & Hsu, 2006; Faullant & Matzler, 2008). However, none of these previous mentioned studies have focused on the effect customer satisfaction on business performance in fitness centers. A study by Howat, Murray and Crilley (1999) did investigate the relationship between customer satisfaction and business performance while focusing on fitness centers. In this study it was found that customer satisfaction in fitness centers is positively linked to increased visits, repeated use and service recommendation. A more recent study by Gonçalves, Meireles and Carvalho (2016) confirmed that also in the fitness industry customer satisfaction is a determinant for retention and profitability.

2.2. Factors influencing customer satisfaction

Due to the important role of customer satisfaction on business performance and customer loyalty, a variety of research has focused on investigating the determinants of customer satisfaction. There have also been numerous studies which have examined attributes that influence customer satisfaction in the service industry. However, there is no clear focus on determinants of customer satisfaction in the

fitness industry. While there is no focus on customer satisfaction, there are a handful of studies which have tried to identify attributes which influence service quality in the fitness industry. However, these studies regarding service quality have mainly focused on larger fitness centers instead of SME's and are often not generalizable for the entire fitness industry due to small samples. In order to develop a model which can explain satisfaction in Dutch SME's with the highest explained variance possible, models and knowledge from general studies on customer satisfaction and knowledge from studies which have focused on customer satisfaction or service quality in the service industry are used in this theoretical framework.

2.2.1. Service quality and satisfaction models

The first satisfaction and service quality models were developed in the 1980s. There were two main theories which were derived from a Nordic and a American ideology (Grönroos, 1984; Parasuraman, Zeitmal, & Berry, 1988). The Nordic ideology suggested functional quality and technical quality as two dimensions of service quality (Grönroos, 1984). While, the American ideology suggests that satisfaction and service quality is the difference between the expected quality and the perceived quality (Parasuraman et al., 1988). Based on their theory, Parasuraman et al. (1988) created the SERVQUAL model in order to measure the quality of services. However, some authors argued that the SERVQUAL model is not realistic since customer expectations are always high. This reasoning has resulted in development of the SERVPERF model, which only measures service quality and satisfaction and excludes customer expectations (Cronin & Taylor, 1992).

Based on the Nordic ideology regarding satisfaction and service quality developed by Grönroos (1984), Rust and Oliver (1994) developed a model to measure service quality which contained three distinct components, namely: service environment, service product and service delivery. Another study by Dabholkar, Thrope, and Rentz (1996) continued investigating service quality and consequently suggested the use of a multilevel conceptualization in order to measure service quality. In a more recent study Brady and Conin (2001) combined the multilevel conceptualization model by Dabholkar et al. (1996) with the three component model introduced by Rust and Oliver (1994). This has resulted in a model where service quality and satisfaction can be explained by three dimensions, namely physical environment quality, interaction quality and outcome quality. However, the model by Brady and Conin (2001) is more complex then it seems, since it actually contains more the three dimensions due to three corresponding sub-dimensions in each dimension. The model created by Brady and Cronin (2001) was deemed very competent according to fellow researchers, however it was found that the model contains contradictions (Martinez & Martinez, 2010). Due to the contradictions Martinez and Martinez (2010) described that the model by Brady and Cronin (2001) can be used, but dimensions of service quality and satisfaction should be identified in the context of a particular industry.

2.2.2. Service quality and satisfaction models in the fitness industry

While there are many different satisfaction models in general, there are no clear models or scales in order to measure customer satisfaction in Dutch fitness centers. In order to find independent variables which could explain customer satisfaction in Dutch fitness centers, service quality models have been reviewed due to the similarity between satisfaction and service quality.

The first models that were used to measure service quality and satisfaction in the sport and recreation industry used scales where quality was measured in a quantitative approach. The models designed for the fitness industry inherited their designs and ideas from generic models previously made to measure service quality and satisfaction. Due to their importance, models such as SERVQUAL

and SERVPERF had a significant influence on the development on the creation of quality indicators in the fitness industry.

Chelladurai, Scott, and Haywood-Farmer (1987) were one of the first to define and describe the dimensions of fitness service attributes. Chelladurai, Scott, and Haywood-Farmer (1987) developed a model and called it the scale of attributes of fitness services (SAFS). The SAFS model contained five dimensions: Primary core professional services, primary core consumer services, primary peripheral services, primary facilitating goods, secondary consumer services and facilitating goods.

Due to the increasing critique on the SERVQUAL model regarding its inadequacy for use in specific industries, other researchers also started to develop models and scales in order to comprehend quality in the fitness industry. Kim And Kim (1995) developed the scale of quality excellence of sports centers (QUESC). The developed scale was created in order to allow fitness centers to identify; the desired service by customers, the level of service required by customers and specific areas which require managerial attention. In order to create the QUESC instrument, Kim and Kim (1995) have analyzed survey data from which they found thirty three significant customer needs that can be classified into twelve distinct dimensions. The dimensions obtained by the QUESC instrument are: environment, employee attitudes, reliability, information, scheduling, personal consideration, price, exclusivity, privacy, convenience, stimulation and social opportunity (Kim & Kim, 1995). According to the authors only 7 out of the 12 dimensions found, show correspondence with SERVQUAL model. However, due to the Korean respondents sample in the study by Kim and Kim (1995), a follow-up study by Papadimitriou and Karteroliotis (2000) with Greek respondents did not support the QUESC instrument. In their study Papadimitriou and Karteroliotis (2000) instead proposed an instrument with the following four dimensions: facility attraction and operation, program availability, instructor quality and delivery and other services.

In a more recent study Chelladurai and Chang (2003) developed a new model based on the SERVQUAL model from Parasuraman et al. (1988), the RECQUAL model by Crompton, MacKay, and Fesenmaier (1991) and the SAFS Model by Chelladurai et al. (1987). The developed model, called SQFS, included dimensions and items from previous similar models and dimensions produced by the researchers themselves (Chelladurai & Chang, 2003). The created model has resulted in a total of nine proposed dimensions in order to measure quality in fitness services, namely: management commitment to service quality, service climate, design of core services, employee's task interaction, employee's interpersonal interaction, contact with physical environment, contact with other clients, service failure and perceived service quality (Chelladurai & Chang, 2003).

Another service quality assessment scale that was developed to evaluate the service quality of fitness centers was created by Lam, Zhang, and Jensen (2005). The researchers created a model called SQAS, with forty items and six dimensions: physical facility, workout facility, program, staff, locker room and child care. The authors suggested that other researchers should re-examine the SQAS model while using a different sample in order to study further factor structure (Lam et al., 2005).

Alexandris, Zahariadis, Tsorbatzoudis, and Grouios (2004) used the model by Brady and Conin (2001) as a basis for their model. The model by Alexandris et al. (2004) is based on three dimension (physical environment, interaction quality, outcome quality) and nine sub-dimensions. The model used items which were adapted from the model by Brady and Cronin (2001), but also excluded several items while introducing new items. However, due to use of a sample which contained 175 participants of a single fitness center in Greece, the study had limitations. Therefore, the authors have suggested that the model should be tested with bigger sample sizes from different fitness centers (Alexandris et al., 2004).

Similar to Alexandris et al. (2004), the researchers Ko and Pastore (2005) have adopted items from the previous models by Dabholkar et al. (1996) and Brady and Cronin (2001). Their research has resulted in a scale called SSQRS, which can be used to measure service quality in recreational sports. The reviewing of literature has resulted in a model which consisted of 4 main dimensions, namely: program quality, interaction quality, outcome quality and physical environment quality. The SSQRS scale has used a total of 77 items which explained the 11 sub-dimensions of service quality. The measures for these sub-dimensions were adopted from existing models and scales which also have been discussed in this literature review (Kim & Kim, 1995; Howat et al., 1996; Chelladurai & Chang, 2000; Papadimitriou & Karteroliotis, 2000; Rust & Oliver, 1994; Brady & Cronin, 2001). However, the SSQRS has the same limitations as the model by Alexandris et al. (2004). For the SSQRS scale a sample of university students from one university in the United States was used. Therefore, the generalizability of the scale may be limited. Due to this issue, the researchers suggested further analysis and testing of the scale in other industry segments and with other samples.

In a more recent study, Morales and Gálvez (2011) developed the Perceived Quality Assessment Questionnaire in Sports Services (CECASDEP). In their study Morales and Gálvez (2011) have used a total of 71 items and found sports facilities, attention to the user, sports spaces, dressing room, program of activities and instructor, as the six dimensions in their model. However, the data was only collected from 2 cities in Spain. Therefore, the developed scale may not be appropriate for use for fitness centers in other countries.

A study by Yıldız (2011) proposed another more recent service quality scale for fitness centers. The scale was named SQS-FC (Service Quality Scale for Fitness Centers) and examines its effectiveness using importance-performance analysis (Yıldız, 2011). The SQS-FC scale has revealed 25 service attributes offered in fitness centers and contained four service quality dimensions, which are: personnel, physical environment, supporting services and program. The author mentions that the scale has similarities to previous scales from Chelladurai and Chang (2003), and Lam et al. (2005). The main difference is the breakdown of dimensions and the introduction of the items personnel's ethical behavior and customer consultation by specialists. However, the scale was developed for a private fitness club in Turkey. Therefore, a relatively small sample from a single fitness club in Turkey was used. Moreover, Yıldız (2011) describes that the market in Turkey is emerging and that therefore the results of this study may not be generalizable for fitness centers in other continents.

Another scale which was developed by Rial, Varela, Rial and Real (2010), The QSport-10 scale consists of a 10-item scale to measure service quality in fitness centers in only two dimensions, namely: staff and installations. However, there was criticism on the scale since many authors in this field address offered programs as a crucial dimension while the model by Riat et al., (2010) does address this aspect, it does not include it. Due to this issue, Yıldız and Kara (2012) re-examined the QSport-10 scale in order to create a modified version. Yıldız and Kara (2012) introduced the QSport14 scale which was developed to measure service quality in physical activity and sports centers. The scale consists of 14 items which were distributed among 3 dimensions (staff, program and installations). A confirmatory factor analysis was conducted and it was found that the dimension of program was the most important factor for the largest segment of customers.

Another model to measure service quality was developed by Uçan (2007) and was called the Sport-Fitness Centers Perceived Service Quality Scale (SFC-PSQS). However, the model has not received many attention, this is most likely due to it being a doctorate thesis which was never published. While the model has not received much attention, Yıldız, Polat and Güzel (2018) have used it in their study in order to investigate sport center members perceived service quality levels. In their study Yıldız et

al., (2018) explain the SFC-PSQS, and describe that it includes 31 items and the following dimensions: interaction quality, outcome quality, physical environment quality, exercise tools and equipment, program quality and environmental conditions quality.

In a recent study, Vieira and Ferreira (2017) have aimed to understand the current strategic framework of fitness centers, advocated by the blue ocean strategy, from the quality dimensions. In order to realize this, the quality dimensions of Papadimitriou and Karteliotis (2000) and Yildiz (2011) have been used for their theoretical models. The study has used a sample of 151 fitness clubs and conducted an exploratory factor analysis. The results from this study indicate that employee competences, facilities, core services and complementary services are the key factors in the strategy of fitness centers.

Several authors have created models, instruments and scales which have aimed to measure service quality in the fitness industry. While many authors found similar dimensions, some studies have suggested or introduced new dimensions. Table 1 summarizes the scales and models which have been created to measure service quality in fitness centers. Many of the studies mentioned in table 1 have used different terminologies for dimensions which are similar or identical. In order to create a model with a limited amount of dimensions but with a high explained variance, the most commonly used dimensions from previously developed scales have been merged, in order to create a new set of dimensions which literature agrees upon. The resulting dimensions have been added to table 1, and have been named: personnel, physical environment, program, social environment, supporting services and outcome. Furthermore, table 1 indicates which developed dimensions have been used in which previously developed scales in order to support the new dimensions and to get an estimate about the importance of each new dimension.

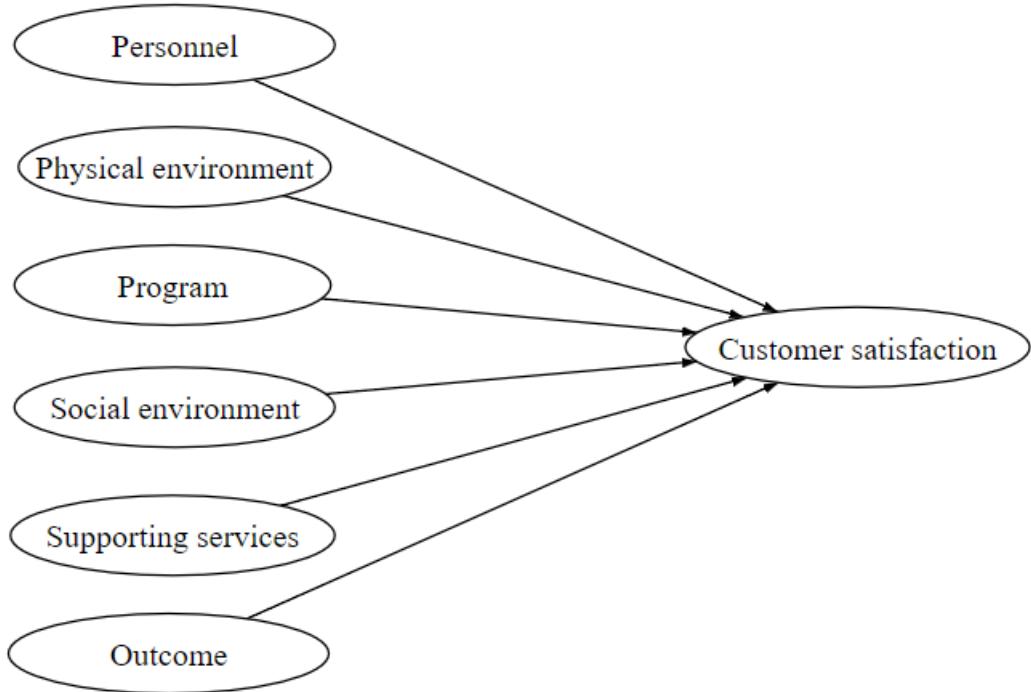
Table 1. Service quality scales for fitness centers and resulting dimensions.

Research (sorting based on publishing year)	Scale	Dimensions					
		Personnel	Physical environment	Program	Social environment	Supporting services	Outcome
Chelladurai et al., 1987	SAFS	Yes	Yes	Yes	No	Yes	No
Kim & Kim, 1995	QUESC	Yes	Yes	Yes	Yes	No	No
Papadimitriou & Karteroliotis, 2000	Four factor solution	Yes	Yes	Yes	No	Yes	No
Chelladurai & Chang, 2003	SQFS	Yes	Yes	Yes	No	No	Yes
Alexandris et al., 2004	-	Yes	Yes	No	Yes	No	Yes
Lam et al., 2005	SQAS	Yes	Yes	Yes	Yes	Yes	No
Ko & Pastore, 2005	SSQRS	Yes	Yes	Yes	No	No	Yes
Uçan, 2007	SFC-PSQS	Yes	Yes	No	No	No	Yes
Rial et al., 2010	QSport-10	Yes	Yes	No	No	No	No
Morales & Gálvez, 2011	CECASDEP	Yes	Yes	Yes	No	No	No
Yildiz, 2011	SQS-FC	Yes	Yes	Yes	Yes	Yes	No
Yildiz & Kara, 2012	QSport-14	Yes	Yes	Yes	No	No	No
Vieira & Ferreira 2017	-	Yes	Yes	Yes	No	Yes	No

3. Model Development and Hypotheses

The theoretical framework has led to the development of the proposed conceptual framework as indicated in Figure 1. The proposed conceptual framework consists of the following independent variables: personnel, physical environment, program, social environment, supporting services and outcome. These independent variables will be used to predict the dependent variable customer satisfaction.

Figure 1. proposed conceptual framework



3.1. Hypothesizes

Based on the proposed conceptual framework as indicated in Figure 1, the following 6 hypotheses are suggested.

- H1: Personnel has a positive effect on satisfaction
- H2: Physical environment has a positive effect on satisfaction
- H3: Program has a positive effect on satisfaction
- H4: Social environment has a positive effect on satisfaction
- H5: Supporting services has a positive effect on satisfaction
- H6: Outcome has a positive effect on satisfaction

4. Research design & Methodology

For this study, both secondary and primary data has been used. Secondary data is information that has been previously collected for another study, but can be used again (Tustin et al., 2010). Secondary data has been used in the theoretical framework in order to develop the proposed conceptual framework and the hypothesizes. On the other hand, primary data is information that is specifically collected through original or first-hand research (Tustin, Lighelm, Martins, & Van Wyk, 2010). Primary data has been collected in order to test the hypothesizes and answer the research question.

In this study a quantitative approach was used. This involves the collection of primary data from a representative sample of the total population with the intention of projecting the results to a larger population. A quantitative approach has been used since it allows researchers to quantify attitudes, opinions and other variables and make generalizations. Based on these statements, the use of a quantitative approach seems to have a better fit to test the hypothesizes and answer the research question than qualitative research.

4.1. Sample & data collection

The data for this study has been collected from three Dutch SME Fitness Centers. The questionnaires have been distributed to members when they entered or left the club, and could be completed in a convenient place. The reasoning for this method of data collection is based on the theory that respondents are more interested in completing the questionnaire and would provide more meaningful answers (Dabholkar, Thorpe, & Rentz, 1996). Participation in this research was anonymous, voluntary and in accordance with the human subject guidelines of the University of Twente.

4.2. Instrument development

In order to collect the required information, a self-administered questionnaire has been developed based on relevant literature regarding service quality and customer satisfaction. In order to ensure content validity the dependent and independent variables used in this study have been based on previous studies. As mentioned before, there are a large amount of conceptualizations regarding customer satisfaction. In this study, customer satisfaction is considered to be based on the expectancy disconfirmation theory, since most conceptualizations are in line with the expectancy disconfirmation theory (McQuitty, Finn and Wiley, 2000). Therefore, customer satisfaction was measured by adopting three items from a relevant study by Hill and Brierley (2017). The independent items in the questionnaire have been adopted from previous studies by Yildiz (2011), Papadimitriou and Karteroliotis (2000), Lam et al. (2005) and Vieira & Ferreira (2017). These studies have been selected based on their similarity with the independent variables suggested in the proposed model of this research. Discussions with a fitness center manager, feedback from a research expert from the UT and a pilot study using 10 participants were used in order to refine the instrument.

The final questionnaire consists of two sections. The first section of questionnaire covers the demographic and general information of respondents such as: age, gender and frequency of visits. The second section of the questionnaire includes 31 measurements in order to measure possible constructs which are identified in the theoretical framework. These 31 items have been measured using a 5-point Likert scale (1 – strongly disagree and 5 – strongly agree). The decision to use a 5-point likert scale has been based upon the finding that it allows researchers to increase the response quality and the response rate among respondents (Buttle, 1996).

4.3. Measurements

The data from the questionnaires has been analyzed by using SPSS (Statistical Package for Social Science) version 25. While there are many other Statistical Software Packages, only SPSS has been used in this research. The choice to use SPSS has been based on the indication that researchers in the field of social sciences often use SPSS (Bala, 2016). Therefore, the use of SPSS may make it easier for other researchers to reproduce the results of this study with another sample to verify the model.

In this study descriptive statistics have been utilized in order to analyze demographic characteristics of the participants. Furthermore, inferential statistics have been used to measure reliability and validity, and in order to answer the developed research question and hypothesizes.

5. Results

While many researchers choose to describe each decision and the rationale for choosing it in the methodology section, there are also researchers who combine their explanations with the results in the result chapter for practical reasons. In this study several assumptions and analyses have been explained, justified and conducted. In order to create a logical, clear and easy interpretable research, the explanation, justification and results of the conducted analyses have been described in the result chapter. This has been done since several decisions regarding tests have been based on described literature or previous results. If a reader would have to go back to the methodology in order to comprehend why certain decisions were made, the research may become harder to interpret.

5.1. Descriptive statistics

In total, the data of 140 participants has been used in this study. As presented in Table 2, an analysis of the demographic characteristics indicates that this study contains a slightly higher percentage of female respondents (55,7%) compared to male respondents (44,3%). Furthermore, the descriptive statistics reveal that a large percentage of respondents (39,3%) were from the age group >60 years, indicating the presence of many older respondents in the sample. In terms of the highest completed education, 9 respondents (6,4%) completed elementary education, 37 respondents (26,4%) completed high school, 39 respondents (27,9%) completed vocational education, 45 respondents (32,1%) have a bachelor's degree and 10 respondents (7,1%) have a master's degree. Regarding the membership length, it was found that a large percentage of respondents (32,1%) has been a member for >10 years. Furthermore, 29 respondents (20,7%) have been a member for 0-1 year, 13 respondents (9,3%) have been a member for 1-2 year, 29 respondents (20,7%) have been a member for 2-5 year and 24 respondents (17,1%) have been a member for 5-10 year. Finally, when analyzing the average visits per week it can be seen that most respondents visit 2 times per week (43,6%) or 3 times per week (35,0%). Furthermore, 4 respondents (2,9%) visit <1 time per week on average, 7 respondents (5,0%) visit 1 time per week on average, 12 respondents (8,6%) visit 4 times per week on average and 7 respondents (5,0%) visit >4 times per week on average.

Table 2. Demographic profile of the respondents

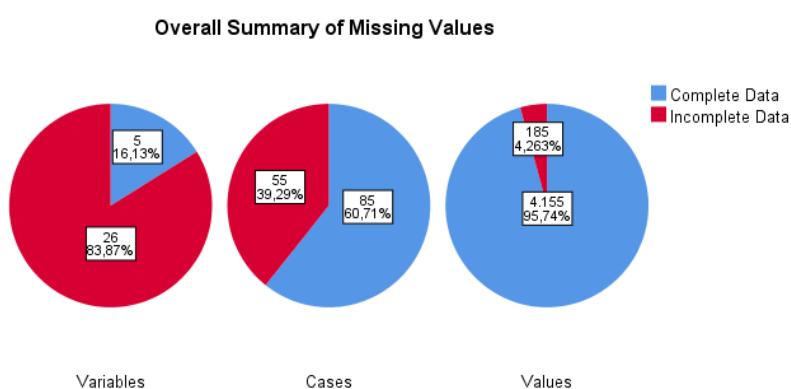
Measures	Category	Frequency	Percentage
Gender	Male	62	44,3
	Female	78	55,7
Age	<20	21	15,0
	20-30	18	12,9
	31-40	10	7,1
	41-50	16	11,4
	51-60	18	12,9

	>60	55	39,3
	Missing	2	1,4
Education	Elementary	9	6,4
	High school	37	26,4
	Vocational education	39	27,9
	Bachelor	45	32,1
	Master	10	7,1
Membership length	0-1 Year	29	20,7
	1-2 Year	13	9,3
	2-5 Year	29	20,7
	5-10 Year	24	17,1
	>10 Year	45	32,1
Visits per week	<1	4	2,9
	1	7	5,0
	2	61	43,6
	3	49	35,0
	4	12	8,6
	>4	7	5,0

5.2. Missing data

Missing data is a common problem for all kinds of research, this study has been no exception. Figure 2 indicates that 185 values (4,3%) are missing and that 4155 values are present (95,74%). The 4,3% missing data are not likely to result in issues with analyses. However, this sample contains 55 incomplete cases (39,29%) and 85 complete cases (60,71%). Due to the use of a multiple regression based on factor scores, the large percentage of incomplete cases could lead to a large percentage of data loss. The large percentage of data loss is the result of factor scores, since these can only be created using complete cases. Therefore, only 60,71% of all cases will be used if basic missing data approaches like pairwise or listwise exclusion are used.

Figure 2. Summary of the missing values using SPSS



Due to the large amount of data that will be lost by the use of pairwise or listwise exclusion, these methods will not be used. Other traditional methods that are used to deal with missing data (i.e., hot-deck imputation, mean imputation and regression imputation), are better than just deleting cases, but these methods still fail to obtain unbiased estimators (Tufis, 2008). Therefore, statisticians suggest using the maximum likelihood method of imputing the missing data or multiple imputation. However,

multiple imputation has a wider applicability and can be used with any type of statistical model (Tufis, 2008).

Due to the advantages in comparison to other methods, multiple imputation has been used in this research. Multiple imputation is a general solution to the missing data problem which can be found in a few statistical programs. Multiple imputation handles missing data by generating several different possible imputed data sets and comparing the results obtained from each (Sterne et al., 2009). Before multiple imputation can be used, it is necessary to check if the missing data is missing at random (MAR) instead of missing completely at random (MCAR) or not missing at random (NMAR). If the missing data is considered to be MCAR or NMAR then Jakobsen et al., (2017) suggest only using observed variables. In order to identify if the missing data is MCAR, Little's MCAR test has been applied. Based on Little's MCAR test it can be observed that the missing data is not MCAR since the P.value is <0.05.

Table 3. Little's MCAR test.

EM Means ^a		
Little's MCAR test	Approx. Chi-Square	1246,271
	df	1135
	Sig.	,011

Unfortunately there are currently no tests to indicate if missing data is MAR or NMAR. In order to determine if missing data is MAR or NMAR, researchers should use their knowledge about the data in order to determine which type of missing data is present. Since no sensitive questions are asked in the questionnaire (i.e. income, IQ), it is assumed that the missing data is not NMAR. Therefore, it is assumed that the missing data is MAR. Since the missing data is assumed to be MAR, multiple imputation seems appropriate in this research.

Before multiple imputation can be applied, it is important to consider how many imputations should be used. Theoretically speaking it would always best to use the highest amount of m (amount of imputations) possible, but this involves large amount of computation, storage and analyzation. According to Schafer and Olsen (1998), there is little advantage to producing and analyzing more than a few imputed datasets. Buuren (2018) compared many different theories and concluded that researchers should set the amount of imputations to the average percentage of missing data. Buuren (2018) also explains that substantive conclusions are unlikely to differ as a result of increasing the amount of imputations beyond 5. Based on the percentage of missing data and previous research, multiple imputation has been applied in this research using 5 imputed datasets.

5.3. Validity

Validity is used to measure the extent to which the research measurements properly represent the research concept (Hair et al., 2014). Unlike reliability, the concept of validity can be divided in several validities. In research, validity is often believed to consist of four types: face, criterion, content and construct validity.

Face validity indicates if a test appears to measure whatever it is supposed to measure (Hecker & Violato, 2009). Face validity is only objective and plays a secondary role to content validity, therefore it is often considered as a weak form of validity (Hecker & Violato, 2009). Since content validity will be adopted, face validity will not be used in this research.

Criterion validity is the extent to which a measure agrees with a gold standard. Criterion validity testing is often difficult due to the general lack of gold standards for questionnaire-based measures (DePoy & Gitlin, 2016). Due to the lack of a gold standard for customer satisfaction in the fitness industry, criterion validity will not be adopted in this research.

Content validity refers to the degree to which a developed model is accurately representative

of the construct it is designed to measure (Sireci, 1998). In order to ensure content validity, the created constructs have been based on current knowledge as described in the theoretical framework and contain items adopted from similar studies.

Besides content validity, this research also takes into account construct validity. Construct validity is generally being applied to determine to which extent a scale measures a variable of interest (Punniyamoorthy & Prasanna, 2007). In order to measure construct validity, discriminant validity and convergent validity have been utilized. Discriminant validity is used to indicate if constructs that theoretically should not be related are indeed not related to each other. Construct validity is used to indicate if constructs that theoretically should be related are indeed related to each other. In order to make statements about the discriminant and convergent validity, a factor analysis and a Pearson Correlation Coefficient Analysis have been applied.

5.4. Factor Analysis independent variables

Based on the research design and data, a factor analysis seems like an appropriate analysis since it reduces the relative large amount of variables to a smaller set of categories (factors), to get an underlying concept, and to facilitate interpretations (Yong & Pearce, 2013). Without the use of a factor analysis the results would be hard to interpret and there would most likely be multicollinearity issues, since questions from questionnaires are often related.

In advance of the factor analysis, certain assumptions have to be met in order to ensure valid results. First, multiple variables that are measured at a continuous level should be used. While Likert type data is not measured at a continuous level, it can be used as continuous without any harm to the analysis if they contain five or more categories (Johnson & Creech, 1983; Norman, 2010; Sullivan & Artino, 2013). Since the applied Likert scale contained 5 categories, we refer to the variables as a ordinal approximation of a continuous variable.

Unfortunately, the use of a data from a likert scale can result in issues when testing assumptions of linearity. In order to test linearity a Pearson Correlation Coefficient Analysis has been applied. This test has been applied since a scatterplot cannot be applied due to the ordinal value of the variables in SPSS. Usually, a Spearman's Rho is preferred when using ordinal variables. However, as mentioned before we refer to our variables as a ordinal approximation of a continuous variable. Therefore, a Pearson Correlation Coefficient Analysis seems appropriate. The result from the Pearson Correlation Coefficient Analysis (Appendix 8.4) indicates that linearity is apparent due to the high significance of correlation.

Furthermore, the sampling adequacy of the data has been investigated by the use of a Kaiser-Meyer-Olkin (KMO) test. The KMO test has measured the sampling adequacy of every variable in the conceptualized model and for the entire model. Therefore, this statistic indicates the proportion of variance in the variables that may be caused due to underlying factors (Dziuban & Shirkey, 1974).

Table 4. Pooled KMO and Barlett's Test (independent variables)

(Pooled) KMO and Barlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		,825
	Approx. Chi-Square	2052,568
Bartlett's Test of Sphericity	df	378
	Sig.	,000

The outcome value is ,825 which indicates that a factor analysis is appropriate and may be useful, since the outcome value exceeds the $>.5$ threshold. For an ideal sampling adequacy, there should also be at least 150 cases with an ratio of at least 5 cases for each variable (Pallant, 2010). This study has 140

cases with a ratio of 4,5 cases per variable. While this is not ideal, it is acceptable since the overall sampling adequacy for the data set is high enough according to the KMO test.

Moreover, a Bartlett's test of sphericity has been applied to measure the suitability for data reduction. The Bartlett's test of sphericity allows researchers to compare the correlation matrix to the identity matrix. This indicates if there is a redundancy amongst variables that can be summarized with some factors (Tobias & Carlson, 1969). The result of the Bartlett's test of sphericity indicates if a factor analysis may be useful with the acquired data. The result of the Bartlett's test of sphericity indicates a significance of ,000 which is lower than the <0.05 threshold. Therefore, the Bartlett's test of sphericity indicates that a factor analysis might be useful.

The final assumption required for a valid result, is the absence of significant outliers. However answering the extremes in the likert scale (1 or 5) questionnaire does not represent general outlier behavior. Therefore, outliers do not really exist in this likert based dataset.

Based on the previous assumptions, a factor analysis seems appropriate. An exploratory factor analysis utilizing the principle component method with a varimax rotation has been conducted in order to enhance interpretability and avoid multicollinearity problems. Since the resulting factors are most likely an outcome of the indicators, principle component analysis (PCA) seems appropriate. For example, we might say that satisfaction about employees is a combination of employee knowledge and employee willingness to help ($z=x_1+x_2$). Furthermore, normality is an assumption for other Exploratory Factor Analysis (EFA) methods, but not for PCA. Since data from a likert scale is not normally distributed, the use of a PCA seems more appropriate than other factor extraction methods. It should be noted that PCA and EFA are both variable reduction techniques and are often seen as the same statistical method, but PCA is technically speaking not an EFA method. The difference of these methods has been acknowledged in this study, but PCA will be considered to be an EFA method since the difference is technical and outside of the scope of this research.

The factor extraction amount has been based on the theoretical framework and proposed model. Therefore, a fixed number of 6 factors have been chosen to be extracted with a threshold of >1 eigenvalue. The threshold of >1 eigenvalue has been based on the Kaiser Criterion which explains that factors with an Eigenvalue >1, should be considered as a factor and that factors with an Eigenvalue <1, should be discarded (Braeken & van Assen, 2017). The resulting factor loadings from the exploratory factor analysis are the correlation coefficients between the variables and factors. Hair et al., (2006) describe that the significance of factor loadings generally depend on the sample size. However, theoretical considerations and latent variable meaning should also be taken into account. Based on the guidelines by Hair et al., (2006) and theoretical considerations, measurement items with a factor loading <0.450 have been removed. The remaining items will contain factor loadings >0.450 which indicates that there is a positive correlation among the items of each component. Furthermore, the factor loadings above 0.450 also indicate discriminant and convergent validity of the data. (Hair et al., 2006).

Table 5. Pooled Total variance explained by factors (independent variables)

Component	(Pooled) Total Variance Explained					
	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8,375	29,912	29,912	4,558	16,280	16,280
2	3,120	11,143	41,055	3,063	10,940	27,221
3	1,942	6,936	47,991	3,035	10,840	38,060
4	1,621	5,791	53,781	2,719	9,712	47,772
5	1,456	5,198	58,980	2,427	8,668	56,440
6	1,373	4,905	63,885	2,084	7,445	63,885
7	1,026	3,665	67,549			
8	,808	2,887	70,437			
9	,765	2,733	73,170			
10	,753	2,690	75,860			
11	,693	2,474	78,334			
12	,671	2,395	80,729			
13	,600	2,141	82,870			
14	,532	1,902	84,772			
15	,503	1,798	86,570			
16	,441	1,575	88,145			
17	,438	1,565	89,710			
18	,410	1,465	91,175			
19	,348	1,242	92,417			
20	,346	1,237	93,655			
21	,337	1,205	94,860			
22	,268	,958	95,818			
23	,264	,943	96,762			
24	,246	,880	97,642			
25	,201	,718	98,360			
26	,175	,625	98,986			
27	,145	,517	99,503			
28	,139	,497	100,000			

Extraction Method: Principal Component Analysis.

Table 6. Pooled Exploratory factor analysis including Cronbach's Alpha & Eigenvalue (independent variables)

			Rotated Component Matrix ^a						Cronbach's alpha
			1	2	3	4	5	6	
Physical environment	Q2	The facilities are modern	,848						
	Q3	The facilities are attractive	,827						
	Q4	The equipment is in good condition	,782						
	Q6	The environment (temperature, light) is pleasant	,719						8,375 ,881
	Q1	The facilities are clean	,701						
	Q5	There is a good variety of equipment	,663						
	Q7	The location is pleasant and sensible	,495						
	Q8	There is a decent locker room and shower	,462						
Program	Q15	The program is diverse enough	,776						
	Q16	The quality / content of the programs is good	,770						
	Q17	There is a convenient program time/schedule	,750						3,120 ,800
	Q18	The class size of the programs is appropriate	,675						
	Q14	Enough secondary services, like childcare, suntanning & physiotherapy are offered	,487						
Personnel	Q20	The employees are willing to help	,842						
	Q21	The employees communicate in a clear and appropriate way	,830						1,942 ,872
	Q22	The employees are responsive to suggestions and complaints	,762						
	Q19	The employees possess the required knowledge/skills	,746						
Outcome	Q27	Visiting can help me to improve my mood	,865						
	Q28	Visiting can help me to increase my energy	,818						1,621 ,880
	Q26	Visiting can help me to improve my health	,781						
Supporting services	Q13	Enough attention is paid to the safety of visitors	,707						
	Q11	The range of food and drinks is good	,702						
	Q12	Appropriate background music is played	,692						1,456 ,708
	Q9	The membership fee is reasonable	,559						
	Q10	The opening hours are long enough	,474						
Social environment	Q23	I have the chance to meet new people	,810						
	Q25	Exercising gives me the chance to socialize	,746						1,373 ,748
	Q24	There is a nice and friendly atmosphere	,574						

Footnote: 5 point scale (1= strongly disagree to 5= strongly agree). Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 6 iterations.

The factor analysis has resulted in 6 factors, which have been named: Physical environment, Program, Personnel, Supporting services, Outcome and Social environment. All 6 factors contain variables which have a factor loading >450, indicating that each variable has enough association with the underlying factor. Furthermore, each factor has a eigenvalue of >1, indicating that a decent amount variance out of the total variance is explained by each factor. Since a PCA instead of EFA has been used, it is important to mention that the resulting factors (constructs) should technically be called Principle components. However, since this research doesn't makes a distinction between PCA and EFA the resulting principle components will also be referred to as factors or constructs.

It is important to note that table 5 indicates that a possible seventh factor also would have had an eigenvalue of >1. Therefore, another exploratory factor analysis has been run where factor extraction is based on eigenvalue >1. The 7th factor extracted has a eigenvalue of 1,026 and consists of the variables safety and background music with factors loading of ,720 and ,632 respectively. The exploratory factor analysis based on eigenvalue >1, results in 19 higher factor loadings and 8 lower factor loadings in comparison to the exploratory factor analysis with 6 fixed factors. However, it also results in the 5th and 7th factor having a Cronbach's alpha of ,635 and ,615 respectively. This is a significant decrease in reliability since the lowest Cronbach's alpha of the 6 factor solution is ,708. Since the 7th factor has a eigenvalue which is close to 1 and since it does not seem logical or add anything useful to the model, it will be discarded. For clarity, the result of the exploratory factor analysis based on eigenvalue >1, has been added in the appendix (8.5).

5.4.1. Factor Analysis dependent variables

It is important to note that an exploratory factor analysis does not distinguish between dependent variables and independent variables (Hoang Trong and Chu Nguyen Mong Ngoc, 2008). Therefore, an exploratory factor analysis considers relationships between all selected variables in the analysis. In order to create a reliable model a separate exploratory factor analysis has been conducted for the dependent variables.

Table 7. Pooled KMO and Barlett's Test (dependent variables)

(Pooled) KMO and Barlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		,766
	Approx. Chi-Square	331,492
Bartlett's Test of Sphericity	df	3
	Sig.	,000

Most assumptions are in line with the previously tested assumptions. However, it still important to conduct a KMO and Barlett's Test since these results may differ. The KMO test regarding the dependent variables has an outcome value of ,766 which indicates that a factor analysis is appropriate and may be useful, since the outcome value exceeds the >.5 threshold. The result of the Bartlett's test of sphericity indicates a significance of ,000 which is lower than the <0.05 threshold. Therefore, the Bartlett's test of sphericity indicates that a factor analysis might be useful for the dependent variables.

Table 8. Pooled total variance explained by factors (dependent variables)

Component	(Pooled) Total Variance Explained					
	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2,631	87,701	87,701	2,631	87,701	87,701
2	,199	6,630	94,331			
3	,170	5,669	100,000			

Based on the KMO and Bartlett's test another factor analysis for the dependent variables seems appropriate. Therefore, a second exploratory factor analysis utilizing the principle component method has been conducted. Based on theory, 1 fixed factor has been chosen to be extracted.

Table 9. Pooled exploratory factor analysis including Cronbach's Alpha & Eigenvalue (dependent variables)

		Component Matrix ^a		
		1	Eigenvalue	Cronbach's alpha
Customer Satisfaction	I am generally satisfied	,931		
	I am pleased to have taken the decision to become a member	,937	2,631	,929
	I would recommend this fitness center to a friend	,941		

Footnote: 5 point scale (1= strongly disagree to 5= strongly agree). Extraction Method: Principal Component Analysis.

The factor analysis containing the dependent variables has resulted in 1 factor, namely: Customer satisfaction. Only one factor was found which indicates that all dependent variables fit onto a single theoretical construct. A varimax rotation has not been used in this factor analysis, since only 1 factor was found. Rotation is often used to see underlying dimensions more clearly. However, with only 1 factor the rotating of the axes is not possible since there is nothing to rotate.

The resulting factor only contains variables with a factor loading $>,450$ indicating that the variables have enough association with the underlying factor. Furthermore, the resulting factor has an eigenvalue of >1 , which indicates that it explains a decent amount of variance.

After analysis of the second exploratory factor analysis it becomes clear that there are no significant differences with the proposed model. Therefore, the research model and hypotheses aren't subject to change.

5.5. Reliability

According to Hair et al., (2014) reliability is an evaluation of the degree of consistency among multiple measurements of a variable. In this research Cronbach's alpha has been applied in order to test the reliability of the measurements. Cortina (1993), describes that Cronbach's alpha is one of the most widespread and important statistic measures in research involving test construction and use. The use of Cronbach's alpha has become so common that its use in research containing multiple-item measurements is being considered routine (Schmitt, 1996). Due to the significance and importance in research, Cronbach's alpha has been chosen as the reliability test statistic.

Cronbach's alpha coefficient ranges from 0 to 1 in providing the overall assessment of a measure's reliability. When all items are independent from another with no correlation or covariance then Cronbach's alpha would be 0. On the contrary, when all items have a high covariances, then Cronbach's alpha will approach 1. Therefore, a high Cronbach's alpha coefficient indicates if items have shared covariance and indicates if they measure the same underlying concept. Cronbach (1951), suggested that that a relative high alpha score is desirable, but he argues that obtained scores from a test instrument should be interpretable. Schmitt (1996), also describes that there is no exact threshold where Cronbach's alpha becomes acceptable. Therefore, Schmitt (1996) explains that some instruments with a relative low Cronbach's alpha can still prove useful in certain circumstances. The idea that there is no exact threshold may confuse some researchers since a general level of 0,70 is often considered to be good or acceptable. However, Cronbach's alpha tends to increase with the size of the used instrument (Cortina, 1993; Schmitt, 1996), therefore researchers can simply add more items to an instrument in order to increase Cronbach's alpha to 0,70. Due to this instrumental weakness, researchers should not only look at the general threshold of 0,70, but also apply their theoretical knowledge in order to determine if something is considered reliable.

To evaluate if the factors derived from both conducted exploratory factor analyses are reliable, Cronbach's alpha has been used as a reliability measure for all resulting factors. The result indicates that all constructs had a Cronbach's alpha coefficients >0,7. Cronbach's alpha coefficient ranged from 0,708 (Supporting services) to 0,929 (Customer satisfaction). Based on theoretical knowledge regarding the scale in question, the resulting Cronbach's alpha coefficients and the resulting factor analysis results, it can be assumed that these factors do indeed tap into underlying constructs among respondents. Therefore, a good reliability is assumed.

5.6. Pearson Correlation Coefficient Analysis

Pearson's Correlation is considered a bivariate analysis which allows researchers to measures the strength of association between variables and the direction of the relationship. From the result shown in Table 10, it can be seen that all identified independent factors show significance correlation with the dependent factor. The Pearson Correlation of ,365 indicates that the dependent variable "Customer Satisfaction" correlates strongest with the independent variable "Personnel". The Pearson Correlation of ,249 reveals that the weakest correlation of the dependent variable "Satisfaction" is with the independent variable "Supporting services". The different Pearson Correlation scores indicate difference in correlated strength. However, all correlations between the dependent and independent variables represent positive signs, which indicates the positive direction of the associations among all tested constructs.

Furthermore, the Pearson Correlation Coefficient Analysis indicates that the independent factors show no significance correlation among them. Therefore, multicollinearity does not seem to cause issues when computing a multiple regression model.

Table 10. Pooled Pearson Correlation Coefficient Analysis

	(Pooled) Correlations						
	1	2	3	4	5	6	7
1. Physical environment	-						
2. Program	-,009	-					
3. Personnel	-,014	-,011	-				
4. Outcome	,014	-,009	-,004	-			
5. Supporting services	-,017	-,001	,003	-,008	-		
6. Social environment	,007	,003	,004	-,013	-,007	-	
7. Customer satisfaction	,337**	,323**	,365**	,357**	,249**	,361**	-

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

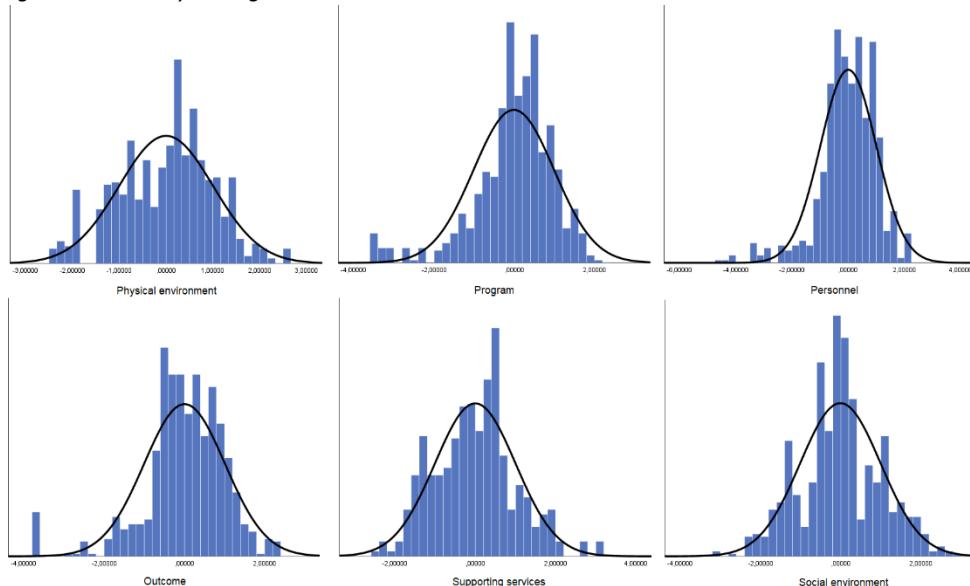
5.7. Multiple Regression Analysis

A multiple regression analysis allows researchers to predict values of the dependent variable, or it can be used to indicate which independent variables have a major effect on the dependent variable (McDonald, 2014). Furthermore, multiple regression is being applied in order to create an understanding of the functional relationships between the independent variables and dependent variable (McDonald, 2014). This allows researchers to see what might cause variation in the dependent variable. Since this research contains one dependent factor and multiple independent factors which are assumed to have an significant effect on the dependent variable, a multiple regression analysis seems suitable. Not only does a multiple regression seem suitable in this research, it is also considered to be important since the results can be used to test the hypothesizes. Before the multiple regression analysis can be conducted, certain assumptions have to be met in order to ensure valid results. It is important to note that this multiple regression is computed based on previously established factors (principal components), therefore not every basic assumption for multiple regression has to be met in order to ensure valid results (NCSS, 2019).

While it is often not considered as an assumption, it is important to check the sample size since sufficient data is required to provide reliable estimates of the correlations. A common rule of thumb to examine the sample size is that researchers should have at least 50 cases and at least 10 - 20 as many cases as there are independent variables (Tabachnick & Fidell, 2007). Since this dataset contains 140 cases and 6 independent variables, the sample size seems appropriate.

The normality assumption is the first assumption that has been tested. Normality has to be tested since regression assumes that variables have normal distributions. Therefore, non-normally distributed variables may result in distorted relationships and significance tests (Osborne & Waters, 2002). Unfortunately, tests of normality (e.g., the Shapiro-Wilk test) are notorious for their sensitivity to sample size (Boedec, 2016). Due to the limited sample size of this research, graphical methods are considered to be a better alternative to evaluate normality than tests of normality.

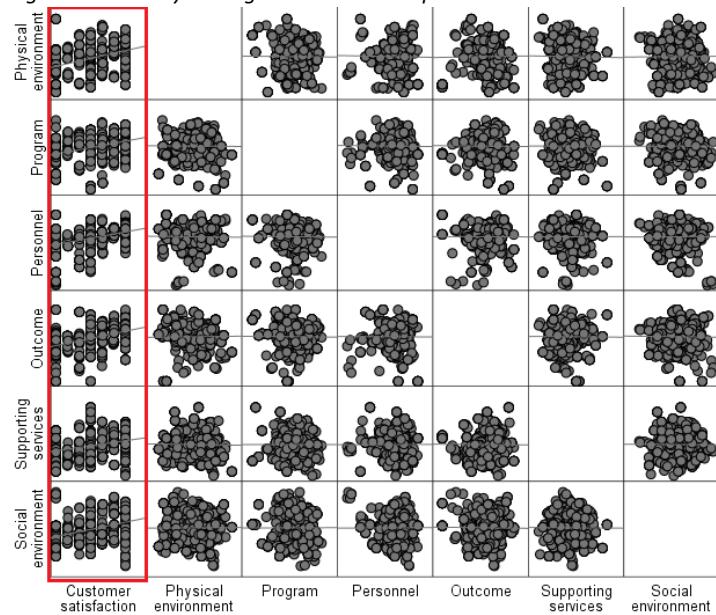
Figure 3. Normality testing.



Based on Figure 3 we can state that the data seems to have a normal distribution to a certain degree. Since the created principle components are based on likert-scale data, the data does not have a clear and perfect normal distribution due to skewness and kurtosis. This is not surprising since likert based data is often considered to be non-normal. However, this does not result in problems since principle component regression does not provide confidence limits, which means that normality can be preferred but does not have to be assumed per se (NCSS ,2019).

Another assumption is the requirement of a linear relationship between the dependent principle component (customer satisfaction) and each independent principle component (Osborne & Waters, 2002). In order to check linearity, scatterplots have been created. Figure 4 indicates that there is indeed a linear relationship between the dependent principle component (customer satisfaction) and each independent principle component.

Figure 4. Linearity testing based on scatterplot



The scatterplot matrix can also be used to test assumptions regarding homoscedasticity. Homoscedasticity refers to the amount of variances along the line of best fit, where the variance should remain similar along the line in order to have homoscedasticity (Osborne & Waters, 2002). Based on the scatterplot matrix in Figure 4, it can be stated that the homoscedasticity is decent. Since this research uses social data it was expected that the data would not indicate perfect homoscedasticity. However, the data seems to indicate more homoscedasticity rather than heteroscedasticity. Therefore, the assumptions regarding homoscedasticity has been met.

The scatterplot matrix also reveals that there are some variables which may be considered as outliers. This may cause issues since the absence of outliers is another assumption which has to be met in order to ensure valid multiple regression results. However, as mentioned before, answering the extremes in the used likert-scale questionnaire (1 or 5) does not represent general outlier behavior. Therefore, we assume that true outliers do not really exist in this dataset.

Lastly, the previously created Pearson Correlation Coefficient Analysis (Table 10) has been used to examine if there is a multicollinearity problem, since the final assumption indicates that independent factors should not be overly correlated with one another. Multicollinearity is a problem since it blurs which independent factor contributes to the variance explained in the dependent factor. The conducted Pearson Correlation Coefficient Analysis indicates that the independent principal components show no significance correlation among them. Therefore, multicollinearity does not seem to cause issues when computing a multiple regression model.

All the assumptions for a multiple regression using principal components have been met. Since the assumptions have been met, and since multiple regression is used by researchers to evaluate the relationship between 1 dependent construct and at least 2 independent constructs, the use of this test seems appropriate.

Table 11. Pooled regression model summary

Model Summary ^b					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,826 ^a	,683	,667	,576266	1,849

a. Predictors: (Constant), Social environment, Program, Supporting services, Personnel, Outcome, Physical environment

b. Dependent Variable: Customer satisfaction

The first results from the multiple regression analysis can be found in table 11. While the resulting R, R Square and Adjusted R Square sound similar, there is a difference in meaning. R indicates the correlation coefficient between the predictors and the outcome and can be considered as a quality measure of the prediction of the dependent variable (Dhakal, 2018). The resulting R of ,826 indicates a good level of prediction.

R^2 (R Square) indicates which amount of variance in the dependent variable can be explained by the independent variables. The resulting R^2 of ,683 indicates that 68,3% of variance can be explained by the independent variables in this research. The remaining 31.7% depends on the impact of others factors which have not been tested in this research.

The adjusted R^2 (Adjusted R Square) is considered to be a modified version of R^2 that has been adjusted for the amount of predictors in a model. The adjusted R^2 only increases if added variables improve the model more than would be expected by chance. Since regular R^2 does not take chance into account, every predictor added to a model increases R^2 while it would never decrease. Researchers aim to have an R^2 close to the adjusted R^2 value, since this indicates that predictors do not only add explained variance due to chance (Miles, 2014). In this research an adjusted R^2 of ,667 has been found. Since R^2 and the adjusted R^2 are fairly close, it can be concluded that the resulting R^2 is only slightly overestimated. Due to the adjusted R^2 attempting to yield a more honest value and due to the slight overestimation of R^2 , it seems appropriate to make statements using the adjusted R^2 . The adjusted R^2 of ,667 indicates that 66,7% of variance in the dependent variable can be explained by the independent variables. The remaining 33,3% can be explained by other factors which have not been used in this research.

Furthermore, a Durbin-Watson test has been applied in order to determine if the data has problems with autocorrelation in residuals. It may be useful to test for autocorrelation, since autocorrelation may result in underestimates of the standard error which can cause researchers to think predictors are significant while they are not. Ideally a Durbin-Watson test results in a value of 2, which indicates that there is no autocorrelation. However, there is almost always some degree of autocorrelation. Therefore, literature suggests that the Durbin-Watson test result should be between the two critical values of 1.5 and 2.5 (Hutcheson & Sofroniou, 1999). Since the outcome value of 1,849 is between the two critical values, it can be assumed that there is no problem with autocorrelation.

Table 12. Pooled ANOVA

ANOVA ^a					
Model		Sum of Squares	df	Mean Square	F
1	Regression	89,087	6	89,312	45,182
	Residual	40,806	123	,320	
	Total	129,893	129		

a. Dependent Variable: Customer satisfaction

b. Predictors: (Constant), Social environment, Program, Supporting services, Personnel, Outcome, Physical environment

The ANOVA results can be used to see if the independent factors reliably predict the dependent factor. Since the resulting p-value (Sig.) is <0.05 we can conclude that the independent factors reliably predict the dependent factor. Which means that Physical environment, Program, Personnel, Outcome, Supporting services and Social environment can be used to reliably predict Customer satisfaction. It is important to realize that ANOVA indicates whether the independent factors combined together reliably predict the dependent factor. The results from the ANOVA test do not indicate the ability of each individual independent factor to predict customer satisfaction.

Table 13. Pooled Coefficients Multiple Regression Analysis results

	Coefficients ^a						
	Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
	B	Std. Error	Beta	t	Sig.	Tolerance	VIF
(Constant)	,007	,020		,347	,729		
Physical environment	,340	,020	,341	16,906	,000	,999	1,001
Program	,331	,020	,333	16,489	,000	1,000	1,000
Personnel	,374	,020	,372	18,432	,000	1,000	1,000
Outcome	,359	,020	,363	17,975	,000	,999	1,001
Supporting services	,258	,020	,259	12,842	,000	1,000	1,000
Social environment	,361	,020	,362	17,936	,000	1,000	1,000

a. Dependent Variable: Customer satisfaction

Before, the resulting coefficients will be discussed, the collinearity statistics will be discussed in order to determine if there is no multicollinearity issue. Both the Variance Inflation Factor (VIF) and tolerance are measures which are used to identify multicollinearity. In order to measure multicollinearity, VIF measures the impact of collinearity among the used variables in a regression model (Adeboye, Fagoyinbo, & Olatayo, 2014). Researchers desire lower levels of VIF, since values >2,50 start to indicate possible issues with multicollinearity (Adeboye, Fagoyinbo, & Olatayo, 2014). Tolerance is estimated by $1-R^2$, where R^2 is measured by regressing a predictor variable to the remaining predictor variables included in the multiple regression analysis (Adeboye, Fagoyinbo, & Olatayo, 2014). If the resulting tolerance levels are <0,40, then problems with multicollinearity should be considered. Since the highest VIF estimate is 1,009 and since the lowest tolerance level is 0,990, there seems to be no problem with multicollinearity, which has also been indicated by the previously conducted Pearson Correlation Coefficient Analysis.

Table 13 also contains the unstandardized and standardized coefficients for each predictor variable. B value estimates are presented which allows us to get an estimate of the contribution of each predictor factor in the model. However, there is an important difference between unstandardized and standardized coefficients. Unstandardized coefficients are often considered as the default value, they indicate the amount by which a dependent variable changes when an independent variable changes by one unit when the other independent variables are kept constant. Standardized regression coefficients indicate how many standard deviations a dependent variable will change per standard deviation increase in an independent variable when the other independent variables are kept constant. Standardized regression coefficients are often used when variables are measured in different units of measurement (King, 1986). Since researchers are voicing concerns that standardized coefficients can be misleading, and since all variables are measured using the same Likert scale, unstandardized coefficients will be used and further discussed (King, 1986).

As shown in table 13, every regression coefficient is considered to be positive and significant since $P<0,05$. This means that each of the 6 independent factors have a positive and significant effect on customer satisfaction. Based on the unstandardized coefficients (B) we can conclude that Personnel has the highest impact ($B = ,374$) on customer satisfaction. Personnel is followed by Social environment ($B = ,361$), outcome ($B = ,359$), Physical environment ($B = ,340$), program ($B = ,331$) and lastly Supporting services ($B = ,258$).

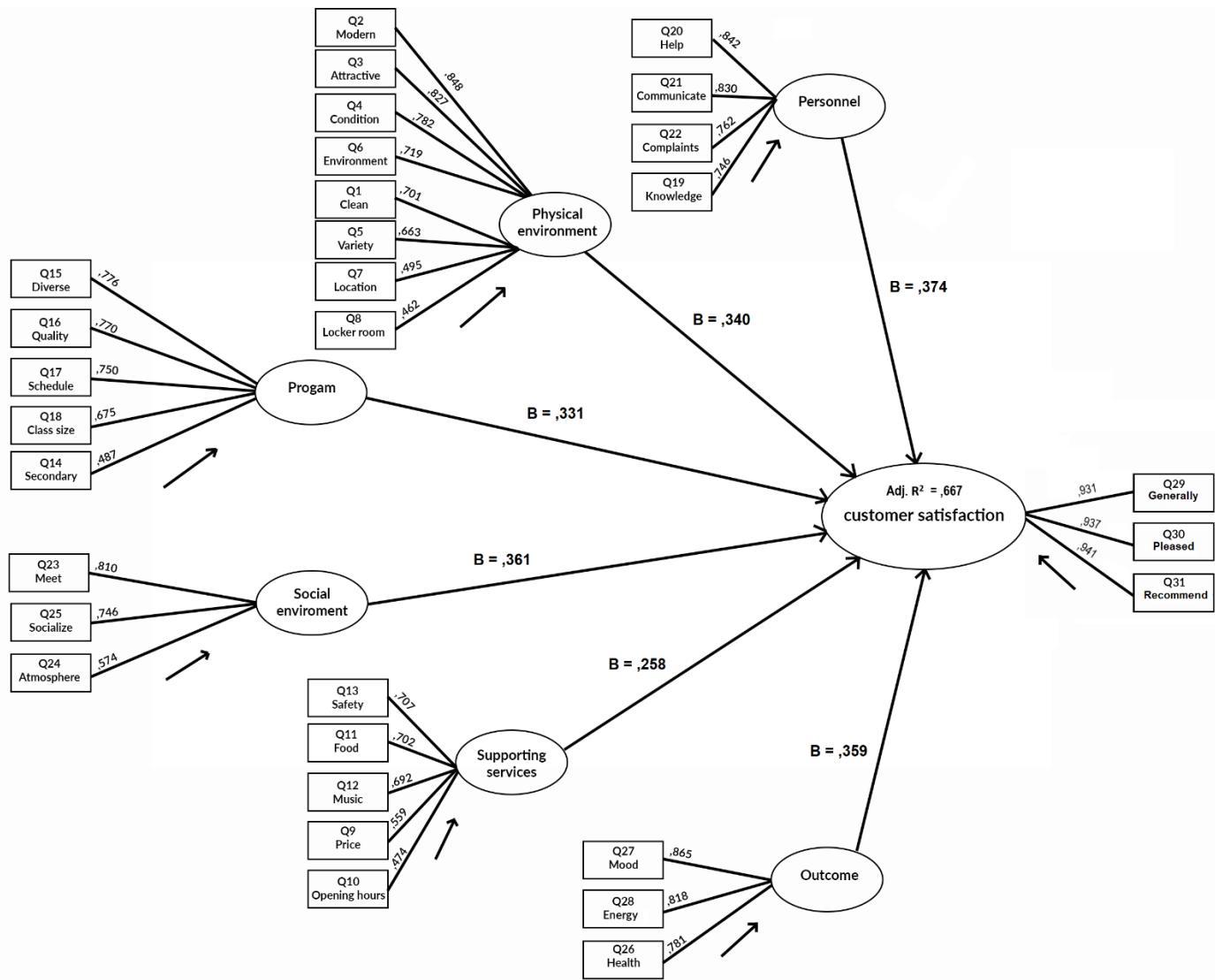
Based on the unstandardized coefficients, the following multiple regression equation was formed: Consumer satisfaction = 0,007 + 0,374 (Personnel) + 0,361 (Social environment) + 0,359 (outcome) + 0,340 (Physical environment) + 0,331 (program) + 0,258 (Supporting services).

Based on the resulting significance and regression coefficients it can be concluded that hypothesis 1, 2, 3, 4, 5 and 6 have been accepted.

Table 13. Hypothesis testing

Hypothesis testing				
	Hypothesis	Result	B	Sig.
H1	Personnel has positive effect on satisfaction	Accepted	,374	,000
H2	Physical environment has positive effect on satisfaction	Accepted	,340	,000
H3	Program has positive effect on satisfaction	Accepted	,331	,000
H4	Social environment has positive effect on satisfaction	Accepted	,361	,000
H5	Supporting services has positive effect on satisfaction	Accepted	,258	,000
H6	Outcome has a positive effect on satisfaction	Accepted	,359	,000

Figure 5. Resulting model using factor scores and unstandardized coefficients



6. Discussion & Conclusion

Currently there is a clear problem with obesity and with the business continuity of Dutch SME fitness centers. Literature indicates that increasing customer satisfaction will improve the business continuity of SME fitness centers which will in turn help to decrease or stabilize the current obesity problem. This research has focused on SME fitness centers since the increasing presence of rivaling low-budget fitness centers has resulted in the business continuity problems. Moreover, these rivaling low-budget fitness centers do not offer much in terms of service and guidance in comparison to many SME's. Therefore, SME's are contributing more towards solving or stabilizing the obesity problem in comparison to the larger low-budget fitness chains. Consequently, the aim of this study was to determine which factors influence customer satisfaction in Dutch SME fitness centers.

In order to find which factors influence customer satisfaction at Dutch SME fitness centers, this study has combined primary and secondary data. Secondary data has been used in order to create a hypothesized model which indicates what factors might influence customer satisfaction. Primary data has been collected in order to test the hypothesized model which was based on opinions and results from relevant studies.

The results of this study indicate that: Personnel, Physical environment, Program, Social environment, Supporting services and Outcome have a significant and positive effect on Customer satisfaction in Dutch SME fitness centers. Based on the multiple regression outcome it can be concluded that Personnel has the highest impact ($B=,374$) on customer satisfaction. Personnel is followed by Social environment ($B=,361$), outcome ($B=,359$), Physical environment ($B=,340$), program ($B=,331$) and lastly Supporting services ($B=,258$). Based on the resulting significance and regression coefficients it can be concluded that hypothesis 1, 2, 3, 4, 5 and 6 have been accepted.

Based on the theoretical framework it was expected that most of the factors would have a significant impact on customer satisfaction. However, it was not expected that every factor would have a significant impact since some factors were only mentioned in a few studies. Furthermore, many people would assume that Physical environment has the largest effect on customer satisfaction, when in reality Personnel, Social environment and Outcome have a larger effect on customer satisfaction. This indicates that fitness center managers should not purely focus on the Physical environment but also pay a decent amount of attention to other factors.

6.1. Theoretical contribution

Current literature provides insights on factors which affect service quality in the fitness industry but does not focus on factors influencing customer satisfaction. This research offers a theoretical contribution since it has focused on customer satisfaction, instead of service quality. Furthermore, authors in this field of research have suggested the development of industry and country specific models since the use of other models may be inappropriate, due to the use of samples which only contain respondents from one country (Lam et al. 2005). Therefore, many relevant models are only representative for fitness centers in specific countries and are not generalizable for the entire fitness industry. Currently there is no research on service quality or satisfaction in the Dutch fitness industry. This research adds to literature by creating a satisfaction model for Dutch SME fitness centers, which may also be applicable for foreign SME fitness centers that operate in a similar fitness industry. Moreover, most relevant models are developed more than ten years ago. However, the fitness industry has changed significantly due to the increasing presence of large budget fitness chains which started in 2008 (Rabobank, 2016). Hence, many of the models may have become outdated. For this study, new data has been collected in order to create a relevant and contemporary model.

6.2. Practical contribution

Besides a theoretical contribution, this research also offers a practical contribution due to the identification of factors which influence the customer satisfaction rate in Dutch SME fitness centers. Due to the use of a factor analysis and multiple regression analysis, the outcome of this research indicates which factors affect customer satisfaction and to what extent. This knowledge may help fitness centers to pay more attention to factors which did not seem as important to them previously. It is likely that when used by fitness centers, this information could result in an increase in customer satisfaction. The increase in customer satisfaction is in turn likely to result in an improvement to the retention rate and binding of customers. These improvements may not only improve the pressured business continuity of many SME fitness centers, but may also aid in the decrease or stabilization of the current obesity problem. This is likely due to the focus on service and guidance by SME's while the increasing amount of low-budget fitness centers focus on price.

6.3. Research limitations

Almost every research has limitations, this study is no exception. Due to the rise of the COVID-19 virus, it is important to acknowledge that data could have been affected. Due to COVID-19, participants might have answered that they don't deem a fitness center to be clean or safe, while they might have considered it to be safe and clean before the COVID-19 outbreak. Furthermore, fitness centers were forced to close by the government due to the COVID-19 outbreak. Therefore, a sample of 140 participants has been used, but a larger sample size could have generated more accurate results. Furthermore, data has been collected from 3 SME fitness centers, this can be considered to be limiting since the entire population is comprised of significantly more SME fitness centers. However, due to time limitations and the COVID-19 outbreak, the possibility to collect data from 3 SME fitness centers can be considered as adequate. Another limitation is the use of convenience sampling, which is considered to be a non-probability sampling technique. Due to the use of this technique, the generalizability of this study is not optimal. However, this sampling method is deemed to be justifiable, since it is not possible to create a list containing every Dutch SME fitness member in order to use probability sampling.

6.4. Suggestions for future research

This study has resulted in a model with an explained variance of 66,7%. The researcher recommends that future research should focus on increasing the amount of explained variance by introducing new variables or factors. These successive studies could, if possible, also use a larger sample size in order to improve the reliability of the model. Furthermore, future studies could also test if this model is applicable in other countries and to which extent.

7. References

- Adeboye, N. O., Fagoyinbo, I. S., & Olatayo, T. O. (2014). Estimation of the effect of multicollinearity on the standard error for regression coefficients. *Journal of Mathematics*, 10(4), 16-20.
- Alexandris, K., Zahariadis, P.N., Tsorbatzoudis, C., & Grouios, G. (2004). An empirical investigation of the relationships among service quality, customer satisfaction and psychological commitment in a health club context. *European Sport Management Quarterly*, 4(1), 36-52.
- Anderson, R.E., & Srinivasan, S.S. (2003). E-satisfaction and E-loyalty: A Contingency Framework. *Psychology and Marketing*, 20(2), 123-138.
- Bala, J. (2016). Contribution of SPSS in Social Sciences Research. *International Journal of Advanced Research in Computer Science*, 7(6), 250-254.
- Boedec, K.L. (2016). Sensitivity and specificity of normality tests and consequences on reference interval accuracy at small sample size: a computer-simulation study. *Veterinary clinical pathology*, 45(4), 648-656 .
- Bodet, G. (2006). Investigating customer satisfaction in a health club context by an application of the tetracalss model. *European Sport Management Quarterly*, 6(2), 149-165.
- Brady, M., & Cronin. J. (2001). Some New Thoughts on Conceptualizing Perceived Service Quality: A Hierarchical Approach. *Journal of Marketing*, 65, 34-49.
- Braeken, J., & van Assen, M. A. L. M. (2017). An empirical Kaiser criterion. *Psychological Methods*, 22(3), 450-466.
- Buttle, F. (1996). SERVQUAL: review, critique, research agenda. *European Journal of Marketing*, 30(1), 8-32.
- Buuren, S. (2018). *Flexible imputation of missing data*. Boca Raton, FL: CRC Press.
- Cardozo, R. (1965). An Experimental Study of Customer Effort, Expectation, and Satisfaction. *Journal of Marketing Research*, 2(3), 244-249.
- CBS (2017). *Fitnesscentra; klanten, abonnementsprijzen en voorzieningen*. Retrieved from: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83822NED/table?fromstatweb>
- Chang, K., & Chelladurai, P. (2003). System-based quality dimensions in fitness services: development of the scale of quality. *The Service Industries Journal*, 23(5), 65-83.
- Chelladurai, P., Scott, F.L., & Haywood-Farmer, J. (1987). Dimensions of fitness services development of a model. *Journal of Sport Management*, 1(2), 159-172.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98–104.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
- Crompton, J.M., MacKay, K.J., & Fesenmaier, D.R. (1991). Identifying dimensions of service quality in public recreation. *Journal of Park and Recreation Administration*, 9(3), 15-28.
- Cronin, J.J., & Taylor, S.A. (1992). Measuring service quality: a re-examination and extension. *Journal of Marketing*, 56(3), 55-68.

- Dabholkar, P.A., Thrope, D.I. & Rentz, J.O. (1996). Measure of service quality for retail stores: scale development and validation. *Journal of the Academy of Marketing Science*, 24(1), 3-16.
- DePoy, E., & Gitlin, L. N. (2016). Introduction to research: Understanding and applying multiple strategies.
- Dhakal, C. P. (2018). Interpreting the Basic Outputs (SPSS) of Multiple Linear Regression. *International Journal of Science and Research*, 8(6), 1448-1452.
- Dimitriades, Z. (2006). Customer satisfaction, loyalty and commitment in service organizations. *Management Research News*, 29(12), 782-800.
- Dziuban, C.D., & Shirkey, E.C. (1974). When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological bulletin*, 81(6), 358.
- Faullant, R., Matzler, K., & Füller, J. (2008). The impact of satisfaction and image on loyalty: The case of Alpine ski resorts. *Managing Service Quality*, 18(2), 163-178.
- Fornell, C. (1992). A National Customer Satisfaction Barometer: The Swedish Experience. *Journal of Marketing*, 56(1), 6-21.
- Gonçalves, C., Meireles, P.E., & Carvalho, M.J. (2016). Consumer Behaviour in Fitness Club: Study of the Weekly Frequency of Use, Expectations, Satisfaction and Retention. *The Open Sports Sciences Journal*, 9, 1-9.
- Grönroos, C. (1984), A service quality model and its marketing implications. *European Journal of Marketing*, 18(4), 36-44.
- Gupta, A.K., & Basumatary, S.J. (2017). The role of customer's satisfaction in the fitness industry development. *International Journal of Advanced Research and Development*, 2(5), 121-124.
- Hair, J.F., Black, W.C., Babin, B.J., & Anderson, R.E. (2014). Multivariate Data Analysis (7th ed.). England, Pearson Education.
- Hair, J.F., Bush, R.P., & Ortinau, D.J. (2006). Marketing research: within a changing information environment (3rd ed.). New York, McGraw-Hill/Irwin.
- Hall, K. D., & Kahan, S. (2018). Maintenance of Lost Weight and Long-Term Management of Obesity. *The Medical clinics of North America*, 102(1), 183–197.
- Hallowell, R. (1996). The relationships of customer satisfaction, customer loyalty, and profitability: an empirical study. *International Journal of Service Industry Management*, 7(4), 27-42.
- Hanspal, S., & Devasagayam, P. R. (2017). Impact of Consumers' Self-Image and Demographics on Preference for Healthy Labeled Foods, *SAGE Open*, 7(1), 1-18.
- Hecker, K., & Violato, C. (2009). Validity, reliability, and defensibility of assessments in veterinary education. *Journal of veterinary medical education*, 36(3), 271-5.
- Helgesen, Ø. (2006). Are Loyal Customers Profitable? Customer Satisfaction, Customer (Action) Loyalty and Customer Profitability at the Individual Level. *Journal of Marketing Management*, 22, 245-266.
- Hill, N., & Brierley, J. (2017). *How to measure customer satisfaction*. Routledge.

Hoang Trong, Chu Nguyen Mong Ngoc. (2008). Analyzing data with SPSS research, Hong Duc publisher, Ho Chi Minh city.

Howat, G., Murray, D., & Crilley, G. (1999). The relationships between service problems and perceptions of service quality, satisfaction, and behavioral intentions of Australian public sports and leisure center customers. *Journal of Park and Recreation Administration*, 17(2), 42-64.

Huiqun, L & Xin, Z. (2019). A Literature Review and Critique on Customer Satisfaction.

Hutcheson, G. D., & Sofroniou, N. (1999). The multivariate social scientist: Introductory statistics using generalized linear models. Sage.

Jakobsen, J. C., Gluud, C., Wetterslev, J., & Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. *BMC medical research methodology*, 17(1), 162.

Jiradilok, T., Malisuwant, S., Madan, N., & Sivaraks, J. (2014). The Impact of Customer Satisfaction on Online Purchasing: A Case Study Analysis in Thailand. *Journal of Economics, Business and Management*, 2(1), 5-11.

Johnson, D.R., & Creech, J.C. (1983). Ordinal measures in multiple indicator models: A simulation study of categorization error. *American Sociological Review*, 48, 398-407.

Ko, Y.J., & Pastore, D.L. (2004). Current issues and conceptualizations of service quality in the recreational sport industry. *Sport Marketing Quarterly*, 13(3), 159-167.

Kotler, P., Cunningham, M.H., & Turner, R.E. (2001) Marketing Management, Pearson Education, Canada.

Kotler, P., & Keller, K.L. (2012). Marketing Management (15th ed.), Prentice Hall, New Jersey.

Kotler, P., & Keller K.L. (2006). Marketing Management, (12th Ed.), Pearson Education Inc, New Jersey.

Kim, D., & Kim, S.Y. (1995). QUESC: an instrument for assessing the service quality of sport centers in Korea. *Journal of Sport Management*, 9(2), 208-220.

King, G. G. (1986). How not to lie with statistics: Avoiding common mistakes in quantitative political science. *American Journal of Political Science*, 30(3), 666-687

KSB (2018). *Fitnessbranche in Nederland /2018*. Retrieved from:
<https://www.kennisbanksportenbewegen.nl/?file=9304&m=1546971033&action=file.download>

Lam, E.T.C., Zhang, J.J., & Jensen, B.E. (2005). Service quality assessment scale (SQAS): an instrument for evaluating service quality of health-fitness clubs. *Measurement in Physical Education and Exercise Science*, 9(2), 79-111.

Martinez, J.A.G., & Martinez, L.C. (2010). Rethinking perceived service quality: an alternative to hierarchical and multidimensional models. *Total Quality Management & Business Excellence*, 21(1), 93-118.

McCarthy, J. (2007). IHRSA'S Guide to membership retention (2.nd Ed.). Boston: IHSA.

McDonald, J. H. (2014). Handbook of Biological Statistics. Maryland: Sparky House Publishing

- McQuitty, S., Finn, A., & Wiley, J. B. (2000). Systematically Varying Customer Satisfaction and its Implications for Product Choice. *Academy of Marketing Science Review, 2000*, 1-16.
- Miles, J. (2014). R squared, adjusted R squared. Wiley StatsRef: Statistics Reference Online.
- Morales, V. S., & Gálvez, P. R. (2011). La percepción del usuario en la evaluación de la calidad de los servicios municipales deportivos. *Cuadernos de Psicología Del Deporte, 11*(2), 147-154.
- NOCNSF (2017). Onderzoeksrapport: *Zo sport Nederland*. Retrieved from: <https://www.nocnsf.nl/ledentallen>
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education, 15*(5), 625-632.
- NCSS Statistical Software (2019). *Principal Components Regression*. Retrieved from: https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Principal_Components_Regression.pdf
- Olorunniwo, F., Hsu, M., & Udo, G. (2006). Service quality, customer satisfaction, and behavioral intentions in the service factory. *Journal of Services Marketing, 20*(1), 59-72.
- Oliver, R.L. (1997). Satisfaction: A Behavioral Perspective on the Consumer, McGraw-Hill, New York.
- Oliver, R.L. (2015). Satisfaction: A Behavioral Perspective on the Consumer, NY: Routledge, New York.
- Osborne, J. W., & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research, and Evaluation, 8*(1), 2.
- Painter, S. L., Ahmed, R., Kushner, R. F., Hill, J. O., Lindquist, R., Brunning, S., & Margulies, A. (2018). Expert Coaching in Weight Loss: Retrospective Analysis. *Journal of medical Internet research, 20*(3), 92.
- Pallant, J. (2010). SPSS survival manual : a step by step guide to data analysis using SPSS, Maidenhead, McGraw-Hill.
- Papadimitriou, D.A., & Karteroliotis, K. (2000). The service quality expectations in private sport and fitness centers: a reexamination of the factor structure. *Sport Marketing Quarterly, 9*(3), 157-164.
- Parasuraman, A., Zeitmal, V.A., & Berry, L.L. (1988). SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing, 64*(1), 12-40.
- Punniyamoorthy, M., & Mohanraj, P. (2007). An empirical model for brand loyalty measurement. *Journal of Targeting, Measurement and Analysis for Marketing, 15*, 222-233.
- Rabobank (2008). *Rabobank Cijfers & Trends*. Retrieved from: https://www.efaa.nl/files/downloads/rabo_ondern_update_mei.pdf
- Rabobank (2016). *Rabobank Cijfers & Trends*. Retrieved from: <https://www.hiddenprofitsmarketing.com/uploads/files/201610%20Fitness%20cijfers%20en%20trends%20Rabobank.pdf>
- Rabobank (2019). *Rabobank Cijfers & Trends*. Retrieved from: <https://www.rabobank.nl/bedrijven/cijfers-en-trends/horeca-en-recreatie/fitnesscentra/>

- Reibstein, D. J. (2002). What Attracts Customers to Online Stores, and What Keeps Them Coming Back?. *Journal of the Academy of Marketing Science*, 30(4), 465–473.
- Reichheld, F.F., & Sasser, W.E. (1990). Zero defections: Quality comes to services. *Harvard Business Review*, 68(5), 105–111.
- Rial, J., Varela, J., Rial, A. & Real, E. (2010) Modeling and measuring perceived quality in sports centres: QSport-10 scale, *International Journal of Sport Science* 18(6), 57-73.
- Robinson, L. (2006). Costumer expectations of sports organizations. *European Sport Management Quarterly*, 6(1), 67–84
- Rust, R., & Oliver, R. (1994). Service quality: insights and managerial implications from the frontier. In Rust, R.T., & Oliver, R.L., (Eds.), *Service quality: New directions in theory and practice* (pp. 1-20). SAGE Publications, Thousand Oaks.
- Schafer, J. L., & Olsen, M.k. (1998). Multiple Imputation for Multivariate Missing-Data Problems: A Data Analyst's Perspective. *Multivariate Behavioral Research* 33(4), 545–71.
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, 8(4), 350–353.
- Sireci, S.G. (1998). The construct of content validity. *Social indicators research*, 45(3), 83-117.
- Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M., & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ (Clinical research ed.)*, 338, b2393.
- Sullivan, G. & Artino Jr., A. R. (2013). Analyzing and Interpreting Data From Likert-Type Scales. *Journal of Graduate Medical Education*. 5(4), 541-542.
- Sumadi, E., & Soliha, E. (2015). The Effect of Bank Image and Trust on Loyalty Mediated by Customer Satisfaction. *Jurnal Dinamika Manajemen*, 6(2), 121-32.
- Suwono, L., & Oktaria, S. (2016). Factors Affecting Customer Loyalty of Fitness Centers: An Empirical Study. *Jurnal Dinamika Manajemen*, 7(1), 45-55.
- Tabachnick, B. G., & Fidell, L. S. (2007). Using multivariate statistics (5th ed.). Allyn & Bacon/Pearson Education.
- Talley, M. (2008). Customer retention: A manager's perspective. In Oakley, B. & Rhys, M. (Eds), The sport and fitness sector: An introduction. London: Routledge.
- Tobias, S., & Carlson, J. E. (1969). Brief report: Bartlett's test of sphericity and chance findings in factor analysis. *Multivariate Behavioral Research*, 4(3), 375-377.
- Tufis, C.D. (2008). Multiple Imputation as a Solution to the Missing Data Problem in the Social Sciences. *Calitatea Vietii*, 19, 199-212.
- Tustin, D. H., Ligthelm. A. A., Martins, J. H., & Van Wyk, H. J. (2010). Marketing research in practice. South Africa: Unisa Press.
- Uçan, Y. (2007). Developing the perceived service quality scale of sport-fitness centers. Doctorate Thesis. Abant İzzet Baysal University, Social Sciences Institute, Bolu, Turkey.
- Vieira, E. R. M., & Ferreira, J. J. (2017). Strategic framework of fitness clubs based on quality dimensions: the blue ocean strategy approach.

- Volksgezondheidenzorg (2017). *Trend overgewicht volwassenen*. Retrieved from:
<https://www.volksgezondheidenzorg.info/onderwerp/overgewicht/cijfers-context/trends#node-trend-overgewicht-volwassenen>
- Yildiz, S.M. (2011), An importance-performance analysis of fitness center service quality: empirical results from fitness centers in Turkey. *African Journal of Business Management*, 5(16), 7031-7041.
- Yildiz, S. M., & Kara, A. (2012). A re-examination and extension of measuring perceived service quality in Physical Activity and Sports Centres (PSC): QSport-14 scale. *International Journal of Sports Marketing and Sponsorship*, 13(3), 190-208.
- Yildiz, K., Polat, E., & Güzel, P. (2018). A Study Investigating the Perceived Service Quality Levels of Sport Center Members: A Kano Model Perspective. *Journal of Education and Training Studies*, 6(4), 177-188.
- Yong, A.G., & Pearce, S. (2013). A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. *Tutorials in Quantitative Methods for Psychology*, 9, 79-94.

8. Appendices

8.1 Questionnaire in English

Questionnaire customer satisfaction

KEI-FIT

What is this questionnaire?

This research is about factors influencing customer satisfaction in fitness centers. This questionnaire consists of 37 multiple choice questions divided in 2 sections. Participation in this study takes approximately 5-7 minutes.

Do I benefit from this research?

Participation will result in a model which explains customer satisfaction in Dutch fitness centers. This will allow fitness centers to pay more attention to factors which influence customer satisfaction.

Data

The research is intended for academic purposes only. Participation is anonymously and participants are not required to provide their name. The information you provide will also be stored and processed anonymously. Furthermore, this research has been approved by the ethics committee of the University of Twente.

Withdraw

Participants may withdraw from the research at any time without explanation/justification

Contact

If you have questions regarding this survey you can contact: marioeklasens@hotmail.com

Permission

I agree to take part in this research as specified above.

Yes

No

Section 1:

In this section general questions are asked regarding demographics and participation

- | | | |
|---|--|------------------------------------|
| 1. What is your gender | <input type="radio"/> Male | <input type="radio"/> Female |
| | <input type="radio"/> Other | |
| 2. What is your age | <input type="radio"/> < 20 | <input type="radio"/> 20 - 30 |
| | <input type="radio"/> 31 - 40 | <input type="radio"/> 41 - 50 |
| | <input type="radio"/> 51 - 60 | <input type="radio"/> > 60 |
| 3. What is your highest completed educational level | <input type="radio"/> Elementary school | <input type="radio"/> High School |
| | <input type="radio"/> Vocational education | <input type="radio"/> Bachelor |
| | <input type="radio"/> Master | |
| 4. At which Kei-Fit fitness center do you mainly train? | <input type="radio"/> Haaksbergen | <input type="radio"/> Gendringen |
| | <input type="radio"/> Lichtenvoorde | <input type="radio"/> Doetinchem |
| 5. Which activity program do you practice (multiple answers can be filled in) | <input type="checkbox"/> Fitness | <input type="checkbox"/> Tennis |
| | <input type="checkbox"/> Squash | <input type="checkbox"/> Other |
| 6. How long have you been a member | <input type="radio"/> 0 - 1 year | <input type="radio"/> 1 - 2 years |
| | <input type="radio"/> 2 - 5 years | <input type="radio"/> 5 - 10 years |
| | <input type="radio"/> > 10 years | |
| 7. On average, how many times a week do you visit the fitness center | <input type="radio"/> < 1 time | <input type="radio"/> 1 time |
| | <input type="radio"/> 2 times | <input type="radio"/> 3 times |
| | <input type="radio"/> 4 times | <input type="radio"/> > 4 times |

Section 2:

In this section 30 statements about Kei-Fit are given. You can answer by selecting 1,2,3,4 or 5 on every statement.
 [1 = Strongly disagree] [2 = Disagree] [3 = Neutral] [4 = Agree] [5 = Strongly agree] [X = Not available]

STRONGLY DISAGREE	DISAGREE	NEUTRAL	AGREE	STRONGLY AGREE	N/A
-------------------	----------	---------	-------	----------------	-----

AT THIS FITNESS CENTER:

1	The facilities are clean	1	2	3	4	5	X
2	The facilities are modern	1	2	3	4	5	X
3	The facilities are attractive	1	2	3	4	5	X
4	The equipment is in good condition	1	2	3	4	5	X
5	There is a good variety of equipment	1	2	3	4	5	X
6	The environment (temperature, light) is pleasant	1	2	3	4	5	X
7	The location is pleasant and sensible	1	2	3	4	5	X
8	There is a decent locker room and shower	1	2	3	4	5	X
9	The membership fee is reasonable	1	2	3	4	5	X
10	The opening hours are long enough	1	2	3	4	5	X
11	The range of food and drinks is good	1	2	3	4	5	X
12	Appropriate background music is played	1	2	3	4	5	X
13	Enough attention is paid to the safety of visitors	1	2	3	4	5	X
14	Enough secondary services, like childcare, suntanning & physiotherapy are offered	1	2	3	4	5	X
15	The program is diverse enough	1	2	3	4	5	X
16	The quality / content of the programs is good	1	2	3	4	5	X
17	There is a convenient program time/schedule	1	2	3	4	5	X
18	The class size of the programs is appropriate	1	2	3	4	5	X
19	The employees possess the required knowledge/skills	1	2	3	4	5	X
20	The employees are willing to help	1	2	3	4	5	X
21	The employees communicate in a clear and appropriate way with members	1	2	3	4	5	X
22	The employees are responsive to suggestions and complaints	1	2	3	4	5	X
23	I have the chance to meet new people	1	2	3	4	5	X
24	There is a nice and friendly atmosphere	1	2	3	4	5	X
25	Exercising gives me the chance to socialize	1	2	3	4	5	X
26	Visiting can help me to improve my health	1	2	3	4	5	X
27	Visiting can help me to improve my mood	1	2	3	4	5	X
28	Visiting can help me to increase my energy	1	2	3	4	5	X
29	I am generally satisfied	1	2	3	4	5	X
30	I am pleased to have taken the decision to become a member	1	2	3	4	5	X
31	I would recommend this fitness center to a friend	1	2	3	4	5	X

8.2 Questionnaire in Dutch

Enquête klanttevredenheid

Waar gaat dit onderzoek / deze enquête over?

Dit onderzoek gaat over factoren die klanttevredenheid beïnvloeden bij fitnesscentra. Deze enquête bestaat uit 37 multiple choice vragen verdeeld in 2 stukken. Deelname aan dit onderzoek duurt circa 5-7 minuten.

Wat is het nut van deze enquête?

Deelname zal resulteren in een model dat de klanttevredenheid in Nederlandse fitnesscentra verklaart. Hierdoor kunnen fitnesscentra meer aandacht besteden aan factoren die de klanttevredenheid beïnvloeden.

Data

Het onderzoek is bedoeld voor academische doeleinden. Deelname is anoniem en u hoeft uw naam niet op te geven. De informatie die u verstrekt, wordt anoniem opgeslagen en verwerkt. Dit onderzoek is tevens goedgekeurd door de ethische commissie van de Universiteit Twente.

Deelname

Deelnemers kunnen zich op elk moment terugtrekken uit het onderzoek zonder uitleg te geven.

Contact

Als u vragen hebt over deze enquête, kunt u contact opnemen met: marioeklasens@hotmail.com

Toestemming

Ik ga ermee akkoord om deel te nemen aan het onderzoek van de Universiteit Twente zoals hierboven aangegeven.

Ja

Nee

Sectie 1:

In dit gedeelte worden algemene vragen gesteld over o.a. demografie

- | | | |
|--|--------------------------------------|--|
| 1. Wat is uw geslacht | <input type="radio"/> Man | <input type="radio"/> Vrouw |
| | <input type="radio"/> Anders | |
| 2. Wat is uw leeftijdsgroep | <input type="radio"/> < 20 | <input type="radio"/> 20 - 30 |
| | <input type="radio"/> 31 - 40 | <input type="radio"/> 41 - 50 |
| | <input type="radio"/> 51 - 60 | <input type="radio"/> > 60 |
| 3. Wat is uw hoogst voltooide opleidingsniveau | <input type="radio"/> Basisonderwijs | <input type="radio"/> Middelbaar onderwijs |
| | <input type="radio"/> MBO | <input type="radio"/> Bachelor (HBO / WO) |
| | <input type="radio"/> Master | |
| 4. Bij welk Kei-Fit fitnesscentrum traint u voornamelijk | <input type="radio"/> Haaksbergen | <input type="radio"/> Gendringen |
| | <input type="radio"/> Lichtenvoorde | <input type="radio"/> Doetinchem |
| 5. Van welke faciliteiten maak je voornamelijk gebruik (meerdere antwoorden zijn mogelijk) | <input type="checkbox"/> Fitness | <input type="checkbox"/> Tennis |
| | <input type="checkbox"/> Squash | <input type="checkbox"/> Anders |
| 6. Hoe lang bent u al lid | <input type="radio"/> 0 - 1 jaar | <input type="radio"/> 1 - 2 jaar |
| | <input type="radio"/> 2 - 5 jaar | <input type="radio"/> 5 - 10 jaar |
| | <input type="radio"/> >10 jaar | |
| 7. Hoe vaak per week bezoekt u gemiddeld het fitnesscentrum | <input type="radio"/> < 1 keer | <input type="radio"/> 1 keer |
| | <input type="radio"/> 2 keer | <input type="radio"/> 3 keer |
| | <input type="radio"/> 4 keer | <input type="radio"/> > 4 keer |

Sectie 2:

In dit gedeelte worden stellingen over Kei-Fit voorgelegd. Je kunt antwoorden door 1,2,3,4 of 5 te omcirkelen bij elke statement.

[1 = Sterk mee oneens] [2 = Oneens] [3 = Neutraal] [4 = Eens] [5 = Sterk mee eens] [X = Niet van toepassing]

STERK MEE ONEENS	ONEENS	NEUTRAAL	EENS	STERK MEE EENS	N.V.T.
------------------	--------	----------	------	----------------	--------

DIT FITNESS CENTRUM:

1	Heeft schone faciliteiten	1	2	3	4	5	X
2	Heeft moderne faciliteiten	1	2	3	4	5	X
3	Heeft aantrekkelijke / prettig ogende faciliteiten	1	2	3	4	5	X
4	Heeft faciliteiten en apparatuur die in goede staat zijn	1	2	3	4	5	X
5	Heeft een goede variëteit / variatie aan apparatuur	1	2	3	4	5	X
6	Heeft een prettige omgeving (temperatuur, licht)	1	2	3	4	5	X
7	Heeft een prettige en handige locatie	1	2	3	4	5	X
8	Heeft een fatsoenlijke kleedkamer en douche	1	2	3	4	5	X
9	Heeft redelijke lidmaatschapskosten	1	2	3	4	5	X
10	Heeft openingstijden die lang genoeg zijn	1	2	3	4	5	X
11	Heeft een goed aanbod aan eten en drinken	1	2	3	4	5	X
12	Heeft passende achtergrondmuziek	1	2	3	4	5	X
13	Besteed voldoende aandacht aan de veiligheid van bezoekers	1	2	3	4	5	X
14	Heeft voldoende secundaire diensten zoals: kinderopvang, zonnebank en fysiotherapie	1	2	3	4	5	X
15	Heeft een programma / totaal aanbod wat divers genoeg is	1	2	3	4	5	X
16	Heeft een programma / totaal aanbod waarvan de kwaliteit en inhoud goed is	1	2	3	4	5	X
17	Heeft een handige en prettige programmatijd / planning voor de aangeboden diensten	1	2	3	4	5	X
18	Heeft geschikte ledenaantallen bij de verschillende activiteiten	1	2	3	4	5	X
19	Heeft medewerkers die beschikken over de vereiste kennis / vaardigheden	1	2	3	4	5	X
20	Heeft medewerkers die bereid zijn om te helpen	1	2	3	4	5	X
21	Heeft medewerkers die communiceren op een duidelijke en gepaste manier	1	2	3	4	5	X
22	Heeft medewerkers die reageren op suggesties en klachten	1	2	3	4	5	X
23	Geeft mij de mogelijkheid om nieuwe mensen te ontmoeten	1	2	3	4	5	X
24	Heeft een prettige en vriendelijke sfeer	1	2	3	4	5	X
25	Biedt mij de mogelijkheid om te socialiseren	1	2	3	4	5	X
26	Bezoeken kan mij helpen mijn gezondheid te verbeteren	1	2	3	4	5	X
27	Bezoeken kan mij helpen mijn stemming te verbeteren	1	2	3	4	5	X
28	Bezoeken kan mij helpen mijn energie te verbeteren	1	2	3	4	5	X
29	Ben ik in het algemeen tevreden over	1	2	3	4	5	X
30	Zou ik aanbevelen aan een vriend	1	2	3	4	5	X
31	Vind ik dusdanig goed, dat ik blij ben dat ik lid ben geworden	1	2	3	4	5	X

8.3 Questionnaire measurement / question sources

1	The facilities are clean	Papadimitriou & Karteroliotis, 2000; Chang & Chelladurai, 2003; Alexandris et al., 2004; Lam et al., 2005; Yildiz, 2011; Yildiz & Kara, 2012; Vieira & Ferreira 2017;
2	The facilities are modern	Papadimitriou & Karteroliotis, 2000; Alexandris et al., 2004; Lam et al., 2005; Yildiz, 2011; Vieira & Ferreira 2017;
3	The facilities are attractive	Chang & Chelladurai, 2003; Alexandris et al., 2004;
4	The equipment is in good condition	Chang & Chelladurai, 2003; Alexandris et al., 2004; Lam et al., 2005;
5	There is a good variety of equipment	Lam et al., 2005; Yildiz, 2011; Vieira & Ferreira 2017
6	The environment (temperature, light) is pleasant	Papadimitriou & Karteroliotis, 2000; Lam et al., 2005; Yildiz, 2011; Vieira & Ferreira 2017
7	The location is pleasant and sensible	Lam et al., 2005; Yildiz, 2011
8	There is a decent locker room and shower	Chang & Chelladurai, 2003; Lam et al., 2005; Yildiz, 2011; Vieira & Ferreira 2017
9	The membership fee is reasonable	Papadimitriou & Karteroliotis, 2000; Yildiz, 2011;
10	The opening hours are long enough	Papadimitriou & Karteroliotis, 2000, Lam et al., 2005;
11	The range of food and drinks is good	Yildiz, 2011; Vieira & Ferreira 2017
12	Appropriate background music is played	Lam et al., 2005; Yildiz, 2011
13	Enough attention is paid to the safety of visitors	Papadimitriou & Karteroliotis, 2000; Lam et al., 2005; Yildiz, 2011; Vieira & Ferreira 2017
14	Enough secondary services, like childcare, suntanning & physiotherapy are offered	Papadimitriou & Karteroliotis, 2000; Yildiz, 2011; Vieira & Ferreira 2017
15	The program is diverse enough	Papadimitriou & Karteroliotis, 2000; Chang & Chelladurai, 2003; Lam et al., 2005; Yildiz, 2011; Yildiz & Kara, 2012; Vieira & Ferreira 2017
16	The quality / content of the programs is good	Lam et al., 2005; Yildiz, 2011; Yildiz & Kara, 2012
17	There is a convenient program time/schedule	Papadimitriou & Karteroliotis, 2000; Chang & Chelladurai, 2003; Lam et al., 2005; Yildiz, 2011; Yildiz & Kara, 2012;
18	The class size of the programs is appropriate	Lam et al., 2005; Yildiz, 2011; Yildiz & Kara, 2012; Vieira & Ferreira 2017
19	The employees possess the required knowledge/skills	Papadimitriou & Karteroliotis, 2000; Alexandris et al., 2004; Lam et al., 2005; Yildiz, 2011; Vieira & Ferreira 2017
20	The employees are willing to help	Papadimitriou & Karteroliotis, 2000; Chang & Chelladurai, 2003; Lam et al., 2005; Vieira & Ferreira 2017
21	The employees communicate in a clear and appropriate way with members	Alexandris et al., 2004; Lam et al., 2005; Yildiz, 2011;
22	The employees are responsive to suggestions and complaints	Papadimitriou & Karteroliotis, 2000; Chang & Chelladurai, 2003; Alexandris et al., 2004; Lam et al., 2005; Yildiz, 2011; Vieira & Ferreira 2017
23	I have the chance to meet new people	Alexandris et al., 2004;
24	There is a nice and friendly atmosphere	Alexandris et al., 2004; Lam et al., 2005;

25	Exercising gives me the chance to socialize	Papadimitriou & Karteroliotis, 2000
26	Visiting can help me to improve my health	Alexandris et al., 2004
27	Visiting can help me to improve my mood	Alexandris et al., 2004
28	Visiting can help me to increase my energy	Alexandris et al., 2004
29	I am generally satisfied	Hill & Brierley, 2017
30	I am pleased to have taken the decision to become a member	Hill & Brierley, 2017
31	I would recommend this fitness center to a friend	Hill & Brierley, 2017

8.4 Pearson Correlation Coefficient Analysis

Due to size/display issues (31X31 variables), only the first 7 variables have been included as an indication.

	(Pooled) Correlations						
	1	2	3	4	5	6	7
1. The facilities are clean	-						
2. The facilities are modern	,597**	-					
3. The facilities look nice	,540**	,813**	-				
4. The equipment is in good condition	,586**	,735**	,640**	-			
5. There is a good variety of equipment	,403**	,594**	,562**	,550**	-		
6. The environment (temperature, light) is pleasant	,467**	,562**	,584**	,553**	,476**	-	
7. The location is pleasant and sensible	,314**	,501**	,498**	,334**	,479**	,444**	-

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

8.5 Factor analysis with 7 factors

Rotated Component Matrix ^a									
	1	2	3	4	5	6	7	Eigenvalue	Cronbach's alpha
Physical environment	The facilities are modern	,868							
	The facilities look nice	,852							
	The equipment is in good condition	,736							
	The environment (temperature, light) is pleasant	,713						8,375	,881
	The facilities are clean	,695							
	There is a good variety of equipment	,625							
	The location is pleasant and sensible	,561							
	There is a decent locker room and shower	,532							
Personnel	The employees are willing to help	,848							
	The employees communicate in a clear and appropriate way	,847							
	The employees possess the required knowledge/skills	,770						3,120	,800
	The employees are responsive to suggestions and complaints	,745							
Program	The program is diverse enough	,785							
	The quality / content of the programs is good	,779							
	There is a convenient program time/schedule	,746						1,942	,872
	The class size of the programs is appropriate	,677							
	Enough secondary services, like childcare, suntanning & physiotherapy are offered	,510							
Outcome	Visiting can help me to improve my mood	,863							
	Visiting can help me to increase my energy	,834						1,621	,880
	Visiting can help me to improve my health	,800							
5th variable	The opening hours are long enough	,672							
	The range of food and drinks is good	,667						1,456	,635
	The membership fee is reasonable	,659							
Social environment	I have the chance to meet new people	,872							
	Exercising gives me the chance to socialize	,789						1,373	,748
	There is a nice and friendly atmosphere	,480							
7th variable	Appropriate background music is played	,720							
	Enough attention is paid to the safety of visitors	,632						1,026	,615

Footnote: 5 point scale (1= strongly disagree to 5= strongly agree). Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 7 iterations.