Erasmus
School of
Economics

Master thesis Economics and Business

Specialization: Marketing

How to strengthen the loyalty of soccer fans during and after the pandemic

Bram ten Barge: 476420

Supervisor: S. L. Malek

Date of final version: 13-08-2021

Abstract

The COVID-19 pandemic has caused the most significant disruption to professional football on a global scale since World War II. Especially, the restrictions on gathering prohibit football clubs to allow fans in their stadiums and organize physical fans events, which are vital sources of revenue for football clubs.

This research aims to determine what the effect of these restrictions is on fan loyalty for football fans. Furthermore, numerous other loyalty-influencing factors were investigated to design a blueprint for sport marketers to find out which factors need to be enhanced and which ones mitigated, valid for during and after the pandemic. In this context, the fan loyalty concept is defined as a double-dimension concept of attitudinal and behavioral loyalty.

To test all the hypotheses of this research, an online survey was distributed to football fans where the majority was Dutch. The survey questionnaire consisted of statements on all the loyalty-influencing factors and both dimensions of fan loyalty to capture their latent attitudes.

The latent constructs were validated by confirmatory factor analysis. Afterwards, ordinary logistic regression was applied on the data to investigate the hypothesized relationships. The results showed a positive, significant effect of fan identification and team-related content (on social media) for attitudinal loyalty, whereas two reasons of initially becoming a fan, team's success and friends' influence had a negative, significant effect.

As for the behavioral dimension of loyalty, attitudinal loyalty, fan identification, social media, lack of fan interaction and parental influence as a reason for initially becoming a fan showed a positive, significant effect. Nostalgic content showed a negative, significant effect.

These results suggest, in contrary to the expectations, that prohibiting stadium attendance and offline fan gatherings did not negatively impact both dimensions of fan loyalty. Furthermore, the designed blueprint advices sport marketers to focus on prioritizing fan identification and social media, while avoiding nostalgic content, in their marketing strategy. As for the long-term strategy, the focus should be on gaining fans through parental influence, possibly in the form of parent-children days and reduced ticket prices for children of all ages.

Table of Contents

1. Introduction

In the beginning of 2021 the Dutch newspaper AD published an article with an alarming title for Dutch football clubs: fan loyalty is finite (Abenhuijs, 2021). Fans and their loyalty are key drivers for the continuity, the future and the revenues of every football club and should never be taken lightly. Fans motivate and stimulate the club and their players to perform to the maximum of their abilities and to win prices (Psychology Educator, 2013).

Furthermore, they are also a very significant value asset. Fans provide direct income for clubs through ticket sales and club merchandise. So, they are directly responsible for additional club incomes (Young, 2021). The size and composition of the fan base provide the basis for the attractiveness for club sponsoring and the sponsor money they can obtain (Sikorski, 2016). This also applies to the broadcasting fees football clubs can acquire (Gazapo, 2020).

Fan loyalty is something which can be taken for granted way too easily. However, especially during this pandemic, this shouldn't be the case. Currently, there are plenty of warning signs that demonstrate the finiteness of fan loyalty. Fans experience less joy when they watch football matches right now and almost half of the fans feels less connected to their favorite club (Abenhuijs, 2021).

The benefits of fan loyalty shouldn't be underestimated and are right now more important than ever. A decrease of fan loyalty will have a significant impact on football clubs. Namely, fan loyalty drives revenue: loyal fans have a significantly higher likeliness to buy and their lifetime value is three times as high as fans without an emotional relationship (Young, 2021).

Moreover, nowadays nearly 50% of fans in the age category 16-24 support a second team (Young, 2021). Only attracting fans isn't sufficient anymore, putting marketing resources into your fan base to achieve a high degree of fan loyalty is a must. The competition has increased and the battle for fans is harder than ever.

The pandemic created a unique situation in the football industry which hasn't occurred since World War II: professional football matches were canceled. Currently, regulations are still valid preventing football clubs allowing to use their full capacity in stadiums. So, football clubs cannot reach their fans as easily as before. These factors are causing a real danger for football clubs of losing the true bond with their fans, which will harm the respective clubs in the maintenance and development of fan loyalty. To prevent this decrease in fan loyalty and therefore revenue, the following research question has been drafted:

'How can professional football clubs strengthen their current degree of fan loyalty during and after the pandemic?'

This paper does not only investigate the influence of six factors on fan loyalty, but also strives to gain a more thorough understanding of the fan loyalty concept by examining its dimensions and their respective relationship. The following sub questions will help in answering the main research question:

1. What is the influence of fan identification on fan loyalty?
2. Does a favorable attitude towards club-related social media influence fan loyalty?
3. Could the addition of an esports department affect fan loyalty?
4. Does the reason of initially becoming a fan matter for the degree of fan loyalty?
5. Did 'ghost games' affect fan loyalty?
6. Did the lack of fan interaction affect fan loyalty?
7. What is the relationship between the two dimensions of fan loyalty?
8. Does the relationship length of a fan with their club moderate the relationship between the two dimensions of fan loyalty?

1.2 Managerial relevance

This research is highly relevant for sport marketers, who are specialized in professional football. The pandemic turned the football world upside down and created a unique situation, of which the consequences are unknown. This paper strives to provide insights into the unknown by providing a blueprint, effective during and after the pandemic, for the marketing strategy of professional football clubs with the end-goal of creating a large, loyal fan base, which will lead to incremental increases in revenue. Altogether, this research will be relevant for sport marketers through pinpointing which loyalty-influencing factors needs to be enhanced and which ones needs to be mitigated.

1.3 Academic relevance

Besides the managerial relevance, the findings of all the key elements covered in this paper offer great value to the current researchers who are active in the academic field of sport marketing. First of all, fan identification was perceived as a strong driver of fan loyalty before the pandemic (Bauer et al., 2008; Gladden and Funk 2001; Tapp 2004). However, there is no guarantee that this relationship remains unchanged during a period with many restrictions where clubs cannot realize identification among fans in the same way as prior to the pandemic.

Also, with the scarce possibilities of organizing physical events, football clubs should benefit from the transition away from product quality and towards service quality (Deng et al., 2010; Hallencreutz and Palmer, 2019). In other words, fan loyalty is not purely driven by the joy derived from football matches. Little to none academic research is conducted to directly investigate the relationship between digital activities (i.e., consuming social media and esports) and the degree of fan loyalty. However, these digital activities trigger multiple loyalty-enhancing emotions, such as entertainment, escapism and nostalgia (Bauer et al., 2008), which makes it a worthwhile relationship to investigate.

Furthermore, there has been quite some research conducted on the reasons of initially becoming a fan (Greenwood, 2001; Jones, 1997; Parker and Stuart, 1997; Wann et al., 1996). However, all these papers were focused on finding the most prevalent reason of becoming a fan, instead of investigating the relationship between initially becoming a fan and fan loyalty.

Finding the most frequently reason of becoming a fan provides less value for sports marketers than identifying which reason has the strongest influence on fan loyalty.

Without a doubt, the restrictions caused by the pandemic affected the hedonic activities fan interaction and watching football matches through the interdiction of physically meeting in large groups and stadium attendance. Nevertheless, these hedonic activities positively influenced fan loyalty prior to the pandemic (Bauer et al., 2008; Sarstedt et al., 2014). Currently, no researchers devoted their work to investigate how the limitations on these hedonic activities affected the degree of fan loyalty. This paper also strives to be the first breakthrough in this unsolved matter.

The final addition of value this research provides to the academic field concerns the double-dimension concept of fan loyalty, which will be explained in subsection 2.3. The relationship between the two dimensions has been thoroughly investigated (Baldinger and Rubinson, 1996; Bandyopadhyay and Martell, 2007; Dick and Basu 1994). However, specifically for fan loyalty, little research has been conducted to confirm this relationship (Heere and Dickson, 2008). This paper will not simply validate the relationship between the two dimensions of fan loyalty, but the influence of the relationship length of a fan with their respective club on this relationship will be investigated.

2. Theory

In this chapter, the main topics and variables of interest discussed in this research will be discussed. The purpose of the literature review is to gain an understanding of the existing research and debated relevant to this topic. First, the literature review will start with gaining an understanding of the concept of loyalty in general and its drivers., followed by a more specific part on fan loyalty for football fans. After the literature review, the designed hypotheses will be shown and visualized by a conceptual map.

2.1 Customer loyalty: what is it and why is it important?

In the current business environment competition is on an all-time high and the battle for every single customer matters even more. Moreover, due to the rapid rise in technological tools, a single customer is reached by a larger set of brands, which causes a decrease in switching costs. Therefore, the traditional marketing strategy focusing solely on customer acquisition is simply not sufficient to achieve long-term profitability nowadays.

Slater and Narver (2000) state that acquiring new customers' costs marketers between five to ten times more than retaining current customers. Alternating your focus from acquisition to retention will also transform switching costs from a liability to an asset by making unique, brand-related loyalty benefits part of the switching costs (Stan et al., 2013).

2.1.1 Definition of customer loyalty

Dick and Basu (1994) define customer loyalty as 'the relationship between relative attitude and repeat patronage', where relative attitude is the predictor of repeat patronage. Multiple brands are compared to measure the relative attitude, which will give a more realistic view of the relationship between attitude and repeat patronage than measuring the attitude by a brand in isolation. Naturally, the higher the relative attitude, the higher the repeat patronage. Moreover, Ehrenberg et al. (2004) state that this relationship isn't moderated by the magnitude of a brand: customers of small and large brands don't differ much in how loyal they are.

2.1.2 Goal I: cost savings

The goal of customer retention is building a loyal customer base by creating a powerful connection between the brand and the customer through value-adding processes. The addition of value will lead to incremental benefits for both the customer and the brand. Subsection 2.2 will dive deeper into the value-added processes and drivers of loyalty. First, the essence of customer loyalty will be elucidated on the basis of the benefits gained by creating a long-term win-win situation.

Developing a loyal customer base will lead to cost savings for all sort of expenses. Griffin (2002) states that increased loyalty is the main driver for cost savings in six critical areas: from reduced marketing expenses through reduced customer turnover expenses. Duffy (2003) provides a logical reasoning for these cost reductions by emphasizing the fact that searching and guiding a new customer costs more time, and thus money, than a loyal customer. Last but not least, Ehrenberg and Goodhardt (2000) concluded that the stated cost savings will increase in a mature, competitive market.

2.1.3 Goal II: positive word-of-mouth

One of Griffins (2002) six critical area's in cost savings was the increase in word-of-mouth. The need for brands to advertise decreases when satisfied customers will act like a brand advocate by promoting the brand to family and friends. Loyal customers are more likely to have your brand at the top of their mind due to unaided awareness, which generates positive word-of-mouth (Duffy, 2003).

Villanueva et al. (2008) acknowledged the effectiveness of word-of-mouth, because it adds twice as much long-term value to a brand compared to marketing strategies executed by the brand themselves. Von Wangenheim and Bayón (2004) stated that this added value is maximized when the communicator has a high degree of expertise and similarity with the receiver.

Brands need to realize that word-of-mouth is a double-edged sword: the benefits that positive word-of-mouth (PWOM) generate are significant, but negative word-of-mouth (NWOM) could seriously harm the profitability of a brand. Customer dissatisfaction was found to be the biggest driver of NWOM (Von Wangenheim, 2005). Luckily for brands, PWOM is more influential on customer's decision making than NWOM (East et al., 2008; Martin, 2017), which refutes the opposite direction findings of Chevalier and Mayzlin (20060.

2.1.4 Goal III: complain rather than defect

Another benefit gained from customer loyalty is the choice of customers to complain rather than defect from a brand. At first sight qualifying complaints as a benefit may seem strange, but Duffy (2003) stresses the importance of receiving a second chance from your customer. As Stan et al. (2013) correctly pointed out, the decrease of switching costs makes the customer more fickle. Thus, getting the chance to handle complaints is a proper benefit for brands.

Besides the benefit of the chance to handle complaints, Umashankar et al. (2017) emphasize the opportunity brands get to increase customer loyalty while handling complaints. Brands should perceive complaints as a tool to strengthen relationships which may be at risk. By showing a genuine interest to listen and willingness to fix, brands prevent coming across as inauthentic and inadequate.

Moreover, handling complaints properly contributes to minimizing dissatisfaction, the biggest driver of NWOM (Von Wangenheim, 2005). Finally, Morgeson et al. (2020) conducted research to identify moderators who affect the recovery-loyalty relationship. Highly satisfied customers and acting in fast-growing, highly competitive industries will strengthen this relationship, whereas high expectations of product reliability and selling manufactured goods will weaken this relationship.

2.1.5 Goal IV: cross-selling through multiple channels

The last incremental benefit which will be discussed in this subsection is cross-selling through multiple channels. Reinartz et al. (2008) conducted research to determine that cross-buying isn't an antecedent, but a consequence of loyalty. Hence, brands should first build a strong relationship with their customer before targeting them with cross-selling strategies. Eventually, revenues will rise and the costs of doing business with your customer decreases (Duffy, 2003).

Ackermann and Von Wangenheim (2014) investigate which effect channel migration has on cross-selling, unlike Reinartz et al. (2008), who only focused on in-store purchases. They concluded that customers migrating from offline to online channels will lead to a significant increase in cross-buying. Moreover, Li et al. (2016) state that existing customers are participating more in cross-buying than new customers through multiple channels, which is in alignment with the findings of Reinartz et al. (2008). Finally, Li et al. (2016) disagree with the importance of being a first-mover to a new online channel, because your competitor has to use a lot of resources to gain channel awareness. In fact, implementing a follow-up strategy in using online channels will result in higher purchase frequencies for your firm and lower purchase frequencies for your competitor, applicable for both for existing and new customers.

2.2 Determinants of customer loyalty

After thoroughly discussing the benefits gained from customer loyalty, the next step is to identify the main drivers of customer loyalty in general. Subsection 2.4 will follow up with an in-depth clarification on the specific drivers of loyalty in a sports environment.

2.2.1 Determinant I: customer satisfaction

There is a common assumption in the current literature that satisfaction is a strong driver of customer loyalty, but there is some disagreement in which fashion. Jones and Suh (2000) concluded that overall satisfaction had a direct effect on repurchase intentions, whereas transaction-specific satisfaction only had a little impact on repurchase intentions when overall satisfaction was high.

However, Chinomona and Dubihlela (2014) found an indirect effect between satisfaction and repurchase intentions through customer trust and loyalty, unlike Jones and Suh (2000). Moreover, Ranaweera and Prabu (2003) state that satisfaction has a stronger effect on customer loyalty than trust in this indirect relationship. Deng et al. (2010) excluded repurchase intentions from their research and sought to determine direct drivers of customer loyalty, dissimilar to previous studies. Satisfaction enhanced loyalty the most, followed by customer trust and increased switching costs.

Bennet and Rundle-Thiele (2004) questioned whether this strong, positive relationship between satisfaction and loyalty didn't imply synonymy. This was not the case, but the risks of overestimating the importance of satisfaction were exemplified. Namely, satisfaction levels do not always translate to high levels of loyalty. Therefore, relying purely on satisfaction to predict repurchase intentions is badly, which again is in disagreement with the findings of Jones and Suh (2000). Singh (2006) also agrees with Bennet and Rundle-Thiele (2004) that satisfaction does not guarantee repurchase, because satisfied customers still have the tendency to defect from the brand, unlike loyal customers.

2.2.2 Determinant II: service quality

The positive relationship between customer satisfaction and loyalty is clarified, but the drivers of this relationship remains unclear. Churchill and Surprenant (1982) identified product performance as the only significant driver of satisfaction for both durable and non-durable goods, whereas disconfirmation and initial expectations only affected satisfaction for non-durable goods.

More recent research concluded that product quality is substituted by service quality as the main driver for satisfaction and loyalty, which is in alignment with the current trend of co-creating value by implementing service-dominant logic, instead of goods-dominant logic, elements in the marketing strategy (Deng et al., 2010; Hallencreutz and Parmler, 2019; Kristensen et al., 2000). Moreover, Deng et al. (2010) and Kristensen et al. (2000) both concluded that perceived customer value directly affects customer satisfaction, which was enhanced by a positive brand image and a high quality of customer interaction (Kristensen et al., 2000).

2.2.3 Determinant III: customer characteristics

To finalize the satisfaction-loyalty link, the influence of customer characteristics on this relationship will be discussed. Homburg and Giering (2000) sought to determine which characteristics moderate this relationship. The results showed that age and income did significantly affect the relationship as a moderator, unlike gender. Older people and people with a lower income strengthened the satisfaction-loyalty link, which is caused by experience-based evaluation and the financial risk of buying a poor-quality product. Also, Anderson et al.

(2008) found higher satisfaction levels for older and low-income customers, but no significant difference was found for gender. This is consistent with the findings of Homburg and Giering (2000).

2.3 The two dimensions of fan loyalty: attitudinal and behavioral

After obtaining knowledge about customer loyalty in general, the concept and dimensions of fan loyalty will be explained to get a better understanding about this unique form of loyalty, prior to discussing the drivers of fan loyalty in the next subsection.

2.3.1 Defining the double-dimension concept of fan loyalty

Dietz-Uhler et al. (2000) define a fan as someone who perceives him- or herself as a fan of a certain team or sport in general. Thus, researchers do not qualify whether someone is a fan or not, but the participants do that for themselves. Levinson and Pfister (2013) complement Dietz-Uhler et al. (2000) by defining fan loyalty as a deeper connection between the fan and the respective sports club. Unlike a bandwagon fan, winning isn't the sole determinant of supporting a club. The more loyal a fan is, the more revenue can be generated, which makes fan loyalty a vital key performance indicator to track and increase over time.

Jacoby and Chestnut (1978) were the first researchers to conclude that loyalty isn't a one-dimension measurement that is fully captured by behavioral criteria, but that adding an attitudinal dimension to the loyalty concept is essential to fully capture and understand this phenomenon. Dick and Basu (1994) acknowledged these findings and defined loyalty as a double-dimension concept, where attitudinal was the predictor of behavioral.

Bandyopadhyay and Martell (2007) supported the theoretical framework of Dick and Basu (1994) by adding validation through empirical research. Besides confirming the attitudinal-behavioral relationship, the importance of targeting non-users who hold favorable attitudes towards your brand became clear. These non-users have the potential to contribute significantly to an increase in revenue by triggering their behavioral intentions through effective marketing strategies.

Finally, Baldinger and Rubinson (1996) strengthen the stated attitudinal-behavioral link even more by finding the predictive characteristics of attitudinal loyalty. Moreover, in alignment with the findings of Bandyopadhyay and Martell (2007), non-users with favorable attitudes have a significant higher conversion rate than non-users without those favorable attitudes.

2.3.2 Dimension I: attitudinal loyalty

With the relationship between the two dimensions being clear, the transition will be made to those dimensions in the context of fan loyalty. Bauer et al. (2008) state that attitudinal loyalty is represented by the psychological commitment of a fan to a team. This commitment is measured by the inner attachment, persistence and resistance of a fan. Thus, a deep inner attachment and showing consistency over time of those favorable attitudes represents a high degree of attitudinal loyalty.

2.3.3 Dimension II: behavioral loyalty

Yim and Kannan (1999) conclude that behavioral loyalty should not only be measured by past behavior, but future behavioral intentions should be included in the concept to fully capture the degree of behavior loyalty. PWOM is also an important indicator of behavioral loyalty, besides purchases. Bauer et al. (2008) classified watching matches of your favorite club on television or in the stadium, consuming club-related social media during leisure time, purchasing club merchandise and being a club advocate in public with PWOM as the four main criteria for behavioral loyalty in a sports context. Fans who score high on these criteria for both past and future (intended) behavior represent a high degree of behavioral loyalty, which is the ultimate goal of clubs to maximize their revenue.

Little research has been conducted for validating the attitudinal-behavioral link in a sports environment. Only Heere & Dickson (2008) have found a significant correlation between attitudinal loyalty and behavioral involvement, which demonstrates that attitudinal loyalty positively influences behavior. This is consistent with the findings of previous researchers (Baldinger and Rubinson, 1996; Bandyopadhyay and Martell, 2007; Dick and Basu, 1994).

2.4 Determinants of fan loyalty

At this stage, the concept of double-dimension fan loyalty and the relationship between the two dimensions has been comprehensively explained. Right now, the transition will be made towards the determinants of fan loyalty to find out which factors are the main drivers of this concept.

2.4.1 Determinant I: brand associations

Gladden and Funk (2001) examined the drivers of fan loyalty through an in-depth study of 13 brand association dimensions, which they defined as anything in the consumer's mind linked to a specific sports team. Surprisingly, the team success and having a star player didn't enhance fan loyalty. Fan identification, nostalgic memories, entertainment and the need to escape the daily rigors of life were all found to be significant drivers of fan loyalty. Thus, marketing strategies in a sports environment need to be more advanced than simply relying on your star player and team performance to create a high degree of fan loyalty.

In contrary of Gladden and Funk (2001), who found loyalty drivers for sports in general, Bauer et al. (2005) specified their research by purely focusing on drivers of fan loyalty for football fans. Brand associations were categorized in either attributes or benefits, where attributes were the predictor of perceived benefits.

Also, the attributes were split up into non-product (NPR) and product related (PR) attributes. PR-attributes consist of aspects who directly impact the product and/or service performance. Star player, team success, head coach and the whole squad can all be considered as PR-attributes, because they directly impact the performance. Stadium, club colors, logo, club history and other fans don't directly influence the performance of a football club, which makes them examples of NPR-attributes.

Bauer et al. (2005) stated that both PR- and NPR-attributes enhanced fan loyalty, but the effect of NPR-attributes was almost triple that of PR-attributes, where fan interaction was the strongest driver of all NPR-attributes. This is inconsistent with the findings of Gladden and Funk (2001), who found a negative relationship between PR-attributes and long-term fan loyalty. Furthermore, Sarstedt et al. (2014) strengthens the findings of Bauer et al. (2005) by identifying the NPR-attributes stadium and fan-based support as the main drivers of fan satisfaction, from which it's importance was explained in subsection 2.2.1.

Three years later, Bauer et al. (2008) conducted the same research with matching conclusions, but in a more extensive fashion with in-depth clarifications. The strong effect of NPR-attributes was explained due to their consistency over time, while PR-attributes change more frequently, which makes them less effective in driving long-term fan loyalty. However, the effect of PR-attributes on loyalty cannot be fully neglected. Multiple researchers have found a significant short-term effect on fan loyalty, but they acknowledge the long-term ineffectiveness of enhancing fan loyalty through focusing on PR-attributes (Gladden and Funk, 2001; Kaynak et al., 2007; Wu et al., 2012).

2.4.2 Determinant II: brand attitude

Currently, only different dimensions of brand associations have been discussed. However, brand image consists of the cumulative product of brand associations (Bauer et al, 2005), which means that this concept is a direct influencer of fan loyalty. Bauer et al. (2008) added brand attitude as an extra layer to the double-layered brand image concept of Bauer et al. (2005), which only consisted of attributes and benefits. This association holds abstract, overall evaluations of a sports team which will be more favorable when the perceived benefits are higher. Figure 1 visualizes this relationship for clarity.

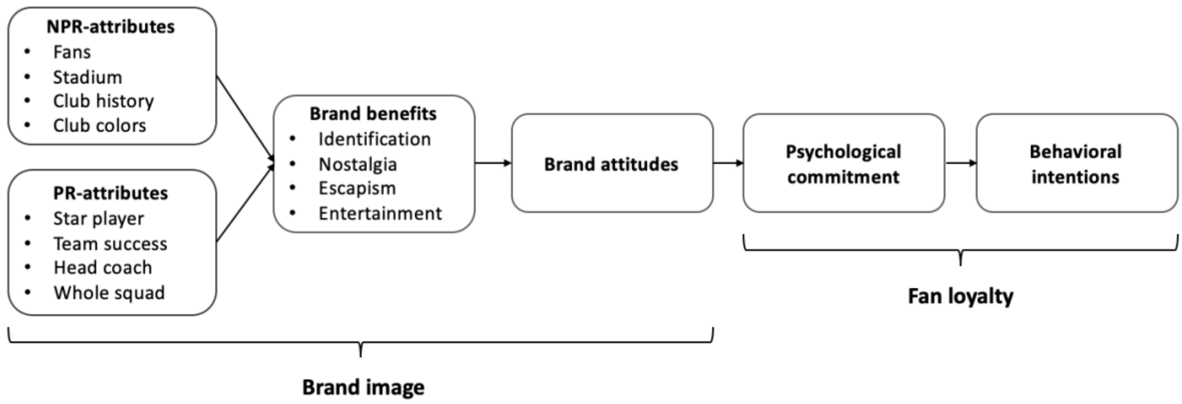

Figure 1: the relationship between brand image and fan loyalty conceptualized (Bauer et al., 2008)

Fan identification is the most influential brand benefit and entertainment the least influential one in the process of forming favorable attitudes, which will ultimately lead to the desired end-goal of creating a fan-base with a high degree of behavioral loyalty (Bauer et al., 2008; Gladden and Funk, 2001; Tapp, 2004). Last but not least, Karjaluto et al. (2016) state that the brand image-loyalty link is valid for both dimensions and stronger for newer fans due to their lack of experiences with the sports team. Competitiveness, authenticity and uniqueness were identified as the three most influential traits on gaining a high degree of fan loyalty.

2.5 The relevance of esports for football clubs

The pandemic has causes all kind of disturbances in the sports industry, but it also offered significant growth opportunities to industries who are highly active on online platforms. People are more at home than they have ever been, which caused an increase in leisure time and digital consumption. However, for football clubs the restrictions have led to a decrease in options to deliver entertaining content (Mastromartino et al., 2020), which can be classified as a loyalty enhancing NPR-attribute (Gladden and Funk, 2001; Bauer et al., 2008).

2.5.1 Defining esports

Therefore, numerous football clubs added esports to their marketing activities to solve this content shortage and keep their entertainment level on a high note (Mastromartino et al., 2020). Electronic sports, or esports, is a form of sports where the primary aspects are facilitated by technical systems and the outcome defining activities always occur in the virtual world (Hamari and Sjöblom, 2017).

2.5.2 Benefit I: enormous reach

There are countless marketing strategies that football clubs could execute to fill up the void of digital content. So, why would integrating an esports department in your firm be the most desired alternative? First of all, esports has an enormous reach with hundreds of millions of people spectating on a yearly basis all over the world (Hamri and Sjöblom, 2017) and during the pandemic a record-breaking growth of 150% was recorded (Mastromartino et al., 2020).

2.5.3 Benefit II: diversified viewership

Moreover, the viewership of esports is shifting away from the generalization of young men as the only sort of spectators. In 2019, only 65% of the viewers were male and 73% was under the age of 35. This shows that the viewership is not only expanding, but also diversifying, which offers marketers opportunities to build strong relationship in a fast-growing market with fans who are traditionally hard to reach (Lehnert et al., 2020).

2.5.4 Benefit III: football videogame qualified for esports

To qualify a digital game as esports, there must be amateur or professional competition within the scope of that game that represents the professionalism and competitiveness of this industry well (Fakazli, 2020). However, being qualified as an esports game does not guarantee favorable results in this market. Increased viewership demand ensues from more people playing the respective game. Games with an above-average level of competitiveness and required skill will receive the most popularity for recreational purposes, which is a necessary demand to succeed in the esports business (Lee and Schoenstedt, 2011).

Last but not least, Lettieri and Orsenigo (2020) concluded that sports-related esports increases football consumption, whereas non-sports related esports reduces it. FIFA, the most popular soccer video game, meets all the previously stated requirements and is steadily manifesting themselves as a big fish in this booming, profitable market.

2.5.5 Requirements for success in the esports scene

After acknowledging the potential of esports for football clubs, the question remains which motivational factors would trigger fans to frequently watch esports. Escapism and excitement were found to be the two biggest antecedents of watching esports (Hamari and Sjöblom, 2017; Xiao, 2020). Not coincidentally, these specific antecedents also enhance fan loyalty, which makes them increasingly important (Gladden and Funk, 2001; Bauer et al., 2008).

Moreover, Lehnert et al. (2020) state that clubs should choose wisely who they choose as the representative of their brand, because characteristics of your esporter moderate this relationship. Wrongful actions, like gender discrimination, will weaken this relationship. They are easily picked up by a large group of viewers due to esporters being under constant scrutiny.

Finally, the applicability of implementing esports in your marketing strategy will be strengthened using the open system theory. This theory purports that organizations are strongly influenced by the environment in which they operate (Bastedo, 2004). Esports is one of the factors causing environmental changes for football clubs. Namely, the ongoing trend of increasing interest in esports does not only count for online spectatorship, but also the desire to physically attend esport events (Jenny et al., 2018).

Currently, these events are not possible due to the restrictions, but this stated desire offers extra tools for football clubs to achieve long-term profitability when allocating marketing budget to an esports department (Fakazli, 2020; Jenny et al., 2018). Thus, the short-term deficits due to the cancellations of physical events will be vastly compensated, because of the increased interest in both offline and online attendance to esports events (Fakazli, 2020).

2.6 Exploring the reasons of initially becoming a fan

Numerous drivers of fan loyalty have been discussed, but to optimize the marketing strategy knowledge about reasons of initially becoming a fan are crucial to maximize the effectiveness of the budget spent. Nevertheless, little research has been conducted to clarify the relationship between the reason of becoming a fan and their respective degree of fan loyalty.

Currently, relationship-based reasons are reported more frequently in becoming a fan than recognition-based reasons (Koch and Wann, 2016). In other words, establishing and maintaining connections are valued more highly by fans than gaining social approval through supporting a successful club. This highlights the importance of focusing on NPR-attributes, such as fan interaction and entertainment. Not only for increasing fan loyalty (Bauer et al., 2008; Gladden and Funk, 2001), but also for growing your fan base.

2.6.1 Discrepancy among researchers in finding the prevalent reason

The main reasons found for initially becoming a fan did not cause discrepancy in the field, but there is some disagreement in determining which is the prevalent reason. Some found geographical reasons as the dominant reason (Greenwood, 2001; Jones 1997), whereas others concluded that parental influence was the strongest driver of becoming a fan (Parker and Stuart, 1997; Wann et al., 1996). Additionally, Parker and Stuart (1997) highlight the importance of the strong bonding between father and sons on becoming a fan and developing loyalty.

2.6.2 Entertainment as the reason of becoming a fan

Drivers related to entertainment, such as fan events and enjoyment derived from matches, can also contribute to becoming a fan (Dietz-Uhler et al., 2000; Greenwood, 2001). However, these factors are significantly stronger for women than men, who in general have weaker ties with their respective club (Dietz-Uhler et al., 2000). Furthermore, women tend to become a fan at a later age due to the influence of third parties like boyfriends influencing club selection, instead of being triggered by marketing activities to become a fan (Parker and Stuart, 1997).

Hence, Greenwood et al. (2006) state that using entertainment as the core value of the marketing strategy may result in fan attraction and short-term profits, but they highly doubt the potential of realizing fan identification and thus long-term loyalty with this strategy. In other words, entertainment should only fulfill a supportive role in the marketing strategy.

2.7 How loyal were football fans before the pandemic?

The consistency over-time of fan loyalty is the final topic that will be discussed in this literature review. Do fans support one club during their entire lifespan? Or is the general tendency of fan loyalty more directed towards relatively easy switching between clubs?

2.7.1 Refuting exclusive fan loyalty

Parker and Stuart (1997) were one of the first researchers who devoted their work to this matter. Exclusive fan loyalty was defined as the norm, whereas club switching was the exception to this rule. Many researchers questioned these findings due to the lack of empirical evidence. Mahony et al. (2000) identified four different loyalty segments, which refutes the concept of exclusive fan loyalty of Parker and Stuart (1997). Namely, not every fan can be automatically classified as a highly-committed fan who will never defect from their respective club. Besides, smaller clubs have the most committed fan base (Newson et al., 2021).

2.7.2 Different segments of fan loyalty

Moreover, Tapp and Clowes (2002) not only supported the findings of Mahony et al. (2000), but also provided added value by pinpointing several sub-groups in the loyalty segments. This offers opportunities to clubs to target your fans more precisely and thus effectively. For example, in the low loyalty segment 'carefree casuals' prefers to see an entertaining game, even when the team loses. The 'committed casuals', also in the low loyalty segment, value winning as much as the fanatics of a club, but unlike them they give equal or greater priority to non-football activities.

Also, Tapp (2004) did acknowledge the existence of exclusive fan loyalty of Stuart and Parker (1997), but stated that the concept of fan loyalty is way more complex than putting all your fans in one basket. Different segments in fan loyalty were identified (Mahony et al., 2000; Tapp and Clowes, 2002), which directed Tapp (2004) towards the conclusion that loyalty is not something to automatically rely on. In other words, the old saying 'we'll support you ever more' is not a given anymore and marketing resources have to be invested in your supporters to achieve this exclusive brand loyalty.

Last but not least, Richardson and O'Dwyer (2003) highlighted the importance of receiving social approval from the supported football club, because it can be a potential driver of defection when fans experience a lack of social approval. This also refutes the findings of Parker and Stuart (1997).

2.8 Development of hypotheses

Following a comprehensive review of the current literature in the area of customer and brand loyalty, the hypotheses will be drafted to reveal which factors have the expectation to drive fan loyalty and which ones have an impeding influence. The concept of double-dimension fan loyalty will be used in this research to get a better understanding in which magnitude and direction each factor enhances or worsens the attitudinal and behavioral dimension of fan loyalty.

Lots of research already has been conducted to classify the attitudinal dimension as the predictor of behavioral intentions (Baldinger and Rubinson, 1996; Bandyopadhyay and Martell, 2007; Dick and Basu, 1994), but the question arises whether this relationship is just as strong in a sports environment. The assumption is that this will be the case, because sports evokes a wide range of strong emotions. These emotions are the basis of forming an, favorable or unfavorable, attitude towards a club, where the direction and strength of the attitude will determine the degree of soccer consumption related to the team.

However, the independent-procedure theory of Snyder and Tanke (1976) states that a potential relationship between the two dimensions does not automatically result in overlapping predictors. Namely, changes in attitude and behavior occur independently of each other. In other words, not every predictor of loyalty influences both dimensions of loyalty, because the two dimensions are not interchangeable.

Moreover, the relationship length of a fan is expected to moderate the double-dimension relationship. Specifically, the longer a fan perceives himself as a fan, the stronger the relationship between the two dimensions. For example, the expectation is that fans who greatly differ in relationship length can have the same level of attitudinal loyalty, but the fan who supports the club way longer will exhibit significantly higher levels of behavioral loyalty due to their increasing desire of soccer consumption formed over-time. These beliefs have led to the following hypotheses:

H_{1a}: In the double-dimension fan loyalty concept, attitudinal loyalty is the predictor of behavioral loyalty.

H_{1b}: Relationship length acts as a moderator and positively influences the attitudinal-behavioral relationship

Small or large, every brand has the potential to gain a loyal customer base (Ehrenberg, 2004). In other words, all football clubs should strive for exclusive brand loyalty to maximize their revenue. Realizing a high degree of fan identification has led to strong levels of behavioral loyalty before the pandemic (Bauer et al., 2008; Gladden and Funk, 2001; Tapp, 2004). The expectation is that identification is one of the main motives to stay engaged with your club during the rough times of the pandemic. Thus, it will not only drive behavioral loyalty, but also the attitudinal dimension. Namely, identification will also transform a generic fan into a psychological committed fan, which is the measurement for attitudinal loyalty. On these bases, the following hypothesis is formulated:

H_2: Fan identification has a strong, positive effect on both dimensions of fan loyalty

Currently, social media offers loads of opportunities for clubs to enhance their service quality, which is becoming increasingly important (Hallencreutz and Palmer, 2019) and reduces the amount of NWOM (Von Wangenheim, 2005). The versatility in the area of content creation is massive, where being the first mover in terms of new content is not even a must to succeed (Li et al., 2016). Furthermore, clubs can reach out to non-users with favorable attitudes, who have the potential to become strong loyal fans (Baldinger and Rubinson, 1996; Bandyopadhay and Martell, 2007).

Implementing an extensive social media strategy has the expectation to trigger numerous loyalty enhancers for fans, such as escapism, nostalgia and entertainment (Gladden and Funk, 2001; Bauer et al., 2008) Team centered content, like behind the scenes footage and interviews, give fans a moment to escape the daily rigors of life and be entertained, while it can also trigger behavioral intentions. Nostalgic content, in the form of old-highlights and interesting facts, will evoke positive emotions and thus favorable attitudes.

The general belief that social media is mainly consumed by younger fans, who attach less value to nostalgic content, leads to the assumption that nostalgic content will be less effective than team centered content. These thoughts lead to the following hypotheses:

H_{3a}: Making use of the versatility of social media will enhance both dimensions of fan loyalty.
H_{3b}: Team centered content will have a stronger, positive effect on fan loyalty than nostalgic content.

Esports activities have the potential to be an effective marketing tool for strengthening the bond with your fans and thus loyalty. At this moment, esports is an attractive market to enter with its enormous reach, exponential growth and an increasing diversification among the esports fanatics (Hamari and Sjöblom, 2017; Lehnert et al., 2020; Mastromartino et al., 2020). Furthermore, the loyalty enhancers escapism and excitement are the result of watching esports (Hamari and Sjöblom, 2017), while it also increases soccer consumption (Lettieri and Orsenigo, 2020).

The digital consumption time of fans increases by watching esports, which will increase the top-of-mind brand awareness of the respective club. In theory, this should lead to an increasing desire to watch more real life matches and potentially buying more merchandise. Thus, this stated domino effect makes participating in esports activities an expected antecedent of behavioral loyalty.

Finally, it is very unlikely that esports will significantly contribute to forming favorable attitudes towards the club due to the little impact in comparison to more traditional factors, such as the club's history and the current team. Hence, the following hypothesis is formulated, based on the independent-procedure theory of Snyder and Tanke (1976):

H_4: The addition of an esports department will positively impact the behavioral dimension of fan loyalty.

Parental influence, friends' influence, team's success and entertainment are the four reasons of initially becoming a fan from which the influence on fan loyalty will be investigated. Koch and Wann (2016) stated that relation-based reasons are reported more often than recognition-based reasons, but it remains unclear whether influence from friends and family is a stronger fundament for a high degree of fan loyalty than the team's success.

The assumption will be made that relationship-based reasons are indeed the stronger predictor of high loyalty levels due to the fact that these reasons channel a strong emotional bonding with the club, which will also cause behavioral intentions. In contrary of relationship-based reasons, team's success as the main reason for becoming a fan tends more towards a bandwagon fan with low levels of loyalty on both dimensions.

Furthermore, this research strives to be the final piece of the puzzle with regards to finding the most prevalent reason of becoming a fan. The effect of parental influence (Parker and Stuart, 1997; Wann et al., 1996) is expected to outweigh geographical reasons (Greenwood, 2001; Jones, 1997). Namely, the club which receives parental support is expected to be less likely to be replaced than clubs followed due to geographical reasons, where moving towards another city can cause fans to defect from the respective club. As a result of these thoughts, the following hypotheses have been drafted:

H_{5a}: Being a fan through the team's success will be the only reason which negatively impacts both dimensions of fan loyalty.
H_{5b}: Parental influence as a reason for becoming a fan has the strongest effect on both dimensions of fan loyalty.

Also, due to the restrictions football games have been played for a long time without fans in the stadium, the so-called ghost games. The joy derived from watching matches was classified as a strong driver of fan satisfaction (Sarstedt et al., 2014). Naturally, not having the opportunity to even watch your favorite team live and being forced to watch it at home will cause dissatisfaction among fans, which worsens the degree of fan loyalty (Chinomona and Dubihlela, 2014; Deng et al., 2010; Ranaweera and Prabu, 2003). The entertainment aspect, a NPR-attribute and known loyalty enhancer, of stadium attendance also disappeared during the pandemic.

The expectation is that ghost games caused a decrease in soccer consumption and thus fan loyalty. Namely, ghost games did not only forbid fans to enter the stadium, but fans were forced to watch a less exciting match at home with few to none people due to a maximum number of visitors. However, the attitudinal dimension is not expected to being impacted, because fans should have the knowledge that these restrictions are external factors and not form negative attitudes or even defect from their favorite club. So, the following hypothesis is formulated, based on the the independent-procedure theory of Snyder and Tanke (1976):

H_6: Ghost games did have a negative impact on the behavioral dimension of fan loyalty.

Additionally, opportunities to create activities that contribute to interaction between fans have been drastically reduced due to the restrictions. Fan interaction was, together with stadium related factors, the biggest driver of fan satisfaction (Sarstedt et al., 2014). Also, Bauer et al. (2005) identified fan interaction as the NPR-attribute with the largest influence on fan loyalty, which only confirms its importance even more.

Community events hosted by the club, attending training sessions and watching football matches with fellow fans in a pub or at someone's home were events where fans interacted. During the pandemic, these fan interactions were less or not possible at all.

Therefore, online events were organized in an attempt to replace those physical meetings, but the assumption has been made that this kind of events cannot evoke the same emotions of joy as offline events. In other words, the lack of fan interaction causes dissatisfaction and worsens only the behavioral dimension of loyalty for the same reason as ghost games. This expected relationship has led to the following hypothesis, based on the independent-procedure theory of Snyder and Tanke (1976):

H_7: The lack of fan interaction weakens the degree of the behavioral dimension of fan loyalty.

2.9 Conceptualization of the hypotheses

A conceptual framework has been designed to visualize the hypotheses in a more simplified manner, which can be seen in figure 2. A side note has to be placed to prevent presenting a distorted picture. Namely, to prevent the idea that some factors will only be tested on one instead of both dimensions, which could be concluded when only seeing one arrow for every factor.

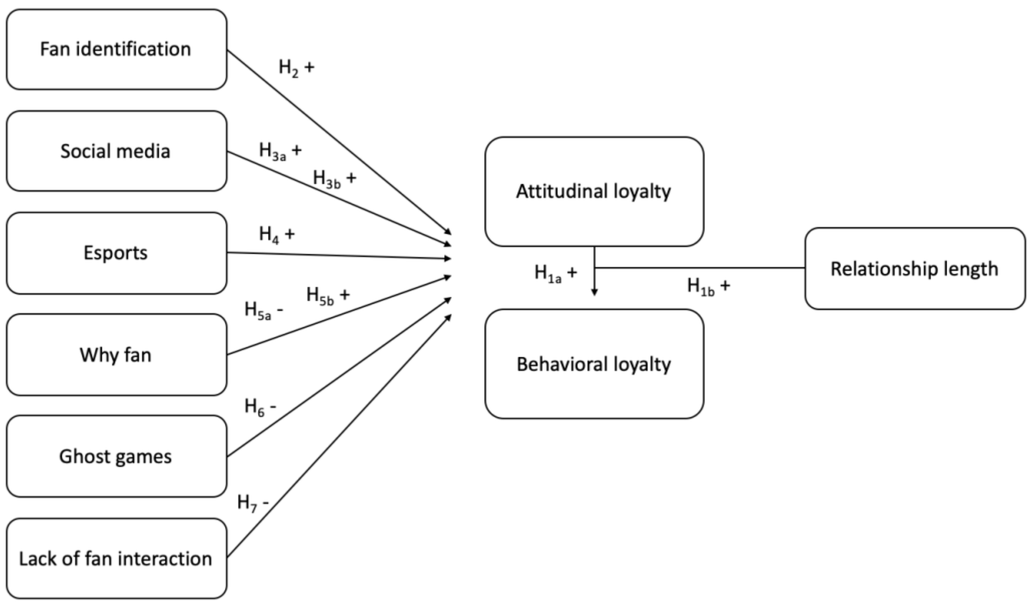

Figure 2: the conceptual framework of this paper, visualizing all the hypotheses with their expected directions.

3. Methodology

In this chapter, various methods will be discussed which are used in this research to answer the research question in the best possible way. In addition to the standard explanation of the used methods, the reasoning behind the choices, backed up by academic literature, will be explained to insure a research paper of high-quality.

3.1 Type of data collection

3.1.1 Online survey research: how, why and what?

The quantitative method online survey research was used to collect data for this research, which resulted in a data set consisting of only primary data. The survey was developed with the online survey tool Qualtrics to benefit from their attractive features, such as offering assistance for survey customization and automatically processing the data, which saved a lot of time during the period of data collection (Wright, 2005).

Also, the cost of reaching out to respondents for online survey research was low (Schmidt, 1997). Namely, the survey could be quickly distributed through numerous platforms, like Reddit, LinkedIn and Instagram to reach out to respondents and achieve the desired amount. The sample size and corresponding method will be discussed shortly in subsection 3.5.

Survey research was preferred over other data collecting methods for a variety of reasons. In 2021, soccer was the most popular sport in the world with an estimated fan base of 3.5 billion (Sourav, 2021). Survey research can easily reach all kinds of soccer fans over the whole world. This makes it possible to provide a representative sample for a large population, unlike time-consuming interviews with their limited sample sizes.

Moreover, concepts like fan identification and loyalty can not only be more effectively measured in survey research due to their abstract nature, but this form of primary data is also more suitable for statistical analyses. Also, the nature of this research, i.e. investigating the influence of six factors on both dimensions of fan loyalty, makes survey research the best possible option. The survey questionnaire has comprehensively captured the attitude of respondents on a wide range of topics, whereas conducting an experiment with manipulation of only one topic would have been too specific and thus not appropriate for this research.

Right now, a closer look will be taken on the design of the survey questionnaire, which can be seen in its entirety in appendix A. Filling in the survey instrument takes approximately 5 minutes and consists of 38 questions, from which the first question is a screening question to guarantee a representative sample. The next five questions can be placed in the category general questions, such as age, gender and education. The other 32 questions are statements where the respondents attitude towards their degree of loyalty and the respective six factors is measured. The corresponding response scale is a 7-point Likert scale, with level of agreement as the predominant scale. The level of importance scale was only used to record the respondent's attitudes towards the reasons of initially becoming a fan (Q7 till Q10, see appendix A).

3.1.2 Likert scale: definition and benefits

Bertram (2007) defines Likert scale as a psychometric response scale to obtain participant's degree of agreement or importance with a set of statements. They are particularly useful for measuring latent variables, i.e. unobservable individual characteristics without a concrete, objective measurement (Bertram, 2007; Willits et al., 2016). These findings make Likert scales very relevant for this research due to the presence of seven latent variables, which are variables that cannot be directly observed. Hence, they are constructed by multiple observed items. Subsection 3.4 will dive deeper into which observed items belong to what latent variable.

Finally, the choice of using a 7-point Likert scale instead of the 5-point alternative will be explained. Needless to say, the 7-point Likert scale provides more options than the 5-point variant. However, including two more options does result in a higher degree of variety and a decreased desire to interpolate (i.e., wanting to choose an unavailable option), which in turn increases the probability of meeting the objective reality of the respondents (Finstad, 2010; Joshi et al., 2015).

3.2 Statistical analysis I: CFA

First, confirmatory factor analysis (CFA), a type of structural equation modeling (SEM), will be conducted in the statistical program SPSS to validate the seven latent constructs of this research, which all consist of at least three factors to prevent any under-identified models. Under-identified models occur when a latent variable is formed by too little observed items (i.e., 2 or less). This causes a negative outcome for degrees of freedom, which means not sufficient values have the freedom to vary and the results cannot be interpreted (Brown and Moore, 2012). Appendix B provides the corresponding calculations to support this theory.

Prior to analysis, the raw data needs to be converted to a correlation matrix to create the proper structure for CFA. This research made use of a polychoric correlation matrix, instead of the more frequently used Pearson matrix, as the input for CFA. Namely, Pearson correlations evaluate the relationships of continuous variables (Lorenzo-Seva and Ferrando, 2014). Needless to say, 7-point Likert scale items are not considered continuous, but ordinal and discrete.

Hence, using Pearson correlations and treating Likert scale items as continuous variables would have caused wrongful path coefficients and thus produce less accurate results (Holgado-Tello et al., 2008; Van der Eijk and Rose, 2015). Therefore, polychoric correlations were used to bypass potential misleading results. Lorenzo-Seva and Ferrando (2014) provided a SPSS syntax named Polymat-C, which made it possible to obtain the polychoric correlation matrix with the raw data and use it as the main input for CFA.

SPSS Amos, known as powerful SEM software, provided the output of CFA. For the factor loadings, the cutoff was set at 0.40 (i.e., items with a factor loading of 0.40 or greater were retained). A side note must be placed that 0.40 is the lowest acceptable threshold, where factor loadings of at least 0.70 or greater were desired for excellent construct validity (Matsunaga, 2010). Sun (2005) defines construct validity as the cohesiveness of a set of items in measuring their corresponding latent factor, which is exactly the kind of validity this research strives for its seven latent variables.

The interpretation of the CFA output will consist of two parts: checking the significance and value of the factor loadings and assessing the goodness of fit of the CFA model. The *p*-value will determine whether the loadings are significant or non-significant. For example, a *p*-value of .05 is equal to a chance of 5% that the stated relationship is not really there and is found by mistake. Table 1 shows three levels of the *p*-value and their respective degree of significance.

***p*-value**	**Level of significance**
<.10	Moderately significant (*)
<.05	Significant (**)
<.01	Highly significant (***)

Table 1: three *p*-values with their respective level of significance

For assessing the goodness of fit, three fit-indices were considered to be the most appropriate for this research: the Tucker-Lewis index (TLI), confirmed fit index (CFI) and the root mean square error of approximation (RMSEA). These fit indices were chosen due to the fact that they are all relatively robust to the large sample inflation effect (i.e., value of fit increases systematically when the sample sizes become larger) (Marsh et al., 1988; Sun, 2005). Moreover, all three of them are qualified as good fit-indices to evaluate the construct validity (Sun, 2005), which supports the choice for these three even more.

Rule of thumb criteria were used for interpreting the three fit-indices: the CFA model was qualified as acceptable when the fit-indices values of TLI and CFI were .90 or higher (Bentler and Bonett, 1980) and the value of RMSEA was below 0.08 (Awang, 2012)\

3.3 Statistical analysis II: OLR

3.3.1 OLR: how, why and what?

After validating the latent variables, the next step was to estimate the strength and direction of all the factors on both the attitudinal and behavioral dimension of fan loyalty. This research did deliberately not opt for a paired sample t-test to compare the pre- and during pandemic attitudes of the fans on both the factors and loyalty. Namely, measuring past behavior is unreliable due to the incompetence of respondents to remember exactly what they did, whereas measuring past attitudes is even harder (Bohte et al., 2009). Therefore, a regression analysis was applied with data consisting of current attitudes and behavior to obtain reliable and valid results.

Specifically, ordinal logistic regression (OLR) was used in SPSS to model and estimate the predictive relationship between the factors and the two dimensions of fan loyalty. Applying linear regression (LR) was not possible due to the vast majority of the data set being ordinal variables (i.e., the 7-point Likert scale items), which violates the LR assumption of having a continuous dependent variable (DV) (Casson et al., 2014).

Also, OLR was preferred over multinomial logistic regression (MLR). MLR is generally used when the categories of the DV are unordered, whereas OLR is used for DV's with ordered categories such as a 7-point Likert scale. Moreover, unlike MLR, OLR yields only a single set of coefficients to estimate the hypothesized relationships between the DV and independent variables (IV's) (Osborne, 2015). Thus, OLR provides a more parsimonious representation (i.e., simplest model with the greatest explanatory power) of the data than MLR.

3.3.2 Model equation I: Attitudinal loyalty

So, two OLR equations were formed. For the first equation, the natural logarithm odds (log odds) of the latent variable attitudinal loyalty were the dependent variable (DV), with numerous loyalty-influencing factors acting as an independent variable (IV) in the model. Below, the first model equation is shown:

$$\begin{aligned}\ln\left(\frac{p}{1-p}\right) = {} & \alpha + \beta_1 Fan\ Identification + \beta_2 Social\ Media + \beta_3 Team\ Content \\ & + \beta_4 Nostalgic\ Content + \beta_5 Parental\ Influence + \beta_6 Friend\ Influence \\ & + \beta_7 Geographical\ Location + \beta_8 Team's\ Success \\ & + \beta_9 \text{ Male} + \beta_{10} \text{ Female} + \beta_{11} \text{ Third Gender} + \beta_{12}\ Age \\ & + \beta_{13}\ Less\ than\ High\ Scool\ Degree + \beta_{14} High\ School + \beta_{15}\ Some\ Degree \\ & + \beta_{16}\ Bachelor + \beta_{17} \text{ Master} + \beta_{18}\ Small\ Club + \beta_{19}\ Medium\ Club + \varepsilon\end{aligned}$$

$\ln\left(\frac{p}{1-p}\right)$ represents the natural logarithm odds for obtaining a higher degree of attitudinal loyalty, where p stands for the probability of falling in a greater level of attitudinal loyalty and 1- p is the probability of not falling in a greater level of attitudinal loyalty. Without using the log odds, regression analysis would not have been possible. Namely, log odds can take any positive or negative number [-∞, ∞], whereas odds are restricted to positive values only [0, ∞] due to the finite, positive range of p [0,1]. When the DV would have been limited to positive outcomes only, interpreting negative coefficients would be impossible and thus lead to invalid results.

β_1 till β_7 represent the coefficients that show the main effect of the regression, while β_9 till β_{17} are coefficients for control variables. Control variables, factors which affect the outcome of the DV, are added to the regression to prevent omitted variable bias (OVB), which results in a more accurate estimation of the coefficients.

The variables β_3 till β_8 were measured directly through one survey question (Q7-Q10 and Q14-Q15, see appendix A). Also, the latent variables in both equations were constructed by adding up the means of their respective observed items, divided by the number of observed items. Statistical significance is determined by the *p*-values, as shown in Table 1.

Older and lower income fans exhibit in general higher levels of loyalty (Anderson et al., 2008; Homburg and Giering, 2000), whereas females have weaker ties with their favorite club than males (Dietz-Uhler et al., 2000). Also, fans who support a smaller club tend to be more loyal than fans of bigger clubs, which have a higher percentage of less-committed bandwagon supporters among their fan base (Newson et al., 2021).

Therefore, the control variables age, education, gender and club magnitude have been added to the equation. Education was preferred over income due to the reluctance of respondents to answer income-related questions. Appendix D shows which football clubs have been placed in what category and why.

Coefficients of the latent ordinal variables (β_1 and β_2), directly measured ordinal variables (β_3 till β_8) and the continuous variable (β_{10}) are all interpreted in the same way: for every one unit increase on an IV, there is a predicted change of the corresponding β in the log odds of falling in a greater level of attitudinal loyalty, holding the remaining IV's constant.

More generally speaking, a positive β indicates an increased probability of obtaining a higher degree of attitudinal loyalty, whereas a negative β indicates a decreased probability. Naturally, the higher/lower the outcome of the log odds, the higher/lower the probability of obtaining a higher degree of attitudinal loyalty. Keep in mind that α is always your starting value where you have to add/detract the changes from. Unlike linear regression, there are multiple values of α, for each unique level of attitudinal loyalty.

The interpretation for coefficients of categorical variables differs slightly from the previous interpretation due to the comparison with their respective reference category (RC). Their removal out of the equation is necessary to prevent multicollinearity. Prefer not to say is the RC for gender (β_9 till β_{11}), PhD is the RC for education (β_{13} till β_{17}) and large club is the RC for club magnitude (β_{18} and β_{19}). So, the interpretation of β can be thought of as the average difference in log odds of falling in a greater level of attitudinal loyalty between a certain category and its respective RC, holding the remaining IV's constant.

Also, raising the coefficient to the power of e (exp β) can provide an alternative interpretation. If the whole equation is raised to the power of e, the log on the DV will be removed. So, only the odds ratio of the DV will remain. So, exp β can be interpreted as the change in odds ratio of the DV. Keep in mind that the odds ratio is $\frac{p}{1-p}$, of which p is the probability of falling in a greater level of attitudinal loyalty.

If exp β is greater than 1, then there is an increased probability of falling in a greater level of attitudinal loyalty. Namely, when the value is greater than 1, the numerator (p) is larger than the denominator $(1-p)$. If exp β is exactly one, the variable of the coefficient has no impact on the probability of falling in a greater level of the DV. If exp β is below one, the respective variable decreases the probability of falling in a greater level of the DV.

Finally, ε represents the error term of the equation. This model tries to predict the value of the DV with the IV's. However, these predictions are rarely precise. Hence, the ε stands for the difference between the predicted and fitted output.

3.3.3 Model equation II: Behavioral loyalty

The second model equation was constructed in the same manner with matching interpretations, but more extensive with six extra IV's for measuring the main effect. This time, the log odds of obtaining a higher degree of the latent variable behavioral loyalty were the DV and one of the new IV's (β_{14}) is an interaction variable. Below, the second equation model is shown:

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta_1 Attitudinal\ Loyalty + \beta_2 Fan\ Identification + \beta_3 Social\ Media$$
$$+ \beta_4 Team\ Content + \beta_5 Nostalgic\ Content + \beta_6 Esports$$
$$+ \beta_7 Parental\ Influence + \beta_8 Friend\ Influence$$
$$+ \beta_9 Geographical\ Location + \beta_{10} Team's\ Success$$
$$+ \beta_{11}\ Lack\ of\ Fan\ Interaction + \beta_{12} Ghost\ Games$$
$$+ \beta_{13} Relationship\ Length$$
$$+ \beta_{14}(\text{Attitudinal Loyalty} * \text{Relationship Length}) + \beta_{15}\text{Male} + \beta_{16}\text{Female}$$
$$+ \beta_{17}\ \text{Third Gender} + \beta_{18}\ Age + \beta_{19} Less\ than\ High\ School\ Degree$$
$$+ \beta_{20} High\ School + \beta_{21}\ Some\ Degree + \beta_{22}\ Bachelor + \beta_{23}\ \text{Master}$$
$$+ \beta_{24}\ Small\ Club + \beta_{25}\ Medium\ Club + \varepsilon$$

Four of the six newly added coefficients represent latent variables (β_1, β_6, $\beta_{11}$and β_{12}), which will be explained in the more detailed next subsection. β_{14} shows the interaction effect, which consists of one ordinal latent variable and one continuous variable. In essence, if the interaction effect is significantly different from zero, there is an underlying dynamic between higher levels of attitudinal loyalty and the relationship length between the fan and the club.

Keep in mind that you have to add/detract the coefficients of both variables in the interaction (β_1 and β_{13}) from the interaction coefficient itself. Only then, the net effect of an increase in both attitudinal loyalty and relationship length on the log odds of behavioral loyalty is found, remaining the other IV's constant. Therefore, the continuous variable relationship length was added to the regression to fully capture the interaction effect.

3.3.4 OLR model fit interpretation

The most commonly used method to assess the model fit for OLR is the likelihood ratio chi-square test. This test compares the -2 Log Likelihood values of the null model (i.e., intercept only model) and the final model (i.e., the shown model equation). The difference between those two is equal to the Chi-square value ($\chi2$). A significant $\chi2$ (p<.05), with taken the amount of degrees of freedom into account, makes the final model a significant improvement in fit over the null model (Petrucci, 2009).

Also, the Pearson chi-square and Deviance test were interpreted for assessing the goodness of fit of both final models. These tests measure how well the chosen models fit the data set. In contrary to the likelihood ratio chi-square test, non-significance ($p>0.05$) is desired. Namely, statistical significance would indicate that there is a discrepancy between the final model and a perfect model (Petrucci, 2009).

For assessing the model fit, the final measurement used was the pseudo R^2. In ordinary least squares (OLS) regression, the R^2 value states how much of the variability of the DV is explained by the model. The higher the R^2, the better the model predicts the DV. However, for OLR, R^2 is an invalid goodness-of-fit-statistic, because of the underlying assumption of fitting a linear model. Therefore, pseudo R^2 values were developed based on the comparison between the log likelihood of the final model and null model to approximate the accounted variability of the final model (McFadden, 1974).

The pseudo R^2 of Nagelkerke was the preferred method for this research due to covering the full probability range from 0 to 1. Namely, the Nagelkerke R^2 is an adjustment of Cox and Snell's R^2, which had a theoretical maximum value of less than one (Nagelkerke, 1991). Also, all pseudo R^2 yield lower estimates than their OLS R^2 counterparts, where Nagelkerke's R^2 was defined as the closest approximation of the OLS R^2 (Smith and McKenna, 2013).

Finally, due to the fact that pseudo R^2 only yields a lower approximation of the R^2 in OLS regression, the rule-of-thumb boundaries of R^2 will be less strict than normal. For models with latent variables in the field of marketing, R^2 values of 0.75, 0.50 or 0.25 can be generally described as substantial, moderate or weak respectively (Hair et al., 2011).

3.4 Explanation of latent variables

This subsection will shed a light on the constructs of the seven latent variables of this research, which all can be classified as reflective measurement models. Appendix A provides insight into all the survey questions with their corresponding variable name showed in brackets, which were the building blocks of the latent constructs.

3.4.1 Fan identification

Fan identification is the first latent construct that will be discussed. Figure 3 shows that four observed items were used for optimally measuring the degree of identification, building further on the work of Gladden and Funk (2001). Namely, fan identification is the latent driver of the following four observed items. First, fans who have a high degree of identification with their favorite team perceive the club as a part of themselves. Therefore, they often say ‘we’ instead of ‘they’ when fans talk about their favorite team (Ide2).

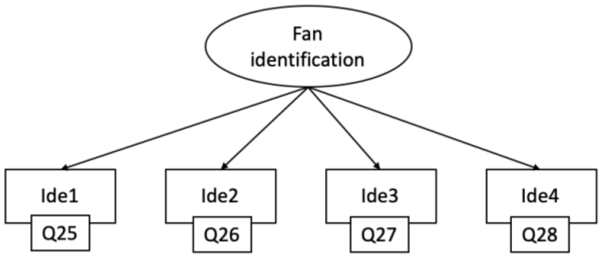

Figure 3: the over-identified latent construct of fan identification

Moreover, highly identified fans are not shy to express their club love towards friends and family (Ide4). Also, being a highly-identified fan makes praise about your team feel like a personal compliment (Ide1). Last but not least, this research extended the fan identification construct of Gladden and Funk (2001) by one item: the feeling of a personal insult when someone criticizes your team (Ide3) to fully capture the degree of emotional reactions.

3.4.2 Esports

The following latent construct for esports does not originate from other literature, but it is constructed by own ideas. Namely, the type of esports this research is interested in is very specific and as followed: club-tied esporters who play the videogame FIFA on a high level. Figure 4 shows this stated latent construct of esports.

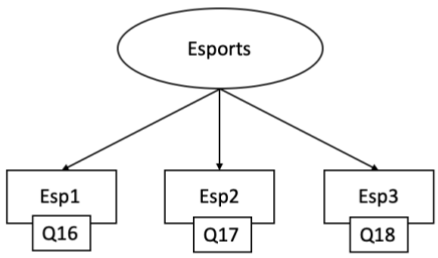

Figure 4: the just-identified latent construct of esports

Thus, if fans like this type of esports, they should not only like the videogame FIFA (Esp1) and gain a feeling of excitement of watching esports (Esp2), but they should also show interest in following club-related esports activities (Esp3). So, the attitude towards this specific kind of esports was correctly formed and measured.

3.4.3 Social media

The latent construct of social media is the predictor of three observed items, which are newly designed items to measure the general attitude of fans towards club-related social media. If this stated attitude is positive, then fans have a high degree of digital consumption for club-related social media content during their leisure time. Hereby, they do not only relax (Soc2), but also become enthusiastic about their favorite club (Soc3). This results in the expectation of a strong presence from both club and players on social media (Soc1).

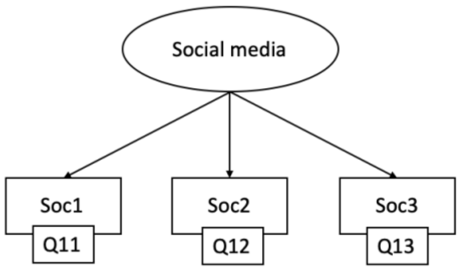

Figure 5: the just-identified latent construct of social media

Figure 5 visualizes this latent construct, which does not include any item related to the type of content. These two questions (Q14 and Q15, see appendix A) were extracted from the construct to directly measure which content type has the strongest effect on both dimensions of fan loyalty.

3.4.4 Lack of fan interaction

The latent variable lack of fan interaction is also a construct with customized items for this research, because previous researchers never used this latent construct. Naturally, before the pandemic there were little reasons to investigate the lack of fan interaction. However, the restrictions on gathering made this reasoning invalid. Figure 6 shows the latent construct of the lack of fan interaction.

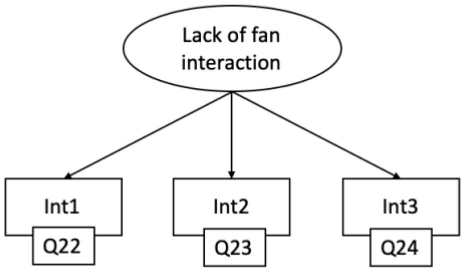

Figure 6: the just-identified latent construct of lack of fan interaction

If fans experience a notable lack of fan interaction, then they do not only miss watching football matches together with friends and/or family (Int1), but also feel more alone and thus less connected with their fellow fans (Int3). Also, the lack of fan interaction caused fans to dislike online events, which could not replace the feeling during offline meetings and events (Int2).

3.4.5 Ghost games

The latent construct of ghost games is also designed by own thoughts and beliefs instead of relying on the literature. This was simply not possible due to the lack of research on this topic. Before the pandemic, the term ghost games did not even exist, because of the constant presence of fans during football matches in the top leagues all over the world. Figure 7 shows the latent variable ghost games and the three observed items belonging to this latent variable.

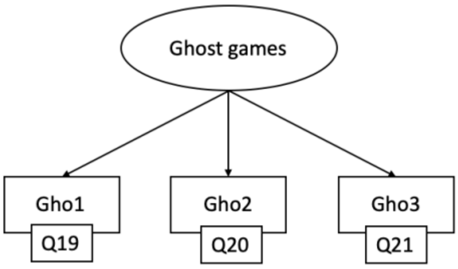

Figure 7: the just-identified latent construct of ghost games

If fans have developed a negative attitude towards ghost games, then they have experienced less joy while watching the football matches (Gho1). Also, the decrease of enjoyment of watching football matches is close to interchangeable with the opinion that ghost games are boring to watch (Gho2) and a general feeling of dissatisfaction because of the ghost games (Gho3).

3.4.6 Attitudinal loyalty

Attitudinal loyalty is the first of the two latent DV's that will be discussed. In contrary to the prior three latent constructs, attitudinal loyalty is a more established constructed in the field of research (Baldinger and Rubinson, 1996; Bauer et al., 2008; Dick and Basu, 1994). Therefore, the observed items were adapted from the existing scales of Bauer et al. (2008). Figure 8 shows the latent construct of attitudinal loyalty.

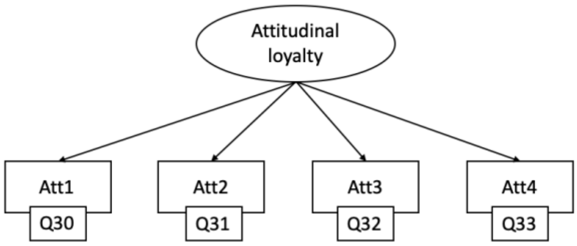

Figure 8: the over-identified latent construct of attitudinal loyalty

The attitudinal loyalty of fans is measured for two categories: their self-perceived commitment and robustness to external factors. If fans exhibit a high degree of attitudinal loyalty, their self-perceived commitment is high. In other words, they see themselves as a highly-committed fan (Att1) who supports the club for the rest of its life (Att4). Also, their commitment won't decrease when the team underperforms massively (Att2) or due to negative opinions of friends and family on their favorite club (Att3), which makes them robust to external factors.

3.4.7 Behavioral loyalty

Behavioral loyalty is the final latent construct which will be discussed in this subsection. Similar to the attitudinal dimension of loyalty, the behavioral dimension is a well-established construct in the marketing research field. Thus, the observed items are also adapted from the existing scales of Bauer et al. (2008).

However, for the behavioral dimension more adjustments were necessary. First, one observed item (Beh1) is transferred from the attitudinal to the behavioral dimension of loyalty for this research, because it is a better fit for this dimension. Namely, PWOM is a consequence of the behavioral dimension of loyalty instead of the attitudinal one (Yim and Kannan, 1999).

Also, one observed item (Beh5) is added to record the future intentions of a fan. Only then, the full concept of behavioral loyalty is captured. Last but not least, the item of physically attending football matches has been excluded from the construct, because it was not possible to attend physically attend matches for the previous season.

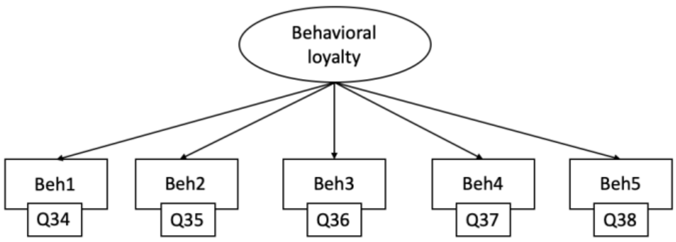

Figure 9: the over-identified latent construct of behavioral loyalty

Figure 9 shows the most comprehensive over-identified latent construct in this research with five observed items, which are all caused by fans who exhibit high levels of behavioral loyalty. First, PWOM will be conducted in public when it is necessary (Beh1). Also, they will regularly watch matches on TV (Beh2), actively read news about their team on social media (Beh3) and purchase club-related merchandise (Beh4). Finally, the future intention of the fan is measured by asking whether they will be more engaged than their club with last year (Beh5). In other words, being planning to be equally or more engaged than last year shows positive signs for future intentions.

3.5 Sample size and method

This subsection will discuss the minimum sample size required for this research to make statistical inferences about the population (i.e., football fans). In addition, the corresponding sampling method will be covered.

3.5.1 Sample size

A sample is defined as a representative subset of a population, which is used to for making statistical inferences about a large population. Namely, unlike a sample, an entire population cannot be observed and studied. However, in order to draw valid conclusions from a sample an adequate size is necessary to prevent sampling errors and bias. The tool this research used for determining the sample size was specifying the three criteria of Miaoulis and Michener (1976): sampling error, confidence level and the degree of variability.

The sampling error, sometimes referred to as the level of precision, is the range in which the true value of the population is estimated to be. Needless to say, a sample cannot exactly mimic the behavior of an entire population. This criterion is expressed in percentage points, which shows the allowed difference this research accepted between this sample and the entire population. The acceptable sampling error percentage of 5% was used for this research (Taherdoost, 2016). In other words, the result section has to be interpreted with a margin of error of +/- 5% respective to the actual population.

The second criterion of Miaouilis and Michener (1976) is the confidence level. The key idea behind this criterion is that if a population were to be sampled repeatedly, the average value of a variable (f.e., attitudinal loyalty) would be equal to the true population value. This idea is based on the Central Limit Theorem, which states that the distribution of the sample means will be approximately normally distributed (Israel, 1992).

The three most common confidence levels for marketing related research are 90%, 95% and 99%. This research opted for a confidence level of 95%, because the 99% level has a wider range to be more confident that the true population value falls within the stated interval, whereas the 90% level gives a narrower, less confident rage. Thus, the confidence level of 95% is the golden mean between the two.

A confidence level of 95% means that 95 out of 100 samples will have the true population value within the +/- 5% range of the sampling error. Furthermore, the statistical value Z, corresponding to the level of confidence required, is for a 95% confidence level equal to 1.96 (Taherdoost, 2016). This statistical value will be used shortly in the formula for determining the sample size.

The final criterion which will be specified is the degree of variability. This criterion is also a percentile scale ranging from 0 to 1 and refers to the distribution of attitudes in the population, which depends on the heterogeneity of the population (Israel, 1992). The more heterogeneous a population, the higher the variability in the attitudes and the larger the sample size required to avoid violating the sampling error of 5%.

Due to the fact that the degree of variability remains unknown prior to the survey research, Bartlett et al. (2001) suggest to use a value of 0.5, which results in the maximization of variance and produces the maximum sample size. Namely, 30% and 70% indicate that a large majority does or does not possess the attitude of interest. A degree of variability of 50% states that the distribution is the most spread out and thus equals the highest degree of variability, which will be used for this research.

Right now, the sample size formula of Israel (1992) for very large population sizes, such as the 3.5 billion football fans all over the world (Sourav, 2021), will be used for determining the minimum sample size necessary for this research. The formula is as followed:

$n_0 = \frac{Z^2pq}{e^2}$, with n_0 as the minimum required sample size, Z as the value corresponding to the level of confidence, p as the estimated proportion of an attitude that is present in the population, q as 1- p and e as the sampling error.

To summarize, a sampling error of 5%, a confidence level of 95% and a degree of variability of 0.5 were the specified three criteria of Miaoulis and Michener (1976) used for this research. Hence, the following calculation with $Z = 1.96, p = 0.5, q = 0.5$ and $e = 0.05$:

$$n_0 = \frac{Z^2pq}{e^2} = \frac{(1.96)^2(0.5)(0.5)}{(0.05)^2} = 385$$

Thus, 385 is the minimum sample size necessary to make statistical inferences about the population (i.e., soccer fans). Furthermore, this sample size does not violate rule-of-thumb assumptions for the desired sample size of the statistical methods used.

Namely, a good general rule of thumb for factor analysis is 50 respondents per factor (Pedhazur & Schmelkin, 1991). So, the seven latent variables of this research equal a minimum of 350 respondents according to Pedhazur and Schmelkin (1991). Also, Comrey and Lee (1992) define a sample size for factor analysis of 300 as good and 500 as very good. Thus, a sample size of 385 can be considered as moderately good according to this rule-of-thumb criterion.

For regression analysis, the rule-of-thumb criterion is that an absolute minimum of 10 respondents per IV is appropriate (Van Voorhis and Morgan, 2007). Both equation models pass this criterion with ease, where the second equation model, the most extensive of the two, has a total of 22 IV's, which equals a minimum sample size of 220. To conclude, the minimum sample size necessary for this research was set at 385.

3.5.2 Sampling method

Determining the minimum necessary sample size is on its own not enough for securing a representative sample and thus make statistical inferences about the population. Namely, there is a mistaken belief that a sufficient or even large sample size does automatically result into representativeness, which is described by Lantz (2012) as the large sample size fallacy. This is definitely not true. In essence, representativeness depends on the chosen sampling method, which will be described in this subsection.

In general, all sampling methods belong to one of the following two types: probability sampling or non-probability sampling. Probability sampling means that every person of the entire population has an equal chance of being included in the sample. This method is not appropriate for this research due to immense population size of approximately 3.5 billion football fans. Naturally, it was simply not possible to apply this sampling method for this research.

Therefore, two non-probability sampling methods were used to gather a sample size of both a sufficient magnitude and representativeness. In contrary to probability sampling, randomization is not important for non-probability sampling. Rather, subjective methods were used to determine which people of the population will be present in the sample. In other words, not everyone of the population has an equal chance of being in the sample.

Convenience sampling, selecting participants who are often easily and readily available, was the primary sampling method used in this research, whereas snowball sampling, participants who were recruited by other participants, fulfilled a supportive role. Convenience sampling was preferred over a similar non-probability method, called purposive sampling.

Namely, this method only seeks participants meeting pre-defined demographic characteristics, such as age, gender and education. This research wanted to obtain these characteristics without any restrictions to fully measure their effect on both dimensions of loyalty. So, OVB was prevented and the main effects were measured more precisely. Also, convenience sampling places more emphasis on generalizability than purposive sampling (Etikan et al., 2016). This is in alignment with the goal of this research of ensuring that the knowledge gained is representative of the population from which the sample is drawn.

Using convenience sampling resulted in two major benefits in the areas of time saving and representativeness (Berndt, 2020; Sharma 2017). First, respondents met numerous practical criteria, like easy accessibility and availability, which decreased the amount of time necessary for finding representative respondents. Also, persons who choose voluntary to participate are more likely to be committed and thus provide truthful responses, which increases the representativeness (Sharma, 2017).

A potential drawback of using convenience sampling is dealing with selection bias (Berndt 2020; Sharma 2017), which is a systematic error that occurs when the sample is not representative of the target population. This bias negatively affects the coefficients and is detrimental for the statistical inference of the population, which is crucial for drawing generalizing conclusions.

Two actions were conducted to minimize the selection bias of this sample. First, a screen-out question was added to the survey to avoid non-football fans filling in the questionnaire and thus prevent unrepresentative answers. Furthermore, snowball sampling is used in a supportive role to increase the number of respondents who are less accessible, such as football fans in the age group 40-65 and football fans outside the Netherlands. So, the representativeness of this sample is maximized.

3.6 Data characteristic I: validity

In the second-last subsection of this chapter, four types of the data characteristic trait validity will be discussed. The final subsection will discuss another trait, namely reliability. The third trait, relevance, is already discussed in subsections 1.2 and 1.3. In today's business environment, these three traits are vital for ensuring high quality data. If the data set does not meet the requirements of these traits, the information it contains is not valuable and thus eventually leads to invalid and irrelevant conclusions.

Validity refers to whether the measuring instrument (i.e., the survey questionnaire) measures the behavior or quality it is intended to measured and is a measure of how well the measuring instrument performs its function (Anastasi and Urbina, 1997). For determining the validity of this research, the following four types of validity will be discussed: face validity, convergent validity, predictive validity and construct validity.

3.6.1 Face Validity

The first of the four types, face validity, is described as the subjective assessment of the presentation and relevance of the survey questionnaire to rate the suitability for its intended use (Oluwatayo, 2012). In other words, are the survey questions appropriate for constructing latent variables and predicting both dimensions of fan loyalty? Based on the face validity criteria of Sürücü and Maslakci (2020), several steps have been taken to satisfy those criteria and achieve a saturated level of face validity, as shown in Table 2.

Criteria	**Step A**	**Step B**	**Step C**
All statements are appropriate for the (latent) variable it measures.	Theoretical evidence	Researcher's input	Peer-review
The response scale is clearly understood by respondents.	Close-ended questions	7-point Likert scale	Peer-review
All statements are clear and understandable.	No double-barreled questions	Short and straightforward	Peer-review
The survey design is attractive.	Mobile-friendly	Grouping questions	Peer-review
The difficulty level of the questions is appropriate for the level of the respondents.	Simple wording	Attitude-related questions	Peer-review

Table 2: five criteria for face validity and its corresponding executed steps

For satisfying the first criteria, a mix of theoretical evidence, researcher's input and peer-review was used. Some constructs relied on theoretical evidence, whereas other were newly designed constructs for this research (see subsection 3.4). Also, pre-testing was a necessary step before distributing the survey and eventually applying statistical analyses on the data set to prevent invalid results. Pre-testing helped in the process of assessing all the five criteria.

One of the reasons that the survey was pre-tested and thus peer-reviewed was to gain some initial insights into the appropriateness of the observed items for measuring the latent constructs. In contrary to the actual respondents, the pre-tested survey respondents exactly saw which observed items would belong to what latent variable. By doing so, they could assess the appropriateness for both types of latent constructs, which contributed to achieving face validity.

The second criterion refers to the used response scale of the survey, which mainly consists of close-ended questions (see appendix A). Namely, close-ended questions are not only easier and faster filled in by respondents, but also provide a more accurate representation of their attitudes (Reja et al., 2003). Moreover, close-ended questions are more suitable for statistical analyses. The reasoning behind the choice for the 7-point Likert scale was already thoroughly explained (see subsection 3.1.2).

The third criterion specifically examines the design of the statements (i.e., survey questions). All the statements were no longer than one sentence and respondents were not forced to answer two questions with a single answer. In other words, the questions were straightforward and double barreled questions were avoided to prevent inaccurate results.

The fourth criterion is, unlike the other four, not related to the survey questions themselves, but to the survey design in general, of which mobile-friendliness is an important theme. One feature, 'click and go', specifically contributed to the attractiveness of the survey on mobile devices. If respondents answered one statement, the statement with its response scale would close and the next statement automatically opens. This feature enhances the scrolling process of the respondents, which is an important indicator of faster response times and a higher completion rate (De Bruijne and Wijnant, 2014; Mavletova and Couper, 2014).

Furthermore, the survey questions were built up in a logical order: starting with an introduction, proceeding with demographic questions and ending with all the statements. Also, the same question sequence was used for every respondent, where the statements were grouped together, without explicitly naming every latent construct, to minimize confusion and maximize a smooth survey experience.

Finally, the difficulty level of the survey will be assessed. For this research, there was no minimum level of education required to take part in this survey. As long as the respondent perceived him- or her selves as a football fan, the survey could be filled in completely. Therefore, simple wording is used to prevent potential ambiguous interpretations by respondents, which could lead to invalid results.

Also, the content of the questions it selves is not overly demanding, because they simply ask the personal attitude of the respondent. There is no right or wrong answer, which lowers the difficulty level significantly.

3.6.2 Convergent validity

The second measurement of validity, convergent validity, can be seen as the prequel of construct validity. Namely, correlation analysis was used to verify that the observed items which measure the same latent construct highly correlate. A more in-depth explanation of the type of correlation used, polychoric correlation, can be found in subsection 3.2.

The output of this analysis are correlation coefficients, which are summary values of a large set of data with a limited range [-1,1] representing the degree of linear association between two measured variables (Taylor, 1990). Absolutely seen, the larger the correlation coefficient, the stronger the correlation and thus relationship between the two variables.

In general, a correlation coefficient lower than 0.3 implies a weak correlation, 0.3-0.5 implies moderate and greater than 0.5 equals a strong correlation (Heale and Twycross, 2015). So, for convergent validity this research strives for values of 0.5 and higher between items which measure the same latent construct. The corresponding correlation analysis can be found in subsection 4.4.1.

3.6.3 Predictive validity

The third type of validity, predictive validity falls under the same category as convergent validity, called criterion validity. Criterion validity analyzes the correlation coefficients to determine whether the data is valid or not, where predictive validity is obtained by interpreting the correlation coefficients between the predicting (IV) and predicted (DV) variable (Sürücü and Maslakci, 2020).

So, for this research, correlation analysis will be applied on both the model equations to see whether the IV's have a strong correlation with the respective DV (i.e., attitudinal or behavioral loyalty). Therefore, predictive validity is seen as the prequel of OLR, because it is the first analysis which investigates the hypothesized relationships. Naturally, the coefficient criteria are the same as those for assessing convergent validity. The corresponding correlation analyses can be found in subsections 4.4.2 and 4.4.3.

3.6.4 Construct validity

The final type of validity which will be assessed for this research is construct validity, which is described as the extent to which the statements of the survey questionnaire measure the intended latent constructs (Heale and Twycross, 2015). CFA is the most suitable statistical technique for assessing this specific type of validity (Oluwatayo, 2012). The seven latent variables and their corresponding statistical analysis in the form of CFA were already thoroughly discussed (see subsections 3.2 and 3.4). In subsection 4.6 the CFA of this paper can be found.

3.7 Data characteristic II: Reliability

The second and also final data characteristic trait which will be discussed is reliability. Even though reliability and validity are closely related, they are not interchangeable. Reliability refers to the consistency over time of the results. A reliable research must demonstrate that if it were to be carried out on a similar group of respondents in a similar context, similar results would be obtained (Oluwatayo, 2012).

Also, reliability does not automatically result in guaranteeing validity for this research (Sürücü and Maslakci, 2020). Even if a survey questionnaire produces reliable (i.e., consistent) outcomes, the result could be invalid. For example, a clock time which is always 30 minutes off the actual time shows an invalid time, but it is classified as reliable due to its consistency over time. On the other hand, a valid survey questionnaire is likely to be reliable.

Heale and Twycross (2015) identified three attributes of reliability: internal consistency, stability and equivalence. The first attribute, internal consistency, tests for the homogeneity of the observed items of the survey questionnaire. In other words, how consistent are the observed items in predicting the latent variable?

This research made use of the most popular method in the field for testing the internal consistency, namely the determination of Cronbach's alpha coefficient. Cronbach (1951) developed a formula for testing the internal consistency of observed items which are not measured by a dichotomous response scale, such as Yes/No or True/False. This makes Cronbach's alpha an excellent fit for the observed items, measured by a 7-point Likert scale, of this research.

Cronbach's alpha usually has a limited range of [0,1] of which a higher value equals higher internal consistency (Sürücü and Maslakci, 2020). Furthermore, Gliem and Gliem (2003) provided rules of thumb for interpreting Cronbach's alpha specifically for Likert scale items, which is shown in Table 3.

Cronbach's alpha	**Interpretation**
>.9	Excellent
>.8	Good
>.7	Acceptable
>.6	Questionable
>.5	Poor
<.5	Unacceptable

Table 3: rules of thumb for interpreting Cronbach's alpha

Even though a higher value of Cronbach’s alpha is partially caused by an increasing number of observed items for measuring one latent variable, the returns are diminishing (Gliem and Gliem, 2003). Therefore, this research strives to achieve a Cronbach’s alpha of at least 0.8 to secure a good level of internal consistency.

The second attribute of reliability, stability, can be assessed by test-retest reliability, which tests the consistency over time of the same respondents (Sürücü and Maslakci, 2020). In other words, will the same respondents give the same answers when they fill in the questionnaire again with an interval of 1-2 weeks? Unfortunately, it was not feasible for this research to identify all the respondents due to the fact that the survey was filled in anonymously. Therefore, the exact sample group could not be formed again, resulting in an inability of executing the test-retest.

The third and also final attribute of reliability, equivalence, refers to the consistency of responses between alternative measuring forms (Heale and Twycross, 2015). In other words, will an alternative survey measuring the same latent constructs but with other statements yield similar results? The alternative forms method is the appropriate analysis for assessing this type of validity (Sürücü and Maslakci, 2020). However, for this method another sample with a minimum of 385 respondents is required. The main sampling method used, convenience sampling, did not have the distribution power of reaching another 385 respondents. Therefore, this method remains also unused for this research.

Even though two of the two of the three reliability attributes could not be directly assessed, two important questions have been added to the survey questionnaire with the goal of maximizing the consistency over time. The survey started with a screen-out question to enhance the data quality (Q1, see appendix A). Namely, the respondent was asked whether he did or did not perceive himself as a fan of a football club.

If the answer on the question was ‘No’, the survey would end immediately. Removing non-football fans from the sample was the first step for securing not only consistency over time, but also getting relevant and valid results. Namely, non-football fans can not relate themselves to the statements in the questionnaire, which would result in random, unreliable answers.

Furthermore, inattentive respondents were also removed from the sample if they wrongfully answered the attention question (Q29, see appendix A). Inattentive respondents were undesired, because they introduce noise into the data set, weaken correlations between the observed items and increase the likelihood of finding insignificant results (Berinsky et al., 2019). Thus, removing inattentive respondents contributed to the minimization of random, unreliable filled in surveys and should increase the consistency over time.

4. Results

After thoroughly explaining the used methods and providing supportive arguments for their usefulness in this research, the result section sets the key results out. The first five subsections will discuss the outcome of all the descriptive analyses. Afterwards, the results of the two-main statistical analysis, CFA and OLR, will be shown and interpreted correctly.

4.1 Descriptive analysis I: screeners

In total the survey questionnaire was filled in completely 525 times in a time span of two weeks. Naturally, not every recorded responded answered the two screening questions in the desired way, which results in a smaller, but more representative sample size. This subsection will analyze the descriptive statistics related to those two screening questions (Q1 and Q29, see appendix A).

4.1.1 Screener I: do you consider yourself as a fan of a football club?

The first screening question asked the respondents whether they perceive themselves as a football fan. If they answered 'Yes', the respondents could continue filling in the survey. However, answering 'No' resulted in an immediate jump to the end of the survey.

	Frequency	**Percent**
Yes	442	84.2
No	83	15.8
Total	525	100.0

Table 4: frequency table of screener I: do you consider yourself as a fan of a football club?

Thus, the unrepresentative respondents who answered 'No' to this question were excluded from the sample. In total 15.8% of the respondents answered 'No', which results in an exclusion of 83 respondents caused by the first screening question, as shown in Table 4.

4.1.2 Screener II: attention question

So, 442 respondents are left who filled in the questionnaire completely. However, not all of them filled in the second screening question correctly, which resulted in a second batch of excluded respondents. This time, 9.5% of the respondents did not answer 'Strongly disagree' to the attention question, which lead to another exclusion of 42 respondents, as shown in Table 4.

	Frequency	Percent
Strongly disagree	400	90.5
Disagree	18	4.1
Neither agree nor disagree	7	1.6
Somewhat agree	7	1.6
Agree	3	0.7
Strongly agree	7	1.6
Total	442	100.0

Table 5: frequency table of Screener II: choose the option 'Strongly disagree' to show that you still pay attention

4.1.3 Response time

Right now, 125 of the 525 respondents are excluded from the sample, which means that the current sample size consists of 400 respondents who passed both screening questions. Their average response time was just over 10 minutes with an extremely wide range of [1.38, 1021], as shown in Table 6. However, the necessary time for filling in this survey was estimated at a maximum of 7 minutes.

	Mean	Std. Dev	Minimum	Maximum
Response time in minutes (N=400)	10.2	63.6	1.38	1024
Response time in minutes (N=395)	4.4	3.6	1.38	1024

Table 6: descriptive statistics of the response time in minutes with and without outliers

Therefore, the response times were categorized to detect any possible outliers. Table 7 shows the seven designed categories for response time, of which the latest category is qualified as the outlier category. Namely, after 32 minutes the next highest response time was just over 200 minutes with the highest value being 1024 minutes.

Also, table 6 shows the standard deviation of the response time. This descriptive statistic tells you how spread out the data is. It is a measure of how far each observed value is from the mean. In any distribution, approximately 95% of all the values are within two standard deviations of the mean (Fisher and Marshall, 2009).

Response time	Frequency	Percent	Cumulative Percent
1-3 minutes	144	36.0	36.0
3-5 minutes	163	40.8	76.8
5-7 minutes	49	12.3	89.0
7-9 minutes	13	3.3	92.3
9-11 minutes	8	2.0	94.3
11-32 minutes	18	4.5	98.8
>200 minutes	5	1.3	100.0
Total	400	100.0	

Table 7: frequency table of the response time in minutes categorized

With the removal of the outliers, the central tendency moves towards a more realistic value with a mean of 4.4, the spread of the responses decreases massively with a more representative standard deviation of 3.6 instead of 63.6 with the outliers (see Table 6). Moreover, 89% of the respondents filled in the survey within the estimated time of 7 minutes (see Table 6).

Even though the response time of the five respondents were classified as outliers, the recorded responses of them were not unrealistic and there was no reason to remove them from the sample. There is a high likeliness they took a break and reopened the survey after a while again to finish it.

4.1.4 Missing values

To conclude, the sample size of 400 respondents will not decrease any further. Namely, partial responses were deleted after one week and thus never present in the data set, which unfortunately lead to the inability of analyzing the non-response rate.

Moreover, without taking into account the respondents who failed to pass the first screening question, the data set did not have any missing values. Making the questions mandatory to fill in led to the removal of the possibility of encountering missing values. Also, no further outliers were detected in the data set, which makes sense since the vast majority of the questions were close-ended and measured by a 7-point Likert scale.

4.2 Descriptive analysis II: control variables

With the sample size set, the second part of the descriptive analyses will shift its focus towards the control variables in the model equations. Keep in mind that every analysis from now will work with a sample size of 400. Age, club magnitude, gender and education are the four control variables which will be analyzed in this subsection.

4.2.1 Control variable I: age

First of all, the age of the respondents was asked in an open-ended fashion with an allowed range of [1,99] to prevent outliers. However, conducting descriptive analysis on the actual age itself will be too extensive. Therefore, the variable is categorized to gain better insights about the age distribution among the respondents, as shown in Table 8.

Age group	Frequency	Percent	Cumulative Percent
12-17	1	0.3	0.3
18-24	168	42.0	42.3
25-34	109	27.3	69.5
35-44	41	10.3	79.8
45-54	45	11.3	91.0
55-64	30	7.5	98.5
65-74	4	1.0	99.5
75-84	2	0.5	100.0
Total	400	100.0	

Table 8: frequency table of the variable age, categorized in seven age groups

Table 8 shows that the age group 18-24 is represented the most in this survey. With 42%, almost half of the respondents of this sample size belonging in this age group. Also, the next age group 25-34 has the second largest frequency with 109 respondents belonging to this age group, which equals a percentage of 27.3%. So, 69.2% of all the respondents belonged in one of these two age groups.

So, it is clear that the respondents are nowhere near a normal distribution of the age groups. The young-adults are overrepresented in this sample size, whereas the middle-aged adults (35-60) and old adults (60+) are underrepresented. This leads to a left skewed distribution, as shown in Appendix E.

	Mode	Median	Mean	Std. Deviation	Min	Max
Age	22	27	32.0	13.3	16	81

Table 9: several descriptive statistics for the variable age

Table 9 shows all of the three central tendency measures for the variable age: mode, median and mean. To be clear, the mode is the age with the highest frequency, the median is the middle age and the mean is the average age. For an even sample size (N=400), the median is the average of the middle value and the next value. Otherwise, the data cannot be split equally in halves.

In general, if the distribution is skewed to the left, the three central tendency measures order themselves in a certain way. Namely, the median will be higher than the mean, but lower than the mode. For the variable age the mode is 22, the median is 27 and the mean is 32 (see Table 9), which confirms a left skewed distribution. Additionally, the left-skewed distribution results in a relatively high standard deviation of 13.3 on a range of [16,81].

4.2.2 Control variable II: club magnitude

Secondly, descriptive statistics of the control variable club magnitude will be shown. During the questionnaire, the respondents were asked what their favorite club is (Q4, see appendix A). All of the clubs which were mentioned in the survey, were placed into one of the three following categories of club magnitude: small club, medium club or large club. Appendix C will show exactly which clubs were placed into what category.

Club Magnitude	Frequency	Percent
Small	54	13.5
Medium	54	13.5
Large	292	73.0
Total	400	100.0

Table 10: frequency table of the variable club magnitude

Most of the respondents (N=292) in the sample size supported a large club, which is to be precise 73% of the total respondents. Coincidentally, there is an equal number of respondents (N=54) who support a small club as respondents who support a medium club. Together, they represent the club magnitude of the remaining 27% of the respondents (see Table 10). Naturally, large clubs have a larger fan base than medium and small clubs, which makes it logical that the vast majority of the respondents support a large club.

4.2.3 Control variable III: gender

Thirdly, the control variable gender and its corresponding descriptive statistics will be analyzed. Respondents could choose between the following four option when asked about their gender (Q2, see Appendix A): male, female, non-binary/third gender and prefer not to say.

Gender	**Frequency**	**Percent**	**Cumulative Percent**
Male	311	77.8	77.8
Female	85	21.3	99.0
Non-binary/Third gender	2	0.5	99.5
Prefer not to say	2	0.5	100.0
Total	400	100.0	

Table 11: frequency table of the variable gender

Table 11 shows the frequency table of the variable gender with an expected outcome. First of all, in a male-focused sport it is no surprise that the majority of the respondents is male. Even though the woman branch of football is rising, the number of practitioners of the sport nor the number of fans come close to the numbers of their male counterparts. Thus, this research reflects reality with 77.8% of sample size consisting of males against 21.3% of the respondents being a female (see Table 11).

Also, 99% of the sample size classified themselves as either a male or female, whereas the other remaining 1% of respondents was equally divided between the two remaining options. Namely, two respondents classified themselves as non-binary/third gender, whereas the other two respondents preferred to keep their gender anonymous (see Table 11).

4.2.4 Control variable IV: education

The fourth and thus final control variable of this research is education, of which the descriptive statistics will be discussed right now. Respondents were asked what their highest obtained level of education is (Q6, see appendix A). They could choose between the following six options: less than high school degree, high school degree, some college degree, bachelor's degree, master's degree and PhD.

Education	Frequency	Percent	Cumulative Percent
Less than high school degree	16	4.0	4.0
High school degree	85	21.3	25.3
Some college degree	110	27.5	52.8
Bachelor's degree	149	37.3	90.0
Master's degree	33	8.3	98.3
PhD	7	1.8	100.0
Total	400	100.0	

Table 12: frequency table of the variable education

The two middle levels of education, some college degree and bachelor's degree, were the two most frequent obtained level of education, as shown in Table 12. In total, 64.8% of the respondents obtained either some college degree or bachelor's degree, where the latter is also the most frequently observed level of education with 149 respondents choosing this level of education, which is equal to 37.3% of all the respondents.

Furthermore, the two lowest obtained levels of education, less than high school degree and high school degree, together more than double the frequency of the two highest levels of education (i.e., master's degree and PhD). Table 12 shows that adding the frequencies of the lowest two levels results in a total of 101 respondents, whereas the frequency sum of the highest two level is no higher than 40.

To conclude, the majority of the total respondents have some college degree or a bachelor's degree, whereas the majority of the remaining minority has a ratio advantage of 2.25:1 in favor of the two lowest levels of education.

4.3 Descriptive analysis III: main effect and DV's

Part three of the descriptive analyses will focus on the descriptive statistics of the main effects and DV's in the regression analysis. Model equation II (see subsection 3.3.3) is the most extensive equation of the two, which makes the choice logical for analyzing the main effects of that equation and its DV (i.e., behavioral loyalty). For latent variables measuring the main effect, the descriptive statistics of their corresponding observed items will also be shown.

Besides, all the main effects of model equation I are also in the second model equation, even the DV of model I (i.e., attitudinal loyalty) is included as an IV measuring the main effect. In other words, descriptive statistics of all the main effects and both DV's of this research will be covered in this subsection.

4.3.1 DV I: behavioral loyalty

First of all, the descriptive statistics DV of model equation II, behavioral loyalty will be analyzed. Behavioral loyalty is a latent variable, which was constructed by five observed items (see subsection 3.4.7). Due to the ordinal nature of the observed items, other descriptive analysis methods have to be used than mean and standard deviation. Namely, these measurement methods for central tendency and variability cannot be applied on ordinal data (Bertram, 2007).

	Median	**Q3**	**Q1**	**IQR**
Behavioral loyalty	5.2	5.8	4.6	1.2
Beh1	5.0	6.0	4.0	2
Beh2	6.0	7.0	5.0	2
Beh3	6.0	6.0	5.0	1
Beh4	4.0	5.0	3.0	2
Beh5	5.0	6.0	4.0	2

Table 13: descriptive statistics of the latent variable behavioral loyalty and its five observed items

Therefore, the mean is replaced by the median and standard deviation is replaced by the interquartile range (IQR), which is another measurement method of statistical dispersion. The IQR is calculated by detracting the latest point of the first quartile (Q1) from the latest point of the third quartile (Q3). Q1 contains the first 25% of the data, whereas Q3 contains the first

75% of the total data. So, detracting Q1 from Q3 results in getting the range of the middle 50% of the data. The lower the IQR, the smaller the range and thus the lower the variability (Fisher and Marshall, 2009).

So, two of the five observed items for behavioral loyalty, Beh2 and Beh3, had the joint highest central tendency value with a median of 6 (see Table 13). In other words, the respondents expressed the most agreement on regularly watching matches of their favorite football club on TV and on actively reading news related to their favorite team on social media. The least agreement was shown on Beh4 (see Table 13), which means that fans are more likely to watch football matches and/or read social media news than purchasing club-related merchandise.

As for the IQR's of the observed items, Beh3 has the lowest of them all with an IQR of 1 against an IRQ of 2 for the other four observed items (see Table 13). To put it another way, the range of the middle 50% of Beh3 is the lowest with [5,6], which means that this observed item has the lowest variability of the five.

To conclude, the latent variable behavioral loyalty, constructed by taking the means of the five observed variables has a median of 5.2 with an IQR of 1.2 (see Table 13), which shows a slight agreement of exhibiting behavioral loyalty. Also, the IQR of 1.2 means that the middle 50% of the respondents exhibit similar levels of behavioral loyalty.

4.3.2 DV II: Attitudinal loyalty

The second main effect which will be discussed is the latent variable attitudinal loyalty, which is constructed by taking the average of its four observed items (see subsection 3.4.6). Also, attitudinal loyalty is measuring the main effect in model equation II, but serves as the predicted variable (DV) in model equation I.

	Median	Q3	Q1	IQR
Attitudinal loyalty	5.75	6.25	5.0	1.25
Att1	5.0	6.0	5.0	1
Att2	6.0	6.0	5.0	1
Att3	6.0	6.0	5.0	1
Att4	6.0	7.0	5.0	2

Table 14: descriptive statistics of the latent variable attitudinal loyalty and its four observed items

Three of the four observed items (Att2, 3 and 4) of attitudinal loyalty have a median of six (see Table 14), which means that there is an overall very strong agreement among the respondents on their robustness to external factors, such as opinions of relatives and/or the team's performance. Also, most respondents indicated strong agreement with the statement of supporting the club for the rest of their lives. Their self-perception of being a real fan was slightly less strong than the other three observed items with a median of five (see Table 14).

Not only a high degree of similarity was found when comparing the medians of the observed items, but Table 14 shows that the IQR values were also the same for three of the four observed items (Att1, 2 and 3). Those three observed items all have an IQR value of 1 and a range of [5,6]. This low IQR value only strengthens the overall consensus on agreeing with the statements.

Even though Att4 is the only observed item with an IQR value of 2, the level of agreement remains strong due to its range of [5,7] (see Table 14). For example, an IQR of 2 with a range of [5-7] is a stronger indicator of agreement than an IQR of 2 with a decrease of the extreme values of 1, resulting in a range of [4,6].

So, the overall strong consensus on agreeing with the statement related to attitudinal loyalty, has naturally led to strong levels of exhibiting attitudinal loyalty with a median of 5.75 and a respectable IQR of 1.25 with a range of [5, 6.25] (see Table 14), which shows that the middle 50% of the respondent have a similar degree of attitudinal loyalty.

4.3.3 Main effect I: fan identification

Thirdly, the descriptive statistics of the latent variable fan identification and its four observed items (see subsection 3.4.1) will be analyzed. Fan identification is a variable present in both model equations with the same functioning: measuring the main effect.

	Median	**Q3**	**Q1**	**IQR**
Fan identification	5.25	6.0	4.5	1.5
Ide1	6.0	6.0	5.0	1
Ide2	5.0	6.0	4.0	2
Ide3	5.0	6.0	4.0	2
Ide4	6.0	6.0	5.0	1

Table 15: descriptive statistics of the latent variable fan identification and its four observed items

For the four observed items of fan identification, two pairs are formed which have the same median, IQR and range, where the pair of Ide1 with Ide4 shows stronger values on their respective levels of agreement than the pair of Ide2 with Ide3. Namely, respondents expressed the highest level of agreement on Ide1 and Ide4 with a median of 6 (see Table 15).

This means that there is an overall consensus of agreeing on the experience of a personal compliment when someone praises your team and most friends and/or family know that the respondents are a committed fan of their team. This is strengthened by an IQR of 1 with a corresponding range of [5,6] (see Table 15), which shows that there is a low degree of variability among the middle 50% of the respondents for these two observed items.

For the other pair, Ide2 with Ide3, there is less overall agreement with a median of 5 (see Table 15). Not only the overall level of agreement was lower for taking criticism about your team as a personal insult and saying 'We' instead of 'They' when talking about your club, but there was also less consensus on this lower level of agreement. Namely, both Ide2 and Ide3 have an IQR of 2 with a range of [4,6] (see Table 15), which shows a moderate degree of variability among the middle 50% of the respondents.

So, despite the process of identifying two pairs for the four observed items, the latent variable fan identification did not simply have the average value derived from the descriptive statistics of the four observed items. Even though the IQR of 1.5 with a range of [4.5,6] is the average value for the degree of variability, the median of 5.25 meant that the latent variable shifted more towards the pair with the lowest level of agreement, as shown in Table 15. Overall, respondents exhibited moderate levels of fan identification combined with the same degree in variability as agreement on their degree of fan identification.

4.3.4 Main effect II: social media

Fourthly, descriptive statistics of the building blocks and their latent construct social media (see subsection 3.4.3), measured by taking the average of its three observed items, will be analyzed. For all three observed items the median is equal, with a value of 6 (see Table 16). This means that the respondents expressed a strong agreement on all the three dimensions of social media. They relaxed and enjoyed during consuming social media content, which leads to an expected strong presence of the club and players on this medium.

	Median	Q3	Q1	IQR
Social Media	6.0	6.3	5.0	1.3
Soc1	6.0	7.0	5.0	2.0
Soc2	6.0	6.0	5.0	1.0
Soc3	6.0	6.0	5.0	1.0

Table 16: descriptive statistics of the latent variable social media and its three observed items

Also, the overall consensus of agreeing on this latent construct is strong. The observed items Soc2 and Soc3 both have a IQR of 1 with a range of [5,6] (see Table 16), which shows a low degree of variability among the middle 50% of the respondents. As for Soc1, there is a higher degree of variability with a IQR of 2 and a range of [5,7] (see Table 16).

However, there is some disagreement on the degree of agreeing with the statement of expecting a strong presence on social media of your favorite team: some extremely agree, whereas other respondents in the middle 50% of the data only somewhat agree.

In contrary to the latent variable fan identification, the latent construct of social media has descriptive statistics exactly averaging the descriptive statistics of its three observed items with a median of 6, an IQR of 1.3 with a range of [5,6.3], as shown in Table 16. So, respondents expressed an overall positive attitude towards club-related social media content with a low degree of variability among the middle 50% of the respondents.

Type of social media content	Median	Q3	Q1	IQR
Team-related content	6.0	6.0	5.0	1.0
Nostalgic content	5.0	6.0	5.0	1.0

Table 17: descriptive statistics of the variables team-related content and nostalgic content

Fifthly, the descriptive statistics of the variables team-related content and nostalgic content will be analyzed, which were both variables measured by one single-item and present in both model equations.

The respondents agreed with liking both types of content, but they expressed a stronger level of agreement on liking team-related content with a median of six against a median of five for the variable nostalgic content, as shown in Table 17. Also, both variables have an equal IQR of 1 with an also identical range of [5,6] (see Table 17), which shows a low level of variability.

However, the medians were not equal, which means that among the middle 50% of the respondents the difference is not in level of variability, but in a stronger level on agreeing with team-related content compared to nostalgic content. In other words, more respondents strongly agreed with liking team-related content than liking nostalgic content.

4.3.5 Main effect V: esports

Esports is the following latent variable of which its descriptive statistics will be analyzed, together with its three observed items (see subsection 3.4.2). The variable esports is only measuring the main effect in model equation II and thus not present in model equation I.

	Median	**Q3**	**Q1**	**IQR**
Esports	4.0	5.3	2.7	2.6
Esp1	5.0	6.0	4.0	2.0
Esp2	4.0	5.0	2.0	3.0
Esp3	4.0	5.0	2.0	3.0

Table 18: descriptive statistics of the latent variable esports and its three observed items

In contrary to the previous (latent) variables, there is no strong consensus on either agreeing or disagreeing with the statements. Esp2 and Esp3 have a median of 4, an IQR of 3 with a range of [2,5], as shown in Table 18. This means that the respondents did express nor a strong level of agreement nor disagreement with a high degree of variability. Namely, the middle 50% of the respondents shows lots of dispersion, varying from disagreement to somewhat agreeing with their interest of following esports activities of their club and getting excited by watching esports.

As for Esp1, the respondents expressed a low degree of agreement on liking the videogame FIFA. Namely, the median of Esp1 was 4 and the IQR was 2 with a range of [4,6], as shown in Table 18. Even though Esp1 has a higher median than its related observed items Esp2 and

Esp3, the level of agreement is still weak and an IQR value of 2 still results in a relatively high degree of variability among the middle 50% of the respondents.

Naturally, the output of its observed items resulted in a neutral median of 4, which shows that the latent construct shifted more towards Esp2 and Esp3, because the median is not the average value of the three medians (see Table 18). So, respondents expressed a neutral feeling towards liking FIFA-related esports activities of their team. Needless to say, a neutral attitude goes together with a relatively high IQR of 2.6 with a range of [2.7,5.3], which means that the dispersion is high among the middle 50% of the respondents towards their attitude on esports.

4.3.6 Main effect VI: reasons of initially becoming a fan

The next descriptive statistics will cover four variables, all measured by a single-item. Namely, these four variables are reasons of initially becoming a fan of a football club. Respondents were asked to share their respective level of importance for each of the four reasons which all can potentially contribute to initially becoming a fan. These four variables measure the main effect in both model equations.

Reason for initially becoming a fan	**Median**	**Q3**	**Q1**	**IQR**
Parental influence	5.0	6.0	3.0	3.0
Friend's influence	4.0	5.0	3.0	2.0
Geographical location	5.0	6.0	3.0	3.0
Team's success	4.0	5.0	3.0	2.0

Table 19: descriptive statistics of the four reasons for initially becoming a fan

As for the observed items of fan identification, it was possible to make two pairs with an identical median and IQR. Right now, parental influence forms the first pair together with geographical location, whereas friend's influence and the team's success form the second pair.

For the first pair, the respondents expressed somewhat of an agreement on finding these two reasons important with their median of five (see Table 19). However, there was no strong consensus on this level of agreement, shown by an IQR of 3 with a range of [3,6] (see Table

19). So, the degree of dispersion was high for these variables, with some respondents of the middle 50% disagreeing with its importance, whereas others agreed with the level of importance.

As for the second pair, the respondents expressed a neutral feeling towards these two reasons of becoming a fan by neither agreeing nor disagreeing with their level of importance. Even though the degree of variability was lower with an IQR of 2 and a range of [3,5] (see Table 19), the variability is still relatively high and the general consensus is low. Namely, some respondents of the middle 50% somewhat agreed with the statements, whereas others disagreed.

4.3.7 Main effect VII: lack of fan interaction

The second-last latent variable of which its descriptive statistics and those of its three observed items will be analyzed, is the lack of fan interaction. In subsection 3.4.4 more in-depth explanation on the latent construct can be bound. The lack of fan interaction is a variable which is only measuring the main effect in model equation II.

	Median	**Q3**	**Q1**	**IQR**
Lack of fan interaction	5.0	5.7	4.0	1.7
Int1	5.0	6.0	4.0	2.0
Int2	5.0	6.0	4.0	2.0
Int3	5.0	5.0	3.3	1.7

Table 20: descriptive statistics of the latent variable lack of fan interaction and its three observed items

All of the three observed items have an equal median of 5, but the IQR values are only equal for Int1 and Int2, as shown in Table 20. In other words, the respondent's expression of somewhat agreeing with all the indicators for the lack of fan interaction did not led to equal levels of dispersions.

Int1 and Int2 had an IQR of 2 with a range of [4,6] (see Table 20), which means that even though the variability is relatively high, no respondents in the middle 50% of the data expressed some form of disagreement on the statements related to the lack of physically meeting with other fans and watching football matches together.

As for Int3, the IQR of 1.7 is lower than the IQR of 2 than the pair of Int1 and Int2, as shown in Table 20. However, this lower level of dispersion does this time not mean that the overall consensus is stronger. Namely, the range is [3.3, 5.0], which shows that the neutral threshold of 4 is passed in the middle 50% of the respondents.

In other words, for this observed item some respondents slightly disagreed with the fact that they felt less connected with other fans than prior to the pandemic. The other two observed items did not pass this threshold in their respective interquartile range (see Table 20).

As for the latent variable lack of fan interaction, the median was obviously equal to the three observed items and the IQR was equal to 1.7, as shown in Table 20. However, even though the IQR's of Int3 and the latent construct itself are equal, the range differs. Namely, the latent construct has a range of [4, 5.7] (see Table 20), which shows that for the middle 50% of the respondents no one expressed a level of disagreement on the current lack of fan interaction during the pandemic.

4.3.8 Main effect VIII: ghost games

The construct of ghost games and its three observed items (see subsection 3.4.5) is the last latent construct of which its descriptive statistics will be analyzed. The latent variable ghost games is only present as an IV in model equation II, where it contributes in measuring the main effect of the regression.

	Median	**Q3**	**Q1**	**IQR**
Ghost games	4.0	5.3	2.3	3.0
Gho1	5.0	6.0	2.0	4.0
Gho2	4.0	5.0	2.0	3.0
Gho3	4.0	5.0	2.0	3.0

Table 21: descriptive statistics of the latent variable ghost games and its three observed items

Out of the three observed items, two form a pair with an identical median, IQR and range. Namely, Gho2 and Gho3 have a median of 4, an IQR of 3 and a range of [2, 5], as shown in Table 21. So, respondents expressed a neutral feeling by neither agreeing nor disagreeing with gaining a general feeling of dissatisfaction and feeling bored while watching ghost games. Also, the range of the IQR goes through the neutral threshold of 4, which shows that there is no general consensus at all among the middle 50% of the respondents combined with a high degree of variability.

As for Gho1, the median of 5 (see Table 21) indicates that respondents expressed some agreement on deriving less joy from watching football matches than prior to the pandemic. However, of all the observed items, Gho1 is the most polarized one with an IQR of 4 and a range of [2, 6] (see Table 21). This means that there is absolutely no strong consensus on slightly agreeing with the statement of Gho1. The variability is extremely high, because of the middle 50% of the respondents some strongly disagree, whereas others strongly agree with the statement.

So, the polarization of the answers for the three observed items of ghost games leads, unsurprisingly, to a neutral median of 4, an IQR of 3 and a range of [2.3, 5.3], as shown in Table 21. To conclude, the overall expression of the respondents towards the ghost games remain neutral with a high variability. Namely, in the middle 50% of the respondents some disliked ghost games, whereas other respondents did not have negative experiences while watching ghost games.

4.3.9 Main effect IX: relationship length

The final variable of which is descriptive statistics will be analyzed is the continuous variable relationship length. Respondents were asked for how many years they supported their favorite club to gain insight into the relationship length between the respondent and their club. Relationship length is a variable only present in model equation II for providing a more precise estimation of the interaction effect between this variable and attitudinal loyalty, which will be discussed in subsection 4.8.4.

Relationship length (in years)	Frequency	Percent	Cumulative Percent
1-10	80	20.0	20.0
11-20	166	41.5	61.5
21-30	58	14.5	76.0
31-40	42	10.5	86.5
41-50	39	9.8	96.3
51-60	13	3.3	99.5
61-70	2	0.5	100.0
Total	400	100.0	

Table 22: frequency table of the variable relationship length, categorized in seven groups

The variable relationship length has been categorized in seven groups to provide a more parsimonious representation of this variable, just like the variable age with Table 22. The category 11-20 is by far the most frequent category (N=166, 41.5%), with doubling the frequency of the second highest category (N=80, 20%), which is 1-10 (see Table 22). With 61.5% of the total respondents having a relationship length which falls between one of the two first categories, the distribution is again skewed to the left, as shown in Appendix E.

	Mode	Median	Mean	Std. Deviation	Min	Max
Relationship length (in years)	15	18.5	21.8	13.8	1	65

Table 23: several descriptive statistics for the variable relationship length

The variable age also had a left-skewed distribution, which led to the mean being the highest value of the three central tendency measurements, followed by the median and mean. Needless to say, it is quite logical that those two variables have the same distribution, because younger respondents generally for a shorter period fan of a club than older respondents. Table 23 shows that that indeed the mean is the highest value of the three with 21.8, followed by a median of 18.5 and a mode of 15, which confirms the left skewed distribution. Just as for age, the data is quite dispersed with a standard deviation of 13.8 (see Table 23)

4.4 Descriptive analysis IV: correlation analysis

This subsection will focus on assessing two types of validity, convergent validity and predictive validity, with descriptive statistics derived from correlation analysis. First of all, the convergent validity will be assessed by interpreting the correlation matrix of the observed items. Secondly, the predictive validity of both model equations will be assessed by interpreting the output of the correlation matrix, of which the input consists of the same variables as those which are in the model equations.

4.4.1 Type of validity I: convergent validity

varname_	Soc1	Soc2	Soc3	Esp1	Esp2	Esp3	Gho1	Gho2	Gho3	Int1	Int2	Int3	Ide1	Ide2	Ide3	Ide4	Att1	Att2	Att3	Att4	Beh1	Beh2	Beh3	Beh4	Beh5
Soc1	1,00																								
Soc2	0,77	1,00																							
Soc3	0,71	0,81	1,00																						
Esp1	0,40	0,40	0,40	1,00																					
Esp2	0,34	0,31	0,25	0,60	1,00																				
Esp3	0,44	0,43	0,36	0,64	0,89	1,00																			
Gho1	-0,17	-0,17	-0,05	-0,18	-0,23	-0,28	1,00																		
Gho2	-0,08	-0,15	-0,04	-0,13	-0,18	-0,21	0,83	1,00																	
Gho3	-0,17	-0,14	-0,06	-0,19	-0,19	-0,23	0,83	0,82	1,00																
Int1	0,16	0,13	0,22	0,05	-0,03	-0,03	0,37	0,36	0,40	1,00															
Int2	-0,08	-0,05	0,01	-0,09	-0,11	-0,15	0,42	0,38	0,43	0,58	1,00														
Int3	0,08	0,05	0,09	-0,10	-0,06	-0,07	0,41	0,47	0,50	0,56	0,62	1,00													
Ide1	0,29	0,32	0,43	0,22	0,13	0,15	0,21	0,12	0,18	0,41	0,26	0,28	1,00												
Ide2	0,20	0,19	0,31	0,18	0,04	0,04	0,26	0,18	0,21	0,36	0,36	0,32	0,63	1,00											
Ide3	0,47	0,43	0,47	0,25	0,22	0,34	0,02	0,03	0,07	0,29	0,03	0,25	0,56	0,46	1,00										
Ide4	0,35	0,36	0,41	0,20	0,00	0,06	0,22	0,13	0,18	0,43	0,28	0,27	0,63	0,63	0,50	1,00									
Att1	0,38	0,42	0,48	0,22	0,05	0,16	0,21	0,14	0,19	0,45	0,25	0,31	0,60	0,57	0,56	0,79	1,00								
Att2	0,22	0,30	0,36	0,17	-0,06	0,02	0,10	0,04	0,09	0,29	0,25	0,23	0,49	0,49	0,42	0,58	0,55	1,00							
Att3	0,26	0,32	0,39	0,21	-0,01	0,06	0,12	0,07	0,09	0,26	0,24	0,15	0,49	0,47	0,36	0,56	0,55	0,55	1,00						
Att4	0,10	0,17	0,29	0,15	-0,04	-0,05	0,27	0,16	0,21	0,34	0,32	0,21	0,58	0,60	0,34	0,65	0,60	0,57	0,63	1,00					
Beh1	0,39	0,42	0,44	0,27	0,10	0,22	0,11	0,08	0,13	0,29	0,16	0,27	0,56	0,57	0,56	0,57	0,61	0,47	0,54	0,52	1,00				
Beh2	0,41	0,44	0,43	0,19	-0,01	0,07	0,08	0,01	0,05	0,29	0,14	0,17	0,49	0,44	0,37	0,62	0,61	0,51	0,56	0,59	0,60	1,00			
Beh3	0,45	0,51	0,51	0,28	0,09	0,17	0,03	0,03	0,04	0,29	0,18	0,21	0,51	0,44	0,41	0,60	0,62	0,47	0,53	0,50	0,54	0,65	1,00		
Beh4	0,41	0,40	0,39	0,18	0,10	0,17	0,06	0,11	0,11	0,32	0,20	0,32	0,34	0,39	0,39	0,49	0,56	0,34	0,37	0,38	0,50	0,48	0,57	1,00	
Beh5	0,42	0,38	0,38	0,18	0,13	0,23	0,02	0,08	0,06	0,32	0,21	0,31	0,44	0,42	0,50	0,45	0,47	0,35	0,40	0,35	0,56	0,48	0,50	0,49	1,00

Table 24: polychoric correlation matrix of all the observed items for the latent constructs

So, convergent validity is assessed by interpreting the correlations between the observed items. Correlation is a statistical measure that expresses the extent to which two variables are linearly related, which means that they change together at a constant rate (Fisher and Marshall, 2009). The rules of thumb for interpretation correlation coefficient can be found in subsection 3.6.2. Keep in mind that strong correlation coefficients which are close to each other equals a strong degree of convergent validity.

First of all, the convergent validity of the latent construct social media is assessed. The lowest correlation between the observed items of social media was between Soc1 and Soc3 with a coefficient of 0.71), whereas the highest correlation was between Soc3 and Soc2 with a coefficient of 0.81 (see Table 24). A correlation coefficient of 0.5 or higher is perceived as strong. In other words, the latent construct social media shows a high degree of convergent validity.

Secondly, the observed items for the latent construct esports also show strong correlations. Namely, the lowest correlation was 0.60 between Esp1 and Esp2 and the highest correlation was 0.89, as shown in Table 24. Therefore, the conclusion can be drawn that the latent construct of esports also has a high degree of convergent validity.

Thirdly, the correlation coefficients for the latent construct ghost games will be analyzed. The correlations of the observed items are not only strong, but also strikingly close to each other. The correlation between Gho1 and Gho2 is exactly the same as the correlation between Gho1 and Gho3 with a coefficient of 0.83, whereas the final and third correlation of this construct between Gho2 and Gho3 is only 0.01 lower than the other two with a coefficient of 0.82 (see Table 24). So, convergent validity is also very strong for this latent construct.

Fourthly, the latent construct lack of fan interaction will be assessed on its degree of convergent validity. Similar to the previous latent construct, the correlations coefficients are really close to each other. Even though the coefficients are close to each other, they are not as high as the previous latent construct with the lowest correlation coefficient being 0.56 between Int1 and Int3, as shown in Table 24. The highest correlation coefficient was 0.62 between Int2 and Int3 (see Table 24). Nevertheless, the correlations were still strong and also close to each other, which means the convergent validity is strong.

Fifthly, the correlations of the four observed items measuring the latent construct fan identification will be analyzed. The correlations are relatively close to each other, but the weakest correlation, which is between Ide2 and Id3 falls just below the threshold of 0.5 for a strong correlation with a coefficient of 0.46 (see Table 24). However, all the other six correlations of this variable are above the threshold of 0.5, where three correlations have the joint highest coefficient of 0.63, as shown in Table 24. So, even though one correlation was below the threshold of 0.5, the convergent validity is still sufficient due to the remaining cluster of strong correlations.

The second-last latent variable of which the correlation of its four observed items will be analyzed, is attitudinal loyalty. Again, the correlations are heavily clustered with three correlations being the joint lowest with a coefficient of 0.55 (see Table 24). The strongest correlation is between Att3 and Att4 with a coefficient of 0.63, as shown in Table 24. So, all the correlations are clustered and above the threshold of 0.5, which equals a strong convergent validity.

The seventh and also final latent variable of this subsection is behavioral loyalty. Even though this latent construct is measured by the most number of observed items, the correlation coefficients are still quite clustered. The correlation between Beh2 and Beh4 and the one between Beh2 and Beh5 are the joint lowest with 0.48, as shown in Table 24. Also, the correlation between Beh4 and Beh5 is just below the threshold of 0.5 with a coefficient of 0.49 (see Table 24).

However, the other seven correlation coefficients are above this threshold of 0.5 and thus interpreted as strong correlations, with the correlation between Beh2 and Beh3 being the highest of the ten with a coefficient of 0.65, as shown in Table 24. Therefore, the convergent validity for this latent construct is still classified as sufficient.

4.4.2 Type of validity II: predictive validity (model equation I)

The second part of this subsection will focus on assessing the predictive validity of this research, which will be done on the basis of the output of the polychoric correlation matrix, shown in Table 25. Remember that for a strong degree of predictive validity, there no necessity is in obtaining only strong correlations. This is due to the fact that the hypotheses of this research express hypothesized relationships. In other words, there is no guarantee that every hypothesized relationship is an actual relationship.

	Relation	Parental	Friends	Geograph	Team suc	Team con	Nostalgic	Social	Ghost	Lack of	Esports	Fan iden	Attitudinal	Behavioral
Relationship length	1,000													
Parental influence	0.218	1,000												
Friend influence	0.174	0.439	1,000											
Geographical location	0.205	0.69	0.588	1,000										
Team success	0.044	-0.004	0.245	0.061	1,000									
Team content	0.151	0.564	0.302	0.443	0.24	1,000								
Nostalgic content	0.137	0.364	0.064	0.289	0.499	0.655	1,000							
Social media	0.201	0.621	0.393	0.538	0.233	0.902	0.608	1,000						
Ghost games	0.226	0.273	0.343	0.321	0.307	0.212	0.177	0.126	1,000					
Lack of fan interaction	0.221	0.494	0.490	0.384	0.162	0.416	0.262	0.351	0.833	1,000				
Esports	-0.029	0.379	0.309	0.289	0.451	0.69	0.567	0.750	0.041	0.229	1,000			
Fan identification	0.235	0.572	0.251	0.415	0.203	0.731	0.438	0.758	0.465	0.669	0.538	1,000		
Attitudinal loyalty	0.240	0.569	0.138	0.355	0.129	0.691	0.426	0.689	0.479	0.646	0.445	0.918	1,000	
Behavioral loyalty	0.252	0.702	0.312	0.462	0.101	0.850	0.416	0.905	0.360	0.612	0.558	0.919	0.719	1,000

Table 25: polychoric correlation matrix of the variables which are measuring the main effect in the regressions

First, the predictive validity will be assessed for model equation I, of which attitudinal is the DV and eight variables act as an IV which measure the main effect. From those eight variables, four showed strong correlations with attitudinal loyalty, two showed moderate correlations and two showed weak correlations.

From the four variables which strongly correlation with attitudinal loyalty, two variables clearly showed the strongest correlations. Namely, fan identification was the strongest with a coefficient of 0.919, just a bit higher than social media with a coefficient of 0.905, as shown in Table 25. After those two, the variable team content was the third highest correlation with a coefficient of 0.699, whereas parental influence was just slightly above the threshold of 0.5 with a coefficient of 0.569 (see Table 25).

Geographical location and nostalgic content were the two variables who correlated moderately with attitudinal loyalty with respective correlations of 0.355 and 0.426. Lastly, the variables friend influence and team success both showed a weak correlation with close respective coefficients of 0.139 and 0.129, as shown in Table 25.

To conclude, model equation I exhibits a more than sufficient level of predictive validity with six out of the eight variables, which measured the main effect, showing at least moderate correlations. Furthermore, three of the four variables easily passed the threshold of 0.5 for strong correlations. These findings validate the predictive validity of model equation I.

4.4.3 Type of validity II: predictive validity (model equation II)

Right now, model equation II will be assessed on predictive validity, just after validating this type of validity for model equation I. As known, model equation II is more comprehensive than model equation I. For this equation, the DV is behavioral loyalty and 13 variables act as an IV and measure the main effect in this regression. From the 13 variables, 7 correlate strong with behavioral loyalty, 4 correlate moderately and only 2 correlate weak.

As for the variables which correlate strongly with behavioral loyalty, the top three of highest correlations is equal to the variables which had the highest correlations with attitudinal loyalty. Thus, fan identification correlated the strongest, followed by social media and team content with respective correlations of 0.919, 0.905 and 0.850 (see Table 25).

Also, the fourth strongest correlation would have been identical between the two dimensions of loyalty, but extra IV's were added to model equation II, which disrupt the correlation order seen for model equation I. Attitudinal loyalty is in model equation I the predicted variable, but in model equation II it acts as an IV with a strong correlation of 0.719, followed by parental influence with a coefficient of 0.702, as shown in Table 25.

Finally, the newly added variable lack of fan interaction has a coefficient of 0.612 and esports is the closest to the threshold of 0.5 from all the strong correlations with a coefficient of 0.558 (see Table 25).

Four variables correlated moderately with behavioral loyalty, of which two were already in this category for their correlation with attitudinal loyalty. This are the variables geographical location and nostalgic content with respective coefficients of 0.462 and 0.416, as shown in Table 25. Also, the variable friend influence correlated stronger with behavioral loyalty than the attitudinal dimension with a coefficient of 0.312, which was only 0.138 in model equation I. The newly added variable ghost games showed a correlation of 0.360 (see Table 25).

Only two variables showed weak correlations in model equation II. Again, team success was found back in this category with a coefficient of 0.101 and the coefficient of relationship length was not much higher with 0.252, as shown in Table 25.

So, out of the seven variables which correlated strongly, only esport came relatively close to the threshold of 0.5 for strong correlations. The other six had a coefficient of at least 0.6 of higher. Also, from the remaining six variables only two showed weak correlation.

In other words, 11 of the 13 variables, which measure the main effect, correlated at least moderately with the DV behavioral loyalty. Thus, assessing the predictive validity for model equation II resulted into concluding that the predictive validity is very strong and definitely the strongest of the two model equations.

4.5 Descriptive analysis V: Cronbach's alpha

The previous subsection validated two types of validity, convergent and predictive validity. In this subsection, also the fifth and final part of the descriptive analyses, the coefficient of Cronbach's alpha will be interpreted to measure the internal consistency of the seven latent variables in this research. Keep in mind that internal consistency, one of the three reliability attributes, checks how closely related a set of observed items is.

4.5.1 Latent variable I: behavioral loyalty

The used rules of thumb for interpreting Cronbach's alpha are displayed in Table 3. The latent construct behavioral loyalty has a Cronbach's alpha of .854 (see Table 26), which is in between good and excellent. So, the latent construct behavioral loyalty shows a more than sufficient level of internal consistency. Also, each item contributes to achieving this level of internal consistency, because Cronbach's alpha will decrease when one of the five items will be removed, as shown in Table 26.

	Cronbach's alpha	**Cronbach's alpha if item was deleted**
Behavioral loyalty	.841	
Beh1		.802
Beh2		.812
Beh3		.799
Beh4		.809
Beh5		.822

Table 26: reliability statistics of the latent variable behavioral loyalty

4.5.2 Latent variable II: attitudinal loyalty

Secondly, the reliability statistics of the latent variable attitudinal loyalty will be analyzed. The latent construct has a Cronbach's alpha of .816 (see Table 27), which is just above the threshold of a good level of internal consistency. Furthermore, the level of internal consistency will decrease for attitudinal loyalty when one of its four observed items will be removed, as shown in Table 27. Namely, the Cronbach's alpha will fall just under the threshold of a good level of internal consistency. So, no observed items will be removed.

	Cronbach's alpha	**Cronbach's alpha if item was deleted**
Attitudinal loyalty	.816	
Att1		.772
Att2		.792
Att3		.763
Att4		.750

Table 27: reliability statistics of the latent variable attitudinal loyalty

4.5.3 Latent variable III: fan identification

Thirdly, the reliability statistics of the latent variable fan identification will be analyzed. Table 28 shows a Cronbach's alpha of .815, which again passes the threshold of a good level of internal consistency, which makes this latent variable reliable.

However, unlike the previous two latent variables, deleting an item can improve the Cronbach's Alpha. Namely, removing Ide3 will improve the Cronbach's alpha with 0.07 (see Table 28). Nevertheless, the item will remain part of the latent construct, because this extremely minor improvement in reliability does not outweigh its moderate correlation with Ide2 and its strong correlations with Ide1 and Ide4 as discussed in the previous subsection (see Table 28), which contributes to strengthening the convergent validity of this research.

	Cronbach's alpha	**Cronbach's alpha if item was deleted**
Fan identification	.815	
Ide1		.744
Ide2		.758
Ide3		.822
Ide4		.753

Table 28: reliability statistics of the latent variable fan identification

4.5.4 Latent variable IV: social media

Fourthly, the reliability statistics of the latent variable social media will be analyzed, which will show that this latent construct has the highest level of internal consistency in comparison with the three previously discussed latent variables. Namely, Table 29 shows that this latent construct has a Cronbach's alpha of 0.899, which is just 0.001 off of reaching the threshold for an excellent internal consistency. Also, one of the three observed items will be removed, because removing one of them will lower the Cronbach's alpha and thus the reliability, as shown in Table 29.

	Cronbach's alpha	Cronbach's alpha if item was deleted
Social media	.899	
Soc1		.867
Soc2		.824
Soc3		.873

Table 29: reliability statistics of the latent variable social media

4.5.5 Latent variable V: esports

Fifth, the reliability statistics of the latent variable esports will be analyzed. The Cronbach's alpha for this latent variable is .863, as shown in Table 30. Interpreting this coefficient results in seeing that the internal consistency of this latent construct is between the levels good and excellent.

As for the observed items, removing Esp2 or Esp3 will only decrease Cronbach's alpha and the level of internal consistency will even fall under the 0.8 threshold of good internal consistency, as shown in Table 30. However, Esp1, related to liking the videogame FIFA, apparently impedes the internal consistency, because removing this observed item will led to passing the threshold of 0.9 for excellent internal consistency (see Table 30).

However, this research will remain Esp1 as an observed item and accept the currently good level of internal consistency with the coefficient being .863. Namely, keeping Esp1 in the data set is incremental for H_4, because this research looks specifically into the esports scene of the videogame FIFA. Furthermore, removing this observed item will make this latent construct under-identified, which is highly undesired. Also, in the next subsection it will be clear that there is no need to remove Esp1 from the data set.

	Cronbach's alpha	Cronbach's alpha if item was deleted
Esports	.863	
Esp1		.907
Esp2		.775
Esp3		.731

Table 30: reliability statistics of the latent variable esports

4.5.6 Latent variable VI: lack of fan interaction

The second-last latent variable of which its reliability statistics will be analyzed is the lack of fan interaction. Table 31 shows a Cronbach's alpha of .820 for this latent construct, which means that its internal consistency is good. Besides, there is no need to remove one of the observed items, because they will all decrease level of internal consistency by going below the threshold of 0.8 for a good level of internal consistency, as shown in Table 31.

	Cronbach's alpha	Cronbach's alpha if item was deleted
Lack of fan interaction	.820	
Int1		.784
Int2		.735
Int3		.737

Table 31: reliability statistics of the latent variable lack of fan interaction

4.5.6 Latent variable VII: ghost games

The seventh and thus last latent variable of this research will now be analyzed to see whether all the latent constructs show internal consistency. Table 32 shows a Cronbach's alpha of .932, which can be interpreted as a more than excellent level of internal consistency. Even when Gho1 or Gho2 will be deleted from the latent construct, the level of internal consistency would still be above the threshold of 0.9 for excellent internal consistency (see Table 32). Also, Cronbach's alpha will decrease if Gho3 is removed (see Table 32), which means all three observed items remain in the data set as the building blocks for this latent construct.

	Cronbach's alpha	Cronbach's alpha if item was deleted
Ghost games	.932	
Gho1		.908
Gho2		.907
Gho3		.890

Table 32: reliability statistics of the latent variable ghost games

4.6 Factor analysis: CFA

This subsection will analyze and interpret the CFA output, generated by SPSS Amos. The type of validity which can be assessed with this statistical technique is construct validity. All of the seven latent variables will be assessed on this type of validity in order to make sure that the latent variables are constructed with appropriate observed items.

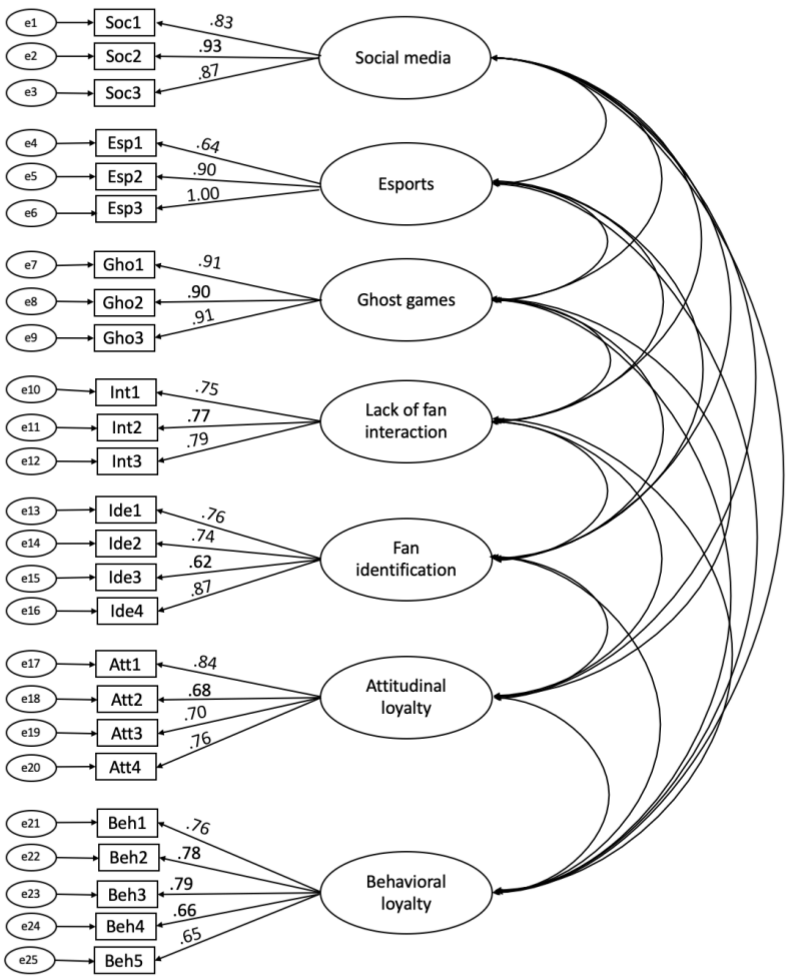

Figure 10:reflective measurement model of the seven latent variables

Figure 10 shows the reflective measurement model of the seven latent variables, where each observed item has its own unique error term and factor loading. All these error terms were added to the model in order to improve the model, because it measures the variation of the observed items which is not captured by the unobserved latent construct. Also, covariance's were drawn between the unobserved latent variables to improve the fit even further, which are the double-arrowed lines between the latent variables in Figure 10.

However, this statistical measurement is not further interpreted, because it is similar to the correlation coefficient, but less preferred. Namely, covariance indicates only the direction of the linear relationship between the variables, whereas correlation measures both the strength and direction (Fisher and Marshall, 2009).

4.6.1 Unstandardized regression weights and factor loadings

Social media is the first latent variable of which its factor structure will be confirmed. The regression weight of Soc1 was set to 1, as shown in Table 33. This action was necessary to undergo, because any unobserved variable in the model has to have a scale. In other words, the regression weights of Soc2 and Soc3 have Soc1 as their benchmark. An estimate higher than 1 means that the latent construct predicts the observed item better than Soc1, whereas an estimate lower than 1 equals a better prediction of the benchmark variable of the two.

Moreover, the critical value is computed by dividing the regression estimate by the standard error. The critical value measures how many standard errors the regression weight is above zero and the *p*-value is derived from this value, which determines whether the factor loading is significant or not.

Namely, in essence, a factor loading is the standardized regression weight. Due to the fact that factor loadings are standardized, they predict more precisely which observed item is predicted the strongest by its latent variable in comparison with the unstandardized estimates. However, unstandardized estimates are still important to assess whether the observed item is significant or not.

IV	DV	Estimate	Standard error	Critical value	*p*-value	Factor loading
Soc1	Social media	1.000				.83***
Soc2	Social media	1.024	0.045	22.946	.000	.93***
Soc3	Social media	0.850	0.040	21.211	.000	.87***

Table 33: unstandardized regression weights and factor loadings for the observed items of the latent variable social media

So, the latent variable social media predicts the observed item Soc2 the best with an estimate of 1.024, where all of the three observed items were found to be significantly different from zero with a *p*-value of .000, as shown in Table 33. For the interpretation of the *p*-value, see Table 1. As for the factor loadings, all three factors were above the threshold of .70 for excellent factor loadings, which means that this latent variable has a strong construct validity (see Table 33).

Table 34 shows the CFA output for the latent variable esports and its three observed items. Again, the regression weight was set to 1 to for the benchmark item Esp1 in order to create a scale. Esp2 and Esp3 have an estimate greater than one, where Esp1 has the lowest factor loading of the three (see Table 34).

IV	DV	Estimate	Standard error	Critical value	*p*-value	Factor loading
Esp1	Esports	1.000				.64***
Esp2	Esports	1.375	0.089	15.419	.000	.90***
Esp3	Esports	1.497	0.097	15.498	.000	1.00***

Table 34: unstandardized regression weights and factor loadings for the observed items of the latent variable esports

Esp3 has a factor loading of .64 which is below the threshold of 0.7 for excellent factor loading, but is still well above the absolute minimum cut-off criteria of 0.4 (see Table 34). A side note must be placed that the exact factor loading of Esp3 is .996, but rounded up to 1 in this table. Furthermore, Esp2 and Esp3 have a critical value of which a *p*-value of .000 is derived (see

Table 34). Thus, all factor loadings are significantly different from zero and thus kept in the model, which shows a sufficient degree of construct validity.

Thirdly, the latent variable ghost games predicted the observed items more than well. As seen in Table 35 with benchmark variable Gho1, both the unstandardized estimates are significantly different from zero with their respective *p*-value of .000. So, the factor loadings of .90 and .91 are not only significant and highly clustered with a difference of only .01 between the three loadings (see Table 35), but also well above the threshold of 0.7 for an excellent factor loading. So, also this latent variable has a very strong degree of construct validity.

IV	**DV**	**Estimate**	**Standard error**	**Critical value**	***p*-value**	**Factor loading**
Gho1	Ghost games	1.000				.91***
Gho2	Ghost games	0.905	0.032	28.030	.000	.90***
Gho3	Ghost games	0.955	0.033	28.984	.000	.91***

Table 35: unstandardized regression weights and factor loadings for the observed items of the latent variable esports

Fourthly, the construct validity will be assessed for the latent variable lack of fan interaction with its three observed items. Again, the unstandardized regression weights of Int2 and Int3 have a critical value which results in a *p*-value of .000 (see Table 36).

Furthermore, the factor loadings are, just as for the previous latent construct, significantly different from zero and highly clustered with a difference of .04 between the highest and lowest factor loading (see Table 36). However, this time the loadings are closer to the threshold of 0.7 for excellent factor loadings, but the construct validity remains also strong for this latent construct.

Fifthly, the CFA output for the latent variable fan identification and its four observed items will be analyzed. Ide1 was the benchmark variable, which allowed the other three observed items to be estimated. Once more, the *p*-values of 0.00 meant that all the factor loadings were significantly different from zero. This time, one factor loading, Ide3 with 0.62 (see Table 37), was below the threshold of 0.7 for an excellent factor loading, but well above the cut-off criteria of 0.4.

IV	DV	Estimate	Standard error	Critical value	*p*-value	Factor loading
Int1	Lack of fan interaction	1.000				.75***
Int2	Lack of fan interaction	1.044	0.076	13.793	.000	.77***
Int3	Lack of fan interaction	1.038	0.074	14.044	.000	.79***

Table 36: unstandardized regression weights and factor loadings for the observed items of the latent variable lack of fan interaction

Furthermore, the other three loadings passed the excellent threshold of 0.7 with .87 being the highest factor loading, which means that the construct validity is still sufficient for this variable and thus all the observed items were kept in the model.

IV	DV	Estimate	Standard error	Critical value	*p*-value	Factor loading
Ide1	Fan identification	1.000				.76***
Ide2	Fan identification	1.251	0.083	15.133	.000	.74***
Ide3	Fan identification	1.189	0.095	12.558	.000	.62***
Ide4	Fan identification	1.357	0.074	18.227	.000	.87***

Table 37: unstandardized regression weights and factor loadings for the observed items of the latent variable fan identification

The second-last latent construct of which its construct validity will be assessed on the basis of the CFA output is attitudinal loyalty. The path of the first observed item, Att1, was again fixed to 1 in order to create a scale. The other three observed items had a critical value, which led to a *p*-value of 0.000, as shown in Table 38.

IV	DV	Estimate	Standard error	Critical value	*p*-value	Factor loading
Att1	Attitudinal loyalty	1.000				.84***
Att2	Attitudinal loyalty	.700	0.046	15.349	.000	.68***
Att3	Attitudinal loyalty	.633	0.040	15.819	.000	.70***
Att4	Attitudinal loyalty	.792	0.044	18.039	.000	.76***

Table 38: unstandardized regression weights and factor loadings for the observed items of the latent variable attitudinal loyalty

As for the factor loadings, only one, Att3 with 0.68, did just not pass the threshold of 0.7 for an excellent factor loading. The other three did, with Att1 having the highest factor loading with a value of .84, as shown in Table 38. So, all factor loadings are significantly different from zero, where one is not classified as an excellent loading, which results in a strong construct validity for this latent variable.

The latent variable behavioral loyalty is the last construct of which its CFA output will be analyzed. Dividing the unstandardized estimate by the standard error result also for this latent construct in critical value's which have a respective *p*-value of .000, as shown in Table 39. So, all the factor loadings were significantly different from zero in this measurement model.

IV	DV	Estimate	Standard error	Critical value	*p*-value	Factor loading
Beh1	Behavioral loyalty	1.000				.76***
Beh2	Behavioral loyalty	.856	0.053	16.224	.000	.78***
Beh3	Behavioral loyalty	.968	0.059	16.385	.000	.79***
Beh4	Behavioral loyalty	.876	0.067	13.064	.000	.66***
Beh5	Behavioral loyalty	.961	0.072	13.377	.000	.65***

Table 39: unstandardized regression weights and factor loadings for the observed items of the latent variable behavioral loyalty

Three out of the five factor loadings of behavioral loyalty passed the threshold of 0.7 for an excellent factor loading, of which Beh3 was loaded the highest with a value of .79 (see Table 39). Beh4 and Beh5 were just under the threshold with respective loadings of .66 and .65, but both were well above the absolute minimum threshold and thus cut-off criteria of 0.4. To conclude, the final latent variable has also a sufficient degree of construct validity, which results in an overall good validity for the entire measurement model shown in Figure 10.

4.6.2 Model fit interpretation

The second part of analyzing the CFA output consists of assessing the model fit of the measurement model, which will be determined by interpreting three fit-indices: TLI, CFI and RMSEA, which were discussed in subsection 3.2

Fit-index	Value
TLI	.880
CFI	.899
RMSEA	.085

Table 40: values for the three model-fit indices

For the first two fit indices, TLI and CFI, the model fit was considered acceptable when the value was .9 or higher. Table 40 shows that the TLI value of this measurement model is .880 and the CFI value is .899, which are both just below the threshold of an acceptable fit. The same holds account for the RMSEA, which is with .085 also just above the threshold of .08 for a reasonable fit (see Table 40). Keep in mind that for RMSEA a low value is desired, whereas a high value was desired for the first two fit-indices.

4.7 Regression analysis: OLR (model equation I)

After successfully assessing the convergent, predictive and construct validity of the variables, the latent construct can be put in the two main model equations of this research. First, model equation I will be analyzed and interpreted correctly and the same analysis and interpretation will be applied on model equation II. Also, Appendix D shows that both model equations pass the four assumptions of OLR.

4.7.1 Choosing the best model

First, the value of adding the control variables to the equation model will be discussed. Table 41 shows the coefficients and its level of significance for two variants for the first equation. A base model (model 1) with the variables of interest measuring the main effect and a model with control variables added (model 2). Despite the addition of more values, the pseudo R^2 of Nagelkerke increases from .631 to .648 (see Table 41), which implies that model 2 explains most of the variation in the DV. So, the conclusions with regards to attitudinal loyalty will be based on the coefficients of model 2, because they are more accurate and less prone to OVB.

	Model 1	Model 2
Fan identification	2.217***	2.145***
Social media	-0.103	-0.107
Team content	0.155	0.220*
Nostalgic content	-0.020	-0.019
Parental influence	0.108*	0.095
Friend influence	-0.246***	-0.240***
Geographical location	-0.102*	-0.096
Team's success	-0.191***	-0.173**
Male		1.080
Female		0.449
Third gender		0.709
Age		0.015**
Less than high school degree		1.499*
High school degree		1.589**
Some degree		1.570**
Bachelor's degree		1.311*
Master's degree		1.480*
Small club		-0.109
Medium club		-0.152
Pseudo R^2 of Nagelkerke	.631	.648

*Table 41: coefficients of the two models for model equation I. *=significant at the 10%-level, **=significant at the 5%-level, ***=significant at the 1%-level*

Also, Table 41 shows that interpreting the coefficients of model 1 would have led to wrongful conclusions. Some variables were significant at the 10%-level for model 1, but not for model 2. For example, geographical location and parental influence cannot be interpreted, because they do not significantly differ from zero. As for the variable team's success, the significance level dropped down from 1% to the 5%-level, as shown in Table 41.

Furthermore, model 1 also suffered slightly from omitted variable bias. Three of the eight variables which measure the main effect had a higher coefficient in model 1 compared to model 2. The biggest difference in coefficients was for the variable fan identification with a delta of 0.072 (see Table 41). In other words, the variables fan identification, social media and parental influence were all upwards biased.

The remaining five variables were downwards biased. The variable team content suffered the most from this bias. Namely, not only the coefficient was higher for model 2, but the significance level also increased from nonsignificant to significantly different from zero at the 10%-level, as shown in Table 41.

4.7.2 Interpreting the main effect coefficients

	β	S.E	exp β	*p*-value	Lower bound 95% C.I	Upper bound 95% C.I
Fan identification	2.145	0.140	8.542	.000	1.871	2.419
Social media	-0.107	0.139	0.899	.440	-0.380	0.165
Team content	0.220	0.129	1.246	.087	-0.032	0.472
Nostalgic content	-0.019	0.079	0.981	.808	-0.174	0.136
Parental influence	0.095	0.059	1.100	.108	-0.021	0.211
Friend influence	-0.240	0.062	0.787	.000	-0.362	-0.118
Geographical location	-0.096	0.067	0.908	.154	-0.228	0.036
Team's success	-0.173	0.068	0.841	.011	-0.306	-0.040
Male	1.080	1.266	2.945	.394	-1.402	3.561
Female	0.449	1.273	1.576	.724	-2.046	2.944
Third gender	0.709	1.832	2.032	.699	-2.882	4.300
Age	0.015	0.008	1.015	.040	0.001	0.030
Less than high school degree	1.499	0.864	4.477	.083	-0.193	3.192
High school	1.589	0.743	4.899	.032	0.133	3.046
Some degree	1.570	0.732	4.807	.032	0.135	3.005
Bachelor's degree	1.311	0.733	3.710	.074	-0.126	2.749
Master's degree	1.480	0.785	4.393	.059	-0.059	3.019
Small club	-0.109	0.309	0.897	0.723	-0.714	0.496
Medium club	-0.152	0.279	0.859	0.586	-0.700	0.395

Table 42: statistics for the regression analysis of model equation I

Table 42 shows the OLR output for model equation I with the following statistics: the coefficient (β), the standard error (S.E), the coefficient raised to the power of e (exp β), the *p*-value and at last the lower and upper bound of the 95% confidence interval for the coefficients.

The standard error measures the precision of the estimate of the coefficient. The smaller, the error, the more precise the estimate (Fisher and Marshall, 2009). Furthermore, the confidence interval gives a range of which it is 95% sure that the respective coefficient falls in. In subsection 3.3.3 the reasoning behind raising the coefficient to the power of e was explained.

As known, only the variables which are significantly different from zero can be interpreted. Fan identification is significantly different from zero at the 1%-level with its *p*-value .000. For every one unit increase of fan identification, the log odds of falling in a greater level of attitudinal loyalty will increase with 2.145, which is equal to an increase in odds ratio of 8.542 (see Table 42), remaining the other variables constant.

The second significant variable, team content has a *p*-value of .087, which means that it is significantly different from zero at the 10%-level. For every one unit increase of team content, the log odds of falling in a greater level of attitudinal loyalty will increase with 0.220, which is equal to an increase in odds ratio of 1.246 (see Table 42), remaining the other variables constant.

The third significant variable in model equation I, friend influence, has a *p*-value of .000, as shown in Table 42. So, friend influence differs significantly from zero at the 1%-level. For every one unit increase of friend influence, the log odds of falling in a greater level of attitudinal loyalty will decrease with 0.240, which is equal to an increase in odds ratio of 0.787, remaining the other variables constant. Keep in mind that an increase in odds ratio lower than one means that the probability of falling in a greater level decreases.

The fourth and final significant variable which measures the main effect is the team's success with a *p*-value of .011 (see Table 42), which means that it is significantly different from zero at the 5-% level. For every one unit increase in team's success, the log odds of falling in a greater level of attitudinal loyalty will decrease with 0.173, which is equal to an increase in the odds ratio of 0.841, remaining the other variables constant.

4.7.3 Interpreting the control variables coefficients

As for the control variables, only two variables were significant: age and education. All category levels of education had a *p*-value below .010, which means that all the five categories differ significantly from zero. Less than high school degree, bachelor's degree and master's degree were significant at the 10%-level, whereas some degree and high school were significant at the 5%-level (see Table 42).

All five categories of education have positive β's which are also quite clustered with Bachelor's degree having the lowest β with a value of 1.311 and High school degree the highest with a value of 1.589, as shown in Table 42. Even though all the categories can be interpreted, only one will be interpreted right now due to their similarity.

High school degree has a *p*-value of .032, which means it significantly differs from zero at the 5%-level. On average, the log odds of falling in a greater level of attitudinal loyalty were 1.589 higher for respondents who obtained a high school degree as their highest level of education compared to respondents who obtained a PhD as their highest level of education, which equals an increase in odds ratio of 4.899, remaining the other variables constant (see Table 42).

The control variable age has a *p*-value of .040, which means that it is significantly different from zero for the 5-% level. For every one unit increase of age (i.e., one year), the log odds of falling in a greater level of attitudinal loyalty will increase with .015, which equals an increase in odds ratio of 1.015, remaining the other variables constant (see Table 42).

4.7.4 Interpreting the constants

Level of attitudinal loyalty	Constant (α)	*p*-value
1.00	2.877	.146
1.50	3.978	.027
3.00	4.759	.006
3.25	5.565	.001
3.50	6.591	.000
3.75	6.856	.000
4.00	7.601	.000
4.25	8.156	.000
4.50	9.776	.000
4.75	10.396	.000
5.00	11.114	.000
5.25	11.858	.000
5.50	12.812	.000
5.75	13.681	.000
6.00	14.605	.000
6.25	15.554	.000
6.50	16.472	.000
6.75	17.467	.000

Table 43: interception estimates with their respective p-values for model equation I

Due to the parallel lines assumption of OLR (see appendix D), the explanatory variables are consistent across the different levels of attitudinal loyalty. That is why every level of attitudinal loyalty has its own starting point, α, which is also known as the intercept of the model equation. Naturally, the higher the level of attitudinal loyalty, the higher the starting point, as seen in Table 43. Namely, if your level of attitudinal loyalty is high, the log odds of falling in a greater level of attitudinal loyalty should also be high.

As for the significance of the constants, only the lowest level of 1 differs not significantly from zero. The second and third level are significantly different from zero at the 5%-level with respective *p*-values of .027 and .006. The remaining 15 levels are all statistically different from zero at the 1%-level. A side note must be placed that an intercept has only to be interpreted when all the predictor values have a value of 0, which is almost impossible.

4.7.5 Interpreting the model fit

With the OLR output analyzed and interpreted when it was possible, the next step in the OLR analysis is to assess whether model equation I is a good fit to the data. This will be done with the help of one model fit test, two goodness-of-fit test and one statistical measure that represents the proportion of the variance of the DV explained by the model. The rules of thumb for interpreting these statistics can be found in subsection 3.3.4

Model	-2 Log Likelihood	Chi-Square (χ2)	Degrees of freedom (df)	*p*-value
Null hypothesis	2045.860			
General	1632.172	413.688	19	.000

Table 44: likelihood chi-square test for model equation I

The test to assess the model fit of both equations models is the likelihood ratio chi-square test. Table 44 shows that a χ2 of 413.688 with a df of 19 results in a *p*-value of .000, which means that the final model is a significant improvement in fit over the null model.

Test	Chi-Square (χ2)	Degrees of freedom (df)	*p*-value
Pearson	7167.307	7163	.483
Deviance	1632.172	7163	1.000

Table 45: Pearson chi-square and Deviance test for model equation I

Secondly, the Pearson chi-square and Deviance test are goodness-of-fit test which, needless to say, measure how well both model equations fit the data set. Table 45 shows that the respective *p*-values of both tests are higher than .05. Namely, the *p*-value for the Pearson chi-square test is .483 and 1.000 for the Deviance test. A *p*-value of >.05 means that the there was no significant discrepancy found between model equation I and the final model.

The final statistical measurement used to assess the model fit of this equation is the pseudo R^2 of Nagelkerke. For model equation I, the pseudo R^2 had a value of 0.648, as shown in Table 41. This means that approximately 64.8% of the variability in the DV is explained by the model, which can be interpreted as an explanatory power of the model that is in between moderate and substantial. However, due to the fact that pseudo R^2 yield lower estimates in general than R^2 for OLR, the explanatory power of this model is most likely closer to substantial than moderate.

4.8 Regression analysis: OLR (model equation II)

The same analysis will now be applied on model equation II to investigate the hypothesized relationships related to the behavioral dimension of fan loyalty. After choosing the best model, the coefficients and constants of model equation II will be interpreted. Finally, the model fit-indices will be interpreted. Appendix D will show that also model equation II did not violate any of the four OLR assumptions.

4.8.1 Choosing the best model

First of all, three different models were run in order to determine how many variables should be added to the regression to maximize the explanatory power of the model, measured by the pseudo R^2 of Nagelkerke. Model 1 is the base model with only the variables measuring the main effect included, for model 2 the control variables were added and model 3 had the most variables with the extra addition of relationship length and the interaction term of relationship length times attitudinal loyalty.

Despite the addition of more variables, the explanatory power kept rising and was the highest for Model 3 with a pseudo R^2 of 0.723 (see Table 46), which means that this model will be used for further analysis and interpretation. Besides increasing the explanatory power of the model, the addition of the interaction variable also resulted into decreasing the OVB and thus providing more accurate estimates of the coefficients.

As for the variables which measured the main effect, three relevant cases of OVB were observed when comparing model 2 with model 3. Attitudinal loyalty and parental influence suffered both from upwards bias, because the estimated coefficient was clearly higher in model 2 than model 3, as shown in Table 46. Social media suffered slightly from the other variant of OVB in the form of downwards bias. Namely, for model 2 the coefficient was estimated at 0.789, where the coefficient rose to 0.828 for model 3, remaining significant (see Table 46).

As for the control variables, all the categories of education were upwards biased with a higher coefficient value in model 2 than model 3. Age was slightly upwards biased and for the variable club magnitude the category small club was downwards biased and medium club upwards biased (see Table 46).

Furthermore, when adding the variable relationship length and the interaction term, some significant levels of the control variables changed drastically. The category master's degree was significant at the 10%-level in model 2, but nonsignificant in model 3, as shown in Table 46. The variable age walked the opposite route: from not significantly differing from zero to becoming a significant variable even at the 5%-level (see Table 46). Last but not least, the categorical variable some degree has a weaker level of significance as compared to model 2, because the level of significance increased from 1% to 5%, as shown in Table 46.

	Model 1	**Model 2**	**Model 3**
Attitudinal loyalty	1.464***	1.435***	1.040***
Fan identification	0.796***	0.800***	0.798***
Social media	0.812***	0.789***	0.828***
Team content	0.029	0.068	0.077
Nostalgic content	-0.245***	-0.257***	-0.252***
Esports	0.054	0.071	0.109
Parental influence	0.262***	0.283***	0.196***
Friend influence	0.051	0.023	0.029
Geographical location	-0.040	-0.039	-0.049
Team's success	-0.059	-0.020	-0.014
Lack of fan interaction	0.236**	0.261***	0.262***
Ghost games	-0.043	-0.082	-0.043
Age		0.013	-0.042**
Male		1.209	1.184
Female		1.204	1.421
Third gender		0.594	0.675
Less than high school degree		2.665***	2.418***
High school degree		1.863**	1.651**
Some degree		2.015***	1.775**
Bachelor's degree		1.835**	1.698**
Master's degree		1.555*	1.334
Small club		-0.151	0.153
Medium club		0.436	0.562
Relationship length			-0.027
Attitudinal loyalty x relationship length			0.016*
Pseudo R^2 of Nagelkerke	.700	.712	0.723

*Table 46: coefficients of the three models for model equation II. *=significant at the 10%-level, **=significant at the 5%-level, ***=significant at the 1%-level*

4.8.2 Interpreting the main effect coefficients

As known, coefficient with a *p*-value higher than .10 do not significantly differ from zero, which mean they cannot be interpreted and thus will not be analyzed. Besides the coefficients of the IV's, their respective standard error is shown to gain insight how close the coefficients came to the fitted output. The lower the standard error, the better the prediction.

Furthermore, the coefficient was again raised to the power of e to compute the change in odds ratio for a one unit increase of an IV and the confidence interval is displayed, which shows a range of which there is a 95% guarantee that the estimated coefficient falls in that range. The wider this range, the higher the uncertainty is about the preciseness of the measured coefficient.

For example, the variables attitudinal loyalty and male both have coefficients which do not differ that much with respective values of 1.040 and 1.184, as shown in Table 47. However, the S.E of male is more than three times as high as the S.E of attitudinal loyalty (see Table 47), which automatically results in a much wider confidence interval for male to guarantee the 95% level of confidence this research strives for. In other words, nearly equal coefficients do not always translate to equal levels of accuracy.

Attitudinal loyalty is the first variable which will be interpreted with its *p*-value of .000 (see Table 47), which means that the variable differs significantly from zero at the 1%-level. So, for every one unite increase in attitudinal loyalty, the log odds of falling at a greater level of behavioral loyalty will increase with 1.040, which equals an increase in odds ratio of 2.829, remaining the other variables constant.

Secondly, the variable fan identification also differs significantly from zero at the 1%-level, because it has a *p*-value .000. So, for every one unit increase in fan identification, the log odds of falling at a greater level of behavioral loyalty will increase with 0.798, which equals an increase in odds ratio of 2.221 (see Table 47), holding the other variables constant.

Thirdly, the latent construct of social media differs significantly from zero at the 1%-level in this model with its respective *p*-value of .000, as shown in Table 47. For every one unit increase in social media, the log odds of falling at a greater level of behavioral loyalty will increase with 0.828, which equals an increase in odds ratio of 2.289, holding the other variables in the model constant.

Nostalgic content is the first main effect variable measured by a one-item that is significant and thus can be interpreted. Namely, the variable has a *p*-value of .002, which means that nostalgic content differs significantly from zero at the 1%-level. For every one unit increase in nostalgic content, the log odds of behavioral loyalty will decrease with .252, which equals an increase in odds ratio of 0.777, keeping the other variables constant (see Table 47). Keep in mind that an increase in odds ratio means a decreased probability of falling in a greater level of the DV, which is in this case behavioral loyalty.

Parental influence, also measured by one-item, has a *p*-value of .003, which means that the variable differs significantly from zero at the 1%-level. So, for every one unit increase in parental influence, the log odds of falling at a greater level of attitudinal loyalty increase with 0.196, which equals an increase in odds ratio of 1.217 (see Table 47), keeping the other variables constant.

The final significant variable which measures the main effect is a latent construct, namely the lack of fan interaction. The variable has a respective *p*-value of .009, which means that it significantly differs from zero at the 1%-level, which is equal to an increase in odds ratio of 1.300, keeping the other variables constant, as shown in Table 47.

4.8.3 Interpreting the control variables coefficients

As for the control variables, age and education are significantly different from zero, except the categorical level master's degree for education with a *p*-value 0.104 just above the threshold of .10 for reaching the 10%-level of significance (see Table 47). Of the other four categories, less than high school degree has the strongest relationship with the DV.

Namely, with its *p*-value of .007 it is not only the sole category which is significantly differing from zero at the 1%-level, but it also has the highest coefficient of the five categories. On average, the log odds of falling in a greater level of behavioral loyalty were 2.418 higher of respondents who had less than a high school degree as their highest level of education compared to respondents who obtained a PhD as their highest level of education, which equals an increase in odds ratio of 11.223, remaining the other variables constant (see Table 47).

The other three coefficients of the significant categories were clustered together with high school being the lowest of the three with a value of 1.651 and some degree was the highest with a value of 1.775, as shown in Table 47. Naturally, all three are significant and can thus be interpreted, but due to the fact that they are highly clustered only one will be interpreted in this subsection.

The categorical variable some degree significantly differs from zero at the 5%-level with its respective *p*-value of .022, as shown in Table 47. So, on average, the log odds of falling in a greater level of behavioral loyalty were 1.775 higher for respondents who obtained some degree as their highest level of education compared to respondents who obtained a PhD as their highest level of education, which equals an increase in odds ratio of 5.212, keeping the other variables constant.

The other significant control variable was age with a respective *p*-value of .011, which means that it is significantly different from zero for the 5-% level. For every one unit increase of age (i.e., one year), the log odds of falling in a greater level of attitudinal loyalty will increase with .042, which equals an increase in odds ratio of 0.959, remaining the other variables constant (see Table 47).

4.8.4 Interpreting the interaction effect

Last but not least, the interaction effect between attitudinal loyalty and relationship length will be discussed. Attitudinal loyalty was already present in the model before adding the interaction, but relationship length was added to the main effect to prevent displaying a confounding effect, which will occur even if one of the variables is not significant. In other words, the interaction coefficient will not be accurate without the addition of both main effects in the regression.

The relationship length variable on its own has a respective *p*-value of .596, which means that it does not significantly differ from zero. However, when interacting the variable relationship length with attitudinal loyalty, a significant coefficient of 0.016 is found with a respective *p*-level of 0.069 (see Table 47). This means that the interaction term is significant at the 10%-level.

Keep in mind that the interaction coefficient is not the net effect of the two variables on the DV. To identify the net effect, the interaction coefficient has to be added on the coefficient of the variables it selves. However, the net effect cannot be derived due to the fact that relationship length does not significantly differ from zero.

If relationship length would differ significantly from zero, the interpretation would be as followed with the net effect in coefficient value being 1.029 (1.040 + 0.016 – 0.027, see Table 47). For every one unit increase in attitudinal loyalty *and* a unit (i.e, one year) increase in relationship length, the log odds of being in a greater level of behavioral loyalty will increase with 1.029, which equals an increase in odds ratio of 2.798, keeping the other variables constant. In other words, as the relationship length increases, the effect of attitudinal loyalty on behavioral loyalty gets greater and greater.

	β	S.E	exp β	*p*-value	Lower bound 95% C.I	Upper bound 95% C.I
Attitudinal loyalty	1.040	0.242	2.829	.000	0.565	1.515
Fan identification	0.798	0.147	2.221	.000	0.510	1.087
Social media	0.828	0.150	2.289	.000	0.534	1.122
Team content	0.077	0.131	1.080	.559	-0.180	0.333
Nostalgic content	-0.252	0.082	0.777	.002	-0.413	-0.092
Esports	0.109	0.075	1.115	.146	-0.038	0.255
Parental influence	0.196	0.065	1.217	.003	0.069	0.324
Friend influence	0.029	0.065	1.029	.658	-0.098	0.155
Geographical location	-0.049	0.067	0.952	.463	-0.180	0.082
Team's success	-0.014	0.068	0.986	.842	-0.148	0.120
Lack of fan interaction	0.262	0.100	1.300	.009	0.066	0.459
Ghost games	-0.043	0.070	0.958	.538	-0.179	0.093
Relationship length	-0.027	0.051	0.973	.596	-0.127	0.073
Att. Loyalty x Rel. Length	0.016	0.009	1.016	.069	-0.001	0.032
Male	1.184	0.878	3.267	.349	-1.292	3.660
Female	1.421	1.249	4.141	.264	-1.071	3.913
Third gender	0.675	1.793	1.964	.707	-2.839	4.189
Age	-0,042	0.016	0.959	.011	-0.074	-0.010
Less than high school degree	2.418	0.904	11.223	.007	0.646	4.190
High school	1.651	0.783	5.212	.035	0.116	3.186
Some degree	1.775	0.773	5.900	.022	0.259	3.291
Bachelor's degree	1.698	0.775	5.463	.028	0.180	3.217
Master's degree	1.334	0.821	3.796	.104	-0.276	2.944
Small club	0.153	0.323	1.165	.636	-0.480	0.786
Medium club	0.562	0.283	1.754	.047	0.008	1.116

Table 47: statistics for the regression analysis of model equation II

4.8.5 Interpreting the constants

For the same reasons as model equation I, the DV has, unlike in linear regressions, multiple constants. Table 48 shows the observed levels of the DV and its respective constant and *p*-value. This DV has more levels than the DV in the previous model equation. Naturally, this can occur due to the latent nature of the DV's. If the DV was measured by a single-item with a 7-point Likert scale, it would have seven levels and thus seven constants.

However, the latent variables of both model equations were constructed by taking the average of five observed item. This results in levels with non-rounded values, which can differ between the DV's, depending on the actual observed averages. So, for this research more unique average were found for the DV in model equation II than model equation I. Similar to the constants of model equation I, the constants gain a higher value when the level of behavioral loyalty increases, as shown in Table 48.

As for the significance of the constants, only the lowest level of 1 differs not significantly from zero. The remaining 26 levels are all statistically different from zero at the 1%-level. Again, an intercept has only to be interpreted when all the predictor values have a value of 0, which is almost impossible.

Level of behavioral loyalty	**Constant (α)**	***p*-value**
1.00	3.300	.189
2.00	9.836	.000
2.20	10.676	.000
2.40	11.459	.000
2.60	11.689	.000
2.80	11.901	.000
3.00	12.431	.000
3.20	13.362	.000
3.40	13.615	.000
3.60	14.192	.000
3.80	14.657	.000
4.00	14.952	.000
4.20	15.588	.000
4.40	16.284	.000
4.60	17.232	.000
4.80	17.956	.000
5.00	18.645	.000
5.20	19.257	.000
5.40	19.938	.000
5.60	20.642	.000
5.80	21.268	.000
6.00	21.708	.000
6.20	22.426	.000
6.40	23.277	.000
6.60	24.176	.000
6.80	24.855	.000

Table 48: interception estimates with their respective p-values for model equation II

4.8.6 Interpreting model fit

As of now, the OLR output is analyzed and interpreted when possible for both model equations. The final part of the results section will focus properly interpreting three model fit tests and one statistical measurement for assessing model fit for model equation II.

Model	**-2 Log Likelihood**	**Chi-Square (χ2)**	**Degrees of freedom (df)**	***p*-value**
Null hypothesis	2369.072			
General	1858.534	510.538	25	.000

Table 49: likelihood chi-square test for model equation II

The test to assess the model fit, the likelihood ratio chi-square test was used. Table 49 shows that a χ2 of 2369.072 with a df of 25 results in a *p*-value of .000, which means that the final model is a significant improvement in fit over the null model.

Test	**Chi-Square (χ2)**	**Degrees of freedom (df)**	***p*-value**
Pearson	8945.098	10349	1.000
Deviance	1632.172	10349	1.000

Table 50: Pearson chi-square and Deviance test for model equation II

Secondly, the Pearson chi-square and Deviance test measured how well the model of the second equation fits the data set. Table 50 shows that the respective *p*-values of both tests are higher than .05. Namely, the *p*-value for the Pearson chi-square test are both 1.000. A *p*-value of >.05 means that the there was no significant discrepancy found between model equation I and the final model.

The final statistical measurement used to assess the model fit of this equation is the pseudo R^2 of Nagelkerke. Model equation II has with a pseudo R^2 of 0.723 a substantially larger value than the previously analyzed model with a pseudo R^2 of 0.648, as shown in Table 41 and Table 46 So, for model equation II, approximately 72.3% of the variability in the DV is explained by the model, which can be interpreted as an explanatory power of the model that is in between moderate and substantial. However, due to the fact that pseudo R^2 yield lower estimates in general than R^2 for OLR, the explanatory power of this model should definitely be more classified as substantial than moderate.

5. Conclusion and discussion

The final chapter of this research will first answer the sub-questions with the help of the hypotheses in order to answer the main research question of this paper in the best way possible. After answering the main research question, the implications for academics and mangers in the field of sport marketing will be covered. Finally, in the discussion section some limitations of this research will be presented and at the end of this chapter, possible interesting directions for future research will be showed.

5.1 Answering the sub-questions

The first sub-question was as followed: what is the influence of fan identification on fan loyalty? To investigate this matter, a positive, hypothesized relationship between both dimensions of fan loyalty and fan identification, resulting in the following hypothesis:

H_2: Fan identification has a strong, positive effect on both dimensions of fan loyalty.

First, the latent construct of fan identification was validated with CFA to show the cohesiveness of the four observed items. Secondly, fan identification was put in both model equations as an IV predicting both the dimensions of loyalty.

The coefficients of fan identification were not only significant at the 1%-level and positive, but also very strong. Fan identification was the strongest predictor in model equation I and the third strongest predictor in model equation II of all the variables which measured a main effect. Hence, H_2 is accepted and is also the answer to the first sub-question.

The second sub-question was: does a favorable attitude towards club-related social media influence fan loyalty? For answering this sub-question, the following hypothesis was drafted:

H_{3a}: Making use of the versatility of social media will enhance both dimensions of fan loyalty.
Social media was also a latent construct in this research, of which the cohesiveness of its three observed items was also confirmed by CFA. Just like fan identification, the variable social media was put in both model equations as an IV in order to test the hypothesis. However, in

contrary to fan identification, a negative, nonsignificant effect was found on attitudinal loyalty.

In model equation II, social media was a highly significant, positive predictor of behavioral loyalty, with an even stronger effect than fan identification. In other words, social media only enhances the behavioral dimension of loyalty, which results in partially accepting H_{3a}.

To gain a more detailed answer on this sub-question, a second hypothesis was drafted in order to determine which type of content will enhance loyalty the most, which is as followed:

H_{3b}: Team centered content will have a stronger, positive effect on fan loyalty than nostalgic content.

To investigate this hypothesized relationship, both types of content were put in model equation I and II in order to fully capture their effect on fan loyalty. As for model equation I, nostalgic content showed no significant effect, whereas team content showed a small, positive effect significant at the 10%-level.

In model equation II, the tables were turned and nostalgic content showed a significant effect, even at the 1%-level. However, this effect was moderate and negative, whereas the variable team content did not significantly differ from zero in model II. Thus, for attitudinal loyalty team content positively influences this dimension, whereas nostalgic content worsened the level of behavioral loyalty. So, H_{3b} is also accepted. To answer the sub-question, social media enhances the behavioral dimension of loyalty, but not when the content type is mainly nostalgic.

The third sub question relates to esports and was as followed: could the addition of an esports department affect fan loyalty? To answer this question properly, the following hypothesis was drafted:

H_4: The addition of an esports department will positively impact the behavioral dimension of fan loyalty.

After sufficiently validating this latent construct with CFA, the variable esports was only an IV in model equation II attempting to predict the levels of behavioral loyalty. The coefficient value of esports was small, but positive. However, the variable did not differ significantly from zero, so the variable could not be interpreted. So, H_4 was rejected, which means that an esports department focused on playing the videogame FIFA did not significantly affect the behavioral dimension of fan loyalty.

The fourth sub question related to four reasons of initially becoming a fan, which was: does the reason of initially becoming a fan matter for the degree of fan loyalty? For answering this sub-question, two hypotheses were drafted:

H_{5a}: Being a fan through the team's success will be the only reason which negatively impacts both dimensions of fan loyalty.
H_{5b}: Parental influence as a reason for becoming a fan has the strongest effect on both dimensions of fan loyalty.

So, the four reasons parental influence, friend influence, geographical location and the team's success were all put in both model equations to predict the levels of both attitudinal and behavioral loyalty. As for model equation I, both friend influence and the team's success significantly differed from zero with a negative coefficient. Furthermore, the effect was stronger for friend influence than the team's success.

In model equation II, only one reason of initially becoming a fan was significantly different from zero, which was parental influence with a moderate, positive effect on the behavioral dimension of loyalty. So, H_{5a} is rejected, because friend influence also worsened the degree of attitudinal loyalty and as for the behavioral dimension, no negative coefficients were significantly different from zero.

The second hypothesis drafted to assist in answering this sub-question, H_{5b} is partially accepted. Namely, as for the attitudinal dimension nor a significant nor a positive effect was found, but for the behavioral dimension the variable parental influence was the only significant reason in the model, which makes it automatically the strongest effect of the four.

The fifth sub-question of this research was related to ghost games, the newborn phenomena during the pandemic, and was as followed: did ghost games affect fan loyalty? For answering this sub-question, the following hypothesized relationship was designed and formulated in the following way:

H_6: Ghost games did have a negative impact on the behavioral dimension of fan loyalty.

So, the variable ghost games was also a latent construct, which was first validated with the help of CFA. Afterwards, the construct was put in model equation II attempting to predict the DV behavioral loyalty as an IV, which measured the main effect. However, the variable did not significantly differ from zero, which means that the small, negative coefficient could not be interpreted. Hence, H_6 was rejected and to answer the sub-question: ghost games did not significantly affect fan loyalty.

The sixth sub-question also related to a new matter which aroused during the pandemic. Namely, did the lack of fan interaction affect fan loyalty? To properly answer this sub-question, the following hypothesis has been drafted:

H_7: The lack of fan interaction weakens the degree of the behavioral dimension of fan loyalty.

So, after again validating the latent construct by assessing the cohesiveness of its three observed items, the latent variable was put in model equation II as an IV measuring the main effect and attempting to predict the DV, which is behavioral loyalty in this case.

Surprisingly, the coefficient did not only significantly differ from zero at the 1%-level, but was of moderate strength in the positive direction. So, H_7 was rejected and to answer the sub-question: the lack of fan interaction affected fan loyalty in a positive way, which was the opposite of the expected direction.

The second-last sub-question strives to seek an answer on the intern relationship between the attitudinal and behavioral dimension of fan loyalty. The sub-question is as followed: what is the relationship between the two dimensions of fan loyalty? The following relationship has been hypothesized in order to give a proper answer to this question:

$H_{1a:}$ In the double-dimension fan loyalty concept, attitudinal loyalty is the predictor of behavioral loyalty.

So, where attitudinal loyalty was the predicted variable in model equation I, it functioned as an IV measures the main effect on the DV in the form of behavioral loyalty. Naturally, the latent construct was first validated with the help of CFA. Attitudinal loyalty was significantly different from zero in model equation II at the 1%-level, where its respective coefficient was strong and positive. It was even the highest coefficient of all the IV's which measured the main effect. So, H_{1a} was accepted, which makes the hypothesized relationship between the two dimensions the answer to this sub-question.

The final sub-question of this research relates to the role of the relationship length between a fan and its favorite club on the relationship between the attitudinal and behavioral dimension of loyalty. So, the sub-question is as followed: does the relationship length of a fan with their club influence the relationship between the two dimensions of fan loyalty? For answering this sub-question, the following hypothesis has been drafted:

H_{1b}: Relationship length acts as a moderator and positively moderates the attitudinal-behavioral relationship.

To test this hypothesis, an interaction variable was computed between the relationship length and the attitudinal loyalty in model equation II to find out whether there is a moderating effect of relationship length on this inter dimension relationship.

Unfortunately, the coefficient of the sole variable relationship length was not significantly different from zero, so it could not be interpreted, which means the net effect of this interaction effect was not identified. However, the interaction effect itself was significantly different from zero at the 10%-level with a small, positive coefficient.

So, even though the net effect could not be computed, the interaction effect showed that it positively influences the relationship, only to which exact degree remains unknown. Hence, H_{1b} is accepted and is also the answer to the final sub-question of this paper.

5.2 Answering the main question

After extensively answer all the eight sub-questions, enough knowledge is obtained to give an extensive, well-argued answer to the main research question of this paper, which was:

'How can professional football clubs strengthen their current degree of fan loyalty during and after the pandemic?'

So, this paper identified two dimensions of fan loyalty, attitudinal and behavioral, of which attitudinal is the predictor of the behavioral dimension. In the end, behavioral loyalty is the desired goal for football clubs due to the revenue generating activities done by fans who show a high degree of behavioral loyalty, such as purchasing club merchandise or attending football matches.

However, behavioral loyalty is significantly predicted by the attitudinal dimension, which means that this dimension should not be overlooked. Namely, attitudinal loyalty enhancing factors, such as fan identification and team-related content on social media, can indirectly positively influence behavioral loyalty through attitudinal loyalty.

To conclude, football clubs can strengthen their current degree of fan loyalty by prioritizing, enhancing, measuring and monitoring the degree of fan identification. Namely, this factor was the only factor which showed a significant, strong effect on both dimensions. Furthermore, allocating more marketing budget to social media campaigns will strengthen the degree of fan loyalty, but the campaigns should not have nostalgia as their central theme.

Furthermore, for increasing the degree of fan loyalty in the long term, football clubs should prioritize gaining new fans through parental influence and not solely rely on their team's success, which negatively influenced the attitudinal dimension of fan loyalty.

5.3 Managerial implications

The goal of this paper was to hand a blueprint to sport marketers on which factors need to be enhanced and which ones mitigated to maximize the degree of fan loyalty and achieve the four benefits of brand loyalty of Griffin (2002), which were cost savings, PWOM, complain rather than defect and cross-selling through multiple channels.

The most important factor was, by far, fan identification. So, sport marketers should transform fan identification into a measurable KPI in order to monitor the progress made on this loyalty enhancing factor. This can be done by conducting randomly a survey questionnaire or in-depth interviews with fans on a three months' basis to track the progress made.

As seen in the conclusion, social media must have a prominent place in the marketing strategy to enhance the degree of behavioral loyalty when using team-related content instead of nostalgic content. Besides team-related content, cross-selling opportunities, like the promotion of club merchandise, should be fulfilled to maximize revenue.

So, this tool of improving service quality and thus loyalty will be optimally used. Thus, tracking social media related KPI's, such as conversation rate of click-through links or the total reach, helps with assessing your success on strengthening the degree of fan loyalty.

Furthermore, due to the vital importance of fan identification in the process of forming strong degrees of fan loyalty, it is a necessity to know the character traits of your fan base. If the marketing communication strategies are not in alignment with the predominant characteristics of your fan base, the loyalty will not be strengthened. For example, a marketing strategy centered around winning for a humble fan base will not enhance identification and thus loyalty.

A combination of general knowledge about the respective fan base and market research should give good insights into which character traits are common under your fan base to determine your main communication message in your marketing activities.

Finally, to ensure a sustainable loyal fan base for the future, sports marketers should focus in convincing young children to support their club through parental influence, as it drives the behavioral dimension of fan loyalty. For example, a parent-children day can be organized with some recreational activities in and around the stadium. Also, clubs could drastically lower the ticket prices for young kids to remove a potential barrier of taking your kid(s) with you to the stadium, especially for the lower educated, but in general more loyal football fans.

5.4 Academic implications

In the beginning of this paper, it was stated that the academic relevance will be guaranteed with breakthrough findings with regards to the changed environment in the football world during the pandemic. In contrary to the expectations of this paper, based on the loyalty enhancing NPR-attributes entertainment and escapism (Bauer et al., 2008; Gladden and Funk, 2001), ghost games did not negatively impact the behavioral dimension of loyalty, whereas the lack of fan interaction even increased the level of behavioral loyalty.

At first sight, this may seem invalid, but it can be explained. Based on the research conducted, the lack of fan interaction did not shift the fans away to other hedonic activities like watching Netflix or playing games, but the fans spent their lost time on interacting with fans on hedonic activities related to their favorite club, such as consuming club-related social media content and watching more football matches.

So, fans showing these strong levels for the behavioral dimension of loyalty during the pandemic leans surprisingly more towards the exclusive fan loyalty concept of Stuart and Parker (1997) than his counterpart in the field of sport marketing, who does not agree with the saying ‘we’ll support you evermore’ (Tapp, 2004).

However, Tapp (2004) is still right that loyalty is not something to rely on, because even if fans do not easily defect from a club, their degree of behavioral loyalty will decrease if not enough of the marketing budget is allocated to loyalty enhancing strategies. So, researchers should definitely not stop with seeking more ways to enhance fan loyalty.

Furthermore, this paper has validated seven latent constructs related to fan loyalty, which researchers can use when conducting research on this concept. Besides newly designed constructs, this paper updated slightly outdates constructs of Bauer et al. (2008) to the current more digitalized world we are living in.

Besides the breakthrough in identifying the effect of the restrictions on fan loyalty, this research also saw consistency in the pre- and during pandemic strength of fan identification (Bauer et al., 2008; Gladden and Funk, 2001). Also, the attitudinal-loyalty link of Dick and Basu (1994) is now validated for the sports environment, with the addition of relationship length as a moderator in this relationship. As for digital activities, social media strategies were shown to be effective, whereas esports activities were not classified as a loyalty-enhancing activity by this research.

5.5 Limitations and bias

The main bias this research suffered from is selection bias, caused by convenience sampling, which was the main sampling method of this research. Due to the selection bias, the continuous variables age and relationship length were not close to a normal distribution, but skewed to the left. In other words, younger fans were overrepresented and older fans underrepresented.

The internal validity was successfully assessed the four respective types face, convergent, construct and predictive validity. So, the results are internally valid, but due to the left-skewed distribution only partially externally valid. In other words, inferences can be drawn of the sample on football fans who are aged between 18-34, but not on the middle aged and older adults.

A side note must be placed that the model fit-indices of the measurement model in CFA, unlike those of OLR, were just above the threshold of an acceptable fit, but that does not reject the internal validity of this research. Namely, the factor loadings were well above the minimum cut-off criterion and also significant.

As for the reliability, only one of the three attributes were tested, which was internal consistency. Even though actions were taken to improve the reliability of this research, there is no guarantee whether the actions were effective or not, because the test-retest for assessing stability and the alternative form method for equivalence were both not used in this research.

5.6 Directions for future research

To strengthen the external validity of this paper, other researchers should also investigate the relationship between phenomena which were caused by the restrictions of the pandemic, such as ghost games and the lack of fan interaction. Probability sampling is advised to make stronger inferences on the whole population of soccer fans, which could potentially be done by randomly drawing a sample from the fan base of one football club.

Also, researchers active in this field could dive deeper into the most effective social media strategies. Social media was found to be a positive predictor of the behavioral dimension of fan loyalty, but there is no strong consensus yet on which type of content is the most effective with only a small indirect effect found of team-related content on behavioral loyalty.

Finally, diving deeper into the drivers of fan identification could be an interesting topic in the field of sports marketing. Namely, its importance was acknowledged in this paper, but it remains unclear what exactly drives fan identification. So, assessing numerous character traits of fans and linking them all to an advised main theme in their marketing communication could strengthen the given blueprint for sport marketers.

Reference list

Abenhuijs, M. (2021, 2nd of January). *Loyaliteit voetbalfan is eindig: 'Stip aan de horizon moet snel komen'.* AD. Retrieved from: https://www.ad.nl/nederlands-voetbal/loyaliteit-voetbalfan-is-eindig-stip-aan-de-horizon-moet-snel-komen~a9b6271a/

Ackermann, S., & Von Wangenheim, F. (2014). Behavioral Consequences of Customer-Initiated Channel Migration. *Journal of Service Research, 17*(3), 262–277.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.) Prentice Hall/Pearson Education.

Awang, Z. (2012). *Research Methodology and Data Analysis* (2nd Ed.) Malaysia: Press UitM. Pp 80-118.

Baldinger, A. L., & Rubinson, J. (1996). Brand loyalty: the link between attitude and behavior. *Journal of Advertising Research, 36*(6), 22-34.

Bandyopadhyay, S., & Martell, M. (2007). Does attitudinal loyalty influence behavioral loyalty? A theoretical and empirical study. *Journal of Retailing and Consumer Services, 14*(1), 35–44.

Bartlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational research: determining appropriate sample size in survey research. *Information Technology, Learning and Performance Journal, 19*(1), 43-50.

Bastedo, M. N. (2004). Open systems theory. *The SAGE Encyclopedia of Educational Leadership and Administration.* Sage Publications.

Bauer, H. H., Stokburger-Sauer, N. E., & Exler, S. (2008). Brand Image and Fan Loyalty in Professional Team Sport: A Refined Model and Empirical Assessment. *Journal of Sport Management, 22*(2), 205–226.

Bauer, H. H., Stokburger-Sauer, N. E., & Exler, S. (2005). The loyalty of German soccer fans: does a team's brand image matter? *International Journal of Sports Marketing and Sponsorship, 7*(1), 8–16.

Bennett, R., & Rundle-Thiele, S. (2004). Customer satisfaction should not be the only goal. *Journal of Services Marketing, 18*(7), 514–523.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588–606.

Berinsky, A. J., Margolis, M. F., Sances, M. W., & Warshaw, C. (2019). Using screeners to measure respondent attention on self-administered surveys: Which items and how many? *Political Science Research and Methods, 9*(2), 430–437.

Berndt, A. E. (2020). Sampling Methods. *Journal of Human Lactation, 36*(2), 224–226.

Bertram, D. (2007). *Likert Scales are the Meaning of Life.* CPSC 681 – Topic Report.

Bohte, W., Maat, K., & Van Wee, B. (2009). Measuring Attitudes in Research on Residential Self-Selection and Travel Behaviour: A Review of Theories and Empirical Research. *Transport Reviews, 29*(3), 325–357.

Brown, T. A., Moore, M. T. (2012). Confirmatory factor analysis. In R. H. Hoyle (Ed.), *Handbook of structural equation modelling*. New York, NY: The Guilford Press. Pp. 361-379.

Casson, R. J., & Farmer, L. D. (2014). Understanding and checking the assumptions of linear regression: a primer for medical researchers. *Clinical & Experimental Ophthalmology, 42*(6), 590–596.

Chevalier, J. A., & Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research, 43*(3), 345–354.

Chinomona, R., & Dubihlela, D. (2014). Does customer satisfaction lead to customer trust, loyalty and repurchase intention of local store brands? The case of Gauteng Province of South Africa. *Mediterranean Journal of Social Sciences, 5*(9), 23-32.

Churchill, G. A., & Surprenant, C. (1982). An Investigation into the Determinants of Customer Satisfaction. *Journal of Marketing Research, 19*(4), 491–504.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.) Hillsdale, NJ: Erlbaum.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334.

De Bruijne, M., & Wijnant, A. (2014). Improving Response Rates and Questionnaire Design for Mobile Web Surveys. *Public Opinion Quarterly, 78*(4), 951–962.

Deng, Z., Lu, Y., Wei, K. K., & Zhang, J. (2010). Understanding customer satisfaction and loyalty: An empirical study of mobile instant messages in China. *International Journal of Information Management, 30*(4), 289–300.

Dick, A. S., & Basu, K. (1994). Customer Loyalty: Toward an Integrated Conceptual Framework. *Journal of the Academy of Marketing Science, 22*(2), 99–113.

Dietz-Uhler, B., Harrick, E.A., End, C., & Jaquemotte, L. (2000). Sex differences in sport fan behavior and reasons for being a sport fan. *Journal of Sport Behavior, 23*(3), 219–232.

Duffy, D. L. (2003). Internal and external factors which affect customer loyalty. *Journal of Consumer Marketing, 20*(5), 480–485.

East, R., Hammond, K., & Lomax, W. (2008). Measuring the impact of positive and negative word of mouth on brand purchase probability. *International Journal of Research in Marketing, 25*(3), 215–224.

Ehrenberg, A., & Goodhardt, G. (2000). New Brands: Near-Instant Loyalty. *Journal of Marketing Management, 16*(6), 607–617.

Ehrenberg, A. S., Uncles, M. D., & Goodhardt, G. J. (2004). Understanding brand performance measures: using Dirichlet benchmarks. *Journal of Business Research, 57*(12), 1307–1325.

Erkan, A., & Yildiz, Z. (2014). Parallel lines assumption in ordinal logistic regression and analysis approachis. *International Interdisciplinary Journal of Scientific Research, 1*(3), 8-23.

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics, 5*(1), 1-4.

Fakazli, A. (2020). The Effect of Covid-19 Pandemic on Digital Games and eSports. *International Journal of Sport Culture and Science, 8*(4), 335-344.

Finstad, K. (2010). Response Interpolation and Scale Sensitivity: Evidence Against 5-Point Scales. *Journal of Usability Studies, 5*(3), 104-110.

Fisher, M. J., & Marshall, A. P. (2009). Understanding descriptive statistics. *Australian Critical Care, 22*(2), 93–97.

Gazapo, C. (2020, 5th of February). *TV Rights in Football - Premier League Analysis.* Sports Business Institute Barcelona.
Retrieved from: https://www.sbibarcelona.com/newsdetails/index/403

Gladden, J. M., & Funk, D. C. (2001). Understanding Brand Loyalty in Professional Sport: Examining the Link Between Brand Associations and Brand Loyalty. *International Journal of Sports Marketing and Sponsorship, 3*(1), 54–81.

Gliem, J. A., & Gliem, R. R. (2003) Calculating, interpreting and reporting Cronbach's alpha reliability coefficient for Likert-type scales. *Midwest research to practice conference in adult, continuing, and community education.* Columbus, OH: The Ohio State University.

Greenwood, P. B. (2001). *Sport Fan Team Identification in a Professional Expansion Setting.* Master's thesis, North Carolina State University, Raleigh, NC.

Greenwood, P. B., Kanters, M. A., & Casper, J. M. (2006). Sport Fan Team Identification Formation in Mid-Level Professional Sport. *European Sport Management Quarterly, 6*(3), 253–265.

Griffin, J. (2002). Customer Loyalty: How to Earn It, How to Keep It. Jossey-Bass.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice, 19*(2), 139–152.

Hallencreutz, J., & Parmler, J. (2019). Important drivers for customer satisfaction – from product focus to image and service quality. *Total Quality Management & Business Excellence, 32*(5–6), 501–510.

Hamari, J., & Sjöblom, M. (2017). What is eSports and why do people watch it? *Internet Research, 27*(2), 211–232.

Harrell, F. E. (2015). Ordinal Logistic Regression. In, Regression *Modeling Strategies: With Application to Linear Models, Logistic and Ordinal Regression, and Survival Analysis* (2nd Ed.), New York, NY: Springer. Pp. 311–325.

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence Based Nursing, 18*(3), 66–67.

Heere, B., & Dickson, G. (2008). Measuring Attitudinal Loyalty: Separating the Terms of Affective Commitment and Attitudinal Loyalty. *Journal of Sport Management, 22*(2), 227-239.

Holgado–Tello, F. P., Chacón–Moscoso, S., Barbero–García, I., & Vila–Abad, E. (2008). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity, 44*(1), 153–166.

Homburg, C., & Giering, A. (2000). Personal characteristics as moderators of the relationship between customer satisfaction and loyalty-an empirical analysis. *Psychology and Marketing, 18*(1), 43–66.

Israel, G. D. (1992). *Determining sample size.* Gainesville, FL: University of Florida Institute of Food and Agricultural Sciences (IFAS) Extension.

Jacoby, J., & Chestnut, R. W. (1978). *Brand Loyalty: Measurement and Management (Wiley Series on Marketing Management)* (1st Edition). John Wiley & Sons Inc.

Jenny, S. E., Keiper, M. C., Taylor, B. J., Williams, D. P., Gawrysiak, J., Manning, R. D., & Tutka, P. M. (2018). eSports Venues: A New Sport Business Opportunity. *Journal of Applied Sport Management, 10*(1), 34–49.

Jones, I. (1997). A further examination of the factors influencing current identification with a sports team, a response to Wann et al. (1996). *Perceptual and Motor Skills, 85*(1), 257-258.

Jones, M. A., & Suh, J. (2000). Transaction-specific satisfaction and overall satisfaction: an empirical analysis. *Journal of Services Marketing, 14*(2), 147–159.

Joshi, A., Kale, S., Chandel, S., & Pal, D. (2015). Likert Scale: Explored and Explained. *British Journal of Applied Science & Technology, 7*(4), 396–403.

Kaynak, E., Salman, G. G., & Tatoglu, E. (2007). An integrative framework linking brand associations and brand loyalty in professional sports. *Journal of Brand Management, 15*(5), 336–357.

Koch, K., & Wann, D. L. (2016). Team identification and sport fandom: gender differences in relationship-based and recognition-based perceived antecedents. *Journal of Sport Behavior, 39*(3), 278-300.

Kristensen, K., Martensen, A., & Gronholdt, L. (2000). Customer satisfaction measurement at Post Denmark: Results of application of the European Customer Satisfaction Index Methodology. *Total Quality Management, 11*(7), 1007–1015.

Lantz, B. (2012). The large sample size fallacy. *Scandinavian Journal of Caring Sciences, 27*(2), 487–492.

Lee, D., & Schoenstedt, L. J. (2011). Comparison of eSports and traditional sports consumption motives. *The ICHPER-SD Journal of Research in Health, Physical Education, Recreation, Sport & Dance, 6*(2), 39-45.

Lehnert, K., Walz, A., & Christianson, R. (2020). The booming eSports market: a field day for fans. *Journal of Business Strategy.* Published ahead-of-print on December 9, 2020.

Lettieri, E., & Orsenigo, C. (2020). Predicting soccer consumption: do eSports matter? Empirical insights from a machine learning approach. *Sport, Business and Management: An International Journal, 10*(5), 523–544.

Levinson, D., & Pfister, G. (2013). *Berkshire Encyclopedia of World Sport.* Van Haren Publishing.

Li, J., Konuş, U., Langerak, F., & Weggeman, M. C. (2016). Customer Channel Migration and Firm Choice: The Effects of Cross-Channel Competition. *International Journal of Electronic Commerce, 21*(1), 8–42.

Lorenzo-Seva, U., & Ferrando, P. J. (2014). POLYMAT-C: a comprehensive SPSS program for computing the polychoric correlation matrix. *Behavior Research Methods, 47*(3), 884–889.

Mahony, D. F., Madrigal R., & Howard, D. R. (2000). Using the psychological commitment to team (PCT) scale to segment sport consumers based on loyalty. *Sport marketing Quarterly, 9*(1), 15-25.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103*(3), 391–410.

Martin, W. C. (2017). Positive versus negative word-of-mouth: Effects on receivers. *Academy of Marketing Studies Journal, 21*(1), 1528–2678.

Mastromartino, B., Ross, W. J., Wear, H., & Naraine, M. L. (2020). Thinking outside the 'box': a discussion of sports fans, teams, and the environment in the context of COVID-19. *Sport in Society, 23*(11), 1707–1723.

Matsunaga, M. (2010). How to factor-analyze your data right: do's, don'ts, and how-to's. *International Journal of Psychological Research, 3*(1), 97–110.

Mavletova, A., & Couper, M. P. (2014). Mobile Web Survey Design: Scrolling versus Paging, SMS versus E-mail Invitations. *Journal of Survey Statistics and Methodology, 2*(4), 498–518.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In F. Zabembka (Ed.), *Frontiers in econometrics*. New York: Academic Press. Pp. 105-142.

Miaoulis, G., & Michener, R. D. (1976). *An Introduction to Sampling*. Dubuque, Iowa: Kendal Hunt Publishing Company.

Morgeson, F. V., Hult, G. T. M., Mithas, S., Keiningham, T., & Fornell, C. (2020). Turning Complaining Customers into Loyal Customers: Moderators of the Complaint Handling–Customer Loyalty Relationship. *Journal of Marketing, 84*(5), 79–99.

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika, 78*(3), 691–692.

Newson, M., Buhrmester, M., & Whitehouse, H. (2021). United in defeat: shared suffering and group bonding among football fans. *Managing Sport and Leisure, 26*(1), 1–18.

Oluwatayo, J. A. (2012). Validity and Reliability Issues in Educational Research. *Journal of Educational and Social Research, 2*(2), 391-400.

Osborne, J. W. (2015). Curvilinear effects in logistic regression. In Osborne, J. W. (Ed.), *Best practices in logistic regression*. Thousand Oaks, CA: Sage Publications. Pp. 201-242.

Parker, K., & Stuart, T. (1997). The West Ham Syndrome. *Journal of the Market Research Society, 39*(3), 509-517.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design and analysis: An intergrated approach.* Hillsdale, NJ: Erlbaum.

Petrucci, C. J. (2009). A Primer for Social Worker Researchers on How to Conduct a Multinomial Logistic Regression. *Journal of Social Service Research, 35*(2), 193–205.

Psychology Educator (2013, 15th of August). *The Psychology of Sports Fans: How Fans Affect Players.* Retrieved from:
https://psychologyeducator.wordpress.com/2013/08/15/the-psychology-of-sports-fans-how-fans-affect-players/

Ranaweera, C., & Prabhu, J. (2003). On the relative importance of customer satisfaction and trust as determinants of customer retention and positive word of mouth. *Journal of Targeting, Measurement and Analysis for Marketing, 12*(1), 82–90.

Reinartz, W., Thomas, J. S., & Bascoul, G. (2008). Investigating cross-buying and customer loyalty. *Journal of Interactive Marketing, 22*(1), 5–20.

Reja, U., Manfreda, K. L., Hlebec, V., & Vehovar, V. (2003). Open-ended vs. close-ended questions in web quesstionnaires. *Development in Applied Statistics, 19*(1), 159-177.

Richardson, B., & O'Dwyer, E. (2003). Football supporters and football team brands: a study in consumer brand loyalty. *Irish marketing review, 16*(1), 43-53.

Sarstedt, M., Ringle, C. M., Raithel, S., & Gudergan, S. P. (2014). In Pursuit of Understanding What Drives Fan Satisfaction*. Journal of Leisure Research, 46*(4), 419–447.

Schmidt, W. C. (1997). World-Wide Web survey research: Benefits, potential problems, and solutions. *Behavior Research Methods, Instruments, & Computers, 29*(2), 274–279.

Sharma, G. (2017). Pros and cons of different sampling techniques. *International Journal of Applied Research, 3*(7), 749-752.

Sikorski, M. (2016, 25nd of May). *Fan Engagement is the Future of Sports Sponsorship.* Huggity. Retrieved from: https://huggity.com/fan-engagement-future-of-sports-sponsorship/

Singh, H. (2006). The Importance of Customer Satisfaction in Relation to Customer Loyalty and Retention. *UCTI Working Paper.* Asia Pacific University College of Technology & Innovation.

Slater, S. F., & Narver, J. C. (2000). Intelligence Generation and Superior Customer Value. *Journal of the Academy of Marketing Science, 28*(1), 120–127.

Smith, T. J., & McKenna, C. M. (2013). A comparison of logistic regression pseudo R^2 indices. *Multiple Linear Regression Viewpoints, 39*(2), 17-26.

Snyder, M., & Tanke, E. D. (1976). Behavior and attitude: Some people are more consistent than others. *Journal of Personality, 44*(3), 501–517.

Sourav, D. (2021, 15th of June). *Top 10 Most Popular Sports In The World | 2021 Power Ranking.* Sports Show. Retrieved from: https://sportsshow.net/top-10-most-popular-sports-in-the-world/

Stan, V., Caemmerer, B., & Cattan-Jallet, R. (2013). Customer Loyalty Development: The Role Of Switching Costs. *Journal of Applied Business Research (JABR), 29*(5), 1541-1554.

Sun, J. (2005). Assessing Goodness of Fit in Confirmatory Factor Analysis. *Measurement and Evaluation in Counseling and Development, 37*(4), 240–256.

Sürücü, L., & Maslakci, A. (2020). Validity and Reliability in Quantitative Research. *Business & Management Studies: An International Journal, 8*(3), 2694-2726.

Taherdoost, H. (2016). Sampling Methods in Research Methodology; How to Choose a Sampling Technique for Research. *International Journal of Academic Research in Management, 5*(4), 18-27.

Tapp, A. (2004). The loyalty of football fans — We'll support you evermore? *Journal of Database Marketing & Customer Strategy Management, 11*(3), 203–215.

Tapp, A., & Clowes, J. (2002). From "carefree casuals" to "professional wanderers". *European Journal of Marketing, 36*(11/12), 1248–1269.

Taylor, R. (1990). Interpretation of the Correlation Coefficient: A Basic Review. *Journal of Diagnostic Medical Sonography, 6*(1), 35–39.

Umashankar, N., Ward, M. K., & Dahl, D. W. (2017). The Benefit of Becoming Friends: Complaining after Service Failures Leads Customers with Strong Ties to Increase Loyalty. *Journal of Marketing, 81*(6), 79–98.

Van der Eijk, C., & Rose, J. (2015). Risky Business: Factor Analysis of Survey Data – Assessing the Probability of Incorrect Dimensionalisation*. PloS One, 10*(3), 1-31.

Van Voorhis, C. R. W., & Morgan, B. L. (2007). Understanding Power and Rules of Thumb for Determining Sample Sizes. *Tutorials in Quantitative Methods for Psychology, 3*(2), 43–50.

Villanueva, J., Yoo, S., & Hanssens, D. M. (2008). The Impact of Marketing-Induced Versus Word-of-Mouth Customer Acquisition on Customer Equity Growth. *Journal of Marketing Research, 45*(1), 48–59.

Von Wangenheim, F. (2005). Postswitching Negative Word of Mouth. *Journal of Service Research, 8*(1), 67–78.

Von Wangenheim, F., & Bayón, T. (2004). The effect of word of mouth on services switching. *European Journal of Marketing, 38*(9/10), 1173–1185.

Wann, D. L., Tucker, K. B., & Schrader, M. P. (1996). An Exploratory Examination of the Factors Influencing the Origination, Continuation, and Cessation of Identification with Sports Teams. *Perceptual and Motor Skills, 82*(3), 995–1001.

Willits, F. K., Theodori, G. L., Luloff, A. E. (2016). Another look at Likert scales. *Journal of Rural Social Sciences ,31*(3), 126-139.

Wright, K. B. (2005). Researching Internet-Based Populations: Advantages and Disadvantages of Online Survey Research, Online Questionnaire Authoring Software Packages, and Web Survey Services. *Journal of Computer-Mediated Communication, 10*(3).

Wu, S. H., Tsai, C. Y. D., & Hung, C. C. (2012). Toward Team or Player? How Trust, Vicarious Achievement Motive, and Identification Affect Fan Loyalty. *Journal of Sport Management, 26*(2), 177–191.

Xiao, M. (2019). Factors Influencing eSports Viewership: An Approach Based on the Theory of Reasoned Action. *Communication & Sport, 8*(1), 92–122.

Yim, C. K., & Kannan, P. (1999). Consumer Behavioral Loyalty: A Segmentation Model and Analysis. *Journal of Business Research, 44*(2), 75–92.

Young, J. (2021, 5th of February). *Why fan loyalty is more important than ever.* Tappit. Retrieved from: https://tappit.com/why-fan-loyalty-is-more-important-than-ever/

Appendix A: Survey questionnaire

This appendix contains a comprehensive reproduction of the survey instrument in its entirety where all the questions are listed down below in the order in which they were presented to the respondents. Also, the respective measurement and source of each question are noted down. Some headings have been added to the questions to create a clear overview of which questions belong to what subject. Besides, after each question the abbreviation for the observed item can be found. Finally, Table 53 shows in detail the two categories of the 7-point Likert scale, which were used as measurements for this survey.

7-point Likert scale: level of agreement (7PLS-A)	**7-point Likert scale: level of importance (7PLS-I)**
Strongly disagree	Not at all important
Disagree	Low importance
Somewhat disagree	Moderately unimportant
Neither agree nor disagree	Neither important nor unimportant
Somewhat agree	Moderately important
Agree	Very important
Strongly agree	Extremely important

Table 51: the two 7-point Likert scale categories level of agreement and importance and their respective answer options.

Survey questionnaire

Dear participant,
My name is Bram ten Barge and I am currently following the master program Economics and Business at the Erasmus University of Rotterdam. The goal of my thesis is to find out what factors impact the loyalty of football supporters.

Therefore, I designed this survey to collect my data, which will be completely confidential and can be filled in anonymously. If you have any questions, please feel free to contact me at 476420bb@student.eur.nl. The survey will take approximately 5 minutes.

Thanks in advance for your help!

Check-up

☐ I agree that I am at least 18 years old or have received parental approval to fill in this survey

Screen-out question

Question	Measurement	Source
Q1: Do you consider yourself as a fan of a football club?	Yes/No	Own

General questions

Question	Measurement	Source
Q2: What is your gender	Male/Female/Third Gender/Prefer not to say	Own
Q3: How old are you?	Numeric open-ended question, range 1-99.	Own
Q4: Which football club do you support?	Open-ended question	Own
Q5: For how many years have you supported your club?	Numeric open-ended question, range 1-99.	Own
Q6: What is your highest attained level of education?	Less than high school degree/High school degree/Some college degree/ Bachelor's degree/ Master's degree/ PhD	Own

How important were the following reasons to you for initially becoming a fan of your club?

Question	Measurement	Source
Q7: Parental influence	7PLS-I	Griffin (2001)
Q8: Friends who support the same team	7PLS-I	Griffin (2001)
Q9: Geographical location (i.e., supporting a team which is close to your residence)	7PLS-I	Griffin (2001)
Q10: The team's success	7PLS-I	Griffin (2001)

Right now, numerous statements are shown related to the digital activities of your favorite team.

Social media

Question	Measurement	Source
Q11: I expect a strong presence on social media from my team and players (Soc1).	7PLS-A	Own
Q12: I relax when watching/reading social media content related to my team and players (Soc2).	7PLS-A	Own
Q13: Watching/reading club-related social media content makes me feel enthusiastic about my favorite team (Soc3).	7PLS-A	Own
Q14: I like seeing team-centered content (f.e, behind the scenes footage or Q&A's) on social media.	7PLS-A	Own
Q15: I like seeing nostalgic content (f.e, watching old highlights or interviews) on social media.	7PLS-A	Own

Esports

Question	Measurement	Source
Q16: I like the videogame FIFA (Esp1).	7PLS-A	Own
Q17: Watching esports is exciting (Esp2).	7PLS-A	Own
Q18: I would be interested in following esports activities of my club (Esp3).	7PLS-A	Own

On the following two pages, statements are shown related to the restrictions due to covid-19 and its impact on fans like you.

Ghost games

Question	Measurement	Source
Q19: I experience less joy while watching matches right now than prior to the pandemic (Gho1).	7PLS-A	Own
Q20: Ghost games were boring to watch (Gho2).	7PLS-A	Own
Q21: Ghost games gave me a general feeling of dissatisfaction (Gho3).	7PLS-A	Own

Fan interaction during the pandemic

Question	Measurement	Source
Q22: I have missed watching football matches together with friend and/or family (Int1).	7PLS-A	Own
Q23: Digital interaction with fellow fans could not directly replace the experience of physically meeting with fans (Int2).	7PLS-A	Own
Q24: I feel less connected with the other fans than prior to the pandemic (Int3).	7PLS-A	Own

The final statements are shown down below, thanks again for filling them in!

Fan identification

Question	Measurement	Source
Q25: When someone praises my favorite team, it feels like a compliment (Ide1).	7PLS-A	Gladden and Funk (2001)
Q26: During a conversation about my team, I often say 'We' instead of 'They' (f.e, 'next year, we will win the league!') (Ide2).	7PLS-A	Gladden and Funk (2001)
Q27: When someone criticizes my favorite team, I experience the criticism as a personal insult (Ide3).	7PLS-A	Own
Q28: Friends and family know that I am a committed supporter of my favorite team (Ide4).	7PLS-A	Gladden and Funk (2001)

Confirmation of attention question

Question	Measurement	Source
Q29: Choose the option strongly disagree to show that you still pay attention.	7PLS-A	Own

Attitudinal loyalty

Question	Measurement	Source
Q30: I perceive myself as a real fan with a high degree of commitment (Att1).	7PLS-A	Bauer et al. (2008)
Q31: My commitment to my favorite club will not change when the team underperforms massively (Att2).	7PLS-A	Bauer et al. (2008)
Q32: The opinions of friends and family on my favorite team will not have any effect on my commitment to them (Att3).	7PLS-A	Bauer et al. (2008)
Q33: I see myself supporting my favorite club for the rest of my life (Att4).	7PLS-A	Bauer et al. (2008)

Behavioral loyalty

Question	Measurement	Source
Q34: I will always defend my club in public even if it causes social disapproval (Beh1).	7PLS-A	Bauer et al. (2008)
Q35: I regularly watch the games of my favorite club on TV (Beh2).	7PLS-A	Bauer et al. (2008)
Q36: I actively read news about my team and their players on social media and/or other platforms (Beh3).	7PLS-A	Bauer et al. (2008)
Q37: I have purchased a lot of club-related merchandise (Beh4).	7PLS-A	Bauer et al. (2008)
Q38: For the upcoming season, I will be more engaged with my club than last year (Beh5).	7PLS-A	Own

Appendix B: Distinguishing under, just and over-identified models

This appendix will provide the corresponsive calculations to support the three-dimension categorization theory for CFA of Brown and Moore (2012). The main conclusion derived from this theory is the necessity of at least three items belonging to a latent variable to get valid results from CFA.

The following formula will be used to calculate the degrees of freedom (df):

$df = dp - p$, with

Data points (dp): $(n\ x\ (n+1))/2$, where n is the number of observed items.

Parameters (p): number of parameters in the model

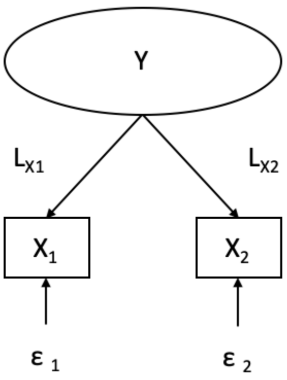

Figure 10: an example of an under-identified model

Figure 10 shows an under-identified model with two observed items (X_1 and X_2), four parameters to estimate (L_{X1}, L_{x2}, ε_1, ε_2) and one latent variable (Y). Hence, the following calculation with $p = 4, n = 2$ and $dp = 2\ x\frac{2+1}{2} = 3$:

$$df = dp - p = 3 - 4 = -1$$

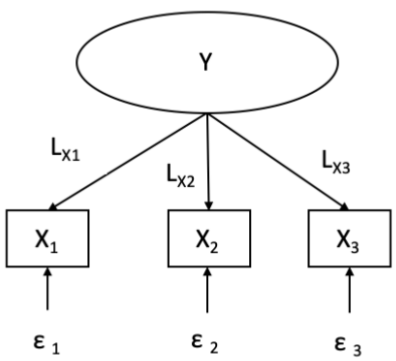

Figure 11: an example of a just-identified model

Figure 11 shows a just-identified model with three observed items (X_1, X_2 and X_3), six parameters to estimate (L_{X1}, L_{x2}, L_{X3}, ε_1, ε_2, ε_3) and one latent variable (Y). Hence, the following calculation with $p = 6, n = 3$ and $dp = 3\ x \frac{3+1}{2} = 6$:

$$df = dp - p = 6 - 6 = \ 0$$

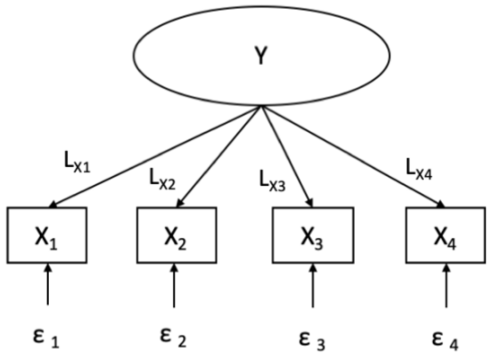

Figure 11: an example of an over-identified model

Finally, figure 12 shows an over-identified model with four observed items (X_1, X_2, X_3, and X_4), eight parameters to estimate (L_{X1}, L_{x2}, L_{X3}, L_{X4}, ε_1, ε_2, ε_3, ε_4) and one latent variable (Y). Hence, the following calculation with $p = 8, n = 4$ and $dp = 4\ x \frac{4+1}{2} = 10$:

$$df = dp - p = 10 - 8 = \ 2$$

Appendix C: the division of the football clubs into the three levels of club magnitude

This appendix will show the division of all the 54 mention clubs in the survey into one of the following three categories of club magnitude: small club, medium club and large club. Besides the division, additional descriptive statistics are provided in the form of frequency and percent.

Furthermore, the division criterion was based on the relative magnitude in their country, instead of an absolute measurement. This division criterion was chosen, because it resulted into a more representative division with the category large clubs being the largest of the three. Otherwise, the large club category would be underrepresented and the other two categories overrepresented, because most of the respondents supported a Dutch club (N=338, 84.5%, see Table 54, 55 and 56).

For example, Feyenoord and AZ Alkmaar are top clubs in the Dutch league, but are not large clubs compared to football club giants such as FC Barcelona, Real Madrid or Paris Saint-Germain. So, for this research Feyenoord and AZ Alkmaar were placed into the category large clubs. For the same reasoning, clubs like FC Twente, Vitesse and FC Utrecht were put in the category medium club instead of small club.

First of all, 54 respondents were placed in the category small club. Table 54 shows that three Dutch clubs formed the top three with regards to their presence in this category: NEC Nijmegen (N=16, 29.6%) was the most frequent club filled in by respondents, which falls under the category small club, followed by FC Den Bosch (N=9, 16.7%) from the 2nd division of the Netherlands and the third most frequent club was Willem II (N=8, 14.8%). Also, the Dutch amateur clubs TEC Tiel, VV Hedel and VV Heerewaarden were included in the sample. Needless to say, they were all placed in the category small club.

Club name	Frequency	Percent
Ado Den Haag	5	9.3
Essendon Bombers	1	1.9
FC Den Bosch	9	16.7
FC Volendam	1	1.9
NEC Nijmegen	16	29.6
RKC Waalwijk	1	1.9
Rotherham United FC	1	1.9
Southend United	1	1.9
Sparta Rotterdam	1	1.9
Sydney Swans	1	1.9
TEC Tiel	4	7.4
VfB Stuttgart	1	1.5
VV Hedel	2	3.7
VV Heerewaarden	1	1.9
Watford FC	1	1.9
Willem II	8	14.8
Total	54	100.0

Table 52: frequency table of the category level small club from the variable club magnitude

Secondly, also 54 respondents were placed in the category medium club. Not only the number of respondents matches with the previous category, but also the origin of the top three clubs with the strongest presence in this category. This time, FC Twente is the most frequent club in this category (N=15, 27.8%), followed by FC Utrecht (N=12, 22.2&) and Vitesse (N=8, 14.8%) (see Table 55).

Club name	Frequency	Percent
Arsenal	10	18.5
FC Groningen	6	11.1
FC Twente	15	27.8
FC Utrecht	12	22.2
Newcastle United	1	1.9
Schalke 04	1	1.9
Tottenham Hotspur	1	1.9
Vitesse	8	14.8
Total	54	100.0

Table 53: frequency table of the category level medium club from the variable club magnitude

Thirdly and finally, Table 56 shows descriptive statistics of the remaining 292 respondents who were placed in the category large club. Not surprisingly, the traditional top three of the Dutch 1st division is equal to the three most frequently mentioned clubs of this entire research. Namely, AFC Ajax was the most frequent club (N=94, 32.2%), followed by Feyenoord (N=81, 27.7%) and PSV Eindhoven (N=59, 20.2%).

Club name	Frequency	Percent
AC Milan	1	0.3
AFC Ajax	94	32.2
Ahly-Egypt	1	0.3
AZ Alkmaar	15	5.1
Bayern Munich	1	0.3
Chelsea FC	8	2.7
FC Barcelona	10	3.4
FC Basel	1	0.3
Fenerbahce	1	0.3
Feyenoord	81	27.7
Galatasaray	1	0.3
Juventus	1	0.3
Liverpool FC	4	1.4
Manchester United	4	1.4
Olympiakos	1	0.3
Paris Saint-Germain	3	1.0
PSV Eindhoven	59	20.2
Real Madrid	4	1.4
São Paulo	1	0.3
SL Benfica	1	0.3
Total	292	100.0

Table 54: frequency table of the category level large club from the variable club magnitude

Appendix D: the four assumptions of OLR

Just like other regression types, it is only appropriate to use OLR if the data set does not violate one of the assumptions. Violating one of its assumptions and still using OLR can results in invalid results and thus conclusions, which is, needless to say, undesired. Therefore, the four OLR assumptions of Harrell (2015) will be assessed for both model equations to prevent invalid results.

The first assumption is that the dependent variable should be measured with an ordinal response scale. Both model equations pass this assumption. Even though the DV's attitudinal and behavioral loyalty are latent variables, they are constructed by observed items which were measured by a 7-point Likert scale.

The second assumption is that one or more of the IV's is either continuous, ordinal or categorical. Also, both model equations pass this assumption with ease. They both have a continuous variable in the form of age, an ordinal variable in the form of fan identification and a categorical variable in the form of education.

Thirdly, there must be no multicollinearity. Harrell (2015) states that multicollinearity occurs when two are more variables are highly correlated to each other, which results in invalid results. For example, multicollinearity can occur when all the categories of a variable are put in the regression. That is why for both model equations, for categorical variables, one category was left out of the regression in order to prevent multicollinearity and pass this assumption.

The fourth and also most fundamental assumption is the proportional odds assumption. The assumption of proportional odds states that each IV has an identical effect on each level of the DV (Harrell, 2015). That is why in the output, there is only one coefficient for every variable, but a unique intercept for every level of the DV.

The proportional odds assumption is tested with the parallel lines test, of which the *p*-value must be higher than .05 in order to not reject the null hypothesis (Erkan and Yildiz, 2014). The null hypothesis in this test states that the slope coefficients are the same across the response categories, which is in alignment with the definition of the proportional odds.

The parallel lines test compares the -2 Log Likelihood values of the estimated model with one set of coefficients for all categories (null hypothesis model) to a model with a separate set of coefficients for each category (general model). If the *p*-value of the test is >.05, the model with one set of coefficients is the better fit and the proportional odds assumption is passed.

Model	**-2 Log Likelihood**	**Chi-Square**	**Degrees of freedom**	***p*-value**
Null hypothesis	1645.145			
General	1527.722	117.423	136	.873

Table 55: parallel lines test for model equation I

Table 57 shows a *p*-value of .873 for the parallel lines test for model equation I, which means that the null hypothesis is not rejected and thus the proportional odds assumption not validated.

Model	**-2 Log Likelihood**	**Chi-Square**	**Degrees of freedom**	***p*-value**
Null hypothesis	1858.534			
General	1353.940	504.595	625	1.000

Table 56: parallel lines test for model equation II

As for model equation II, Table 58 shows a *p*-value of 1.000 for the parallel lines test, which also means that the null hypothesis is not rejected and thus the proportional odds assumption is not violated for both model equations.

Appendix E: SPSS output of the histograms of age and relationship length

Appendix E shows the frequency histograms of the variables age and relationship length. Clearly seen and visualized by the distribution line, the data is not symmetrical and thus normally distributed, but skewed to the left for both variables.