Does content really matter?

How the content of Instagram posts affects consumers' engagement to a company and its credibility.

Bachelor Thesis

Economics and Business Economics programme

Erasmus University Rotterdam

Erasmus School of Economics

Student:	Loïs Schipper
Student number:	443675
Date:	13 July 2020
Supervisor:	S. Lim
Second assessor:	M.J.L. van Hasselt

Abstract

As social media marketing gets more common, it is important to find out to what users respond. This study researches the influence of the content of Instagram posts on perceived credibility and consumer engagement. By using a cross-sectional survey with nine different combinations of post content categories and real-life Instagram data, the dependent variables can be researched. This study finds that the two variables are correlated but that perceived credibility does not have a significant influence on consumer engagement. Also, content category does not significantly influence perceived credibility. Consumer engagement, however, is significantly dependent on the content category and the content creator. This study finally shows that there are significant differences between the restaurant and clothing brand industry, when estimating perceived credibility and consumer engagement.

Contents

List of Figures

List of Tables

1 Introduction

Right now, the fourth big industrial revolution is taking place. The age of computers, the internet and electronics (Schwab, 2017). Simultaneously to the rise of these products, social media emerged as one of the biggest developments in this era. In our Western culture, it is almost impossible to imagine the world without these platforms that constantly update you on the current developments in the world. The current top three social media platforms in the Netherlands are WhatsApp, Facebook and Instagram, with 9.3, 7.1 and 3.4 million users, respectively (de Best, 2020b; Ortiz-Ospina, n.d.).

The rise of these social media platforms also gives companies access to everyday life of consumers. By advertising on these platforms, consumers are able to get in contact with companies and brands more easily. By following the consumers through so-called cookies, advertisements can be tailored to consumers based on their preferences (Ghosh et al., 2015). The consumers then get to see the advertisements intertwined with their social media platform: through pop-ups, advertisement blocks or via their browser. This process makes targeting the ideal consumers more effective, as consumers tend to spend a fair amount online (de Best, 2019). Using social media advertising makes it thus easier to reach the targeted consumers. Companies are therefore eager to add social media usage to their promotional mix (Chu and Kim, 2011).

Yet social media advertising brings companies more: where normal advertising used to be seen as an one-way interaction (company to consumer), advertising on social media can be seen as a two-way interaction (Luten, 2008). Consumers can react and reach out to the company or business, which makes it easier for the company to get feedback and interaction. Besides seeing advertising as a two-way interaction between company and consumer, social media makes it relatively easy for consumers to interact with other consumers as well. This constant interacting between consumers mostly results in online word of mouth (WOM), also known as electronic word of mouth, or eWOM. This makes it possible to give recommendations to others or to express your opinion on certain brands or products (Mangold and Faulds, 2009).

In addition to being able to reach consumers through advertisements, companies use social media platforms to keep their customer engaged and up-to-date with the company. By creating social media pages for their businesses, companies can share posts, products and information with their clientele. These social media posts, and in particular shared photographs, can contain and display multiple objects. These can be divided into a few different categories. The posts mostly display (1) the companies' product; (2) staff or models, interior design or company culture; (3) actual users or; (4) a combination of these. The categories can also contain so-called 'reposts'. This is when a company re-shares content created by one of their real-time customers or users. The original post is a form of user-generated content (UGC) and is created and shared on social media independently of the business. Published content that is created by the company itself or has been created on the company's order, is called brand-generated content (BGC) (Kaplan and Haenlein, 2010; Lipsman, Mudd, Rich and Bruich, 2012).

On the other end of the line are the consumers. They are able to follow the business' page and its content. Doing so, the consumer explicitly likes the brand and adds the business to its social media circle. This will result in seeing a more constant stream of information from the business on their social media. This content can then be liked, commented on and shared[1]. For consumers to perform one of these actions, is to engage with the company. Knowing this, consumer engagement can thus be seen as the number of likes, comments or shares. This consumer engagement tends to be dependent on the context and is therefore dependent on the industry or sort of brand or business (Dessart, Veloutsou and Morgan-Thomas, 2016). Another interesting thing is the perceived credibility of the various posts on Instagram, as this affects how consumers see a company.

Summarizing these statements, it is known that Instagram-post content is divided into three main categories (product, user and company), which can occur simultaneously. Consumer engagement is dependent on the context of the business, thus industry. Perceived credibility is dependent on the source and contents of an Instagram posts. Combining these facts, the question rises if consumer engagement and perceived credibility are then dependent on the used content category in the post and if this differs by industry or kind of business. Also, it is questionable if one influences the other. Therefore, this study aims to investigate the effect of used contents on perceived credibility and the engagement of consumers and whether or not this effect differs in various industries. Thus, the main research question formulated for this study is as follows:

[1] Consumers that have not explicitly liked a company by following their page can still like, comment or share the published content.

What is the impact of the used content category in an Instagram post on the consumer engagement and perceived credibility of customers/followers and does this differ between the Dutch hospitality-sector and the Dutch clothing-industry?

This thesis is of significant relevance since it researches statements that have received little attention. There has been research on the effect of UGC on consumer engagement (Chu and Kim, 2011) and research on credibility on Instagram (Rebelo, 2017), as well as consumer engagement through social media (Lipsman, Mudd, Rich and Bruich, 2012). However, the combination of these subjects has not been researched yet. This study thus sheds a light on this combination and helps us understand how consumers act on social media. The focus in this study lies on the Instagram-platform. This platform has been growing consistently and is the third most-used platform in the Netherlands (de Best, 2020b). It is also a platform used by multiple generations and can thus be analyzed for all age groups (de Best, 2020a)[2]. Lastly, consumer engagement can be measured by other users and is not only visible to the ones that performed the action, which occurs for media like WhatsApp. Even though this study only focuses on one social media platform, this study is relevant for it looks at consumer behaviour of multiple generations in the Dutch society.

This study also tells us if there are differences between industries. Furthermore, the societal relevance can be found on the side of businesses as social media marketing is very relevant for various business-owners and it is being added to promotional mixes all around the world (Vinerean, 2017). By understanding the consumer behaviour in these situations and being able to comprehend which content tends to work best in either industry, this study will help business-owners and brands creating the most efficient content for their Instagram-account.

In the next section, the theoretical framework of this study will be discussed. As well as the definition of consumer engagement used in this study. Figure 1 shows a model of the working of consumer engagement that summarizes the hypotheses that can be found in section 2. That section also explains the different forms of content researched. Section 3 will elaborate on the methodology and models for this study and will give an

[2]Whereas other social media platforms are used more by certain generations. LinkedIn, for instance, is mostly used by people over the age of 25, as the platform is mostly directed to business people.

overview of the chosen businesses. It also explains the data collection. Next, the data will be described shortly. Section 4 consists of the analyses of the data and the found results. After, conclusions follow accompanied by discussion and limitations. This thesis will then be concluded by managerial implications and recommendations for future research.

2 Theoretical framework

During the last decade, Social Media Marketing (SMM) has been of more influence than ever before and influences the way consumers gather information about brands and businesses. With social media, companies and consumers can participate in discussion about the companies' products as well as empowering consumers to share (positive) feelings about the brand with others (Vinerean, 2017). With the rise of eWOM, marketing on social media has been of bigger importance than ever, as eWOM has an important influence on brand image and attitude (Chu and Kim, 2011). The importance of social media on marketing, consumer engagement and perceived credibility will be researched in this study. To ensure the clearness of the explanation and findings, this section will give a theoretical framework and a proposed model with hypotheses to support the main research question.

2.1 Content categories

Social media posts are created in one of two ways. First, content designed by consumers and users themselves. We call this user-generated content or UGC. UGC is a collection of all the content consumers, or end-users, are able to upload, share and create on any social media-platform around the globe. This includes, but is not limited to, reviews, pictures and written text. For content to be called UGC, there are three requirements. These include: (1) publishing the content on a publicly accessible platform; (2) showing a certain amount of creativity and; (3) created outside of professional routines and thereby independent of the shown company (Kaplan and Haenlein, 2010). UGC on the Instagram accounts of brands or companies will most often be a post that the user posted on their personal account and is then shared by the brand. This is thus called a *repost* and can often be detected by the use of tagging the original user in the picture or mentioning the account of the user in the comments or description.

On the other hand, there is the content that has been created by the company itself, or has been ordered to be created by the brand. This form of content is called brand-generated content (BGC) and is used to show a brand in a certain way. The main goal of BGC is to advertise and generate profit[3] (Vinerean, 2017).

[3]Except for non-profit organisations, which mostly have as aim to create awareness or receive dona-

Looking at Instagram posts, there is a clear pattern in the content shown in these post. There are three distinct categories that will be addressed in this study. First and foremost, one of these is the product itself. This category is seen in most business Instagram accounts and shows the product the business sells. For restaurants this will be a dish or a drink. Next, there is the category that contains actual users. These posts sometimes are user-generated and thus re-posted on the Instagram account of the business, as previously discussed. Actual users are also able to be used in brand-generated content. These posts often show a user in the establishment or wearing the brand's products. Lastly, there is the category that includes staff, interior design and the company culture. In case of restaurants, these posts, for instance, contain staff or a table setting. For clothing companies, this category includes owners, staff, models used for photo shoots and culture related posts. These categories can occur simultaneously and can be shown together in one post.

2.2 Measuring consumer engagement

Over the course of the last decades various studies tried to explain consumer engagement. Remarkably, there are many varieties of perspectives that are used and addressed in these studies. A commonly used theory is the (combination of) utilization of the components of attitudes towards brands: dimensions of behaviour, cognition and affect (Beckler, 1984; Dessart et al., 2016; Leckie, Nyadzayo and Johnson, 2016). Combining the different studies and research, Leckie *et al.* (2016) established that these three dimensions are represented as follows: affect engagement is expressed in enthusiasm and enjoyment, behavioural engagement through sharing information and learning. Cognitive engagement is found in capturing of attention of the consumer. Combining these findings, it can be said that consumer engagement results from feeling an association between the consumer and the brand. The result ends in an interaction which extents beyond purchase.

Another reasoning involving consumer engagement can be found in the model designed by Barger, Peltier and Schultz in 2016. Consumer engagement is subject to five different groups of factors, which may or may not lead to a form of engagement. Consumer engagement is driven by different factors concerning the brand, the product,

tions.

the consumer itself, the content and the social medium used. These factors tell us that consumers engage with brand content for entertainment, for incentives and promotions, for information acquisition and for bonding (Barger et al., 2016).

Besides the division in the components of attitudes, the perspectives also differ in terminology when addressing consumer engagement. This makes it difficult to see if researchers are discussing different concepts, or if they are discussing the same concept but use different terms. However, in most studies the following are included in the definition and description of consumer engagement: (1) the consumer as subject; (2) the brand, its community or brand-related content as focus[4]; (3) an interaction between subject and focus that extents beyond purchase; (4) context specific engagement and; (5) different degrees of consumer engagement (van Doorn, Lemon, Mittal, Nass, Pick, Pirner and Verhoef, 2010; Hollebeek, 2011a, 2011b; Patterson, Yu and de Ruyter, 2006). An Instagram-user will be considered as the first one: the consumer as subject. The second one, the brand, will thus be the Instagram-account of the brand or restaurant itself and its posted content.

The more difficult ones to establish are items three through five. On Instagram, consumers can respond to a brand page with the following features: following the page, liking a post, commenting on a post, responding on a story by quick response or text, privately sending a direct message, responding on a long-form video by quick response or text, going to a website or webshop through a swipe-up in stories, going to a webshop featured in a post. Apart from the first two, these actions are not visible for other consumers/Instagram-users (Instagram, n.d.). Nevertheless, these actions all act as the interaction between the consumer and the brand. The context specific engagement is visible since different uploads (a post, story, long-form video, or direct message) evoke different reactions. The different degrees of consumer engagement can also be seen in the different reactions to the different contents. All the different types of uploads have multiple reaction possibilities. These various reactions can be seen as different degrees of consumer engagement, as certain actions require more effort than others, and therefore can be seen as a higher degree of consumer engagement (Hollebeek, Glynn and Brodie, 2014). There are thus multiple actions that can measure consumer engagement.

Lee, Hosanagar and Nair (2018) used multiple of these actions to measure consumer

[4] The subject focuses the engagement on this point of focus.

engagement on Facebook. They used the number of likes, comments and shares of a single post. Similar measurements were used by Dessart, Veloutsou and Morgan-Thomas (2015) in interviews with online community members, discussing several social media-platforms. When looking at Instagram, a difference is that there are actions that are not visible for the rest of the users, such as number of shares.

Therefore, this study measures consumer engagement only by looking at the number of likes and comments from Instagram-users, who represent consumers, on posts of the Instagram-account of the brand or restaurant. The latter will represent the brand or restaurant and its posted content will be considered as the focus. The different degrees are visible in liking the post and comments on the post, where the latter demands more effort and thus is seen as a higher degree of engagement. The number of followers can also be seen as a level of consumer engagement. However, this action can also be done once, where liking and commenting can occur every post. As the number of followers result in a rather small data set, this variable will not be looked at as a measurement of consumer engagement in this study.

Another aspect that has to be looked upon is the activity of the brand on its Instagram account. As uploading a post counts as an activity, brand activity on Instagram can be expressed in the number of posts uploaded per day. Assuming that consumer engagement is an interaction that extents beyond purchase, we assume that this interaction can only occur once an Instagram post is uploaded. This action can occur more often when the level of activity of the brand is higher, meaning that users can like and comment more, if there are more posts uploaded. A higher level of activity also generates a more constant stream of visible information for the user. A user will then notice an Instagram-account more easily and might be more eager to engage. Consumer engagement may thus be higher once the brand is more active on social media (Ashley and Tuten, 2015). This gives us the first hypothesis:

> **H1**: The level of social media activity influences consumer engagement positively.

2.3 Perceived credibility

How consumers see Instagram content is partly determined by how they perceive the credibility of the shown content. Is the content real and credible or is it just show?

Studies in different areas show that credibility is dependent of certain factors. First of all, it seems that perceived credibility is higher when the content is user-generated, instead of brand-generated (Li and Suh, 2015). Simultaneously, perceived credibility appears to be dependent on attractiveness, trustworthiness and competence (Edwards, Stoll, Faculak and Karman, 2015; Rebelo, 2017). These characteristics also contribute to making a post more relatable to users (Edwards et al., 2015). As stated before: a strong association between brand and user, results in an interaction between the two parties. As this extends beyond purchase, this is a form of consumer engagement . Having a better relation, can be caused by being more relatable. A more relatable post thus makes consumer engagement easier (Leckie et al., 2016). As the fourth point in the definition of consumer engagement states, there is context specific engagement. This means that consumer engagement depends on the content in an Instagram post (van Doorn et al., 2010). For perceived credibility, this appears to be the same (Djafarova and Rushworth, 2017; Klassen et al., 2018). The drivers for consumer engagement and perceived credibility seem to overlap in this. Therefore it can be assumed that perceived credibility also increases when a post is more relatable. Lastly, the level of trustworthiness affects perceived credibility and consumer engagement positively. If the post seems more trustworthy, consumer engagement seems to increase (Djafarova and Rushworth, 2017).

2.4 Credibility and engagement

Given the aforementioned theory, it seems that consumer engagement and perceived credibility are more or less driven by the same predictors. They both rely on trustworthiness and relation to the post in their determination. Rebelo (2017) argues that a consumer purchase intention is dependent on the level of perceived crebility. As consumer engagement is a relation that extends purchase, we may assume that the latter does not happen without the first (van Doorn et al., 2010). This leads to the assumption that perceived credibility and consumer engagement are correlated. This is also supported by the fact that they are likely to depend on the used content in the post, as this study researches. The two appear to be driven by the same variables, but not to the same extend, and are thus most likely correlated to one another.

H2: Perceived credibility and consumer engagement are correlated.

Besides the content, consumer engagement seems to be dependent on more. It appears to differ in different contexts and settings, as mentioned before (Dessart et al., 2016). Given the fact that brands are divided in different industries, we can assume that consumer engagement differs depending on the industry. As this study also tests contents' influence on perceived credibility, the influence of the industry on this dependent variable will be tested. This theoretical framework results in seven hypotheses, which will help answering the main research question. Figure 1 summarizes the working of consumer engagement in the environment of Instagram. The hypotheses all question or test a relation between two variables in the model of Figure 1. Therefore, the number of the hypothesis is stated next to the corresponding relation in the figure.

H3: Perceived credibility depends on the content of an Instagram post.

H4: Perceived credibility differs per industry.

H5: Consumer engagement depends on the creator of the content.

H6: Consumer engagement depends on the used category of content.

H7: Consumer engagement differs per industry.

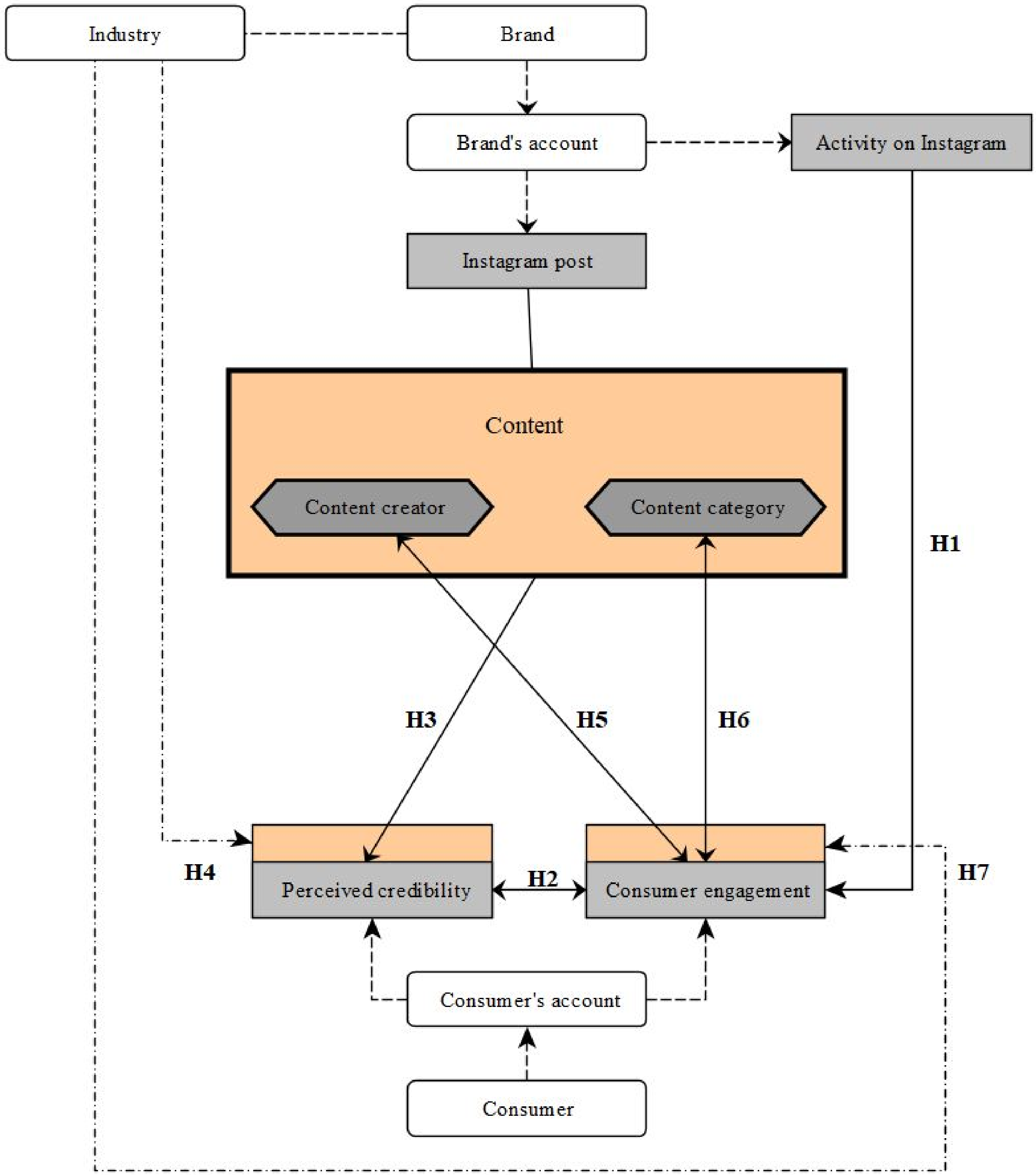


Figure 1: Proposed model of working of a consumer's engagement with a brand through Instagram.

3 Research methodology

This study uses two different data-sets for testing the hypotheses. The first consists of collected data from a survey and is used to measure perceived credibility and self-reported consumer engagement. Secondly, open data from Instagram accounts measures actual consumer engagement. This is addressed as 'open data' in the rest of this study. The quantitative research in this study helps giving a better insight in the total picture, since perceived credibility and consumer engagement might be correlated to one another (Rebelo, 2017).

3.1 Industries and chosen businesses

For this study, two unrelated industries were chosen: Dutch restaurants and Dutch clothing brands. The chosen businesses are all handpicked based on the requirement that they all showed a certain amount of posts with content from all the different categories[5], which required screening the individual Instagram pages to see if all categories are represented. The chosen businesses are stated below, along with their location, segment[6] or type of brand and Instagram account name.

1. Restaurants
   (a) **1nul8**, Rotterdam, middle class restaurant/bar, @1nul8;
   (b) **Supermercado**, Rotterdam, lower class restaurant/bar, @hola.supermercado;
   (c) **Fred**, Rotterdam, 2-Michelin star restaurant, @restaurantfred;
   (d) **Loetje**, multiple locations in the Netherlands, middle class restaurant, @loetje_restaurant;
   (e) **Coffeelicious**, family-business with multiple locations in the Netherlands, lunchroom, @coffeeliciousnederland

[5]These are selected using a combination of multiple websites (van Velzen, 2019; MonStyle, n.d.; Tripadvisor, n.d.)

[6]Segments are based on their relative price range and their segmentation given on Tripadvisor. Segments are given to ensure that the whole industry, not just one segment, is represented. For the clothing brands, there are only two segments. This is chosen, as lower segments may represent fast-fashion, which are more internationally located.

2. Clothing brands[7]

   (a) **Kings of Indigo(KOI)**, luxury brand with focus on sustainability, webshop and more than 200 selling point worldwide (mostly in Europe) @kingsofindigo;

   (b) **Goosecraft**, luxury brand, webshop and more than 250 selling points worldwide (mostly in Europe), @goosecraft;

   (c) **Guts & Gusto**, middle class brand, webshop and five boutiques in the Netherlands, @gutsgusto;

   (d) **Most Wanted**, middle class brand, webshop and two boutiques in the Netherlands, @mostwantednl;

   (e) **My Jewellery**, middle class brand, webshop and thirteen (in-store) boutiques in the Netherlands, @myjewellery

3.2 Open data collection

The collection of open data has been done by gathering data from the Instagram accounts of the selected brands listed above. This data contains approximately hundred Instagram posts per account and its characteristics. These characteristics can be separated into two groups. First, characteristics of the Instagram account included the number of followers and posts. Secondly, Instagram posts were reviewed individually. The information gathered about the Instagram posts includes the main- and subcategory of the posts, whether or not there are people in the picture, if the post is a repost and if the post consists of more than one picture. Most importantly, the number of likes and comments are noted, as these are seen as the dependent variable consumer engagement. Also, the days needed to upload the collected posts have been noted, as this is used to conduct the activity of the Instagram account.

For this data collection, there has been an important decision to only use social media content from the period before the closing of the hospitality branch caused by the COVID-19 virus. As of the closing, restaurants did not have access to the usage of reposts from customers or creating content with customers in their restaurant, therefore the distribution of content is not representative for a normal situation. Also, there may

[7] All are from Dutch origin

be a possibility that users were more active during the quarantine as they were having time on their hands. The latest date used for Instagram-posts is therefore 14 March 2020.

3.3 Experiment data

Researching the perceived credibility and reported consumer engagement of certain Instagram posts has been done by a cross section survey. This survey has been distributed with the use of several social media. These include Facebook, Instagram, LinkedIn and WhatsApp. Furthermore, the survey has been distributed by other users through sharing the survey on their own social media. The survey was open for respondents from 11 May up until 24 May and eventually yielded 296 responses.

The survey has been conducted as follows. Respondents answered questions about Instagram posts that show one of the determined content categories for both of the researched industries. By randomly assigning the categories over the respondents, it was possible to derive the correlation between the categories and the dependent variables. There are three main categories for the used content in Instagram-posts, as established in Section 2.1. There are also two researched industries. The respondents all got a post from each industry. Moreover, the posts are distributed randomly over the respondents. As there are three possible options in both the categories, this gives nine different combinations of the posts. Respondents thus were divided into either one of the nine groups, each with a unique combination of the posts. In addition, this data was used to test if there is a difference between the industries.

For the testing of perceived credibility and self-reported consumer engagement, the Instagram accounts of Supermercado and Kings of Indigo were used. These accounts show a good representation of posts in all categories. Also, Kings of Indigo is used as it is one of the two brands with clothing for both men and women, which makes it more relatable to all respondents as opposed to a women-only brand. The distribution of the Instagram post amongst the respondents can be found in Table 1. The corresponding Instagram posts are listed in Appendix B.

The complete survey (Appendix A) consists of thirteen questions where questions 5 up until 7 are equal to 8 up until 10, but correspond to the two different posts. The survey has been subjected to a pre-test with 15 respondents from different age groups, gender and educational levels, to see if the survey was comprehensive. The results of this

Table 1: Shown Instagram posts per group

	1	2	3	4	5	6	7	8	9
Kings of Indigo									
Content category	product	product	product	user	user	user	company	company	company
Content creator	brand	brand	brand	user	user	user	brand	brand	brand
Supermercado									
Content category	product	user	company	product	user	company	product	user	company
Content creator	brand	user	brand	brand	user	brand	brand	user	brand

pretest were positive which resulted in the publication of the survey (Reips, 2002). Since this study addresses the impact on Dutch restaurants and brands, and therefore would be questioning Dutch people, the survey was conducted in Dutch.

The survey holds questions that ask respondents about certain control variables. The first questions to do that are questions 1 up until 4. Questions 1 and 2 asked respondents about their Instagram usage, as people familiar with Instagram might review post differently. Also, respondents were asked about their familiarity with the brands used in the survey in questions 3 and 4, as they might have already formed an opinion on perceived credibility. The last two questions of the survey also asked respondents about control variables: their age and gender. These control variables are used as Instagram is one of the platforms that are used by multiple generations (de Best, 2020b).

The remaining questions asked respondents about the variables researched in this study. Question 5 and 8 asked respondents what their reported degree of liking or commenting on the respective post would be. This has been answered on a 5-point Likert scale and has been used to determine the consumer engagement.

Edwards *et al.* (2015) argue that positive feelings toward a picture impacts both perceived credibility and consumer engagement. This is tested using questions 6 and 9, where positive feelings towards the shown Instagram posts have been recorded by three sub-questions. The results of these questions were added into a scale with scores between 3 and 15.

The same method has been applied to the overall feeling of perceived credibility. The four sub-questions of questions 7 and 10 tell the perceived credibility of the respondents by asking the respondents to choose if they think the more credible characteristic fits the post better than the unreliable one. The four comparisons all contribute to an overall feeling of perceived credibility[8]. The answers were added into a scale, scoring between 5 and 20 points, where a higher score represents a higher perceived credibility.

The latent variable model sum scoring is used in both situations, as opposed to taking an average of the answers given. The reasoning is as follows. Both the variables for 'positive feeling' and 'perceived credibility' are dependent on three and four characteristics respectively. These characteristics cause their respective variable. Nonetheless, there is no saying in the weighting of these variables or about other characteristics that also may have an influence. The used characteristics are chosen, as they are repeatedly mentioned in the theory and thus seem to indicate having a positive feeling and perceived credibility well. As a good factor needs at least three variables, it is not possible to determine their weights using factor analysis (Tabachnick and Fidell, 2007). Therefore, I chose to weigh all the characteristics evenly. As the ratios are equal, it does not matter if the variables would be created using sum scoring or using averages.

The results of question 11 show if answers given to questions 7 and 10 are in balance. Question 11 measured whether the respondents feel if either one of the posts is more thrustworthy. This question has been scored the same way as questions 7 and 10 and makes it possible to see if people actually value one over the other.

3.4 Testing hypotheses

The results of the survey were used in a multiple regression with the results of questions 7 and 11 as measurements of perceived credibility and thus the dependent variable (**H3**). Researching consumer engagement (**H5**, **H6** and **H7**) has been done in two ways. The results of the survey were used to test reported consumer engagement of the respondents (**H6**). This has been done with an ordered probit model where questions 5 and 8 were be used as dependent variables.

An ordered probit is one of the possible models that can be used when the dependent variable is based on a ordinal variable with multiple levels. It is similar to a regular probit model that uses the cumulative distribution of the standard normal distribution to estimate the probability of performing an action. The regular probit model, however, is used to re-scale any number to a probability, as a number between 0 and 1. The outcomes

[8]The used comparisons are all derived from theory in previously mentioned studies (Djafarova and Rushworth, 2017 Edwards et al., 2015; Rebelo, 2017).

of questions 5 and 8, on the other hand, are based on five different levels.

The ordered probit transforms the observed ordinal variable into the latent continuous variable, even tough there are negative or large values. This variable, depicted as y^*, shows a linear combination of predictors $\mathbf{x}$ and an error-term[9]. Equation 1 shows the general mathematical equation of this relation. Variable j takes on values 0 through 5, as these are the levels in the questions 5 and 8, according to Equation 2.

$$y_i^* = \mathbf{x}_i\beta + e_i, \quad e_i \sim N(0,1), \forall i = 1, ..., N \tag{1}$$

$$y_i = j \Leftrightarrow \mu_{j-1} < y_i^* \leq \mu_j \tag{2}$$

Using the observed ordinal variables, the ordered probit model provides us latent continuous variables. It gives us the probability that person i selects alternative j, according to Equation 3.

$$P[y_i = j] = \phi(\mu_j - \mathbf{x}_i\beta) - \phi(\mu_{j-1} - \mathbf{x}_i\beta) \tag{3}$$

With j taking on a value from 0 through 5 (Jackman, 2000).

This equation has been transformed into a table with level-cuts[10]. These cutoff points tell us on which of the levels the outcome of the model falls. Level one corresponds to the first level used in the analyses. In this case, the first level is 'strongly disagree'. The second corresponds with 'disagree' and so on. The outcome of the ordered probit tells us what the respondent's action will be, on average. If a person scores at the second level, they will, on average, disagree with the statement.

Liking and commenting are asked about in two separate sub questions. To see if the action of liking and commenting can be used as one combined variable, a factor analysis has been conducted[11].

In all models control variables for age and gender were added, as well as Instagram usage and having an Instagram account. These variables are added as different generations spend a different amount of time on social media and on different social media platforms. There are also differences between gender. Therefore it seems best to include these variables and see if there is a difference between older and younger generations,

[9]The predictors are given in Table 7.

[10]This has been automatically done by performing the statistical test in STATA

[11]Addressed in section 2.2

who use different social media platforms and spend a different amount of time on these platforms (de Best, 2019, 2020a, 2020b).

Lastly, this data is used for testing **H2**. This has been done by computing correlations and looking at significant values.

Consumer engagement has also been tested using the open data. This data has been used in a multiple regression analysis where consumer engagement is seen as the dependent variable. The independent variables contain the used category content, content creator, number of followers, industry and other relevant independent variables. Also, other characteristics of the post are included in the regression, such as: whether there are people in the picture and if there is only one picture used or multiple. These characteristics may contribute to the consumer engagement. This analysis also includes the activity of an Instagram account (expressed in the number of post per day). This was done to see if activity influences consumer engagement (**H1**). The previously mentioned analyses of the survey data and the reported consumer engagement were also used to see if there is a difference between reported consumer engagement and actual recorded consumer engagement.

Finally, the dependent variables for consumer engagement and perceived credibility were used in t-tests to see if there are differences between industries (**H4** and **H7**). To fully answer the main research question of this study, I combined the results of the hypothesis tests and conclude in what way the content impacts consumer engagement and perceived credibility.

4 Data description

As the respondents got divided into nine different groups, it is crucial that each group is represented equally. Table 2 shows the demographic variables of the survey data. The χ^2-goodness of fit test (Appendix C) in Table 10 shows that the observed distribution of respondents in the groups is equal to the expected distribution.

The last column of Table 2 shows the descriptive data extracted from the survey data for the complete sample. Table 11a shows the χ^2-test for the observed values for age. It is clear that the sample is not representative for the population based on age (de Best, 2020a). There has been an over-representation for respondents aged 17 and under and aged between 18 and 24. The age-group 25 till 34 seems to be underrepresented. Table 11b shows that there is an over-representation for women. Tables 12 and 13 show the χ^2-goodness of fit tests. These show that age and gender of respondents are distributed equally given the respondents in the sample.

The sample not being representative for the population holds consequences for the conclusions in this study. This is addressed in chapter 6.1. Nonetheless, the respondents are distributed among the groups according to the overall sample. This shows that there are no significant differences in age and gender between groups

Table 3 shows the basic data of the Instagram accounts. It immediately stands out that the number of days needed to post the collected posts are lower for the clothing brands. Also the clothing brands seem to have a bigger number of followers and total number of posts than the restaurants. An explanation could be that clothing brands can

Table 2: Descriptive data per survey group

	1	2	3	4	5	6	7	8	9	Total
Age in years (%)										
17 and under	5 (15,6%)	7 (21,2%)	8 (22,9%)	3 (9,7%)	6 (17,6%)	7 (20,2%)	4 (12,1%)	5 (16,7%)	4 (12,1%)	49 (16,6%)
18-24	11 (34,3%)	7 (21,2%)	12 (34,3%)	9 (29,0%)	16 (47,1%)	10 (28,6%)	14 (42,4%)	9 (30,3%)	9 (27,3%)	97 (32,8%)
25-34	6 (18,8%)	6 (18,2%)	5 (14,3%)	10 (32,3%)	5 (14,7%)	5 (14,3%)	6 (18,2%)	7 (23,3%)	7 (21,2%)	57(19,3%)
35-44	3 (9,4%)	4 (12,1%)	5 (14,3%)	3 (9,7%)	4 (11,8%)	3 (8,6%)	3 (9,1%)	4 (13,3%)	6 (18,2%)	35 (11,8%)
45-54	5 (15,6%)	8 (24,2%)	4 (11,4%)	5 (16,1%)	2 (5,9%)	8 (22,9%)	5 (15,2%)	4 (13,3%)	6 (18,2%)	47 (15,9%)
55-64	2 (6,3%)	1 (3,0%)	1 (2,9%)	1 (3,2%)	0 (0,0%)	2 (5,7%)	1 (3,0%)	1 (3,3%)	1 (3,0%)	10 (3,4%)
65+	0 (0,0%)	0 (0,0%)	0 (0,0%)	0 (0,0%)	1 (2,9%)	0 (0,0%)	0 (0,0%)	0 (0,0%)	0 (0,0%)	1 (0,3%)
Gender (%)										
female	19 (59,4%)	21 (63,6%)	27 (77,1%)	22 (71,0%)	24 (70,6%)	26 (74,3%)	20 (60,6%)	25 (83,3%)	28 (84,8%)	212 (71,6%)
male	13 (40,6%)	12 (36,4%)	8 (22,9%)	7 (22,6%)	10 (29,4%)	9 (25,7%)	13 (39,4%)	5 (16,7%)	5 (15,2%)	82 (27,7%)
undefined	0 (0,0%)	0 (0,0%)	0 (0,0%)	2 (6,5%)	0 (0,0%)	0 (0,0%)	0 (0,0%)	0 (0,0%)	0 (0,0%)	2 (0,7%)
Sample size	32	33	35	31	34	35	33	30	33	296

have consumers throughout the country, as they are able to deliver their products through postal services. Restaurants cannot and are thus more locally focused. Appendix D shows further description of the data of the Instagram accounts.

Figure 8 (Appendix D) shows the distribution of all combinations of main- and subcategories of the collected Instagram data. It immediately stands out that for both industries there are far more posts with product as main category, as opposed to the other categories. Table 14 (Appendix D) shows the distribution of content creators of the collected Instagram data. Both industries post more brand generated content than user generated content. However, in the restaurant industry the amount of UGC is double the amount of the clothing industry.

Table 3: Account data of Instagram accounts

Brand	Total no. of posts*	No. of followers	No. of following	No. of collected posts**	No. of days between collected posts
Restaurants					
1nul8	302	6445	8	101 (1)	299
Supermercado	367	5024	857	101 (1)	244
Fred	172	11900	5945	103 (3)	492
Loetje	353	18900	229	107 (7)	241
Coffeelicious	1245	13000	7474	106 (6)	130
Clothing brands					
Kings of Indigo	1458	34500	690	101 (1)	130
Goosecraft	1587	15700	1054	101 (1)	191
Guts & Gusto	10001	213000	1330	115 (14)	38
Most Wanted	6582	305000	1543	100 (0)	25
My Jewellery	5394	390000	210	101 (0)	45

*Notes: * This number represents the number of posts up until 14 March, as this is the cut-off point for the collection of the data.*

*** The number of collected posts includes a number of posted videos. These posts do not show a number of likes on Instagram and are therefore not used in analyses. The number of posts that can not be used are stated between brackets.*

5 Analysis and results

This section holds the analyses and results for this study. The section is divided into four subsections. These all consist of data-analysis and results, where the first considers the analyses considering perceived credibility and the second looks into analyses regarding consumer engagement. The next subsection looks at the differences between the industries for both perceived credibility and consumer engagement. Lastly, the correlation between the variables is addressed.

5.1 Perceived credibility

As discussed in section 3.4, perceived credibility has been tested using ordinary least square regressions. These regressions can be found in Table 4. The regressions have been conducted for one industry at the time. As explained in section 3.4, perceived credibility is being measured by the sum of the four comparisons regarding trustworthiness and credibility. This summation resulted in a scale from 4 to 20. For both of the industries the maximum scored is 20. For the clothing industry the minimum score is 7, whereas the restaurant industry scored as low as 5. Both the regressions had a sample size of 296. The data used for these tests seems to be a bit left skewed, but not shockingly. For both tests, there are no issues with multicollinearity. There is, however, heteroscedasticity in the data of the clothing brand. This has been solved, using robust standard errors. There is no heteroscedasticity in the data of the restaurant.

The independent variables used, are all the variables conducted in the survey and will now be explained. The variables 'post category' are dummy variables and are scored 1 if the respondent saw the corresponding content category. The category 'user' acts as a reference category as this is the only category that is UGC. The other two categories are brand generated.

The variable 'positive feeling' is also a scale, but scored from 3 to 15, as previously explained[12]. Interesting to see is that this variable is highly significant for the determination of perceived credibility in both industries. The more positive the feeling is towards the post, the higher perceived credibility is rated. For restaurants, this effect is twice as big as for the clothing industry. This can be related to the fact that people tend to value

[12] Conducted from questions 6 and 9 in survey.

perceived credibility based on attractiveness, competence and, most importantly, feeling (Edwards et al., 2015).

The next variable measures the familiarity with the brands. This categorical variable consists of three levels, where higher means more familiar. The lowest level means that the respondent did not know the clothing brand or restaurant at all. The second level means that the respondent is familiar with the business, but has never visited the restaurant or has never worn any clothes of the brand. The highest level means then that the respondent knows the business and has visited the restaurant or has worn the clothing brand before.

The following variables express some information about the respondents, such as age and gender. Age is measured as a continuous variable and gender as a dummy variable. 'Having an Instagram account' is also measured as a dummy variable. Instagram usage is, just as familiarity, measured as a categorical variable. For both of the industries, it is visible that the more the person uses Instagram, the higher the rating of perceived credibility is. People that use Instagram more often, score perceived credibility significantly higher. The levels of the variable correspond with the possible answers in the survey (Appendix A, question 2), where 'Never' corresponds with the lowest level and 'Multiple times a day' corresponds with the highest.

Lastly, a F-test has been performed to analyze the content categories. This test thus measures if the effect of categories differ significantly from each other. The P-value for both industries is lower than 0,10. This shows that the influences of content categories differ significantly on a 90% confidence interval, for both industries.

The only significant variables are, for both industries, the positive feeling towards the post, the Instagram usage of the respondent and the constant value. The first two result in an increase of perceived credibility for both industries. The constant value for the clothing industry seems higher than that for the restaurant industry, meaning that perceived credibility is rated higher for restaurants than clothing brands. On average, however not significantly, following the brand, having an Instagram-account, being female and being older results in a decrease of perceived credibility in both industries. Even tough these effects are not significant, the coefficients of the independent variables also state the following: In the clothing industry the product category scores a higher perceived credibility than the user category, which scores higher compared to the company category.

Table 4: Perceived credibility of Instagram posts *(OLS)* through survey data

Dependent variable: *Perceived credibility (per industry)*	Clothing		Restaurants	
Post category: Company	-0,223	(0,347)	0,203	(0,343)
Post category: Product	0,565	(0,352)	-0,520	(0,372)
Positive feeling	0,310*	(0,091)	0,655*	(0,742)
Familiarity with brand	0,551	(0,398)	-0,067	(0,202)
Following brand	-1,413	(2,546)	-0,274	(0,485)
Having an Instagram account	-0,414	(0,716)	-0,385	(0,690)
Instagram usage	0,272*	(0,118)	0,241*	(0,120)
Age	-0,010	(0,013)	-0,001	(0,013)
Female	-0,205	(0,362)	-0,452	0,309)
Constant	10,011*	(1,226)	7,648*	(0,948)
R^2	0,134		0,274	
F-test *Post categories the same***	2,53		2,39	
P-value	0,081		0,093	
N	294		294	

*Notes: * Significant on a 95% confidence interval*
*** F-test as follows: post category: company = post category: product = 0, with post category: user as the reference category.*

For the restaurant industry, this seems to be the other way around. Lastly, familiarity with the brand, on average, increases for the clothing industry, but decreases credibility for restaurants.

The aforementioned analysis addressed **H3**. As is visible from these statements, there is no significant influence of content category on perceived credibility. The F-tests tell us that there is a significant[13] difference between the influences of the content categories, for both industries. However, it is easily visible that the content category is not of significant influence to the perceived credibility of respondents. Therefore **H3 is rejected**.

[13] On a 90% confidence interval

5.2 Consumer engagement

Before the tests concerning consumer engagement could be executed, the way of measuring had to be addressed. As stated in section 2.2, consumer engagement will be expressed in 'likes' and 'comments'. Whether or not these have to be used in a joined variable has to be tested beforehand. With the use of factor analyses, it is possible to see if the used variables are driven by the same factor. This method shows interrelations among the used variables. If a set of variables are driven by the same factor, they have an interrelation and appear to measure the same thing to some extend. The option to merge these into one is then available. For the dependent variables that are researched in this study, it is important to see if these can be used in one joined variable. Table 15 Appendix E shows the factor analyses of the survey data, for each industry separately, whereas Table 16 shows the analysis for the open data. These different factor analyses indicated that merging the two variables was not desirable. In two of the three cases, the variables ended up being supported by the same factor. These cases, however, resulted in undesirable Cronbach's alphas. The other analysis did not put the two variables together, which made it even more clear not to merge the variables[14]. Therefore, all analyses will be conducted twice: once with likes as dependent variable and once with commenting as such. In the rest of this study consumer engagement will be used as a term for both likes and comments.

5.2.1 Actual consumer engagement through open data

This section focuses on measuring consumer engagement. First, the analyses with the open data is addressed. For both liking and commenting, the data showed heteroscedasticity, which is resolved by performing the regression using robust standard errors. Both data sets also show a small right skew and do not have multicollinearity issues.

In the analyses, the consumer engagement is not measured per industry, but a dummy variable was added that represents the industry. The ordinary least square regressions are shown in Table 6. This table shows the regressions for all data where there is no giveaway promoted in the Instagram posts. The disregarded posts mostly consist of giveaway promotion where users are able to enter the giveaway by liking, but mostly by

[14] See Appendix E for the complete analyses.

commenting on the post. Users are able to win anything from gift-cards of the restaurant to a year of free clothing. These giveaways have a big influence on the number of likes, but even more so on the number of comments, as people are eager to enter such promotions. Moreover, the impact was of such size that it did not feel correct to use these posts in the analyses. Table 5 shows the descriptive data of the Instagram accounts without the giveaways, where Table 17 (Appendix F) shows the data with the giveaways.

The first three independent variables of the OLS are dummy variables that measure the main content category of the post. Instagram post with an undefinable or very divergent category where categorized as 'other', which is the reference category. For consumer engagement through liking, this is of significant influence. On average, consumer engagement will be highest for posts with an actual user, followed by posts with a product and posts with the company. A post with a main category different from the 'other' category gains at least two hundred likes on average. For commenting the 'other' category

Table 5: Descriptive data of Instagram accounts

		Likes			Comments		
Brand	N	Mean	SD	Median	Mean	SD	Median
Restaurants							
1nul8	100	110,05	49,58	103,50	4,26	5,15	3,00
Supermercado	100	94,37	32,78	87,50	4,49	4,88	3,00
Fred	100	613,18	250,77	598,50	14,80	14,24	12,00
Loetje	96	336,79	122,56	324,00	33,52	45,16	19,00
Coffeelicious	96	184,39	85,89	164,50	7,76	17,53	4,00
Clothing brands							
Kings of Indigo	99	274,65	150,15	229,00	7,12	7,28	5,00
Goosecraft	100	85,65	31,42	82,00	1,12	1,54	1,00
Guts & Gusto	98	1577,30	2054,65	1129,50	14,98	24,66	7,50
Most Wanted	100	3781,99	950,47	3683,00	7,63	6,66	5,50
My Jewellery	100	9455,31	2497,72	8826,00	39,12	40,92	25,50
Total	989	1664,159	3029,06	283	13,426	25,246	6
Interval [min, max]		[24, 20161]			[0, 336]		

get most reactions, followed by posts with products and posts with users. These last variables are not significant.

The subcategory variables are also three dummy variables, but these are optional. There is no reference category, as it is possible that there is no subcategory at all. Even though it is rare, it is also possible that there are two subcategories in the picture. An example of a combination of main and subcategory is as follows: a user with a coffee in her hand would get the categorization of user as main category and product as subcategory. The subcategories company and product are of significant influence for both measures of consumer engagement, with product receiving a higher consumer engagement then company.

The other independent variables consist of several dummy variables and two continuous variables. The first two are dummy's and measure if the brand uses more than one picture per post and if there are people in the picture. Both increase the consumer engagement, yet only people are of significant value. The usage of people in Instagram posts increases the number of likes by 71. Repost is a dummy variable that measures if the post is a repost, as explained in Section 2.1. If the post is a repost, this also means that it is user generated, instead of brand generated. This variable seems to be of importance for users as it is of significant value and the number of likes increases by 195,638. UGC seems to increase consumer engagement through likes. For commenting however, it seems to decrease the consumer engagement.

Next is the variable 'number of followers', which is of significance for the consumer engagement, both liking and commenting. Number of followers increased consumer engagement, which is not very surprising, as more followers mean that more people see your post. The dummy-variable restaurant tells the industry and will be equal to 1 for restaurant and equal to 0 for clothing brands. This variable is also significant for both the measurements of consumer engagement. Restaurants, on average, get 137,752 more likes and 5,467 more comments compared to clothing brands.

The last variable used is the number of post per day and is conducted as follows: the number of posts collected divided by the number of days between the first and last posts. This means that number of posts per day is the second-last column of Table 3 divided by the last column.

Table 6: Influence of Instagram content on consumer engagement (*OLS*) through Instagram data

Dependent variable:	Liking		Commenting	
Main category				
Company	200,047*	(100,721)	-13,606	(8,682)
Product	262,262*	(90,191)	-12,054	(8,501)
User	311,140*	(156,786)	-12,183	(8,484)
Sub category				
Company	-274,786*	(131,867)	-6,170*	(2,248)
Product	208,229**	(116,703)	-5,118**	(2,620)
User	-77,616	(221,845)	2,160	(10,495)
Other independent variables				
Multiple pictures in post	180,279	(209,330)	6,776	(4,590)
People in picture	71,785*	(118,930)	1,633	(2,438)
Repost	195,638*	(61,477)	-3,108*	(1,263)
Number of followers	0,033*	(0,001)	0,0001*	(0,00002)
Restaurant	136,752*	(62,294)	5,467*	(1,500)
Number of posts per day	-1731,992*	(97,790)	-9,998*	(1,523)
Constant	176,269	(111,519)	23,132*	(8,248)
R^2	0,870		0,181	
F-test *Main categories the same****	2,95		0,86	
P-value	0,032		0,460	
F-test *Sub categories the same*****	4,89		0,36	
P-value	0,008		0,695	
N	989		989	

*Notes: * Significant on a 95% confidence interval*

*** Significant on a 90% confidence interval*

***F-test as follows: main category-company = main category-product = main category-user = 0, with main category: other as the reference category

****F-test as follows: sub category-company = sub category-product = sub category-user = 0, as sub categories are optional and do not occur in every post.

The average number of posts per day is 1,274 with a minimum of 0,209 and a maximum of 4 posts per day. It significantly decreases consumer engagement by 1731,992 likes and 9,998 comments per extra post per day.

Lastly, the R^2s for liking and commenting are 0,870 and 0,181, respectively. For both models, the same F-tests as in Table 4 have been conducted. The main categories and subcategories are tested separately. The models both have a sample size of 989. This table addresses hypotheses 1, 5 and 6. H5 and H6 will be addressed in Section 5.3. As for H1: The variable that measures activity is of great significant value in estimating consumer engagement. Although theory pointed out that a higher activity most likely results in a higher level of consumer engagement (Ashley and Tuten, 2015), the coefficients for number of posts per day is negative. Consequently, I **reject H1**.

5.2.2 Self-reported consumer engagement

Below, the analyses of the survey data will be discussed. The model looks similar to Table 4 as the variables are identical, except perceived credibility is added as independent variable. The biggest difference however is the used model. Consumer engagement is measured through a Likert-scale. This creates a categorical dependent variable. Therefore consumer engagement is estimated as an ordered probit-model and can be found in Table 7. There have been four separate regressions, which can be found in the second up until fifth column. As mentioned before, the variables are identical to those used in Table 4 and will therefore not be explained that extensively again. The same F-test has been conducted for the four regressions as well.

The interpretation of the regression is slightly different however. When answering the questions about consumer engagement, respondents were able to pick either of the following five levels: strongly disagree, disagree, neutral, agree and strongly agree. The outcome of the ordered probit-models are corresponding to these levels. The level cutoff points, as stated in Table 7, correspond to the five levels. A score lower than 0,894 leads to believe that that person would strongly disagree with liking the shown post of the clothing brand. A score between 2,637 and 3,942 forecasts that this person would probably agree with liking the post of the clothing brand and so on.

Table 7: Consumer engagement through survey data *(ordered probit-models)*

Dependent variable:	Clothing		Restaurants	
Consumer engagement	Liking	Commenting	Liking	Commenting
Perceived credibility	0,031	0,046	-0,018	0,027
	(0,029)	(0,030)	(0,029)	(0,031)
Post category: Company	0,271**	0,205	0,225	0,201
	(0,158)	(0,170)	(0,156)	(0,158)
Post category: Product	-0,175	0,155	0,373*	0,203
	(0,152)	(0,165)	(0,174)	(0,167)
Positive feeling	0,215*	0,179*	0,382*	0,247*
	(0,040)	(0,040)	(0,043)	(0,044)
Familiarity with brand	-0,194	-0,385*	-0,282*	-0,293*
	(0,154)	(0,175)	(0,086)	(0,093)
Following brand	-0,047	1,719*	0,342	0,398
	(0,833)	(0,342)	(0,391)	(0,374)
Having an Instagram account	0,076	0,044	0,483	0,169
	(0,353)	(0,343)	(0,304)	(0,302)
Instagram usage	-0,017	-0,105**	0,088*	0,011
	(0,053)	(0,054)	(0,042)	(0,051)
Age	-0,008	0,015*	0,002	0,019*
	(0,005)	(0,006)	(0,005)	(0,005)
Female	-0,020	-0,148	0,059	-0,205
	(0,156)	(0,158)	(0,162)	(0,151)
Pseudo R^2	0,076	0,082	0,164	0,090
F-test *Post categories the same***	8,10	1,61	4,77	2,01
P-value	0,017	0,4479	0,092	0,366
N	294	294	294	294
Level 1 cutoff point	0,893	1,726	2,965	2,880
Level 2 cutoff point	1,865	2,763	3,860	3,899
Level 3 cutoff point	2,636	3,953	4,534	4,867
Level 4 cutoff point	3,942	4,678	6,210	6,097

*Notes: * Significant on a 95% confidence interval ** Significant on a 90% confidence interval*

****F-test identical to F-test in Table 4*

The influence of the category-variables is only significant in two cases. The category 'company' only has a significant influence on likes for clothing brands. For the category 'company' this applies for the number of likes for restaurants. Positive feelings toward the post is a significant influence in all four models, where perceived credibility is insignificant for all. Perceived credibility does not drive consumer engagement in this model significantly.

Further, familiarity has significant negative influence in all models, except liking for clothing brands. Familiarity thus decreases consumer engagement. This is quite unexpected, as more famous companies have more followers and more consumer engagement. The reason for this negative influence is not clear. One of the explanations could be that the familiarity for the used clothing brand was very low and the familiarity with the restaurant was very high in comparison. Possibly, this inequality holds an underlying variable or explanation. Also, the used brands differ in the operating scale, which could be of influence. Lastly, familiarity does not always influence consumer behaviour in an expected way, which also might have happened here (Wanick, Stallwood, Ranchhod and Willis, 2018).

Instagram usage also has some significant impact, however this differs for all the four models. It decreases consumer engagement for clothing brands, but is only significant when measuring the number of comments. For restaurants the opposite occurs. An increase in Instagram usage increases consumer engagement, however only significantly for the number of likes. The age of respondents significantly increases the number of comments in both industries. Younger generations thus do not comment as much as older generations. The control variable for gender decreases consumer engagement, but not significantly.

5.2.3 Differences and results of consumer engagement

Tables 6 and 7 both address the hypotheses **H5** and **H6**. The F-tests show that, for liking, the categories significantly differ in their influence[15]. Yet, the coefficients of Table 7 do not show any significant value for the content categories, except for company in one of the four models. Table 6, however, shows something very different. When looking

[15]In Table 7: For the clothing brand with a 95% confidence interval. For the restaurant with a 90% confidence interval. In Table 6 both main and subcategory on a 95% confidence interval.

at Instagram data, the main content categories are of significant influence in estimating consumer engagement through likes. The subcategories are partially significant. Also, the variable for repost is significant for both measures of consumer engagement. As repost represents the creator of content and this influences consumer engagement significantly, **H5 is appected**. The coefficients in Table 6 shows that re-posting a picture, increases likes by 195,638. UGC increases consumer engagement through likes. For comments, this is the other way around with a decrease of 3,108 comments when using a repost.

As for the influence of content category, I see two different things. Table 7 looks at self reported consumer engagement. It shows that, apart from company in the first model and product in the second-last model, content category does not seem to be of significant influence. The categories do differ significantly from each other. When looking at the OLS in Table 6, content categories are of great significant influence when estimating actual consumer engagement. Eight of the twelve variables for content categories are highly significant in estimating consumer engagement through likes. The F-tests also state that the categories significantly differ. For comments, these results are not visible. Nonetheless, **H6 is accepted** as liking occurs more often, has a higher mean and a higher range than commenting[16].

5.3 Industry differences

The next analyses look at the differences between industries. It has been previously discussed that perceived credibility and consumer engagement might depend on industries. This can be seen in Table 6, as the variable 'Restaurant' has a significant influence on the number of likes and the number of comments. Being a restaurant account ensures an increase in consumer engagement. Also, the coefficients differ in the two industries when looking at self-reported consumer engagement in Table 7. Moreover, the coefficients tend to have different signs for some of the same variables. Furthermore, the significant variables are not the same for both the industries. There are also differences in coefficients in Table 4. However, here the same variables are significant for either industry. Looking deeper at the industry differences, Table 8 shows the analyses regarding this matter. Using survey data, perceived credibility and self reported consumer engagement are measured using the same scales for each industry. Therefore, they are able to be compared using

[16] See Table 5 and Table 17

paired t-test. Table 8a addresses **H4**. The hypothesis stated that perceived credibility is dependent on industry. The alternative hypotheses in this test is thus that the mean of the difference between the industries is not equal to 0. Given the stated P-value, the null hypothesis is rejected and states that the means differ significantly. Therefore, **H4 is appected**.

Tables 8b and 8c show the paired t-tests of consumer engagement in both industries. Similar to the results for **H4**, the H_a's of these t-test are that the means of differences

Table 8: Differences between industries (*Paired t-tests*) through survey data

(a) Perceived credibility

	N	Mean	Std. Error	Std. Deviation
Clothing industry	296	14,267	0,148	2,554
Restaurant industry	296	14,909	0,160	2,761
Difference	296	-0,642	0,166	2,852

H_0: Mean of differences = 0 t-test = -3,872

H_a: Degrees of freedom = 295

mean <0	mean ≠ 0	mean >0
P = 0,000	P = 0,000	P = 1,000

(b) Consumer engagement expressed in likes

	N	Mean	Std. Error	Std. Deviation
Clothing industry	296	2,588	0,065	1,126
Restaurant industry	296	3,311	0,068	1,181
Difference	296	-0,723	0,074	1,266

H_0: Mean of differences = 0 t-test = -9,820

H_a: Degrees of freedom = 295

mean <0	mean ≠ 0	mean >0
P = 0,000	P = 0,000	P = 1,000

(c) Consumer engagement expressed in comments

	N	Mean	Std. Error	Std. Deviation
Clothing industry	296	1,679	0,048	0,820
Restaurant industry	296	2,030	0,057	0,976
Difference	296	-0,351	0,058	1,001

H_0: Mean of differences = 0 t-test = -6,040

H_a: Degrees of freedom = 295

mean <0	mean ≠ 0	mean >0
P = 0,000	P = 0,000	P = 1,000

Notes: The mean of differences is in all analyses measured as the variable for the clothing industry- the variable for the restaurant industry.

are not equal to 0 and are thus significantly different. The stated P-values for these tests both inform that the H_0 are rejected at a 95% confidence interval. The industries differ significantly in their means for both likes and comments. Hence **H7 is accepted**.

The mean of differences is measured by the variable for the clothing industry minus that for the restaurant industry. As the P-values for the H_a's that measure that the mean of differences is smaller than 0, all lets me reject the H_0, it can be said that the mean of the restaurant industry is higher than that of the clothing industry. Consumer engagement and perceived industry are thus higher for the restaurant industry.

5.4 Correlation

Table 9 shows the correlations of perceived credibility and consumer engagement. The used data for the calculations are that of the survey. Therefore the perceived credibility is measured in the aforementioned scale. Consumer engagement is measured by the self reported degree of liking or commenting of the respondents[17]. Table 9 shows that for both industries perceived credibility and liking are significantly correlated, as well as liking and commenting. Perceived credibility and commenting are also significantly correlated, but only for the restaurant industry.

Table 9: Correlations between perceived credibility and consumer engagement through survey data

(a) Clothing brands

	Perceived credibility	Liking	Commenting
Perceived credibility	1,0000		
Liking	0,1681*	1,0000	
Commenting	0,0815	0,3923*	1,0000

(b) Restaurants

	Perceived credibility	Liking	Commenting
Perceived credibility	1,0000		
Liking	0,2333*	1,0000	
Commenting	0,2037*	0,4596*	1,0000

*Notes: * Significant on a 95% confidence interval.*

[17]Survey question 5 for clothing brands, and question 8 for the restaurants.

Nevertheless, Table 7 also shows correlations between the two variables. Here the variables are accompanied by coexisting independent variables. The coefficient of perceived credibility does not show a significant value. Table 9 thus may show a significant correlation when looking at just the two variables. Table 7, however, shows that the variables are not significantly correlated when there are other coexisting variables to behold. Given these statements, **H2 is accepted** because of the significant pairwise correlation. However, the hypothesis is accepted with *caution* that the two variables are not significantly correlated in a way that one can predict the other.

6 Conclusion, limitations and recommendations

6.1 Conclusion and discussion

Before addressing the results of this study, there are limitations that have to be mentioned. One of the bigger limitations is the access to all data considering consumer engagement, such as shares and direct messages. These measurements could have made the analyses more complete. It could also show whether the same user liked posts more often or only once. This way a 'fan base' of reoccurring customers could be recognized. This could help brands to better recognize the target customers. Another limitation is the reaction to the COVID-19 virus. As restaurants got closed and people spend more time at home and online, it is possible that the more recent pictures showed more consumer engagement than before. This could also have been a result of gaining followers throughout time. This is also one of the statistics that are not public and accessible.

When looking at the survey, there are a limitations as well. As discussed in section 4, the sample is not representative for the population. There seems to be an over representation for the younger generation. There also is a big over representation of females over males in this survey. The insignificance of both of these variables can thus be a result of the incorrect distribution. There thus also may be different coefficients in the regression analyses. A recommendation for future research would be to have a wider survey, with more respondents. This should clear the vagueness around the influence of demographics on the researched variables.

Another limitation of the survey were the used posts and in particular the user generated posts. As the posts used for the user category where the only user generates posts, it was not possible to test the influence of UGC on perceived credibility and self reported consumer engagement. Lastly, the unfamiliarity with Kings of Indigo and familiarity with Supermercado was very unbalanced and might have resulted in underlying and unnoticeable differences.

Below, the main findings of this study will be addressed, as well as the findings regarding the main research question.

H1 stated that social media activity positively influences consumer engagement. However, it could not be accepted and is thus rejected. Even tough theory states that activity will likely result in a higher degree of consumer engagement, this study showed otherwise. The influence of activity of an Instagram account is significantly negative, as discovered in Table 6.

Tables 7 and 9 address the second hypotheses of this study, that states that perceived credibility and consumer engagement are correlated. Table 7 showed no correlation when there are other coexisting variables to behold. Nonetheless, given the significant pairwise correlations in both 9a and 9b, **H2 is accepted** with caution.

The influences of perceived credibility have been researched with the use of two different hypotheses. **H3** stated that content of a post influences perceived credibility, where **H4** stated that perceived credibility differs per industry. The first is tested using an ordinary least square regression in Table 4. This regression and the corresponding F-tests revealed that there are significant differences between the content categories. Nevertheless, there are no significant influences on the dependent variable, perceived credibility. Therefore **H3** has been rejected. Given the limitations of the survey, as mentioned above, there could have been a different conclusion when more posts were used in the survey.

The other hypothesis regarding perceived credibility, however, has been accepted. Table 8a showed that the null hypothesis (the difference of the means of the industries are equal to 0) was rejected at a 95% confidence level. Accordingly, there is a significant difference between the two industries. **H4** is accepted.

Hypotheses 5, 6 and 7 all regard consumer engagement. Tables 6 and 7 test these hypotheses. The creator of content cannot be tested with the observed data from the survey. Hypotheses **H5** has thus only been tested using the open data. The hypothesis states that consumer engagement depends on the creator of content. Table 6 shows that the variable for 'repost' is of significant value for both measures of consumer engagement. Reposts represent UGC. Therefore, reposts are user-generated and posts that are not, are brand-generated. Given the significant influence of this variable, it is stated that consumer engagement depends on the creator of the content and, consequently, **H5 is accepted**.

The idea that consumer engagement depends on used category of content is noted in **H6**. This hypothesis has been tested using both survey data and open data. Looking at self-reported consumer engagement in Table 7, it is visible that content category is not of significant value for consumer engagement. However, when looking at actual consumer engagement (Table 6), the opposite occurs. Content categories, both as main as subcategories) are of significant influence in estimating consumer engagement measured by the number of likes. This does not occur for the consumer engagement measured by the number of comments. However, as liking is more common, has a higher mean and a higher range, **H6 is accepted** and states that consumer engagement is indeed dependent on content category.

Lastly, **H7** states that consumer engagement differs between industries. As seen in Table 6, being a restaurant has a significant positive influence in consumer engagement. Moreover, the paired t-test in Tables 8c and 8b shows that there is a significant difference in the means of the industries. Hence, **H7 is accepted**: there is a difference in consumer engagement between the industries.

Summarizing, this study shows that both content category and content creator significantly impact consumer engagement. However, this differs when looking at the measure of consumer engagement. For likes, it is clear to see that a post consisting UGC results in a higher level of consumer engagement. Posts with people in the picture, whether staff or users, also influence the estimated number of likes significantly. The influence of content categories is dependent on the main and subcategory for the post. For the main category, users score the highest level of consumer engagement as it shows followers a more realistic image. The same cannot be said about comments. Subcategories are significant for comments but in different degrees. Reposts also seem to decrease consumer engagement slightly. Self-reported consumer engagement characterises itself by not being significantly influenced by content categories. A more positive feeling towards the post responds in a significant increase of consumer engagement, as supported by Hollebeek *et al.* (2014). The same can be seen in the regressions regarding perceived credibility. Nevertheless, content categories differ significantly from each other.

Lastly, the industry has a significant influence on both consumer engagement and perceived credibility. Restaurants tend to receive a higher degree of consumer engagement compared to clothing brands. This also follows from all of the t-tests in Table 8. Both consumer engagement and perceived credibility thus dependent on industry. Also the two variables are significantly correlated. Nonetheless, perceived credibility does not generate a significant influence when estimating consumer engagement.

6.2 Managerial implications

For companies it is important to increase consumer engagement. It is a way to invest in the relation with their users, in a way that extends beyond purchase. This relation with users makes a new purchase more easy, as consumers keep getting updates from the company. A more constant showing of the brand to the consumers can also be seem as a form of marketing (SMM) and advertising of the company and its products. This makes it more likely that your products will get into the consumers consideration set, which then makes a purchase more likely. A higher level of consumer engagement may thus be followed by purchase actions. Also, a constant interacting with consumers is likely to result in eWOM, which makes it possible to quickly and effectively increase the number of interested consumers. Increasing consumer engagement thus is an effective marketing tool (Chu and Kim, 2011; Vinerean, 2017).

As perceived credibility is not of significant influence, there is no immediate need to try to increase this. Recommendations for increasing consumer engagement will be given now. Familiarity with the brand has a significant influence on self-reported consumer engagement (Table 7) and not all respondents knew the brands or followed them. In reality, people that engage with the company probably already know the company or business and maybe even follow the brand on Instagram. The only significant variable out of the survey data that is of great importance here, is the feeling the Instagram post gives the user. This study found that positive feelings increases both perceived credibility and consumer engagement significantly.

The most important finding of this study is that it showed that the categories of the content are of significant value for the consumer engagement. Table 6 shows the coefficients of this finding. Consumer engagement, through liking, is highly affected by the content categories. Most importantly, the post should always have either of the three main categories (company, product or user). Indistinguishable posts have a much lower level of consumer engagement. Posts that either feature company, product or user gain, on average, between 200 and 311 likes, opposed to a post that does not really fit in one of these categories. The subcategory also has an influence on the number of likes. Having a product as subcategory will score the highest number of likes, but having no subcategory outscores company and user. The highest possible combination of the categories will thus be a picture of a user with a product. For restaurants, this could be a user with a drink or a dish. For clothing brands, an actual user wearing a top of the brand will be sufficient.

Besides the impact of the content category, the creator of the content has a big significant influence as well. A repost equals UGC in this study. Reposts gain 195 likes on average. This shows that UGC appeals more to users than BGC and it increases the consumer engagement. Reposts show how the product or company really is, opposed to some BGC. Posting this UGC results in a sort of eWOM, where users can see how other users viewed the product and brand. If a person is posting about the brand, it must be good and therefore it creates a positive word-of-mouth. This then results in a higher level of consumer engagement.

Also, having people in pictures increases consumer engagement significantly, with almost 72 likes on average. A higher number of followers also results in a significant higher level of consumer engagement. An extra follower gives an increase of 0,033 likes on average. This also immediately shows that not every follower will engage constantly with your pictures. This does not, however, mean that the consumer does not see the post. Seeing the post, but not liking it, contributes to getting your products in the users awareness set.

Another interesting finding is the influence of posting multiple times a day. This study showed that posting more times a day results in significant lesser number of likes and comments. Thus posting too frequent might tire users. However, this problem may be solved by the following. Posting multiple times a day is done because there is more information to be shown. Instead of posting these separately, this could be done using one post with multiple pictures. Posting multiple pictures in one post increases consumer engagement, however not significant. Nonetheless, this could be a solution for brands and businesses that wish to post multiple posts per day, but fear the negative influence of activity.

Next, there are some differences in consumer engagement through liking and commenting. The recommendation above are based on liking, as this occurs more frequently than commenting. Also, commenting seems to be highly affected by other factors. Table 6 shows that all main categories decrease the number of comments. The 'other' category contained posts that, for instance, asked users a question. Users tend to comment more on these posts than on others. Companies that want to increase consumer engagement in comments, should thus post questions and statements that their users want to respond to. This, however, does interfere with the number of likes, which will decrease when using that category.

Lastly, there are differences between industries. Restaurants have a significant higher consumer engagement through liking and commenting, opposed to clothing-brands. Also, self-reported consumer engagement is very differently affected in the industries. Table 7 shows these findings. To summarize the table: For restaurants, it is best to not post brand generated pictures of products. These may spark the positive feeling, but lack a feeling of trustworthiness and score low on perceived credibility. On the contrary it is best for clothing brands to post pictures featuring their product. Users prefer seeing other users over models, as the company category in Table 7 decreases credibility.

Concluding, to increase consumer engagement, this study suggests that companies should consider the following. The number of posts per day should not be high. Multiple pictures per post might help to avert that issue. The posts should include a person, for instance a user. Combined with a product, this post will likely yield the highest number of likes, but this is shown to be different per industry. If this post is also UGC, the number of likes will be highest as reposts increase the number of consumer engagement.

6.3 Recommendations for future research

Future research should try looking at the, for me, inaccessible data and statistics, as this might influence the conclusions. Also there might be differences between different social media platforms. Instagram is one of the more quick look-through platforms, where comments are more frequently used on platforms such as Facebook. This might also be age-dependent, as Facebook is used more by older generations and this study showed that older generations comment more often. It would also be interesting to see if the difference between the industries can be generalized to different countries and cultures. This could be done by looking at international companies that have a separate Instagram account for every country they operate in. Examples are: Levi's, Sony Music and Red-bull. Also looking at more companies at once or at different industries could provide a better insight in the influence of content. Lastly, conducting a survey with more post combinations of content category and creator could result in a more detailed and complete insight of the influences of these variables on consumer engagement.

To amplify the given managerial implications, a bigger research should be conducted. This research should include more Instagram accounts and should include more brands from different segments in the industry. Also, the Instagram accounts used were selected based on representativeness of the content categories. For an extensive research, the Instagram accounts should be selected more randomly. Lastly, future research should also focus on the positive feeling toward posts and their content. This could include more qualitative research to grasp what makes respondents get a positive feeling[18].

[18] This however may be more of interest for behavioural economists.

References

Ashley, C., & Tuten, T. (2015). Creative Strategies in Social Media Marketing: An Exploratory Study of Branded Social Content and Consumer Engagement. *Psychology and Marketing*, *32*(1), 15–27. https://doi.org/https://doi.org/10.1002/mar.20761

Barger, V., Peltier, J. W., & Schultz, D. E. (2016). Social media and consumer engagement: a review and research agenda. *Journal of Research in Interactive Marketing*, *10*(4), 268–287. https://doi.org/https://doi.org/10.1108/JRIM-06-2016-0065

Beckler, S. J. (1984). Empirical validation of affect, behavior, and cognition as distinct components of attitude. *Journal of Personality and Social Psychology*, *47*(6), 1191–1205. https://doi.org/https://doi.org/10.1037/0022-3514.47.6.1191

Chu, S.-C., & Kim, Y. (2011). Determinants of consumer engagement in electronic word-of-mouth (eWOM) in social networking sites. *International Journal of Advertising*, *30*(1), 47–75. https://doi.org/https://doi-org.eur.idm.oclc.org/10.2501/IJA-30-1-047-075

de Best, R. (2019). *Distribution of social media users in the netherlands in 2017, by hours spent per day on social media and age group*. https://www.statista.com/statistics/861834/hours-per-day-on-social-media-in-the-netherlands-by-age-group/

de Best, R. (2020a). *Distribution of total number of users of instagram in the netherlands as of april 2020, by age group and gender*. https://www.statista.com/statistics/946096/instagram-user-distribution-in-the-netherlands-by-age-group-and-gender/

de Best, R. (2020b). *Number of daily active users (dau) of the leading social media platforms in the netherlands from 2016 to 2020, by social network*. https://www.statista.com/statistics/639790/social-media-active-user-penetration-in-the-netherlands-by-social-network/

Dessart, L., Veloutsou, C., & Morgan-Thomas, A. (2015). Consumer engagement in online brand communities: a social media perspective. *Journal of Product & Brand Management*, *24*(1), 28–42. https://doi.org/10.1108/JPBM-06-2014-0635

Dessart, L., Veloutsou, C., & Morgan-Thomas, A. (2016). Capturing consumer engagement: duality, dimensionality and measurement. *Journal of Marketing Management*, *32*(5-6), 399–426. https://doi.org/https://doi-org.eur.idm.oclc.org/10.1080/0267257X.2015.1130738

Djafarova, E., & Rushworth, C. (2017). Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users. *Computers in Human Behavior*, *68*(1), 1–7. https://doi.org/https://doi.org/10.1016/j.chb.2016.11.009

Edwards, C., Stoll, B., Faculak, N., & Karman, S. (2015). Social Presence on LinkedIn: Perceived Credibility and Interpersonal Attractiveness Based on User Profile Picture. *Online Journal of Communication and Media Technologies*, *5*(4), 102–115. https://doi.org/https://doi.org/10.29333/ojcmt/2528

Ghosh, A., Mahdian, M., McAfee, R. P., & Vassilvitskii, S. (2015). To Match or Not to Match: Economics of Cookie Matching in Online Advertising. *ACM Transactions on Economics and Computation*, *3*(2). https://doi.org/https://doi-org.eur.idm.oclc.org/10.1145/2745801

Hollebeek, L. D. (2011a). Demystifying customer brand engagement: Exploring the loyalty nexus. *Journal of Marketing Management*, *27*(7-8), 785–807. https://doi.org/https://doi.org/10.1080/0267257X.2010.500132

Hollebeek, L. D. (2011b). Exploring customer brand engagement: Definition and themes. *Journal of Strategic Marketing*, *19*(7), 555–573. https://doi.org/https://doi.org/10.1080/0965254X.2011.599493

Hollebeek, L. D., Glynn, M. S., & Brodie, R. J. (2014). Consumer Brand Engagement in Social Media: Conceptualization, Scale Development and Validation. *Journal of Interactive Marketing*, *28*(2), 149–165. https://doi.org/https://doi.org/10.10.1016/j.intmar.2013.12.002

Instagram. (n.d.). *Instagram — Features*. https://about.instagram.com/featuresl

Jackman, S. (2000). Models for unordered outcomes political science 150c/350c.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*(1), 59–68. https://doi.org/https://doi.org/10.1016/j.bushor.2009.09.003

Klassen, K. M., Borleis, E. S., Brennan, L., Reid, M., McCaffrey, T. A., & Lim, M. S. (2018). What People "Like": Analysis of Social Media Strategies Used by Food Industry Brands, Lifestyle Brands, and Health Promotion Organizations on Facebook and Instagram. *Journal of Medical Internet Research*, *20*(6). https://doi.org/10.2196/10227

Leckie, C., Nyadzayo, M. W., & Johnson, L. W. (2016). Antecedents of consumer brand engagement and brand loyalty. *Journal of Marketing Management*, *32*(5-6), 558–578. https://doi.org/https://doi.org/10.1080/0267257X.2015.1131735

Li, R., & Suh, A. (2015). Factors Influencing Information credibility on Social Media Platforms: Evidence from Facebook Pages. *Procedia Computer Science*, *75*(1), 314–328. https://doi.org/https://doi.org/10.1016/j.procs.2015.12.146

Lipsman, A., Mudd, G., Rich, M., & Bruich, S. (2012). The Power of "Like" - How brand reach (and influence) fans through social-media marketing. *Journal of Advertising Research*, *52*(1), 40–52. https://doi.org/https://doi.org/110.2501/jar-52-1-040-052

Luten, T. L. (2008). *Advertising 2.0 - social media marketing in a web 2.0 world*. Praeger Publishers.

Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. *Business Horizons*, *52*(4), 357–365. https://doi.org/https://doi.org/10.1016/j.bushor.2009.03.002

MonStyle. (n.d.). *Proud Dutchies!* https://monstyle.nl/nederlandse-merken-waar-we-trots-op-mogen-zijn/

Ortiz-Ospina, E. (n.d.). *The rise of social media*. https://ourworldindata.org/rise-of-social-media

Patterson, P., Yu, T., & de Ruyter, K. (2006). Understanding customer engagement in services. In Y. Ali and M. van Dessel (Eds.) *Advancing theory, maintaining relevance: Proceedings of ANZMAC 2006 conference.*

Rebelo, M. F. (2017). *How influencers' credibility on instagram is perceived by consumers and its impact on purchase intention*. http://hdl.handle.net/10400.14/23360

Reips, U.-D. (2002). Standards for Internet-Based Experimenting. *Experimental Psychology*, *49*(4), 243–256.

Schwab, K. (2017). *The fourth industrial revolution*. Crown Business.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics (5th ed.)* Allyn & Bacon.

Tripadvisor. (n.d.). *Restaurants in Rotterdam*. https://www.tripadvisor.nl/Restaurants-g188632-Rotterdam_South_Holland_Province.html

Tsai, W.-H. S., & Men, L. R. (2013). Motivations and Antecedents of Consumer Engagement With Brand Pages on Social Networking Sites. *Journal of Interactive Advertising*, *13*(2), 76–87. https://doi.org/https://doi-org.eur.idm.oclc.org/10.1080/15252019.2013.826549

van Doorn, J., Lemon, K. N., Mittal, V., Nass, S., Pick, D., Pirner, P., & Verhoef, P. C. (2010). Customer engagement behavior: Theoretical foundations and research directions. *Journal of Service Research*, *13*(3), 253–266. https://doi.org/https://doi.org/10.1177/1094670510375599

van Velzen, J. (2019). *15 toonaangevende kledingmerken van Nederlandse bodem*. https://man-man.nl/15-kledingmerken-uit-nederland/

Vinerean, S. (2017). Importance of Strategic Social Media Marketing. *Expert Journal of Marketing*, *5*(1), 28–35.

Wanick, V., Stallwood, J., Ranchhod, A., & Willis, G. (2018). Can visual familiarity influence attitudes towards brands? An exploratory study of advergame design and cross-cultural consumer behaviour. *Enterainment Computing*, *27*(1), 194–208. https://doi.org/https://doi.org/10.1016/j.entcom.2018.07.002

A Survey questionnaire

The survey consisted of the following questions, all asked in Dutch as I have been only questioning Dutch people. Questions 5 through 10 are all answered using a 5-scale Likert-type scale.

Q1 Do you have an Instagram account?

- Yes
- No

Q2 How often do you use Instagram?

- Multiple times a day
- Once a day
- 4 to 6 times a week
- 2 to 3 times a week
- Once a week
- Once a month
- Never

Q3A Are you familiar with the clothing brand [...]?

- Yes I know the brand and own or wear this brand (or used to).
- Yes I know the brand, but have never owned or worn this brand.
- No

Q3B Are you familiar with the restaurant [...]?

- Yes I know the restaurant and have eaten or drunk there once or multiple times.
- Yes I know the restaurant, but have not been there (yet).
- No

Q4 Do you 'follow' one of the next Instagram-accounts? If so, check the box.

- Kings of Indigo
- Supermercado

Q5 State to what degree you agree with the following statements after seeing the post or the clothing brand.

- I would 'like' this post on Instagram.
- I would 'comment' on this post on Instagram

Q6 State to what degree you agree with the following statements after seeing the post of [clothing brand].
After seeing the post I. . .

- think I will get a good feeling wearing this product.
- have a positive feeling towards the company.
- would recommend this to someone else.

Q7 For each of the following comparisons of characteristics, state which one you feel fits the post best. For example: If the post give you a very positive feeling, the comparison of negative and positive would be:
negative . . . . x positive

- genuine/fake
- thrustworthy/unreliable
- realistic/unrealistic
- professional/unprofessional

Q8 State to what degree you agree with the following statements after seeing the post of [resto]

- I would 'like' this post on Instagram.
- I would 'comment' on this post on Instagram

Q9 State to what degree you agree with the following statements after seeing the post of [clothing brand].
After seeing the post I. . .

- would want to go there to eat or drink.
- have a positive feeling towards the company.
- would recommend this to someone else.

Q10 For each of the following comparisons of characteristics, state which one you feel fits the post best. For example: If the post give you a very positive feeling, the comparison of negative and positive would be:
negative . . . . x positive

- genuine/fake
- thrustworthy/unreliable
- realistic/unrealistic
- professional/unprofessional

Q11 In this question you will compare the two Instagram posts you just saw. State in what degree you feel that the characteristic is shown in the posts: more in post 1 (clothing brand) or post 2 (restaurant).
For example: If you get a slightly more positive feeling when looking at post 1, compared to post 2:
Positive: Post 1 . x . . . Post 2
If you feel that post 2 is way more fun than post 1:
Fun Post 1 . . . . x Post 2.

- genuine
- thrustworthy
- realistic
- professional

Q12 Age

Q13 Gender

- Male
- Female
- Other/Wishes not to specify

B Used Instagram posts

The following posts were used in the survey.

B.1 Kings of Indigo

Figure 2: Kings of Indigo, category: product

Figure 3: Kings of Indigo, category: user

Figure 4: Kings of Indigo, category: company/staff/interior/models

B.2 Supermercado

Figure 5: Supermercado, category: product

Figure 6: Supermercado, category: user

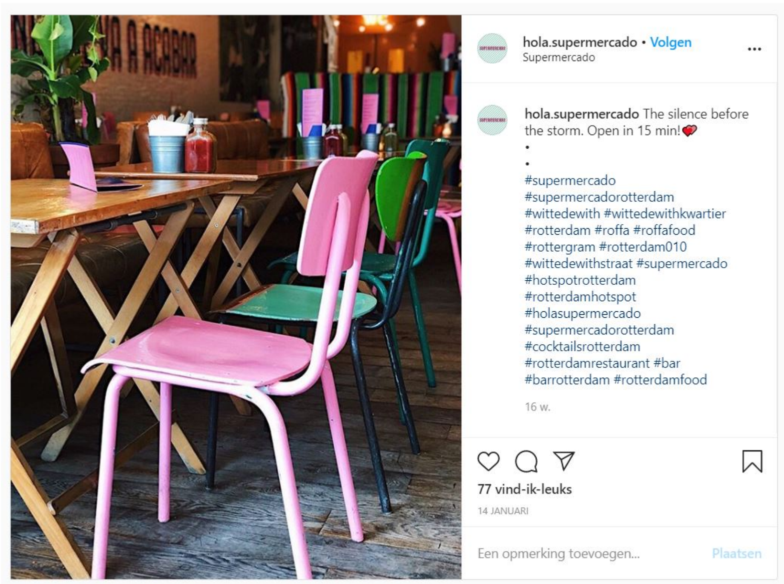


Figure 7: Supermercado, category: company/staff/interior/models

C X^2-goodness of fit tests

Table 10: X^2-goodness of fit test for group distribution

Group	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
1	32	32,890	-0,890	-0,155
2	33	32,890	0,110	0,019
3	35	32,890	2,110	0,368
4	31	32,890	-1,890	-0,330
5	34	32,890	1,110	0,194
6	35	32,890	2,110	0,368
7	33	32,890	0,110	0,019
8	30	32,890	-2,890	-0,504
9	33	32,890	0,110	0,019
Pearson χ^2 (8 degrees of freedom) = 0,696, P = 1,000				
Likelihood-ratio χ^2 (8 degrees of freedom) = 0,680, P = 1,000				

Notes: The expected value of the number of respondents per group has been computed by the total of respondents (296) divided by the number of groups (9)

Table 11: X^2-goodness of fit test for sample from population

(a) Age representation of population

Age in years	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
17 and under	49	16,302	32.700	8.099
18-24	97	71,5	25.500	3.016
25-34	57	82,368	-25.370	-2.795
35-44	35	47,476	-12.480	-1.811
45-54	47	37,752	9.250	1.506
55-64	0	18,304	-18.300	-4.278
65+	1	8,866	-7.870	-2.642
Pearson χ^2 (6 degrees of freedom) = 113,339, P = 0,000				
Likelihood-ratio χ^2 (6 degrees of freedom) = 119,954, P = 0,000				

(b) Gender representation of population

Gender	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
Female	212	164,580	47,420	3,696
Male	82	131,420	-49,420	-4,311
Undefined	2	0,001	1,999	63,214
Pearson χ^2 (2 degrees of freedom) = 0,004, P = 0,000				
Likelihood-ratio χ^2 (2 degrees of freedom) = 60,401, P = 0,000				

Notes: The mean of differences is in all analyses measured as the variable for the clothing industry- the variable for the restaurant industry.

Table 12: X^2-goodness of fit test for distribution of age among individual group from sample

(a) Group 1

Age in years	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
17 and under 5	5,297	-0,297	-0.129	
18-24	11	10,486	0,514	0,159
25-34	6	6,162	-0,162	-0,065
35-44	3	3,784	-0,784	-0,403
45-54	5	5,081	-0,081	-0,036
55-64	2	1,081	0,919	0,884
65+	0	0,108	-0,108	-0,329

Pearson χ^2 (6 degrees of freedom) = 1,099, P = 0,982
Likelihood-ratio χ^2 (6 degrees of freedom) = 1,063, P = 0,983

(b) Group 2

Age in years	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
17 and under	7	5,463	1,537	0,658
18-24	7	10,814	-3,814	-1,160
25-34	6	6,355	-0,355	-0,141
35-44	4	3,902	0,098	0,050
45-54	8	5,240	2,760	1,206
55-64	1	1,115	-0,115	-0,109
65+	0	0,111	-0,111	-0,333

Pearson χ^2 (6 degrees of freedom) = 3,377, P = 0,760
Likelihood-ratio χ^2 (6 degrees of freedom) = 3,443, P = 0,752

(c) Group 3

Age in years	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
17 and under	8	5,794	2,206	0,916
18-24	12	11,470	0,530	0,156
25-34	5	6,740	-1,740	-0,670
35-44	5	4,139	0,8615	0,423
45-54	4	5,557	-1,557	-0,660
55-64	1	1,182	-0,182	-0,167
65+	0	0,118	-0,118	-0,344

Pearson χ^2 (6 degrees of freedom) = 2,075, P = 0,913
Likelihood-ratio χ^2 (6 degrees of freedom) = 2,185, P = 0,902

(d) Group 4

Age in years	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
17 and under	3	5,132	-2,132	-0,941
18-24	9	10,159	-1,159	-0,364
25-34	10	5,970	4,030	1,649
35-44	3	3,666	-0,666	-0,348
45-54	5	4,922	0,078703	0,035
55-64	1	1,0477	-0,047	-0,046
65+	0	0,105	-0,105	-0,324

Pearson χ^2 (6 degrees of freedom) = 9,968, P = 0,681
Likelihood-ratio χ^2 (6 degrees of freedom) = 3,778, P = 0,707

(e) Group 5

Age in years	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
17 and under	6	5,628	0,372	0,157
18-24	16	11,142	4,858	1,455
25-34	5	6,547	-1,547	-0,605
35-44	4	4,020	-0,020	-0,010
45-54	2	5,399	-3,399	-1,463
55-64	0	1,149	-1,149	-1,072
65+	1	0,114	0,885	2,610

Pearson χ^2 (6 degrees of freedom) = 12,608, P = 0,050
Likelihood-ratio χ^2 (6 degrees of freedom) = 9,966, P = 0,126

(f) Group 6

Age in years	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
17 and under	7	5,794	1,206	0,501
18-24	10	11,470	-1,470	-0,434
25-34	5	6,740	-1,740	-0,670
35-44	3	4,139	-1,139	-0,560
45-54	8	5,557	2,443	1,036
55-64	2	1,182	0,818	-0,752
65+	0	0,118	-0,118	0,344

Pearson χ^2 (6 degrees of freedom) = 2,960, P = 0,814
Likelihood-ratio χ^2 (6 degrees of freedom) = 2,921, P = 0,819

Table 12: X^2-goodness of fit test for distribution of age among individual group from sample, continued.

(g) Group 7

Age in years	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
17 and under	4	5,463	-1,463	-0,626
18-24	14	10,814	3,186	0,969
25-34	6	6,355	-0,355	-0,141
35-44	3	3,902	-0,902	-0,457
45-54	5	5,240	-0,240	-0,105
55-64	1	1,115	-0,115	-0,109
65+	0	0,111	-0,111	-0,333
Pearson χ^2 (6 degrees of freedom) = 1,693, P = 0,946				
Likelihood-ratio χ^2 (6 degrees of freedom) = 1,783, P = 0,939				

(h) Group 8

Age in years	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
17 and under	5	4,966	0,034	0,015
18-24	9	9,831	-0,831	-0,265
25-34	7	5,777	1,223	0,509
35-44	4	3,547	0,453	0,241
45-54	4	4,764	-0,764	-0,350
55-64	1	1,014	-0,014	-0,014
65+	0	0,101	-0,101	-0,318
Pearson χ^2 (6 degrees of freedom) = 0,611, P = 0,996				
Likelihood-ratio χ^2 (6 degrees of freedom) = 0,702, P = 0,994				

(i) Group 9

Age in years	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
17 and under	4	5,463	-1,463	-0,626
18-24	9	10,814	-1,814	-0,552
25-34	7	6,355	0,645	0,256
35-44	6	3,902	2,098	1,062
45-54	6	5,240	0,760	0,332
55-64	1	1,115	-0,115	-0,109
65+	0	0,111	-0,111	-0,339
Pearson χ^2 (6 degrees of freedom) = 2,127, P = 0,908				
Likelihood-ratio χ^2 (6 degrees of freedom) = 2,125, P = 0,908				

Table 13: X^2-goodness of fit test for distribution of gender among individual group from sample

(a) Group 1

Gender	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
female	19	22,919	-3,919	-0,819
male	13	8,865	4,135	1,389
undefined	0	0,216	-0,216	-0,465

Pearson χ^2 (2 degrees of freedom) = 2,815, P = 0,245
Likelihood-ratio χ^2 (2 degrees of freedom) = 2,828, P = 0,243

(b) Group 2

Gender	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
female	21	23,635	-2,635	-0,542
male	12	9,142	2,858	0,945
undefined	0	0,223	-0,223	-0,472

Pearson χ^2 (2 degrees of freedom) = 1,410, P = 0,494
Likelihood-ratio χ^2 (2 degrees of freedom) = 1,564, P = 0,457

(c) Group 3

Gender	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
female	27	25,068	1,932	0,386
male	8	9,696	-1,696	-0,545
undefined	0	0,236	-0,236	-0,486

Pearson χ^2 (2 degrees of freedom) = 0,682, P = 0,711
Likelihood-ratio χ^2 (2 degrees of freedom) = 0,933, P = 0,627

(d) Group 4

Gender	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
female	22	22,203	-0,203	-13,431
male	7	8,588	-1,588	-0,542
undefined	2	0,209	1,791	3,918

Pearson χ^2 (2 degrees of freedom) = 196,023, P = 0,000
Likelihood-ratio χ^2 (2 degrees of freedom) = -95,580, P = 1,000

(e) Group 5

Gender	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
female	24	24,351	-0,351	-0,071
male	10	9,419	0,581	0,189
undefined	0	0,230	-0,230	-0,480

Pearson χ^2 (2 degrees of freedom) = 0,271, P = 0,873
Likelihood-ratio χ^2 (2 degrees of freedom) = 0,500, P = 0,779

(f) Group 6

Gender	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
female	26	25,068	0,932	0,186
male	9	9,696	-0,696	-0,224
undefined	0	0,236	-0,236	-0,486

Pearson χ^2 (2 degrees of freedom) = 0,321, P = 0,852
Likelihood-ratio χ^2 (2 degrees of freedom) = 0,557, P = 0,757

(g) Group 7

Gender	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
female	20	23,635	-3,635	-0,748
male	13	9,142	3,858	1,276
undefined	0	0,223	-0,223	-0,472

Pearson χ^2 (2 degrees of freedom) = 2,410, P = 0,300
Likelihood-ratio χ^2 (2 degrees of freedom) = 2,474, P = 0,290

(h) Group 8

Gender	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
female	25	21,486	3,514	0,758
male	5	8,311	-3,311	-1,149
undefined	0	0,203	-0,203	-0,451

Pearson χ^2 (2 degrees of freedom) = 2,097, P = 0,351
Likelihood-ratio χ^2 (2 degrees of freedom) = 2,492, P = 0,288

(i) Group 9

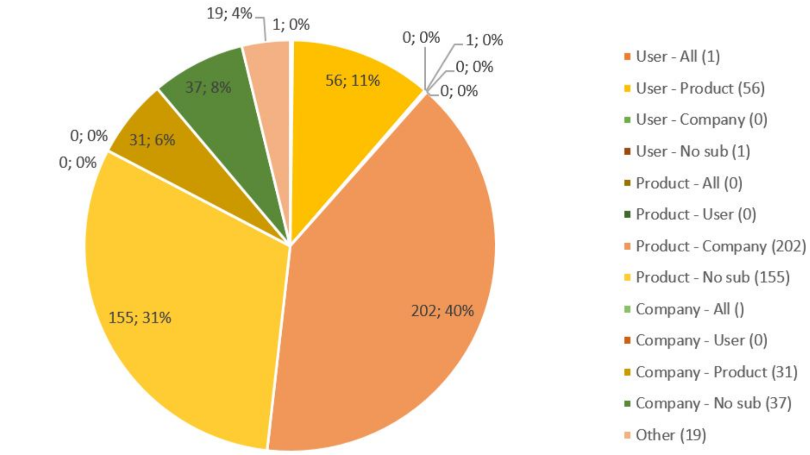
Gender	Observed (1)	Expected (2)	Difference (1)-(2)	Pearson
female	28	23,635	4,365	0,8989
male	5	9,142	-4,142	-1,370
undefined	0	0,223	-0,223	-0,472

Pearson χ^2 (2 degrees of freedom) = 2,906, P = 0,234
Likelihood-ratio χ^2 (2 degrees of freedom) = 3,456, P = 0,178

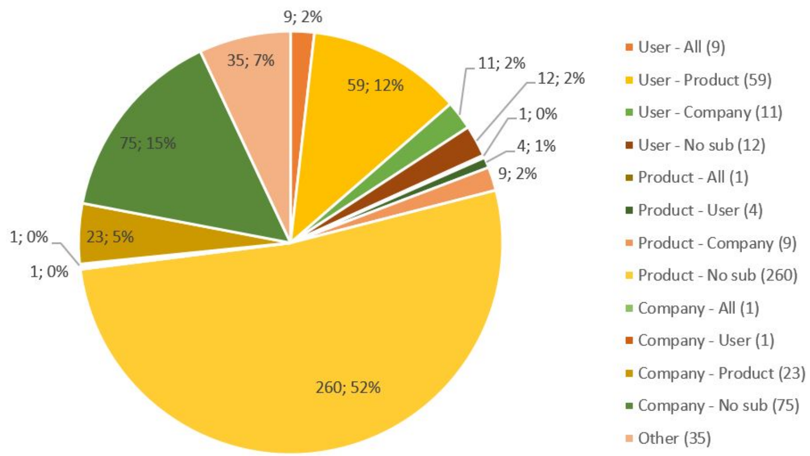
Notes: Group 4 is the only group where there is significant evidence that the distribution of the group is different from the overall sample. This is caused by the two respondents that did not define their gender. Nonetheless, the rest of the groups show that they are distributed according to the sample.

D Content distributions in Instagram data

Figure 8: Content category distribution of collected Instagram data



(a) Clothing brands



(b) Restaurants

Table 14: Content creator distributions of collected Instagram data

	UGC	BGC	Total observations
Restaurants	110	390	500
Clothing brands	58	444	502

E Factor analyses

Likes and comments on Instagram posts both measure consumer engagement. Whether or not to combine the two into one variable, has been tested by the use of factor analyses. This has been done with both the survey data, as well as with the data gathered from the Instagram posts.

E.1 Survey data

The factor analyses performed with the survey data resulted in Table 15[19]. The two variables for likes and comments are both explained by a conjoining variable. However, the Cronbach's alpha for both of the industries are not acceptable: 0,544 and 0,622 for the clothing industry and restaurants respectively.

[19]Both Table 15 as Table 16 only show the coefficients for likes and comments in the tables, as these are the only relevant coefficients for this study. X's mark the coefficients where variables that are grouped together by a factor. Variables that do not have an x in the table did not appear to be grouped in a factor and only showed coefficients below 0,400.

Table 15: Factor analyses for consumer engagement - survey data

(a) Clothing brands - survey data

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Instagram usage	x						
Having an Instagram account	x						
Age	x						
Post category: company		x					
Post category: user			x				
Post category: product			x				
Commenting				0,568			
Liking				0,561			
Perceived credibility							
Following brand							
Familiarity with brand							
Gender							

(b) Restaurants - survey data

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Instagram usage	x						
Having an Instagram account	x						
Age	x						
Post category: user		x					
Post category: company		x					
Post category: product			x				
Commenting				0,611			
Liking				0,596			
Perceived credibility				0,306			
Familiarity with brand					x		
Following brand					x		
Gender							

E.2 Instagram data

The same factor analysis has been performed for the gathered Instagram data. Table 16 is the corresponding output of the analysis. In this case the two variables did not end up in the same factor and are therefore not allowed to be used in a joined variable. Given the three analyses done, it became clear that it is not desirable to join the variables. Therefore all analyses in this study will be run twice: once with likes as a dependent variable and once with commenting as such.

Table 16: Factor analyses for consumer engagement - Instagram data

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Main category: user	x						
Repost	x						
Sub category: product	x						
Number of followers		x					
Likes		0,876					
Industry		x					
Multiple pictures in post		x					
Main category: company			x				
Main category: product			x				
Sub category: company				x			
People on picture				x			
Main category: other					x		
Sub category: user						x	
Giveaway in post							x
Commenting							0,572

F Descriptive data including giveaway-posts

Table 17: Descriptive data of Instagram accounts with giveaway-posts

		Likes			Comments		
Brand	N	Mean	SD	Median	Mean	SD	Median
Restaurants							
1nul8	100	110,05	49,58	103,50	4,26	5,15	3,00
Supermercado	100	94,37	32,78	87,50	4,49	4,88	3,00
Fred	100	613,18	250,77	598,50	14,80	14,24	12,00
Loetje	100	341,86	132,05	324,00	42,85	74,77	21,00
Coffeelicious	100	189,84	92,01	168,00	14,72	45,34	4,00
Clothing brands							
Kings of Indigo	100	273,10	150,19	228,50	7,42	7,84	5,50
Goosecraft	100	85,65	31,42	82,00	1,12	1,54	1,00
Guts & Gusto	101	1661,87	2130,28	1140,00	207,85	1385,37	8,00
Most Wanted	100	3781,99	950,47	3683,00	7,63	6,66	5,50
My Jewellery	101	9450,93	2485,59	8830,00	87,34	486,28	26,00
Total	1002	1168,061	3028,064	286	39,464	469,000	6
Interval [min, max]			[24, 20161]			[0, 12900]	