

Hashtags in NWOM tweets

A study about the possible consequences of hashtags in negative eWOM tweets
for webcare strategies, companies' reputation and the perceived negativity of
these tweets according to consumers

Martine Korthals, Radboud University Nijmegen

s4489780

Bachelor Werkstuk

Thema 4. Negatieve e-WOM vraagt om webcare

Eerste begeleider: R. Le Pair

Tweede begeleider: B. Hilberink

m.korthals@let.ru.nl – 0646307263

Number of words: 5280

ABSTRACT

Research about the negative (electronic) word-of-mouth has become more important and relevant the past couple of years. Lui (2006), among others, found that negative messages about products or services can have a negative influence on the consumer buying decision. Especially social media sites such as Twitter pose a threat to companies because of their high sparing potential. Webcare teams aim to minimize the damage caused by these shared negative messages. This study aims to find a relation between the most reoccurring hashtags on Twitter and the extent to which companies react to these complaint tweets, the type of complaint and the extent to which a dialogue exists. An accompanying pilot study aims to find out how consumers perceive negative tweets and if these hashtags can have any significant influence on a company's reputation. Among other results, it was found that hashtags, such as used on Twitter, seem not to have any significant influence on either webcare strategies and on the reputation. It was however found that consumers perceive negative tweets about products as being harmful for the company addressed in the tweet.

INTRODUCTION

Word-of-mouth has been a relatively trustworthy way in which consumers could obtain information about a product or service. In the past, word-of-mouth, or short: WOM, was usually received through a friend or family member (close ties). In contrast, in this day and age the sender of WOM does not have to be a close tie anymore (so called weak ties; Granovetter, 1983). Through internet consumers can send messages to everyone in their network with the potential of the message being shared with more than just the network of the sender. Although researchers have been aware of the existence of WOM since 1954 (Whyte, 1954), the relevance of studies about this subject has only been proven in the early 2000s. WOM is now proven to have influence on for example consumer buyer decisions and attitude towards a company (for more information: Lui, 2006; Mangold, Miller & Brockway, 1999; Anderson & Salisbury, 2003). Now that the internet is involved, the risks for companies to be faced with negative WOM and a potential negative impact on their reputation as a consequence of this occurrence has only grown bigger. Because of for example the anonymity, online consumers turn increasingly to the internet as a place to express their opinions on, and experiences with products or services (Gruen, Osmonbekov & Czaplewski, 2005). Both good and bad experiences end up online and can have a significant impact on the reputations of companies (Pfeffer, Zorbach & Carley, 2014).

The impact electronic WOM (hereafter: eWOM) can have on a company is demonstrated by Bickart and Schindler (2001). They state that negative eWOM is a form of consumer-generated content and is therefore perceived as more trustworthy by consumers than market-generated content. The potential for a message to be shared is therefore many times higher than when the message is sent by a company. The speed and anonymity of the internet attributes to the sharing potential of especially negative messages (van Noort & Willemsen, 2011). A study by Jansen, Zhang, Sobel and Chowdury (2009) investigating negative eWOM on social media, found results concerning the amount of negative messages. They found that approximately 50% of all eWOM messages about companies were positive and about 35% were negative. These negative eWOM messages often described negative experiences with a product or service. And even though the positive messages outnumber the negative ones, the negative outweigh the positive in effect (Park & Lee, 2009; Doh & Hwang, 2009). Cheung and Lee (2008) support this statement. In their study they aimed to research what the effects of negative as well as positive online customer reviews were on the purchasing decisions.

They found that negative eWOM had a larger impact on customers' intention to not buy a certain product or service, than positive eWOM had of increasing the intention to buy a product or service. These findings are especially significant since Nielson (2007) found that in 2007 over 90% of US consumers based their buying decisions largely on recommendations made by friends and family. These results underline the findings in studies by Lui (2006) and Anderson and Salisbury (2003) which state that negative (e)WOM can have an (negative) impact on consumer buying decisions and attitudes towards the company.

Negative eWOM has become the focus of many studies due to the importance of webcare strategies for companies. Especially social media has become a focal point for studies researching (negative) eWOM. One of the social media sites which have been the focus of several studies concerning negative eWOM is Twitter, a microblogging site. Twitter has already proved to be the fastest social network site and the medium on which consumers ask the most questions about a product or service (Jansen et al., 2009). In their study they aimed to gain insight into microblogging as eWOM communication, with negative eWOM as a focus, and its consequences for branding on Twitter. These authors also found that nowadays consumers use online platforms (social media sites such as Twitter) as trusted sources for information about products or services.

The study by Jansen (2009) mentioned above indicates the importance for companies to understand and anticipate to negative eWOM, especially with regard to Twitter. On this microblogging site the possibilities for a message to spread and be shared are enormous, as the site has millions of subscribers (Jansen et al., 2009). It is this easy way of sharing eWOM that poses a great threat to companies and their reputations (as demonstrated by Anderson and Salisbury, 2003). One of the consequences of this sharing potential is increased attention for webcare. Webcare is briefly described as an intervention in (often negative) eWOM (Kerkhof, Beukebom & Utz, 2010). A study by van Noort and Willemsen (2011) shows that when webcare teams respond to consumers' negative eWOM (especially when asked to) they can evoke sympathy and therefore create a more favorable brand evaluation and a more positive attitude towards that company. Van Noort and Willensem (2011) divide webcare into two categories: (1) reactive webcare and (2) proactive webcare. If a consumer asks a company for an answer the company engages in reactive webcare whereas when a company posts something without a consumer inviting them to do so they engage in proactive webcare. Huibers and Verhoeven (2014) researched the effects of webcare on Twitter. They also maintained a difference between reactive and proactive webcare. They found that when a

company reacts to a complaint on Twitter this often evokes positive feelings with the consumer. These authors also suggest that a reaction by the addressed company is always better than no reaction at all and has predominantly positive effects for the reputation of that company.

According to Derksen, Kelders and Keuning (2015) companies on Twitter invite their consumers increasingly to ask questions: these companies actively engage in dialogue with their consumers, therefore supporting statements made by Jansen et al. (2009). Creating a dialogue with consumers on Twitter is categorized as reactive webcare. Creating a dialogue can have a strong (positive) impact on consumers' evaluation of the company and can increase consumer loyalty to the company (Lee & Song, 2010). Studies show that consumers who have experienced a webcare reaction by the addressed company are more satisfied than consumers who did not receive webcare (van Noort & Willemsen, 2011). Approximately 20% of the companies who have a Twitter account have a separate account especially designed for webcare (Derksen et al., 2015). Huibers and Verhoeven (2014) even go as far as stating that staying silent when a consumer (or consumers) has a complaint about their product or service, is the worst webcare strategy a company can have. According to these authors the question should not be if the company should react but rather how the company should react.

Negative tweets, however, require a different approach than negative eWOM on other social network sites. This different approach is needed because of the hashtags, which are used in most, if not all tweets. Hashtags are often signals of a tweet's meaning, its topic or its target audience (Efron, 2010). Hashtags are used on Twitter more frequently than on any other social media site, on which they might also occur (such as Facebook and Instagram).

Hashtags are also present in negative tweets towards a company to increase its meaning and the gravity of the complaint (for example “@Ziggo the network is down... AGAIN! If this happens one more time, I will transfer to another company #fail” – Twitter, 2015). Le Pair (2014) found that one particular hashtag is used more often than any other hashtag when evaluating negative tweets. In his study he researched Dutch negative tweets that were critical of a product, a service or a company itself. Among other results it was found that #fail was used most often. 80% of all negative tweets contained this particular hashtag. Other reoccurring hashtags were: #jammer (6.7%), #zucht (5.4%), #faal (3.5%) and #slecht (2.3%). Few literature exists about the influence these hashtags might have, how negative tweets that

contain these hashtags are perceived by consumers and how it can damage a companies' reputation. This information gap is fundamental to this study.

The overall aim of this study is to investigate whether hashtags can have any (negative) influence on the reputation of a company and their webcare strategies. This corpus study focuses on the possible consequences hashtags can have on webcare strategies. Additional research was done to find out whether these hashtags can also negatively influence a company's reputation. The main study was a corpus study and the additional study was an experimental one.

#fail was chosen as a focal point for le Pair (2014) found that this hashtag was most often used in negative eWOM on Twitter. This raises the question how influential this hashtag is or has become over the years. In this corpus study #fail will be compared to other reoccurring hashtags to find out whether they are the cause of any (negative) consequences for webcare strategies.

First, this corpus study aims to research whether webcare teams react to tweets that contain #fail more often than they react to tweets that use another hashtag (either #jammer, #zucht, #faal or #slecht). Second, it will be researched which type of complaint is most often used in these negative tweets. Third, this study aims to research whether webcare reactions by companies to #fail tweets create a dialogue more often than webcare reactions to tweets that use another hashtag. This study is followed by a small experimental pilot study. This pilot study aims to complement the results of this corpus study and to research whether hashtags can have any influence on companies' reputation.

RQ1: To what extent do webcare teams react to negative tweets that use hashtags?

RQ1a: How often do webcare teams react to negative tweets containing #fail?

RQ1b: How often do webcare teams react to negative tweets containing #jammer, #zucht, #faal or #slecht?

In the second research question the different types of complaints are investigated. The type of complaint can be anything from a product or a service that does not function like the consumer expects it to or the lack of communication from a company when the consumer addressed their problem several times already. It can also be a consumer disagreeing with an idea or a policy executed by a government.

RQ2: Which type of complaint is most often used in negative tweets?

RQ3: To what extent do companies react to different types of complaints?

The corpus for the third research question will be considerably smaller than the corpus for the first and second research question as only the tweets which got a reaction from the addressed company are relevant. There can only be a dialogue between company and consumer if the consumer gets a reaction to his or her initial complaint tweet.

RQ4: To what extent do webcare reactions of companies to negative tweets create a dialogue between the consumer and the company?

RQ4a: To what extent does a webcare reaction to a negative #fail tweet create a dialogue between the consumer and the company?

RQ4b: To what extent does a webcare reaction to a negative tweet containing #jammer, #zucht, #faal or #slecht create a dialogue between the consumer and the company?

PILOT STUDY

To support statements about the possible influence of #fail, a small experimental pilot study will be executed. The main aim of this pilot study is to find out how consumers perceive negative tweets. This pilot study also aims to research whether hashtags can have a (negative) influence on the reputation of the addressed company. In this pilot study #fail will be compared to the two hashtags most reoccurring after #fail, which are #jammer and #faal.

Several studies suggest the importance of insight in how consumers perceive complaint tweets. Pfeffer et al. (2014) found that exposure to many negative eWOM messages can have a high affective nature and can thus be able to influence a consumers' buying decisions. Le Pair and van Dongen (2013) found that exposure to negative eWOM messages can cause a decrease in company trust. Hence, the research questions of this pilot study is as follows:

RQ5: How do consumers perceive #fail tweets in comparison to negative tweets that use #jammer or #faal?

RQ6: To what extent do different types of hashtags affect a company's reputation?

METHOD SECTION CORPUS STUDY

MATERIAL

For this corpus study several thousands of tweets were analyzed. These tweets, from microblogging site Twitter, were the population of this study. All tweets were collected between August 23rd and September 22nd 2015. Only tweets that were written in the Dutch language were eligible for the corpus study. The data were collected using a program called data crawler which selected tweets automatically according to several conditions (for example: the tweet should contain one of the three hashtags).

The unit of analysis for this study was Twitter. The units of data collection were the tweets collected for this study. The unit of sampling is three sided: (1) do webcare teams react more often to tweets that contain #fail than to negative tweets that contain another hashtag? (2) What type of complaint is most often used in combination with which hashtag? (3) Does a reaction by a webcare team to a #fail tweet create a dialogue more often than a webcare reaction to tweets that contain another hashtag?

The total amount of tweets before the initial coding was 10.742. In the first stage thirteen coders coded these tweets on whether they were real negative eWOM or not. The total amount was eventually reduced to 3287 tweets that were true negative eWOM tweets. All these 3287 tweets were used in this corpus study.

PROCEDURE

After it was decided which tweets would qualify as real negative eWOM by the thirteen coders, it was decided which general variables would be coded by the same thirteen coders. Among these were type of complaint, sector of the company addressed in the tweet, if the company reacted and if there was a dialogue between company and consumer. However, before these variables could be coded a meeting was called to discuss what criteria would be used for which variable. Among other things it was decided which types of complaints would be used for the coding or when an online conversation was categorized as a dialogue. For example: it was decided that when the person who sent the initial tweet reacted to the first reaction by the company it was categorized as a dialogue (from two turn changes on). The codebook the coders used to code the variables can be found in Appendix I.

After all the tweets were coded the researcher sat down with one other coder for a reliability analysis. The reliability analysis was done for four variables as there was a possibility for confusion about these variables among the coders. The interrater reliability of

the variable type of complaint was acceptable ($\kappa = .725, p < .001$). The interrater reliability of the variable sector was also found to be acceptable ($\kappa = .738, p < .001$). The interrater reliability for the variables existence of a dialogue ($\kappa = .961, p < .001$) and turn changes ($\kappa = .951, p < .001$) were found to be excellent.

STATISTICAL TREATMENT

To answer the research questions one statistical test was used. To compare different variables to each other a Chi-square test was executed. In the procedure section the interrater reliability was calculated for several variables as most variables were coded by thirteen coders.

EXPERIMENTAL PILOT STUDY

DESIGN

This experiment had a 3 (types of hashtag: #fail, #jammer and #faal) x 1 (perceived overall negativity of the tweet) between subjects design. Each participant filled in one of three different questionnaires, in which the hashtags used in the (fictional) example tweets differed. The questionnaires were randomly distributed among participants.

MATERIAL

In each questionnaire only one type of hashtag was used. These hashtags (#fail, #jammer or #faal) were manipulated in this study for the participants were assigned only one of these three versions and they were assigned randomly. In this questionnaire the participants evaluated several tweets after answering some general questions about their gender, age, education and if they were familiar with Twitter. The questionnaire as distributed among participants can be found in Appendix II.

The hashtags used in this questionnaire were chosen from a corpus in which several other hashtags were present. For this experimental pilot study it was decided to compare the hashtags that were most often used in the tweets from the corpus. These hashtags were #fail, #faal and #jammer with #fail occurring most often.

The tweets and companies mentioned in the tweets were both fictional to prevent possible positive feelings by consumers towards the company or its products or services to influence the results in any way. The questionnaire was also entirely in Dutch for most, if not

all, participants were expected to be Dutch. This was done to eliminate the possibility for any confusion towards what the tweets or the questions might mean or how they might be interpreted.

PARTICIPANTS

A total of 61 participants took part in this experiment. 60 participants were Dutch, one was Swiss (mean age: 27-36; range 17 – 47; 60,7% female) and most participants had at least a Bachelor's degree (72,1%). A Chi-square test showed no significant relation between the type of hashtag and gender ($\chi^2 (2) = 4.67, p = .097$), and neither did the F-test for type of hashtag and age ($F (3) = 3,33, p = .706$).

INSTRUMENTATION

Participants each filled in a hand out copy of one of the three versions of the questionnaire. After answering some questions about their age, gender and education participants had to answer five questions about a fictional tweet. The participants had to evaluate the seriousness of the complaint, the perceived negativity of the tweet and the perceived effect on the companies' reputation.

The participants evaluated the tweets on a 5-point Likert scale (based on Bayard et al., 2001) following a statement starting with "I think...", anchored by "completely disagree – completely agree" (in Dutch "helemaal mee oneens – helemaal mee eens). Seriousness of the complaint was measured with two items: 'the customer has a serious complaint' and 'the company should take this complaint seriously'. The reliability of the two items measuring seriousness of the complaint was not acceptable ($\alpha = .53$).

The perceived negativity of the tweet was measured with the statement: "Ik vind dat deze tweet de organisatie in een negatief daglicht stelt", anchored by "completely disagree – completely agree". The reliability of the variable perceived negativity was good ($\alpha = .91$).

The perceived effect on the company's reputation was measured with two items: 'negative effects as a consequence and possible damage for the addressed company's reputation'. The reliability of the two items measuring perceived effect on the company's reputation was not acceptable ($\alpha = .54$).

Because two of the computed variables were not reliable these variables were evaluated separately in the result section.

PROCEDURE

The questionnaire was distributed personally among participants. The questionnaire started with short instructions which explained the further procedure. After the instructions participants were first asked to fill in some questions about their background, such as age, gender, education level and lastly some questions about their familiarity with Twitter. Next, the questionnaire focused on the evaluation of five fictional tweets.

The participants were gathered via personal contact. The participants were asked to fill in the questionnaire on the spot and hand it in to the researcher afterwards.

STATISTICAL TREATMENT

To answer the research question, different statistical tests were used. First Cronbach's alpha was calculated to determine the reliability for each (computed) variable. The results of the reliability analysis can be found under the instrument section. Furthermore a Chi-square test and a F-test were executed to test the relation between two variables. For example, the relation between type of hashtag and initial reaction from the company on the consumers complaint tweet. For the pilot study a multivariate analysis was executed to find out what the effects of several variables were on the dependent variable.

RESULTS CORPUS STUDY

In this chapter the results of the corpus study will be discussed. The results for each research question will be discussed separately. The first research question, which was split in two parts – a and b – focused on the extent to which companies responded to negative eWOM on Twitter that used certain hashtags. The hashtags used in these tweets were #fail, #faal, #jammer, #slecht and #zucht. This research question aimed to find out whether companies reacted more to tweets that used #fail in comparison to tweets that used another hashtag.

Generally the number of tweets that did not get a response from companies was higher than the number of tweets that did get a response. 32% of the tweets in this corpus got a response and 68% did not.

To answer this research question a Chi-square test was used. The Chi-square test showed a significant relation between type of hashtag and the extent to which companies

responded to negative eWOM on Twitter ($\chi^2(4) = 36.40, p < .001$). The results of this Chi-square test are presented in Table 1.

Table 1. Chi-square test for type of hashtag and response by the addressed company (n=3287)

hashtag	#faal	Webcare reaction			TOTAL
		No reaction	Reaction		
#fail	Count	117	32	149	2728
	Adjusted residual	3.2	-3.2		
#jammer	Count	1838	890	276	276
	Adjusted residual	2.9	-2.9		
#slecht	Count	150	126	93	93
	Adjusted residual	-4.4	4.4		
#zucht	Count	51	42	41	41
	Adjusted residual	-2.4	2.4		
TOTAL	Count	23	18		3287
	Adjusted residual	-1.4	1.4		

As can be seen in Table 1, companies seem to react significantly less to tweets that contain #fail and #faal than to tweets that contain #jammer, #slecht or #zucht. #zucht was found to not have any significant result. Companies did not react significantly more or less to this hashtag. It was however found that companies reacted more often to tweets that contain #jammer or #slecht.

The second research question focused on which type of complaint was most often found in negative eWOM. It was found that consumers most often complain about a service, such as a telephone or television network. Over half of all tweets studied were complaints about services (53.3%). 17.5% of the tweets were complaints about communication of companies or the lack thereof. For example, a complaint often found was the lack of communication from a company when the consumer had problems with their products or services. 16.9% of the tweets contained a complaint about an idea or policy. Only 7.5% had a complaint about an actual product.

Next it was investigated to which complaints companies tended to react most often. A Chi-square test showed a significant relation between type of complaint and whether these tweets got a response by the companies they were addressed to ($\chi^2(4) = 225.39, p < .001$). The results of this Chi-square test can be found in Table 2.

Table 2. Chi-square test for type of complaint and response by the addressed company
(n=3287)

Type of complaint	Product	Count	Webcare reaction		
			No reaction	Reaction	TOTAL
			139	107	247
Service	Adjusted residual	-3.4	3.4		
	Count	1079	671	1752	
Communication	Adjusted residual	-6.0	6.0		
	Count	324	253	577	
Idea or policy	Adjusted residual	-5.7	5.7		
	Count	506	51	557	
Not clear	Adjusted residual	13.5	-13.5		
	Count	131	26	157	
	Adjusted residual	4.7	-4.7		
TOTAL		2180	1108		3287

As can be seen from the Table, companies reacted least to complaints about ideas or policies. However, it can be difficult for companies to react to complaints about their policy as they usually cannot give away this information to consumers. The average response rate for complaints about products, services and communication is almost equal to each other at around 40%.

The third researched question aimed to investigate whether companies' reaction to negative eWOM evoked a dialogue more often when #fail was used in a tweet than when other hashtags were used. A Chi-square test was used to compare the different types hashtags. First off, it should be mentioned that 'only' 1108 of the 3287 tweets were used for this research question as those 1108 tweets got an initial response by the addressed company in their complaint tweets. A dialogue is described as at least two turn changes between a consumer and a company. For this research question the amount of tweets that was evaluated as having a dialogue is interesting. 63% of the 1108 tweets was evaluated as a dialogue.

A Chi-square test between type of hashtag and the existence of a dialogue between company and consumer showed no significant relation ($\chi^2(4) = 7.73, p = .102$). Type of hashtags seems to have no effect on whether there was an existing dialogue between company and consumer. The results of this Chi-square test are presented in Table 3.

Table 3. Chi-square test for type of hashtag (#faal, #fail, #jammer, #slecht, #zucht) and the existence of a dialogue between company and consumer (n=1108)

Hashtag	#faal		Existence of a dialogue		
			No dialogue	Dialogue	TOTAL
#fail	Count	10	22	32	
	Adjusted residual	-.7	.7		
#jammer	Count	330	560	890	
	Adjusted residual	.1	-.1		
#slecht	Count	54	72	126	
	Adjusted residual	1.4	-1.4		
#zucht	Count	14	28	42	
	Adjusted residual	-.5	.5		
TOTAL	Count	2	16	18	
	Adjusted residual	-2.3	2.3		
	Count	410	698	1108	

RESULTS PILOT STUDY

To complement the results in the corpus study an experimental pilot study was executed. This pilot study aimed to research if a certain type of hashtags (#fail, #faal or #jammer) had any influence on the level of perceived negativity of a complaint tweet when evaluated by consumers. Secondly, this pilot study aimed to investigate whether hashtags can have a (negative) influence on a company's reputation. 61 questionnaires were admitted among consumers. In this chapter the results will be discussed.

To test for significance a one-way multivariate analyses was executed. In the method section it was found that the five variables used for this study could not be computed because they would not be reliable. Thus, all variables were evaluated separately. The test for seriousness of the complaint, extent to which companies should take the complaint serious, perceived extent of negative light on the company, perceived consequences for the company and perceived consequences for the company's reputation with type of hashtag as factor, found no significant multivariate effect ($F(10,11) = .919, p = .519$). It did not seem to matter to consumers which hashtag was used in the negative tweet. In fact, it did not seem to matter at all if hashtags were used in the tweets. The negativity of the tweet itself and its complaint seems to have more effect than the hashtags.

However, after these results were gathered and found not to be significant, the researcher asked some of the participants why they had answered as they did. The participants reported that they hadn't actually given any attention to the hashtags used in the tweet. They

focused mostly on the complaint itself and evaluated the tweet according to what their feelings were towards that complaint. As a consequence of these results it was decided to do some additional research. It was researched what type of complaint was evaluated most negatively. Every questionnaire contained three types of complaints. Because there were not enough results to execute a valid test the means were compared to each other. The three types of complaints compared to each other were: (1) complaints about a service ($M = 3.20$, $SD = .74$), (2) complaints about communication or the lack thereof ($M = 3.16$, $SD = .89$) and (3) complaints about a product ($M = 4.23$, $SD = .66$). When looking at the means for every complaint it seems that the complaint about the product was evaluated as being the most negative.

CONCLUSION CORPUS STUDY

The main aim of the corpus study was to research whether company webcare teams react more often to complaint tweets with #fail than to tweets with another hashtag, which type of complaint was most often present, which type of complaint companies reacted to the most and to what extent reactions by webcare teams created a dialogue with the consumer.

It was found that the number of tweets that did not get a response from webcare teams to their initial tweet was much bigger than the number of tweets that did get a reaction. 32% of all tweets got a reaction to their initial complaint tweet.

By executing a Chi-square test it was found that the response rate for the different types of hashtags was significantly different. It was found that companies reacted least to tweets that contained #faal or #fail. This is striking because Le Pair (2014) found that #fail was the hashtags used most often in complaint tweets. He found that over 80% of all complaint tweets contained this particular hashtag. A possible explanation for this occurrence could be the observation that #faal and #fail are often used in tweets that describe a failed action by most often the user himself. It was also found that companies reacted most often to tweets that used the hashtag #jammer or #slecht.

In the second research question the types of complaints occurring in these negative eWOM tweets were investigated. It was found that most consumers posted complaint tweets about services online. Over half of all tweets were found to be complaining about a service (53.3%).

In line with these findings it was also found that the response rate for complaint tweets about services are fairly high at about 40%. Reactions to complaints about ideas or policies are least often found as it is difficult for companies to react to such tweets.

The fourth and final research question of the corpus study was if a webcare reaction to a #fail complaint tweet created more dialogue than a webcare reaction to a tweet that contained another hashtag. It was found that the chances of creating a dialogue with a consumer are higher when #fail was used in the initial tweet than when #faal, #slecht or #zucht was used. It was also found that there seems to be more dialogue between company and consumer when #jammer was used in a tweet than when #fail was used in the initial complaint tweet. Thus, it seems that #jammer is the hashtag which is taken most seriously by company webcare teams.

CONCLUSION EXPERIMENTAL PILOT STUDY

The main purpose of this experimental pilot study was to complement the results obtained in the results of the corpus study mentioned above. This was done by researching how consumers evaluate complaint tweets with different hashtags in each (of the three) versions. It was also investigated whether hashtags can have any influence on companies' reputation.

No effect was found for the pilot study. The type of hashtag seems to have no effect on the evaluation by consumers. The overall perceived negativity was not different for any type of hashtag. After the questionnaires were gathered and the results turned out to be negative the researcher asked some of the participants why they chose for the answers they had given previously in the questionnaire. The participants answered that they had evaluated the tweets solemnly on the complaint itself and mostly did not notice the added hashtags. Thus suggesting that type of hashtag has no effect whatsoever on the perceived negativity of a complaint tweet and cannot influence a company's reputation in any way.

GENERAL DISCUSSION

The main goal of this study was to find out whether hashtags have any influence on perceptions of negativity when evaluated by consumers and if the results of this study can have (negative) influence on a companies' reputation. In the pilot study it was found that

hashtags have no direct influence on the perceptions of negativity when consumers evaluate a complaint tweet. This would implicate that different types of hashtags, or hashtags in general, have no negative influence on the reputation of companies. A more interesting result concerning companies' reputation was what participants had to say after they filled in the questionnaire. Participants declared that not the hashtag but the complaint itself was the reason they evaluated the tweet as they did. This was the cause for some additional research. It was investigated which type of complaint was evaluated most negatively by consumers, as consumers had previously declared the actual complaint was the reason for how they evaluated the tweets as they did. It was found that the complaint tweet about a product was evaluated most negative. With these results in mind it is interesting to notice that the response rate for tweets about products is around 40%, even though the amount of tweets sent about products is 'only' 7.5% of the total amount of negative eWOM tweets. Even though the response rate is above the average, it should be noted that negative tweets about products can have the most devastating effect on a company's reputation according to the participants who filled in the questionnaire. Therefore it could be interesting to execute some additional research which focuses on the effect tweets about products can have on a companies' reputation and how this can be translated into an effective webcare strategy. However, such a study is possibly only important for companies that sell products. Companies that do not sell products should at least always try to communicate with their consumers. Actively seeking dialogue seems to be the best online strategy for content consumers according to Huibers and Verhoeven (2014) and van Noort and Willemsen (2011).

LITERATURE

- AC Nielson (2007). Trust in Advertising: A Global Nielsen Consumer Report – October 2007.
- Anderson, E.W., & Salisbury, L.C. (2003). The formation of market-level expectations and its covariates. *Journal of Consumer Research*, 30, 115–124.
- Bayard, D., Weatherall, A., Gallois, C., & Pittam J. (2001). Pax Americana? Accent attitudinal evaluations in New Zealand, Australia and America. *Sociolinguist* 5, 22-49.
- Cheung, C.M.K., & Lee, M.K.O. (2008). Online consumer reviews: does negative electronic word-of-mouth hurt more? *Americas Conference on Information Systems (AMCIS)* 1–10.
- DerkSEN, M., Kelders, M., & Keuning, A. (2015). De stand van webcare 2015. Upsteam.
- Doh, S., & Hwang, J. (2009) How consumers evaluate eWOM (electronic word-of-mouth) messages. *Cyberpsychology & behavior*, 12, 193–197.
- Efron, M. (2010). Hashtag retrieval in a microblogging environment. *Graduate School of Library and Information Science*, 1–2.
- Granovetter, M. (1983). The strength of weak ties: a network theory revisited. *Sociological Theory*, 1, 201–233.
- Gruen, T.W., Osmonbekov, T., & Czaplewski, A.J. (2005). eWOM: The impact of customer-to-customer online know-how exchange on customer value and loyalty. *Journal of Business Research*, 59, 449–456.
- Huibers, J. & Verhoeven, J. (2014). Webcare als online reputatiemanagement. Het gebruik van webcare-strategieën en conversational human voice in Nederland, en de effecten hiervan op de corporate reputatie. *Tijdschrift voor Communicatiewetenschap*, 42, 165–189.
- Jansen, B.J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of American Society for information science and technology*, 60, 2169–2188.
- Kerkhof, P., Beukebook, C., & Utz, S. (2010). The Humanization of a Company: Effects of personal vs. impersonal organizational reactions to negative online consumer reviews, paper presented at the Etmaal voor de Communicatiewetenschap, Gent, Belgium.
- Lee, Y.L., & Song, S. (2010). An empirical investigation of electronic word-of-mouth: informational motive and corporate response strategy. *Computers in Human Behavior*, 26, 1073–1080.
- Le Pair, R. (2014). *Negatieve word-of-mouth tweets en webcare: welke hashtags?* Consulted at September 22nd 2015.
- Le Pair, R., & van Dongen, M. (2013). *Als een deskundige 't zegt... effecten van negatieve word-of-mouth tweets.* Consulted at September 22nd 2015.
- Lui, Y. (2006). Word of Mouth for movies: its dynamics and impact on box office revenue. *Journal of marketing*, 70, 74–89.
- Mangold, W.G., Miller, F., & Brockway, G.R. (1999). Word-of-mouth communications in the service marketplace. *Journal of Services Marketing*, 13, 73–89.
- Park, C., & Lee, T.M. (2009). Information direction, website reputation and eWOM effect: a moderating role of product type. *Journal of Business Research*, 62, 61–67.
- Pfeffer, J., Zorbach, T., & Carley, K.M. (2014). Understanding online firestorms: Negative word-of-mouth dynamics in social networks. *Journal of Marketing Communications*, 20, 117–128.
- Van Noort, G., & Willemsen, L.M. (2011). Online Damage Control: The effects of proactive versus reactive webcare interventions in consumer-generated and brand-generated platforms. *Journal of interactive marketing*, 131–140.
- Whyte, W.H., Jr. (1954). The web of word of mouth. *Fortune*, 50, 140–143.

APPENDIX II: CODEBOOK

GENERAL VARIABLES

Type of hashtag

- #faal
- #fail
- #jammer
- #slecht
- #zucht

Number of followers (of the user who sent the initial complaint tweet)

Date of publication (of the complaint tweet)

Time of publication (of the complaint tweet)

Tweet text (of the complaint tweet)

SPECIFIC VARIABLES

Is the tweet a true negative eWOM tweet?

- No
- Yes

What type of complaint was used in the tweet?

- An actual product
- The product is a service
- Communication (or lack thereof)
- An idea or policy
- Unclear or something else

In which sector did the addressed company operate?

- Government organizations
- Transportation
- Financial companies
- Retail
- Media
- Something else

Was the company profit or non-profit?

- Profit
- Non-profit

What was the name of the company?

What place did the @-mention take in the complaint tweet?

- At the absolute beginning
- Somewhere else except for the absolute beginning
- There was no @-mention

How did the consumers address the company in the complaint tweet?

- Via #
- Via @
- Just the name

Was there a webcare reaction?

- No
- Yes

Was there a dialogue between company and consumer?

- Not applicable
- Yes
- No

How many turn changes occurred in the dialogue between company and consumer?

- No applicable
- (any observed number)

Does the addressed company have a separate Twitter account for webcare purposes solemnly?

- No
- Yes

Did the reaction tweet by the company contain human voice in the sense of a name of the person who wrote the tweet to the consumer?

- Not applicable
- Yes, via [^](initials)
- Yes, via [^](name)

Did the reaction tweet by the company contain human voice in the sense of the use of I, me or mine by the person who wrote the tweet?

- Not applicable
- Yes they used I, me or mine
- No they did not use I, me or mine

APPENDIX I: QUESTIONNAIRE

Beste participant,
hartelijk dank voor het invullen van deze enquête. Deze enquête zal zo'n 5 minuten van uw tijd in beslag nemen. Na het invullen van een paar algemene vragen zult u een vijftal fictieve tweets beoordelen. Lees deze tweets eerst zorgvuldig door en vul daarna de bijbehorende vragen in.

1. Wat is uw geslacht?

- Vrouw
 Man

2. Wat is uw leeftijd?

- 16 en jonger
 17 tot 26
 27 tot 36
 37 tot 46
 47 en ouder

3. Wat is uw nationaliteit?

4. Wat is uw huidig opleidingsniveau of hoogst afgeronde opleidingsniveau?

- Middelbare school
 MBO
 HBO
 Universiteit

5. Bent u bekend met Twitter?

- Ja
 Nee

6. Heeft u een Twitter account?

- Ja
 Nee

Heeft u ja geantwoord op de bovenstaande vraag, beantwoord dan vraag 7 en 8. Indien u deze vragen met nee heeft beantwoord, ga dan verder naar het beoordelen van de tweets.

7. Bent u actief op Twitter? (u plaatst tenminste 1 tweet per week)

- Ja
 Nee

8. Leest u regelmatig berichten op Twitter? (tenminste 1 keer per week)

- Ja
 Nee

Tweet 1: “@TopTV mijn tv ligt er WEER uit! Als dit nog één keer gebeurt stap ik over #fail”

Ik vind dat de consument een serieus probleem heeft.

Helemaal mee oneens

O

O

O

O

Helemaal mee eens

O

Ik vind dat de organisatie deze klacht serieus moet nemen.

Helemaal mee oneens

O

O

O

O

Helemaal mee eens

O

Ik vind dat deze tweet de organisatie in een negatief daglicht stelt.

Helemaal mee oneens

O

O

O

O

Helemaal mee eens

O

Ik denk dat deze tweet negatieve gevolgen heeft voor het aangesproken bedrijf.

Helemaal mee oneens

O

O

O

O

Helemaal mee eens

O

Ik denk dat deze tweet de reputatie van het aangesproken bedrijf kan beschadigen.

Helemaal mee oneens

O

O

O

O

Helemaal mee eens

O

Tweet 2: “Pff, dit is nou al de 4^e keer dat ik doorverbonden wordt, bedankt voor niks!

@fashionvictim #geenklantenservice #klantkwijt #fail”

Ik vind dat de consument een serieus probleem heeft.

Helemaal mee oneens

O

O

O

O

Helemaal mee eens

O

Ik vind dat de organisatie deze klacht serieus moet nemen.

Helemaal mee oneens

O

O

O

O

Helemaal mee eens

O

Ik vind dat deze tweet de organisatie in een negatief daglicht stelt.

Helemaal mee oneens

O

O

O

O

Helemaal mee eens

O

Ik denk dat deze tweet negatieve gevolgen heeft voor het aangesproken bedrijf.

Helemaal mee oneens

O

O

O

O

Helemaal mee eens

O

Ik denk dat deze tweet de reputatie van het aangesproken bedrijf kan beschadigen.

Helemaal mee oneens

O

O

O

O

Helemaal mee eens

O

Tweet 3: "Dit kan echt niet!! @FoodzSupermarkt dit is de zoveelste keer dat ik schimmel aantref in JULLIE producten #fail #overstappennaarJumbo"

Ik vind dat de consument een serieus probleem heeft.

Ik vind dat de organisatie deze klacht serieus moet nemen.

Ik vind dat deze tweet de organisatie in een negatief daglicht stelt.

Ik denk dat deze tweet negatieve gevolgen heeft voor het aangesproken bedrijf.

Ik denk dat deze tweet de reputatie van het aangesproken bedrijf kan beschadigen.

Tweet 4: "Hoe vaak moet ik bellen voordat jullie begrijpen dat ik jullie spammails niet wil ontvangen!?!? @SpijkermanBouwmarkt #spam #teveelmail #FAIL #geencommunicatie"

Ik vind dat de consument een serieus probleem heeft.

Ik vind dat de organisatie deze klacht serieus moet nemen.

Ik vind dat deze tweet de organisatie in een negatief daglicht stelt.

Ik denk dat deze tweet negatieve gevolgen heeft voor het aangesproken bedrijf.

Ik denk dat deze tweet de reputatie van het aangesproken bedrijf kan beschadigen.
Helaas
Helaas

