

# THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

The Influence of Text-image Relations in Internet memes

Xinyao Zhang

Anr 1275887

Master Thesis

Communication and Information Sciences

Business Communication and Digital Media

Faculty of Humanities

Tilburg University, Tilburg

Supervisor: Dr. N.T. Cohn

Second reader: Dr. J. Schilperoord

July 2017

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

### Abstract

In the current digital era, Internet memes are becoming prevalent in an online environment. Internet memes inherently include the text and image, which forms a text-image relation. Understanding Internet memes can gain insight into the digital culture and society, and text-image relations in new media. Study 1 explored whether people can distinguish different text-image relations in Internet memes by a rating test for 24 pairs of Internet memes and suggested that the elaboration relation was the easiest one for people to notice. Based on the results of Study 1, a further exploration of the influence of each text-image relation was performed in Study 2. The self-paced viewing paradigm was used to compare the response time and a rating test was applied to discover the differences of understanding and enjoyment among diverse types of Internet memes. It was found that the types of text-image relation indeed influenced response time. Also, the differences were shown regarding response time, ratings of understanding and enjoyment between different modality types. Future research could shed more light on the reasons of the different influences caused by types of text-image relations. Additionally, it will also be interesting to discover the text-image relations in the Internet memes under a more realistic context.

*Key words:* Internet memes, multimodal interaction, text-image relations, comprehension

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

### Introduction

Recently, *Internet memes* have become an integral component of digital culture, drawing both public and academic attention (Nissenbaum & Shifman, 2017). Therefore, understanding Internet memes is important to gain insight into digital culture (Börzsei, 2013). An increasing number of individuals enjoy viewing and using Internet memes through several online platforms, (i.e. 9GAG) or sharing them with friends during an online conversation. Miltner (2014) proposed that Internet memes provide new inspirations for many fields, including advertising, political campaigns and TV shows. Another study by (Chen, 2012) claimed that Internet memes might even break into the mainstream culture, and have value in the political ethos. However, though studies have looked at the impact of Internet memes on diverse aspects of the society, their inherent properties can also be explored. A typical Internet meme intrinsically combines text and image. Therefore, this study aims to explore the semantic relations of text-image interactions in Internet memes (in Study 1) and how this information is combined (Study 2).

### Definition of Internet memes

The term “meme” was initially offered by Richard Dawkins (1976). In his book, *The Selfish Gene*, he proposed that the majority of human behaviors are formed in culture, which are not inherited by genes but rather from imitating others, and called these elements memes (according to the Ancient Greek word *mīmēma* ‘something imitated’) (Dawkins, 1976). Thus, memes are non-genetic behaviors or cultural ideas that are passed from person to person, and they can determine how an organism behaves (Davison, 2012). Different from the genetic behaviors, the main characteristics of memes are the speed of transmission and the relative fidelity of form. In general, memes refer to all imitated behaviors or ideas that are transmitted online or offline, such as fashion, language, religion, etc. In addition, Dawkins (1976) summarized three main factors of memes, *fidelity*, *fecundity* and *longevity*, which determine

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

the inherently imitative nature of memes. These factors emphasized that memes sought to consist of a strong core idea and they can be replicated rapidly, also the pattern of the replication should be sustainable. (Dawkins, 1976). However, since the evolution of the Internet technology, memes on the Internet are transmitted and modified faster than offline memes (Davison, 2012), and some new, colloquial meanings are added into the original memes during the fast spread online. Many sources (e.g. Wikipedia, Urban Dictionary) have described the public perception of the term “Internet memes” and there has not been an academically rigorous definition of Internet memes yet. Davison (2012) reapplied Dawkins’ main characters of offline memes to the digital age and originally defined Internet memes as “a piece of culture, typically a joke, which gains influence through online transmission”. Moreover, Börzsei (2013) expanded this definition to “a form of visual entertainment, which can manifest in many different formats, such as a still image, an animated GIF, or even a video”. Based on the definitions above, the current study considers an Internet meme as *a static form of visual entertainment which gains influence through online transmission (i.e. social media or messaging apps), and can be quickly replicated via evolution, adaption or transformations of the original meme vehicle*. Taking Figure 1 as an example, *the beauty and beast* is a classical romantic Disney story, the original moral of the story is *inner beauty is greater than physical beauty*. However, with the change in the way of thinking, nowadays, people endue this story with a new networked meaning, which is “the unfortunate truth”. That is, the reason that a beautiful girl married with an ugly beast is not because of love rather than money. Hence, in this Internet meme, the background image intends to express the major meaning “unfortunate truth”, and the text expresses similar meaning to “unfortunate truth”. Also, people are able to replace the text on a meme with many new creative pieces of text that convey the similar meaning, and as a result of fast spreading by more and more people, a new Internet meme is formed.



Figure 1. An example of Internet memes (<https://www.kontraband.com/post/9ae3ed0f>)

While before the emergence of Internet memes, emoticons (see figure 2 as an example) serve as the previous form of Internet memes were widely used by individuals. Emoticons can enrich the text-based communication through e-mails or computer mediated communication during the early period of the Internet era (Park, Kim & Lee, 2014) and provide individuals' opportunities to express the emotion that they intend to convey. Due to the memetic behavior of emoticons, namely using and viewing, the meaning of the emoticons has been extended after repeated modifications and use. Internet memes, to some extent, can be seen as emoticons with text. Therefore, they also possess the same nature as the emoticons.

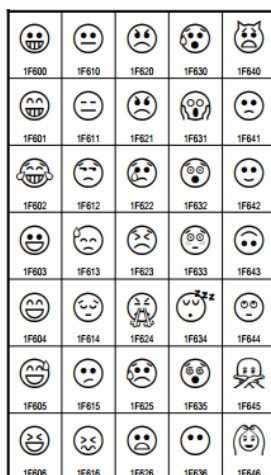


Figure 2. Examples of emoticons (<http://www.unicode.org/charts/PDF/U1F600.pdf>)

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

### **Impact of Internet memes**

Internet memes are becoming ubiquitous in networked environments, and increasing number of people are keen to use this newest type of online inside joke (Miltner, 2014). From this trend, a large number of websites are produced, on which the personalized memes can be generated quickly (ex. <http://knowyourmeme.com/>, a website that provides the possibility to create and download original Internet memes for free). These kinds of websites attract more and more people to pay attention to Internet memes.

Previous studies regarding Internet memes have investigated several aspects of their use. Börzsei (2013) proposed that the invention of the Internet accelerates the Internet memes' propagation. Internet memes have already become one of the most widespread modes of online communication in 2012, and their evolution has led to several new perspectives on society, culture, and technology. Internet memes provide new inspirations and create new ideas in several fields (Miltner, 2014). For instance, novel advertisements often emulate Internet memes, political campaigns involve Internet memes against competitors, and even the traditional media and the popular TV shows adopt Internet memes to capture the *zeitgeist*. Internet memes also function significantly in the society as cultural capital, as well as a kind of inherently unstable cultural forms (Nissenbaum & Shifman, 2017). They also function in several unique ways concerning society, culture and language.

### **Structure of Internet memes**

All online and offline memes are existing in layers. For example, language is a meme which does not only belong to the larger language meme but also include submemes, such as dialects and jargon (Davison, 2012). Similarly, an Internet meme is an "image macro," which contains a set of rules for adding the same text to different images or applying different text to the same image (Davison, 2012). In this situation, Internet memes intrinsically contain the interaction of text and image, where the image in the Internet memes shows the main

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

characters of the inside joke, and the text emphasizes the content that it intends to express. However, the multimodal interaction of the text and image in Internet memes has not been explored and it is therefore of interest and potentially relevant to investigating multimodality in general.

### **Multimodal interaction**

Humans naturally communicate with each other through multiple modes, including text, speech, body language, or a combination of these expressive modes together (Cohn, 2016). This is generally known as a *multimodal interaction*. Especially, after the emergence of web-based communication, the role of visuals has been considered more and more essential to the communication of messages. In addition, owing to the wide use of visuals on Internet, the relationship between the verbal language and visual language has reached a new stage of development. Consequently, the traditional concepts and notions of communication have become out-of-date. (Kong, 2006). On this account, examining the interaction between the verbal and the visual components in an online environment is necessary and relevant for understanding the new concepts of online communication. To be specific, this typical multimodal interaction is a *text-image relation*.

Text-image relation, to a certain extent, can be simply comprehended as where text and image co-occur and both convey the information (Martinec & Salway, 2005). Theories of text-image relations have built on linguistic models of cohesion, which is essentially a non-structural resource for text constructions (Halliday & Hasan, 1976). The cohesion analysis in linguistics has focused on how text organizes the integrated structure within and across the sentences. Similarly, this theory can also be applied to other nonlinguistic elements. Multimodal cohesive relations follow the similar configuration, that is, the linguistic text can be brought into connection with the image somewhere else on the page (Bateman, 2014). For example, the comic strips usually contain both images and text within one panel. The image

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

in each panel provides precise information, meanwhile, the text in the same panel might offer more specific information to repeat the meaning that is conveyed by the image. These two modes create a text-image relation in comics.

Many scholars have studied text-image relations in diverse contexts, including comic strips, newspapers, advertisements, etc. (Phillips, 2000; Royce, 2007; Cohn, 2016). Phillips (2000) examined whether the headlines in advertisements increase people's comprehension of the ads, and showed that by adding headlines of the ads, readers better understood the meaning of a pictorial metaphor because of the additional textual clue. Also, Royce (2007) examined the features of visual and verbal modes co-occurring in the page-based multimodal text in financial magazines. In his study, he found that the visual or the verbal mode in one page mutually complement each other, and produce a single textual relationship which can be referred as *intersemiotic complementarity*. He suggested that there are three intersemiotic features, including the visual-to-verbal interface within the text, the visual-to-visual interface, and the intravisual interface. Cohn (2016) found that words or/and images in comics can guide the meaning, and both modalities follow a particular structure. From this point, he characterized multimodal interactions in terms of their component cognitive structure and integrated a theoretical model of multimodal interaction underlying each component. His model offered a good starting point to discover how the components within a multimodal structure interact with each other.

Despite this work, multimodal interaction specifically in Internet memes has yet to be explored. Therefore, the current study applies theories of multimodality to Internet memes. This study aims to examine different types of text-image relations in Internet memes and how they impact understanding, enjoyment and viewing times. The first study (Study 1) examines text-image relations in Internet memes using a rating test. These scores will be used to verify whether people differentiate various types of text-image relations. Next, based on the results



## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

of Study 1, the second study (Study 2) investigates people's response time, comprehension and enjoyment of text-image relations.

### Study 1

Categorization of text and image relationships has been studied for decades, since Roland Barthes' *Elements of* (1967). He argued that verbal text mainly guides the meanings and perceptions of the content, and he identified two relationships between the verbal and visual modalities: elaboration and extension. He suggested that verbal components are able to elaborate on the meaning of the visual components or extend new information. In response to Barthes's notions, several arguments about systems of text-image relations have been proposed. For example, Martinec & Salway (2005), divide text-image relations into two subsystems, *status* and *logico-semantic*. The status relations of text-image consist of two subcategories, equal and unequal (see Figure 3). The definitions of these two categories are:

*"Images and texts are considered to be unequal in status when one of them modifies the other. Equal status between images and text is further divided into independent and complementary. An image and a text are considered independent and their status equal when they are joined on an equal footing and there are no signs of one modifying the other. When an image and a text are joined equally and modify one another, their status is considered complementary. (p 343)."*

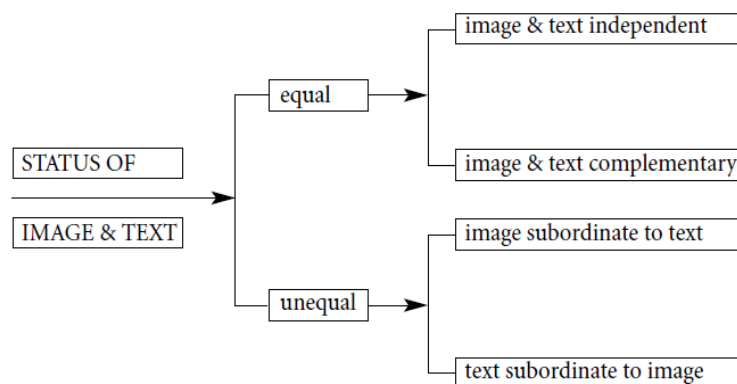


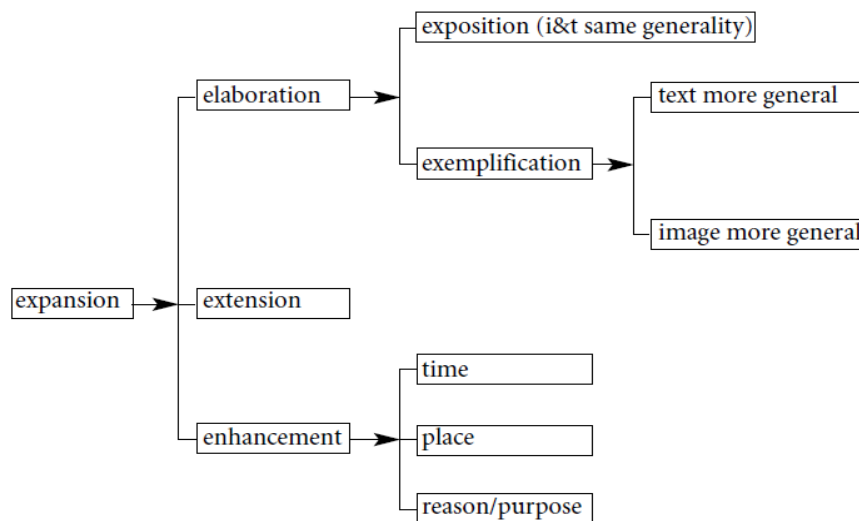
Figure 3. System of image-text relations

It is believed that the unequal status frequently appears in old media when one element has been copied by the other (Martinec & Salway, 2005). However, the current study

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

aims at the Internet memes which are commonly used in new media, in this case, all status relations that will be discussed in this study are treated as equal status relations.

Logico-semantic text-image relations also contain two subcategories: projection and expansion. The projection image-text relation frequently occurs in comic strips where one mode has been re-represented by the other mode (Martinec & Salway, 2005). However, in Internet memes, the image or the text generally adds new meaning to compensate for the other. Therefore, the projection relation will not be discussed in this study. Expansion text-image relations, include three main types: elaboration, extension and enhancement (see Figure 4).



*Figure 4.* System of expansion for image-text relations

In general, elaboration can be understood as the text and the image conveying similar information but with the same or different level of generality (Martinec & Salway, 2005). Figure 5(a) shows an Internet meme with a cartoon character crying and a suppliant facial expression. To some extent, this image expresses asking for a request or pity. At the same time, the word “Please” is in line with the similar meaning of the cartoon character. Altogether, this Internet meme forms an elaboration text-image relation. As for extension text-image relation, it concentrates on new information added by either of the text or the

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

image (Bateman, 2014). As shown in Figure 5(b), “no homework” offers a more explicit request, compared to Figure 5 (a), and adds new but relevant information not provided by the image. Finally, within an enhancement text-image relation, one modality qualifies the other by giving reasons of time, place, or purpose (Bateman, 2014). In Figure 5 (c), the text “I cannot live without the Wi-Fi, please turn the Wi-Fi on” offers a straightforward reason that can explain why the cartoon character is crying (Martinec & Salway, 2005; Bateman, 2014).



Figure 5 (a) Elaboration



Figure 5 (b) Extension



Figure 5 (c) Enhancement

*Figure 5.* The examples of different text-image relations in Internet memes

According to the aforementioned theory, this classification is modeled by linguistic experience and summarized based on the observations of the text-image relations in both new and old media. Yet, these text-image relations have not been applied to Internet memes. The research question of the current study is:

**RQ:** Will people differentiate these three types of logico-semantic text-image relations in Internet memes?

These prior studies have been corpus analyses or observations. However, no experimental studies have examined these types of coherence relations between text-image relations, nor have they been analyzed in Internet memes. As introduced before, the information load in each text-image relation is proposed as being different. Elaboration is predicted as being the easiest one to be noticed because the text and the image in elaboration relations refer to the same participants, processes, and circumstances (Martinec & Salway, 2005). Extension would be the second easiest since even if new information is added, the

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

additional information is related to the meaning provided by the other mode. Finally, the enhancement relation may be more complicated since the temporal, spatial, or causal information takes longer for people to process (Martinec & Salway, 2005). The hypothesis is formulated as follow:

**H1:** Regardless of the condition, the elaboration text-image relation will be easier to recognize than extension text-image relation, and enhancement text-image relation will be the hardest.

### Method

#### Stimuli

24 Internet memes were selected from a self-generating Internet memes website (<https://memegenerator.net/>). Because the main purpose of this study concerns the text-image relations in static images, only still Internet memes were selected. Stimuli were chosen using several selection criteria:

- 1) The image or the text in the Internet memes must *contain a certain meaning*, such as an obviously delightful emotion, angry facial expressions, or meaningful words and sentences.
- 2) To minimize the impacts of the cultural and political background, the selected Internet memes should *include an as simple as possible cultural and political background*.
- 3) Internet memes which contain racial discriminations, policy issues or negative emotions, etc. were excluded.

These selected Internet memes were then manipulated in two ways. Firstly, the logico-semantic text-image relation in each selected Internet meme was identified. Secondly, based on each original Internet meme, two variants were created with new text-image relations. To be specific, this was done by changing the original text in the Internet memes. The manipulated text-image relations of the stimuli were created based on the author's

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

comprehension of the Internet memes and the definitions of each image-text relations (elaboration, extension and enhancement). Based on Bateman (2014), three types of logico-semantic text-image relations are defined as follows:

1. *“Elaboration restates, adds further information at the same level of generality, or exemplifies, in which case either the text is more ‘general’ than the image or vice versa.”*
2. *“Extension adds further, semantically intrinsically unrelated information.”*
3. *“Enhancement offers qualifying information specifically to do with time, place, manner, reason, purpose and other generally ‘circumstantial’ restrictions.”*  
(p196)

In this circumstance, one group of stimuli contained three Internet memes with the same image but different text as well as text-image relations (see figure 5 as an example). In total, 72 stimuli were utilized in this study (as can be seen in Appendix 1).

To control the exposure effect, each participant only viewed the Internet memes with one of the three text-image relation types, and these stimuli were assigned into a Latin square design (Saville & Wood, 1991) as shown in Table 1. In this case, each participant was exposed to different Internet memes and with various text-image relations during the experiment, an example is shown in column 1 of Table 1. There are three text-image relations in total, consequently, 24 Internet memes were divided into eight groups. In order to increase the reliability of the study, the sequence of the groups and the Internet memes in each group were both randomly assigned to each participant.

Table 1.  
*The example of the Latin Square design*

	Column 1	Column 2	Column 3
Internet memes 1	Elaboration	Extension	Enhancement
Internet memes 2	Extension	Enhancement	Elaboration
Internet memes 3	Enhancement	Elaboration	Extension

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

### **Procedure**

In an online survey program (Qualtrics, Provo, Utah, USA), participants were first shown an introduction prior to the actual experiment which included a consent form and some background information of the experiment. Next, a brief instruction showed the participant about how to think about the meaning of the text and the image in Internet memes. Then, four demographic questions were asked, being the age, gender, educational level, and nationality. After that, the participants viewed each of the 24 stimuli with three rating questions with a 7-point bipolar scale, which measured how participants perceived the text-image relations (as shown in Appendix 2). Each rating question was formulated according to the definition by Bateman (2014), such as *“to what extent does the text make the image meaning more specific?”* (*1 = adds no specificity; 7 = makes it much more specific*). Two additional questions with a 7-point bipolar scale were asked to obtain the opinions about the interest and the understanding of the Internet memes.

### **Participants**

In total, 52 participants (20 males, 32 females, mean age: 24.88) from different countries participated in the study via the online questionnaire. The questionnaire was distributed through several social media platforms (i.e. Facebook, WeChat and WhatsApp). Networking and snowball sampling were used for the acquisition of the participants. Most of the participants received a higher degree than Bachelor of Science (92.3%). Chinese participants (38.5%) and Dutch participants (30.8%) occupied the majority of the participants' population.

### **Data analysis**

Before analysis of the data, a screening was performed to exclude incomplete questionnaires from the study. Questionnaires where the response time fell below or above the mean plus/minus  $2.5 \times$  the standard deviation, were defined as outliers which were also excluded from the analysis.

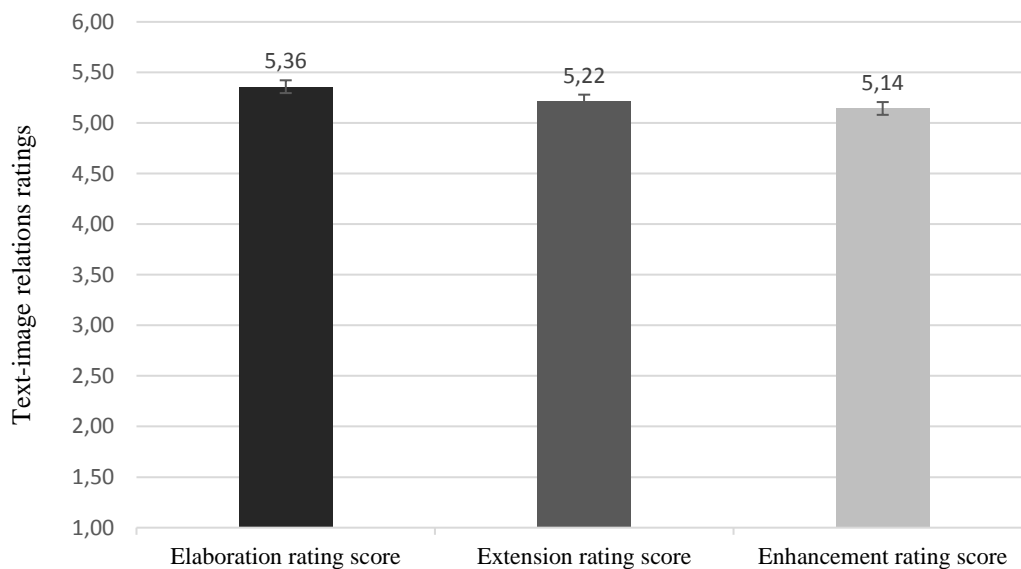
## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

In order to test participants' assessment of text-image relations, a one-way ANOVA was performed to compare the differences among the ratings for each condition with post-hoc comparisons. A Pearson correlation tested for further relationships between the ratings.

### Results

#### Ratings

Figure 6 shows the mean ratings for each of the assessments of the three text-image relations. Participants rated the assessments of three text-image relation differently,  $F(2, 69) = 29.48$ ,  $p < .005$ . Specifically, the people gave higher ratings for elaboration text-image relation and extension text-image relation than the enhancement text-image relation ( $p < .005$ ), but elaboration and extension ratings did not differ ( $p = .26$ ).



*Figure 6.* Ratings on a 1-7 scale for how participants consider the different text-image relations for Internet memes.

One of the purposes of Study 1 is to test whether the intended text-image relations for Internet memes were recognized with their corresponding ratings (i.e., the meme intended as an elaboration relation received highest ratings for the elaboration measurement score, but lower scores for extension and enhancement). Therefore, the average text-image ratings



## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

within one type of Internet memes were compared. Only one type of original Internet memes received the highest ratings for the intended text-image relations. That is, when people viewed an elaboration Internet meme, the elaboration rating ( $M = 4.64$ ,  $SD = .75$ ) was the highest, compared to the ratings of other text-image relations. This match did not occur for the other types of Internet memes.

### Correlation

A Pearson correlation test was performed to explore the relations between ratings of each text-image relation. All three text-image relations ratings highly correlated ( $p < .001$ ) with each other (see Figure 7). This suggests that high ratings of any type of relation co-occurred with high ratings of the other relations.

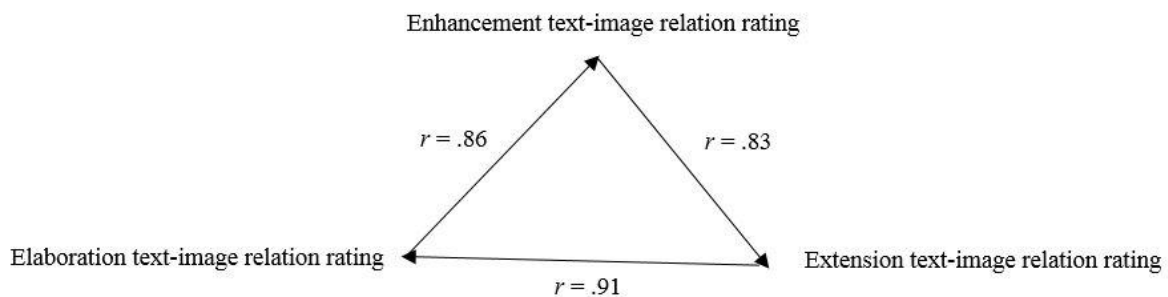


Figure 7. The correlation among the ratings of all types of text-image relations

### Discussion

The primary purpose of Study 1 was to investigate whether people distinguish the various text-image relations by rating their qualities (elaboration, extension and enhancement) for the same Internet memes. First, individuals gave higher elaboration scores than extension and enhancement scores, but the scores for extension and enhancement did not differ, which partially supported Hypothesis 1. This implies that people tend to consider the text-image relations in Internet memes as elaborations, rather than extensions or enhancements. This can be explained by the features of each text-image relation. No new information becomes added in elaboration relation, that is, the information that is given by

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

one mode in elaboration relation is restated by the other mode, which makes easier for people to notice. Conversely, new information and causal information (Martinec & Salway, 2005) do appear in extension relation and enhancement relation, which might cause people not to recognize the appearance of new information, consequently, they did not give higher scores for these two text-image relations.

Second, the rating scores of three types of Internet memes are all highly correlated, which indicates that even though the text-image relations were indeed existing in each Internet memes, various text-image relations are not discrete. This could be interpreted in several ways. There was no second coder verified the text-image relations that were identified and manipulated by the experimenter, which could result in a bias of the manipulation of stimuli. Also, the text-image relations that were discussed in several studies (Kong, 2006; Bateman, 2004; Martinec & Salway, 2005) were retrieved from the website, textbooks or newspapers where contain more complex and larger amount of information than Internet memes. For instance, an example of elaboration text-image relation that was offered by Martinec & Salway (2005) concentrates on a sign of dangerous creatures. In this sign, the image includes a skull with crossed bones, which expresses a dangerous signal in people common sense. Also, the text “Kills by biting prey with jagged teeth” intends to convey the dangerous signal, and further elaborates the meaning of the image. These narrow down the range of meaning in each mode, leading to one appropriate relation. However, in Internet memes, only one or maximal two straightforward sentences are shown to participants, which might not provide sufficient information for participants to assess the text-image relation. Additionally, the image in an Internet meme expresses a more general and broader meaning and there are several interpretations for the meaning of the same image. Therefore, the information that is given by the text and the image might not be concrete for people to fully understand the Internet memes.

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

Some limitations of this study should be noticed. First, Internet memes are popular under diverse cultural environments, which might lead to people from different cultures or societies have various perspectives of the same Internet memes. Ma (2016) proposed that Internet memes have been widely used in China during an online conversation, even though people are from the same background, the senders and the receivers tend to understand the meaning of the Internet memes reflected upon their own experience. The stimuli that were used in this study were retrieved from an English website, which belongs to Western culture, but nearly 40% of participants were Chinese from an Eastern culture. In this situation, the cultural background could be a moderate variable in this study. Next, in this study, only one coder identified the text-image relation in each Internet meme, the reliability could be an issue. This could be improved in Study 2, because all stimuli have been rated by participants. Study 2 stimuli will be selected based on the ratings of each text-image relation in Study 1. Therefore, the selection of stimuli will be controlled better with the participants as additional coders. Last, the duration time of each image was not recorded respectively in the current study, therefore, it is impossible to compare the different response time among three text-image relations. The further study could take the response time of each image into account and provide more aspects to analyze the results of the experiment.

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

### Study 2

Study 1 suggested that people indeed rated text-image relations differently, with most emphasis distinguishing elaboration relation in Internet memes. However, the rating scores of each text-image relation were highly correlated which cannot give a specific direction of how differently each text-image relation influences people. Study 2 sought to better examine these influences. Also, to make the comparison more explicit, these multimodal memes are compared to monomodal memes with just images, which provides the possibility to measure the different influences that are caused by the addition of text.

Comparison between the multimodal images and monomodal images have been studied in terms of several features. As mentioned earlier, Philips (2000) suggested that in an advertisement, adding an explanatory headline offers a clue to the meaning of the advertisement and increases the liking of the metaphor. In addition, adding headlines in ads increases the understanding and the appreciation of them (Ketelaar, van Gisbergen, & Bosman, 2004). Besides advertisements, many microblog posts using images were found to make the posts more attractive and beautiful, compared to non-image tweets (Chen, Lu, Kan, & Cui, 2013). Moreover, an experimental study suggested that adding modality of text would motivate students to read and increase the comprehension of reading. In this study, children rated significantly higher scores of perceived understanding for graphic novels (both image and text are equally important for narrative) compared to traditional novels (contains fewer illustrations per chapter). Also, the graphic novels received the highest score concerning enjoyment, compared to heavily-illustrated novels (includes at least one illustration for each two pages) and traditional novels (Jennings, Rule & Vander Zanden, 2014). However, the difference of enjoyment among text-image relations have not been compared yet.

Overall, these studies provide a general idea about how types of modality have an effect on understanding and enjoyment. Furthermore, Study 1 showed that different text-

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

image relations are harder to distinguish. In elaboration relation, one mode restates the information that is provided by the other mode, which is assumed easier to be understood. However, the new or the casual information is shown in extension and enhancement relations might be more difficult to be perceived. Therefore, the hypotheses are formulated below:

**H2:** People will understand multimodal images (Internet memes) better than monomodal images, and the text-image relations will also influence on the understanding.

**H3:** People will consider that multimodal images (Internet memes) are more enjoyable compared to monomodal images, but the type of text-image relation will not influence on the enjoyment.

In the early studies, self-paced viewing paradigms were widely used to measure the comprehension of several aspects of linguistics. In a self-paced viewing task, people are required to move gradually through the materials at their own pace, and a timing device automatically records the comprehension time during people viewing the materials (Brisard, Frisson, & Sandra, 2001). Haviland & Clark (1974) used Given-New Strategy to measure the understanding of an integrated sentence of two sub-sentences. The comprehension time was measured by pressing the 'Next Button' when readers felt that they understood the sentences. The comprehension time in this study indicated how direct that the information was given in presupposing existence. It was found that less viewing time appeared when the original sentences provided direct information. Moreover, this approach has also been successfully used to measure the comprehension in visual narratives (Cohn, 2014). In his study, self-paced viewing paradigms were applied to distinguish the different categories (e.g. Establishers, Peaks) in visual narratives grammar (VNG). The viewing time provided evidence to compare the difference of each category at each position. The findings of his study implied that canonical narrative patterns take people shorter time to view each panel of a strip.

Altogether, the viewing time that is measured by self-paced viewing paradigms

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

reflects the difficulty of the information for people to comprehend or the consistency of visual narratives for people to process, which offers a systematic approach for the current study. This study will also utilize self-paced viewing paradigms to compare the difference between viewing monomodal images and multimodal images. Brisard, Frisson, & Sandra (2001) argued that when readers do not have familiar sentences that they can use to interpret the metaphor, they spend longer time on processing metaphorical expressions than literal ones. Similarly, a new text-image relation in multimodal images is created by adding additional text into the monomodal images, which people do not prepare for it. Therefore, it is assumed that processing the text-image relation in multimodal memes takes people longer time. Additionally, the results of Study 1 provided evidence for the difficulty of each text-image relation for people to distinguish. Thus, it is predicted that the response time would differ among different text-image relations, and the hypotheses are formulated below:

**H4:** People will spend longer time viewing multimodal images (Internet memes) than monomodal images.

**H5:** The type of text-image relation (elaboration, extension, and enhancement) will influence the response time on each image in different ways.

### Method

#### Stimuli

Study 2 used the same Internet memes as in Study 1, but some of the images were excluded from this study according to their ratings of Study 1. Text-image relations were chosen only if Internet memes of one type had significantly higher ratings than the other types. According to this inclusion criterion, a total of 15 Internet memes were selected from Study 1. Then, in order to compare monomodal memes (an Internet meme without text) and multimodal memes (with text), the original images without text were also used as stimuli. In this case, a total of 60 images were included in Study 2 (see Appendix 3).

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

### **Procedure**

A pre-test was performed to ensure that no common mistakes or misunderstandings regarding the questions arose from the participants in this survey. Common misunderstandings that were found based on feedback from the participants ( $n = 5$ ) were modified accordingly. Afterwards the final study was conducted using the online survey tool (Qualtrics), similar to study 1.

An introduction first explained the aim of the study and a consent form was shown to participants, followed by six demographic questions asking about their age, gender, nationality, educational level, language, and English proficiency. Next, a notification was displayed on the screen instructing participants how to successfully complete the survey. Then, stimuli were displayed on the screen, first the monomodal image and then the multimodal Internet meme of the same type. Three rating questions were asked to measure the understanding, the amusement and the enjoyment of the stimuli (Appendix 4) on the next page of the survey. The viewing time of each stimulus was recorded automatically by Qualtrics and each pair of stimuli were randomly assigned to participants.

### **Participants**

146 individuals [49 males, 97 females, mean age: 24.41 ( $SD = 3.48$ )] participated in the study. This survey was distributed online through several social media platforms (i.e. Facebook and WeChat). The participants were from 19 different countries, but most of them were Chinese (56.8%), and the second largest proportion of the participants was Dutch (27.4%). Most of the participants had higher than a Bachelor educational level (89.1%); 52.1% spoke Chinese most often, followed by Dutch (25.4%) native speakers, and 15.8% of the participants were English proficiencies.

### **Data analysis**

Outliers were removed from the dataset if their response time fell below or above the mean

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

$\pm 2.5 \times$  the standard deviation. This amounted to 2.74% discarded trials. A one-way repeated measure ANOVA was used to analyze the response times and ratings. Additional correlations were performed for further investigation of (potential) relationships between Study 1 and Study 2.

### Results

#### Response time

A one-way repeated-measures ANOVA revealed that the types of Internet memes had a significant effect on participants' response time [ $F(2.74, 388.81) = 21.02, p < .005$ ]. That is, people differed significantly in their viewing times of various types of Internet memes. Post-hoc tests showed that a significantly shorter viewing time appeared when people viewed monomodal memes than extension memes and enhancement memes ( $p < .005$ ). Also, a significant longer time occurred for viewing enhancement memes than elaboration memes ( $p < .005$ ) and extension memes than elaboration memes ( $p = .026$ ). Finally, no difference occurred on response time between viewing elaboration memes and monomodal memes, suggesting that people took similar amount of time to view these two kinds of Internet memes ( $p = .70$ ).

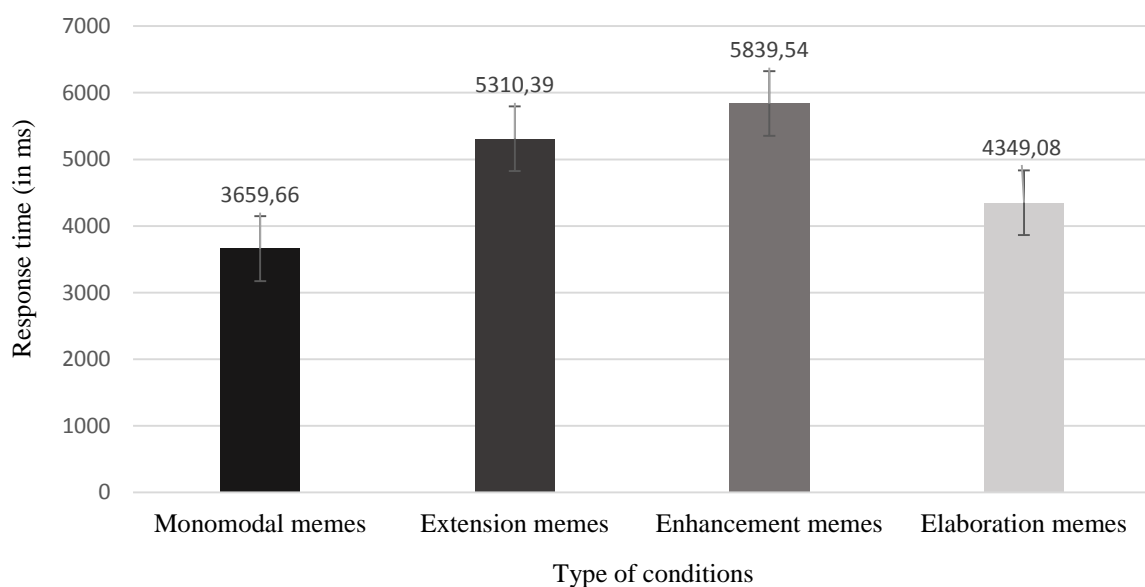


Figure 8. Response time in different text-image relations conditions



## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

### Understanding

A significant difference of the understanding ratings among four conditions as determined by a one-way repeated-measures ANOVA [ $F(2.42, 346.04) = 105.39, p < .005$ ], indicating the type of Internet memes had a significant effect on people's understanding. Additionally, a post-hoc test revealed that a significant lower ratings of understanding occurred for the monomodal memes than other three types of Internet memes ( $p < .005$ ). Also, people also gave significantly higher understanding rating for extension memes than elaboration memes ( $p = .02$ ). However, no significant difference was found for the ratings of understanding between extension and enhancement ( $p = .34$ ), enhancement and elaboration ( $p = .53$ ), can be seen in Figure 9.

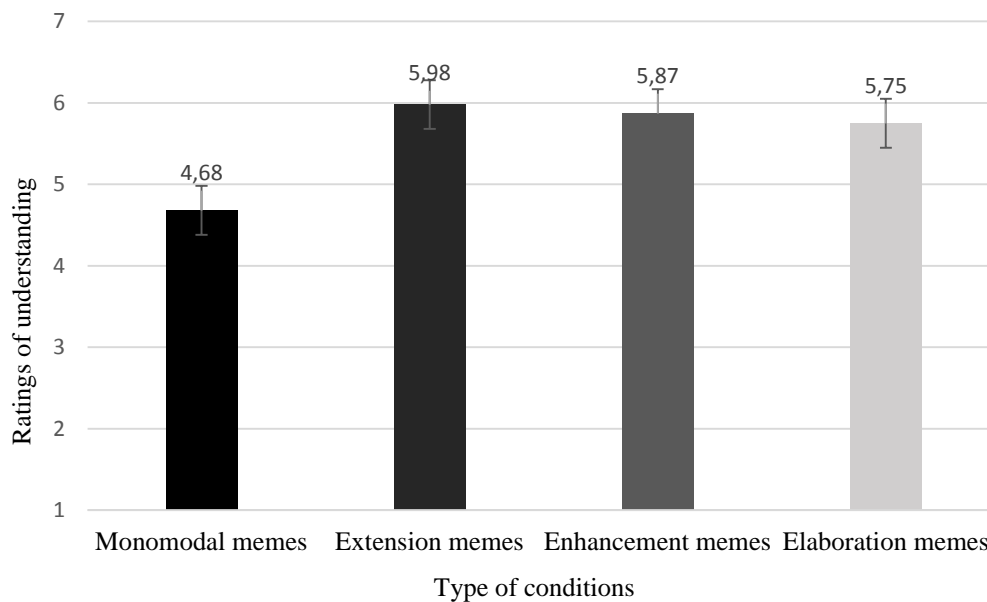


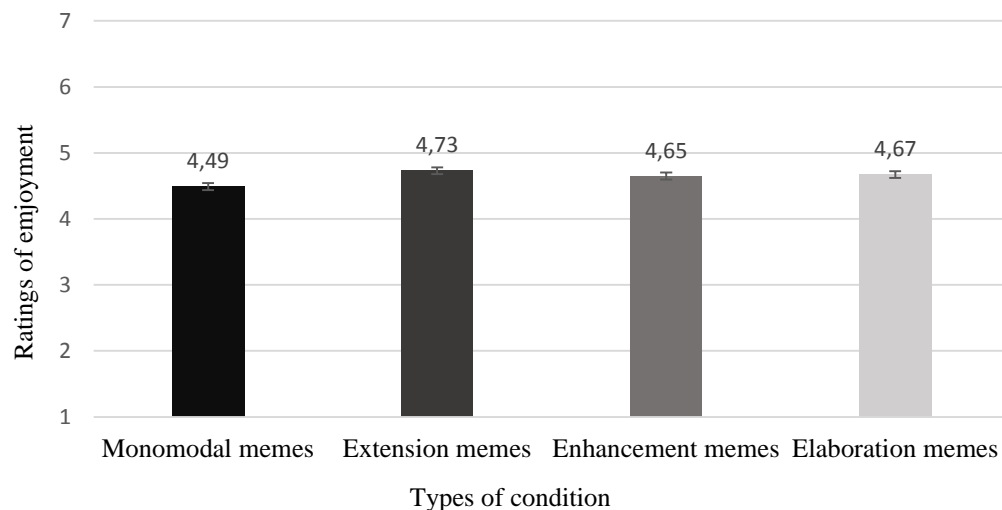
Figure 9. Ratings on a 1-7 scale for how participants understand different types of Internet memes

### Amusement and enjoyment

The mean score of each condition for amusement and enjoyment ratings looked similar. Therefore, the reliability of amusement ratings and enjoyment ratings was tested using a Cronbach's alpha to see if they were measuring the same content. An excellent

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

reliability ( $\alpha = .95$ ) showed this was the case, and thus these two scales were collapsed into one measurement of “enjoyment.” A one-way repeated measure ANOVA on enjoyment ratings showed that the type of Internet memes had a significant effect on enjoyment ratings [ $F(3, 429) = 2.95, p = .03$ ]. This indicates that people rated enjoyment differently for various types of Internet memes. However, the post-hoc test revealed that this effect only appeared between the monomodal memes condition and extension memes condition ( $p = .05$ ). This indicates that people gave significantly higher ratings for enjoyment when people viewed extension memes than monomodal memes, see Figure 10.



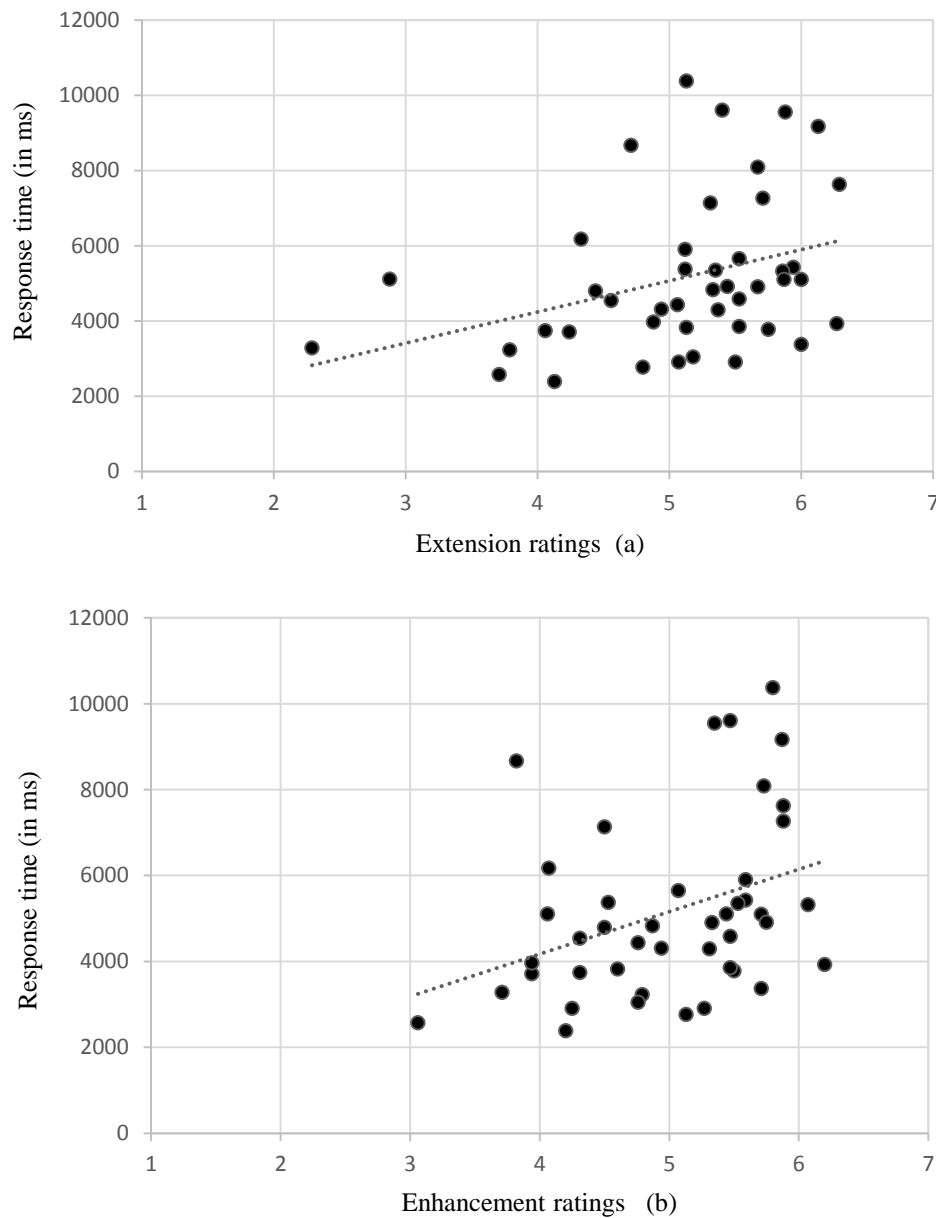
*Figure 10.* Ratings on a 1-7 scale for how participants enjoy viewing different types of Internet memes

### **The correlation between text-image relations ratings and the response time**

The rating scores of each text-image relation from Study 1 were correlated with rating scores and response times from Study 2. A moderate, positive correlations were observed between the ratings of extension and response times, which was statistically significant ( $r = .40, p = .012$ ), see Figure 11(a). Ratings of enhancement also significantly correlated with response times ( $r = .30, p = .05$ ), as shown in Figure 11(b). That is, for multimodal interactions, if people gave higher ratings of extension or enhancement text-image relations for an image,

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

they spent longer viewing the images. No correlation was found from the rating of elaboration and the response time.



(a). Correlation between extension rating and response time

(b). Correlation between enhancement rating and response time

*Figure 11.* The correlation between text-image ratings and the response time

## Discussion

Study 2 explored how modality types and text-image relations influence response time, and ratings of understanding and enjoyment. The main finding of this study is that the

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

types of Internet memes have an effect on the people's response time. More specifically, people spent significantly longer viewing enhancement memes and extension memes, than monomodal images (the exactly same image but without text) which occurred before the multimodal ones. This also supported partially Hypothesis 4. The explanation of this finding could be that viewing multimodal images requires people to process the text, the image as well as the relation between them, which would take longer than simply process the single image. However, the difference of viewing time did not appeared between the elaboration memes and monomodal memes. This could be explained by the results of Study 1. In Study 1, it showed that elaboration memes were the easiest one to be distinguished by participants, in consequence, the additional information that were provided by elaboration memes was assumed to be easier for people to understand. Therefore, even though, the additional information was indeed added into elaboration memes, but the viewing time between these two types of Internet memes did not differ.

Multimodal images were also easier to understand than monomodal images, which also supported Hypothesis 2. This is in line with research showing that adding explanatory headlines in advertisement increases people's comprehension (Philips, 2000; Ketelaar, van Gisbergen, & Bosman, 2004). Analogously, in Internet memes, the text offers a clue that specifies the meaning of the images, which assists people to understand the meaning of the entire Internet memes. Normally, Internet memes use simple or easy images to express their meanings, and thus the same image could have several interpretations. For instance, Figure 5 shows a crying cartoon character, but this face can be interpreted several ways. Adding text restricts the range of meanings, such as Figure 5(a) is of begging, Figure 5(b) is asking for "no homework", and Figure 5(c) is asking for usage of Wi-Fi. Additionally, Jennings, Rule & Vander Zanden, (2014) suggested that children process the information more deeply when they read graphic novels, which leads to a better performance and higher ratings of

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

comprehension than traditional novels. A plausible interpretation of the finding might be that the added text modality and manipulated text-image relation in Internet memes motive people to process the information more deeply and make more effort to comprehend the multimodal images. This also results in a better understanding.

Enjoyment rating scores also showed that people considered extension memes as more enjoyable than monomodal images, and Hypothesis 3 was partially supported. This could be interpreted by processing the multimodal interaction makes people feel engaged in viewing the images, and thus they appreciate viewing multimodal images rather than monomodal images. Another explanation might be that the enjoyment is correlated to the difficulty of understanding. This is in line with Ketelaar, van Gisbergen, & Bosman (2004), who suggested that the ads which contain headlines are easier for people to understand than the ads without any headlines. The headlines that are shown in ads can steer people's comprehension in a certain direction which reduces the difficulty of understanding. When people feel easier to understand, they would feel more enjoyable when they read. The results of understanding scores showed that monomodal images are more difficult for people to understand, and thus people think they are less enjoyable than multimodal images.

Furthermore, response times differ among various text-image relations. The shortest viewing times occurred on elaboration Internet memes, followed by extension images and enhancement images took people the longest to response. This provided evidence to support Hypothesis 5. This finding is also in consistent with the results of Study1, that is, if the text-image relation in the Internet memes were easier to be noticed, people tended to spend less time viewing these kind of Internet memes. Another plausible explanation might be that the amount of information included in each text-image relation varies in difficulty for people to perceive. According to Bateman (2014), elaboration relation adds the least amount of information, but more information is added into extension and enhancement relation. The

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

small amount of information leads to easier processing effort, which results in less response time.

### **General Discussion**

These two studies investigated the semantic relations of text-image in Internet memes and the various effects of each text-image relation on people's response time, ratings of understanding and enjoyment. Study 1 examines whether people perceive the text-image relations in Internet memes differently by giving the rating scores. Based on the ratings in Study 1, the Internet memes which contain more apparent text-image relation were selected for further exploration. Study 2 investigated the influences which are caused by different types of text-image relations and modality types.

First, these results suggest that participants are able to distinguish various text-image relations in Internet memes, especially they would perform better when they are exposed to the elaboration relation. Additionally, the shorter response time occurs when the text-image relation is easier to be noticed, verse vice. The types of text-image relation indeed influence people's response time of Internet memes. This is also in consistent with discoveries in several studies based on self-paced viewing paradigms. The more difficult information that people view, the longer time that people spend on understanding. Besides, a more explicit comparison is performed in Study 2 which compare the influences between different modality types. The results reveal that the longer response times occurred when people viewed multimodal images, despite participants have already viewed the exactly same image, suggesting that text-image relations motive people to deeply process the multimodal images rather than the monomodal images. Moreover, participants show more appreciations and a better understanding of multimodal images than monomodal images, which is in line with Jennings, Rule & Vander Zanden (2014).

This study is the first attempt to discover the text-image relations in Internet memes. Even though only elaboration memes received the highest rating for its category, which provides evidence that existing categories of text-image relations which were put forward

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

based on observations or corpus analysis (Bateman, 2004; Martinec & Salway, 2005; Kong, 2006) can also be applied to Internet memes. However, the ratings of three text-image relations are highly correlated, which indicates that more concrete categories of text-image relations in Internet memes or a small size of online media are also worthy to be explored. The findings of this study also provide a general understanding of multimodal processing in Internet memes. Adding text modality will assist people to comprehend and increases the enjoyment of the entire content. This is also useful in several fields, such as improving the textbooks, designing a more attractive advertisement. However, the different understanding and enjoyment did not appear among various text-image relations, therefore, the further study could explore the effect of each text-image relation more specifically.

There are some limitations should also be noticed in this article. First, the understanding and enjoyment of Internet memes were measured in isolation in. This means, participants only perceived the Internet memes on their own, rather than in an informative context (e.g. an online conversation). However, as stated in the introduction, Internet memes are widely used in several fields, such as advertisements or political campaigns (Miltner, 2014) where the contents are more dominant rather than the meaning of Internet memes. Also, the previous studies suggested that the usage of emoticons would increase the understanding in instant messages (Park, Kim, & Lee, 2014). Nowadays, Internet memes are also widely appeared in online conversation. In this situation, a new multimodal interaction arises where the Internet memes are treated as the image and the text mode provides the dominant meaningful information (Cohn, 2016). Therefore, it would be interesting to discover whether the text-image relations in Internet memes will change when the visual mode is dominant to the verbal. It is suggested that further research could be conducted in a more practical situation, for instance, in a computer-mediated conversation. Also, whether Internet memes in an online conversation will increase the understanding and enjoyment is also worthy to



## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

discover. Second, since the Internet memes have a significant effect on the culture and policy (Chen, 2012) and they are more involved in individuals' online life (Nissenbaum & Shifman, 2017). Therefore, the cultural background of the participants could be a significant moderate variable of the understanding of Internet memes. In this case, the future study could be conducted based on one cultural background and explore whether the same results will be shown when the cultural background is controlled.

In conclusion, the different influences of text-image relations in Internet memes are investigated in this research. The collected data provided evidence for the fact that elaboration text-image relation is the easiest one to be noticed and the response time, ratings of understanding and enjoyment differ between types of modality. Altogether, this research offers the initial insight into the text-image relations in Internet memes, which is beneficial for people to understand the Internet memes better in the digital environment. Additionally, since Internet memes are widely used in advertisements and other fields, understanding the text-image relation in Internet memes is also useful for designing better ads or making more interesting online conversations. Besides, Internet memes also serve as a kind of cultural form, therefore, gaining insight into Internet memes also helps people to understand the online even offline culture and the society.

### **Acknowledgment**

I would first like to thank my supervisor Dr. N.T. Cohn of Humanity School at Tilburg University. He is always welcome to my questions and answers all my question patiently. Also, he provided me many suggestions regarding the analysis of the data and the structure of writing. Besides, I am grateful that my friends and all participants were willing to participate this study. Without their participations, I cannot complete this work in time.

# THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

## References:

- Barthes, R. (1977). *Elements of semiology*. London: Macmillan.
- Bateman, J. (2014). *Text and image: A critical introduction to the visual/verbal divide*. London: Routledge.
- Börzsei, L. K. (2013). Makes a meme instead: A concise history of internet memes. *New Media Studies Magazine*, 7, 152-189.
- Brisard, F., Frisson, S., & Sandra, D. (2001). Processing unfamiliar metaphors in a self-paced reading task. *Metaphor and Symbol*, 16 (1-2), 87-108.
- Chen, C. (2012). The creation and meaning of internet memes in 4chan: Popular internet culture in the age of online digital reproduction. *Habitus*, (3), 6-19.
- Chen, T., Lu, D., Kan, M., & Cui, P. (2013). Understanding and classifying image tweets. *Proceedings of the 21st ACM international conference on Multimedia - MM 13*. doi:10.1145/2502081.2502203
- Cohn, N. (2014). You're a good structure, Charlie Brown: The distribution of narrative categories in comic strips. *Cognitive Science*, 38(7), 1317-1359. doi:10.1111/cogs.12116
- Cohn, N. (2016). A multimodal parallel architecture: A cognitive framework for multimodal interactions. *Cognition*, 146, 304-323.
- Davison, P. (2012). The language of internet memes. In: Mandiberg M (ed.) *The Social Media Reader*. New York: New York University Press, 120-136.
- Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in cognitive sciences*, 3(11), 419-429.
- Halliday, M. A.K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of verbal learning and verbal behavior*, 13(5), 512-521.
- Jennings, K. A., Rule, A. C., & Vander Zanden, S. M. (2014). Fifth Graders' Enjoyment,

# THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

- Interest, and Comprehension of Graphic Novels Compared to Heavily-Illustrated and Traditional Novels. *International Electronic Journal of Elementary Education*, 6(2), 257.
- Ketelaar, P. E., Van Gisbergen, M. S., & Bosman, J. A. M. (2004). Open and closed advertisements: Moderating effects of comprehension on appreciation. *Neijens, PC; Hess, C.; Putte, SJHM van den; Smit, EG (ed.), Content and media factors in advertising*, 51-62.
- Knobel, M., & Lankshear, C. (2005, November). Memes and affinities: Cultural replication and literacy education. *Annual NRC, Miami, November*, 30.
- Kong, K. C. (2006). A taxonomy of the discourse relations between words and visuals. *Information Design Journal*, 14(3), 207-230.
- Ma, X. (2016, July). From Internet Memes to Emoticon Engineering: Insights from the Baozou Comic Phenomenon in China. In *International Conference on Human-Computer Interaction* (pp. 15-27). Springer International Publishing.
- Martinec, R., & Salway, A. (2005). A system for image–text relations in new (and old) media. *Visual communication*, 4(3), 337-371.
- Miltner, K. M. (2014). “There’s no place for lulz on LOLCats”: The role of genre, gender, and group identity in the interpretation and enjoyment of an Internet meme. *First Monday*, 19(8). doi:10.5210/fm.v19i8.5391
- Nissenbaum, A., & Shifman, L. (2015). Internet memes as contested cultural capital: The case of 4chan’s /b/ board. *New Media & Society*, 19(4), 483-501. doi:10.1177/1461444815609313
- Park, T. W., Kim, S. J., & Lee, G. (2014, June). A Study of Emoticon Use in Instant Messaging from Smartphone. In *International Conference on Human-Computer Interaction* (pp. 155-165). Springer, Cham.

## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

Phillips, B. J. (2000). The impact of verbal anchoring on consumer response to image ads.

*Journal of advertising*, 29(1), 15-24.

Royce, T. D. (2007). Intersemiotic complementarity: A framework for multimodal discourse













analysis. In T. D. Royce & W. L. Bowcher (Eds.), *New directions in the analysis of multimodal discourse* (pp. 63).

Saville, D. J., & Wood, G. R. (1991). Latin square design. In *Statistical Methods: The*

*Geometric Approach* (pp. 340-353). Springer New York.

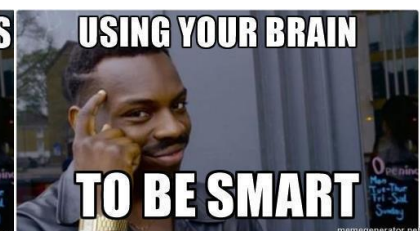
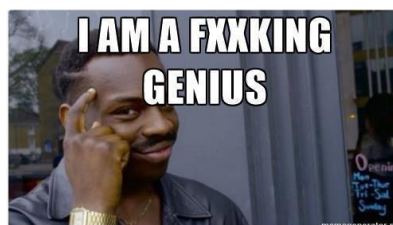
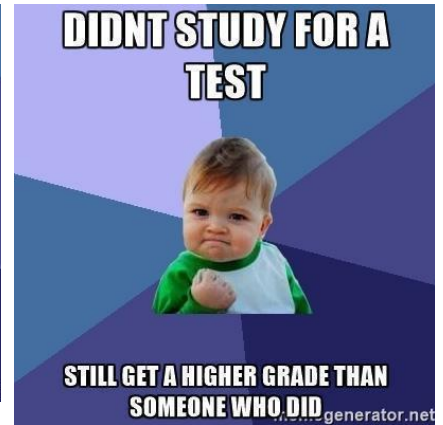
## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

## Appendix 1

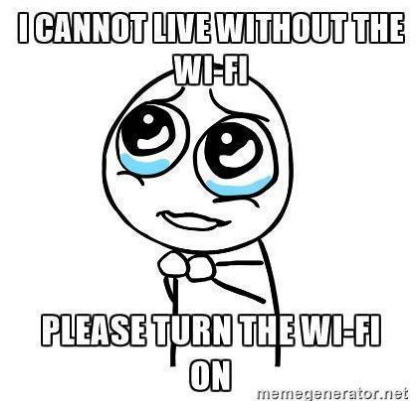
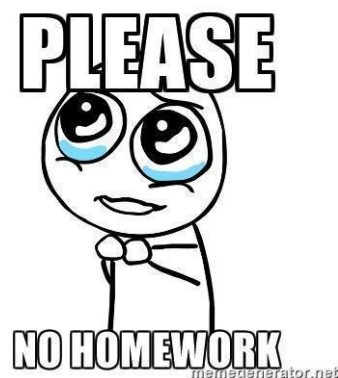
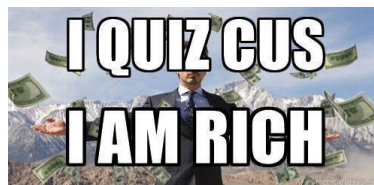
Elaboration	Extension	Enhancement
		
		
		
		



## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

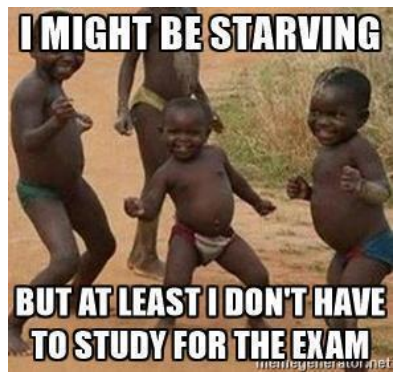


## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES



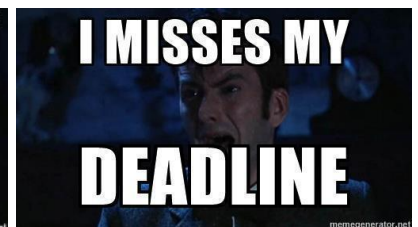
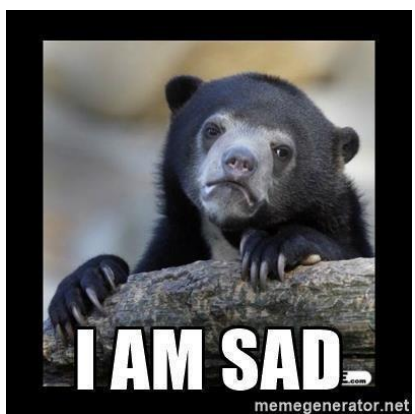
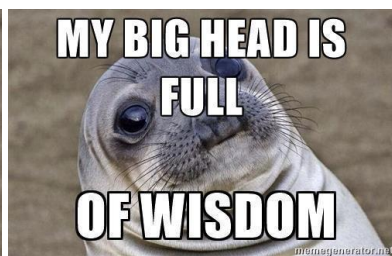
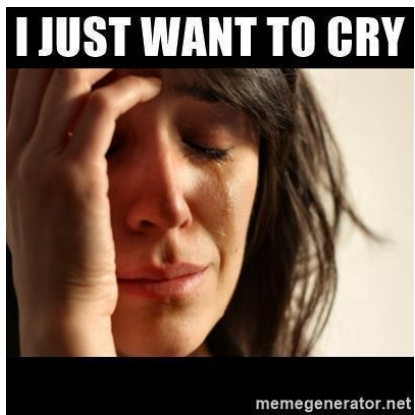


## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

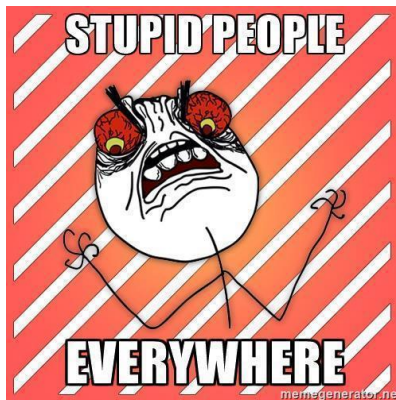




## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES



## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES



## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

## Appendix 2

To what extent does the text make the image meaning more specific?

adds no specificity

☐ ☐ ☐ ☐ ☐ ☐ ☐

makes it much more specific

To what extent does the text indicate meaning that is not expressed by the image?

adds no meaning

☐ ☐ ☐ ☐ ☐ ☐ ☐

adds new meaning

To what extent does the text clarify the reason/intent of the image?

doesn't clarify at all

☐ ☐ ☐ ☐ ☐ ☐ ☐

totally clarifies

To what extent do you think this meme is interesting?

Not interesting at all

☐ ☐ ☐ ☐ ☐ ☐ ☐

Very interesting

To what extent do you understand this meme?

Not understand at all


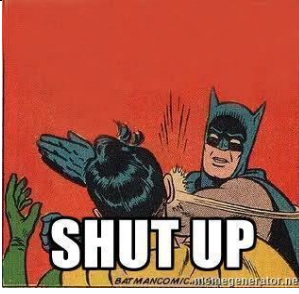









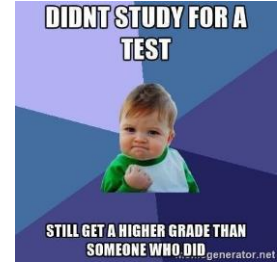





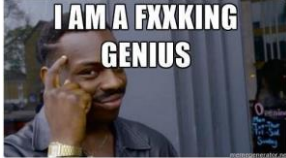


☐ ☐ ☐ ☐ ☐ ☐ ☐

Fully understand



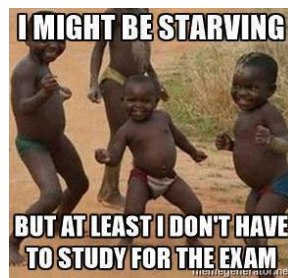
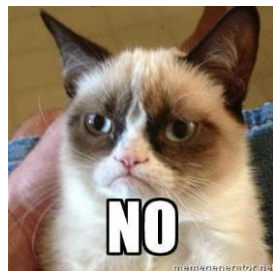
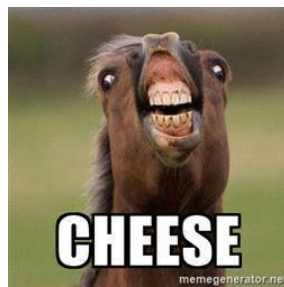
## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

## Appendix 3

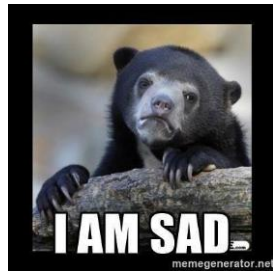
Original	Elaboration	Extension	Enhancement
			
			
			
			
			



## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES



## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES



## THE INFLUENCE OF TEXT-IMAGE RELATIONS IN INTERNET MEMES

**Appendix 4**

To what extent do you understand this image?

not at all



fully understand

To what extent do you think this image is amusing?

not at all



very amusing

To what extent do you enjoy viewing this image?

not at all



very enjoyable