Women Are From 😊, Men Are From 😎

- A study on the sentiment and gender of Emoji

Mikkel Werling(mikkel.werling@post.au.dk)

School of Communication and Cognition, University of Aarhus,

Jens Chr. Skous Vej 2, 8000 Aarhus, Denmark

Carl-Magnus Søgaard Christiansen(carl-magnus.soegaard.christiansen@post.au.dk)

School of Communication and Cognition, University of Aarhus,

Jens Chr. Skous Vej 2, 8000 Aarhus, Denmark

Abstract:

The invention of emoji has created a new communicative tool for emotional expressiveness in computer mediated communication (CMC). Even with our relatively limited understanding of the properties of the emoji, they have been shown to be insightful items in fields such as psychology, linguistics and computer science. This study tries to increase our knowledge about the attributes of emoji. This is done with two methodological approaches: an experimental study on how emoji modulate the sentiment of a sentence, and a more exploratory study, investigating the interpreted gender of 20 emoji.

The sample of the study is composed by 142 participants which were gathered online (age: M = 23.93, SD = 6.48; 52% female). The participants of the study rated sentences in three different conditions: positive, negative or completely without emoji. The results show a support for the hypothesis, that emoji significantly modulate the sentiment of a sentence (p = < .001, R^2 = .26.).

Furthermore, participants were asked to interpret the gender of the emoji. This data was explored, and the emoji were mapped in accordance to their interpreted gender. From an exploratory approach, the interpreted gender seem to be in accordance to stereotypes on gender differences concerning emotion. More affectionate emoji was deemed more feminine and more prideful emoji was deemed more masculine.

Keywords: Emoji, CMC, Gender, Sentiment, Emotion

Introduction (MW, CC)

In the course of human history, language has been ever evolving (Christiansen & Kirby, 2003). As an essential element to human communication, it plays a key role in sharing ideas and doing so in a manner that ones interlocutors can understand. But this is not always possible - we misinterpret what others are trying to share with us. In physical conversations, facial cues help us interpret the intended meaning of an utterance (Niedenthal, 2007). A vast array of modalities play a role in interpreting communication such as lip movement, gaze and facial expression (Haxby, Hoffman, & Gobbini, 2002).

With the emergence of the internet, and non-facial communication, the missing facial cues could result in more frequent misunderstandings when using communication forms such as CMC. This was the concern that led to invention of the emoji by Shigetaka Kurita - vitalizing non-facial communication with a new modality (Schenker, 2016). This is also in line with some studies indicating that emoji have a similar function in CMC as facial expressions have in physical conversations (Derks, Bos, & von Grumbkow, 2008). This might be the reason why the little icons are so popular, with 92% of the young people (13-18) and 77% of the older people (56-64) using emoji (Messenger, 2017).

Although emoji are getting more and more frequent in CMC (Messenger, 2017), the investigation on the properties and characteristics of emoji is still a very new field with much left to discover. The implications, however, could prove pertinent to a wide array of faculties.

Existing implications and future possibilities of emoji (MW, CC)

First of all, emoji have been an area of study in linguistics. Emoji are changing the way we communicate (Barbieri, Ronzano, & Saggion, 2016). In the Facebook Messenger Database, consisting of 1.3 billion people, 53% have sent a message only containing emoji (Messenger, 2017). Studies also reveal that emoji sometimes act as a replacement for words in language, and other studies are pointing to emoji as a new language of visual information. (Barbieri et al., 2016) (Hamza Alshenqeeti, 2016). Comparisons have been made between hieroglyphic and other ancient communication forms, and the emerging new language of emoji (McIntyre, 2016). Another resemblance to language, is that emoji-usage differ between different cultures (Lu et al., 2016).

From a Whorfian point of view, this creates some interesting implications. In essence, Whorf's theory states, that if two people speak different languages, they will think and act differently (Whorf, 1956). Although this strict causal link between language and thought has been widely abandoned in cognitive science, the question concerning whether language shapes thought has been widely debated in the community (Boroditsky, Schmidt, & Phillips, 2003). If emoji is found to be analogous to language, whorfian ideas could also apply to emoji. This could open the debate on whether or not emoji are changing not only the way we communicate, but also the way we think.

This also pertains to the long interest of the difference in discourse between men and women. Newman, Groom, Handelman, & Pennebaker (2008) found a difference in the vocabulary usage between men and women - women using words in relation to psychological and social processes more frequently than men. Men on the other hand, used words pertaining to properties and impersonal topics. If emoji is similar to natural language, we could expect gender difference in emoji usage, just as was found in use of natural language. In fact, there is a clear difference between how the genders communicate digitally. The change of CMC between

groups consisting only of males, only of females and of mixed-gender is substantial. This is especially the case for males, as they start using emoticons much more frequently when communicating in groups of mixed gender (Wolf, 2000).

To delve further into the possible gender difference in the perception of emoji, one could turn one's attention to that of research on the perception of faces. (Dimberg & Lundquist, 1990) studied how facial muscles react when shown stimuli in the form of facial expressions. The participants were shown happy or angry faces, and their muscle response was measured using facial electromyographic (EMG) activity. The theory behind this is observations that EMG could function as a measurement of emotional response. When presented with emotional stimuli, the hedonic value of the stimuli corresponded with muscular activity in the corrugator and zygomatic muscle regions. The corrugator muscle region is used in making a frown, and the zygomatic muscle region is used in forming a smile. Dimberg & Lundquist (1990) found a correspondence between being presented with an angry face and the activation in participants corrugator muscle region, and being presented with a smiling face and the activation of the zygomatic muscle region. In addition, the activation was more profound in female participants than in male participants, indicating that females are more receptive in terms of facial stimuli (Dimberg & Lundquist, 1990).

As the emoji studied in this paper only consists of faces, the emoji included could be viewed as a facial stimuli. Following the implications of the research of Dimberg & Lundquist (1990), this could indicate an increased sensitivity to emoji valence in females.

Secondly, emoji has also been the area of study in psychology. Emoji have been used in examining the relation between emoji-usage and the Big-Five personality traits (Marengo, Giannotta, & Settanni, 2017). Out of 91 examined emoji, 36 of them related significantly with three Big-Five personality traits - agreeableness, emotional stability and extraversion (Marengo et al., 2017). With increased knowledge on the properties of emoji, one might make more predictions on what the use of emoji tells about the user.

Emoji have been proven to be an important factor in predicting information about the sender of messages. In a study by Yoshimura, Rothfus, Liu, & Zheng (2017), students from the Pepperdine University were presented with screenshots of different CMC, and were asked to guess the gender of the sender of the message. They found that emoji-usage was the third most common deciding factor for the participant's judgement following vocabulary and message length (Yoshimura et al., 2017). Building on the idea of recipients gathering information about the sender by looking at emoji-usage, a study has shown an interaction effect concerning jealousy. The genders differ in jealousy according to the emoji-usage; men being more jealous when a winking emoji is presented (😉), and women being more jealous when no emoji was presented (Hudson et al., 2015).

Thirdly, emoji are surfacing in the area of data mining and textual analysis. Emoji has taken center stage in CMC language - especially for the sentiment of texts (Barbieri et al., 2016). Emoji has the potential to increase the precision of textual analysis using computers. It is not only because of the previously stated argument, that emoji hold a vast amount of information. It is also because of the potential to create a model, that would work across languages. One of the first attempts of making such a model was done by Novak et. al in 2015. The study was done with emoji being viewed in relation to the sentiment of the text they accompanied on Twitter. The data consisted of 1.4 million tweets, from 13 different European languages, which

was analyzed by 83 human annotators. As a result, Novak et. al made a Emoji Sentiment Lexicon, designed to be used as a sentiment scoring system across European languages. The study found, that the general sentiment of tweets with emoji was perceived as higher than those without emoji. Furthermore, they found the human annotators agreement on the sentiment of tweets to be significantly higher when emoji were included (Kralj Novak, Smailović, Sluban, & Mozetič, 2015). One could suspect, that even more sophisticated and precise models could be made, with an increased knowledge of how emoji influence the sentiment of a sentence. Right now, the influence emoji have on the sentiment of a sentence is still a topic of discussion in the scientific community, as evidenced by the studies of Walther & D'Addario (2001) and Derks, Bos, & Grumbkow (2007).

The debate on the importance of emoji (MW, CC)

The disagreement pertains to what degree emoticons influence the sentiment of a sentence. Walther & D'Addario (2001) concluded that the verbal content of a message outweighed the valence of the emoticons. Derks et al. (2007) concluded the opposite: Emoticons are to a large extent similar to the non-verbal cues in conversation - they influence the interpreted meaning of a sentence (Derks et al., 2007). This debate is not without importance; it pertains to how much emoji and emoticons influence the valence landscape of CMC. Derks et al. (2007) also looked at the different uses and effects that emoticon-usage can have. The study found three main uses of emoticons: the strengthening of a verbal message, creating ambiguity and expressing sarcasm. Derks et al. (2007) replicated their findings in 2008 with a similar study. The paper manipulated the participant, by having them interact with either a friend or a stranger, and also manipulated the sentiment of the context (negative or positive). Derks et al. (2008) concluded that emoticons were mostly used for the strengthening of a message, expressing emotion or expressing humor - a similar description to his previous from 2007.

What was also found was that emoticons were more frequent in CMC with friends than with strangers, and more frequent in positive than negative context. These findings, Derks et al. (2008) claimed, seemed to point towards a similar usage of emoticons and facial behaviour. This was concluded in regards to the accounting for an interaction partner i.e. the difference in usage between strangers and friends and their increased frequency in positive contexts. Since both of these studies use emoticons, it can be debated whether or not they are relevant in a study of emoji, since emoticons are made up of keyboard characters such as parentheses and colons, while emoji are special characters or pictures. But in modern CMC, the lines between these have become blurry, especially with how emoji, and primarily smilies, are accessed. Most social media platforms transform emoticons to emoji. An example is on Twitch, where writing " :) " produces a smiling robot emote, " R) " creates a robot-pirate emote etc., as Twitch calls their emoji for emotes("Twitch Emotes" n.d.). Another program that converts emoticons is Skype, which accepts a couple of different versions of the smiling emoticon ," :) ", " :-) ", and " :=) " , and turns that into an animated emoticon ("What is the full list of emoticons", n.d.). Since the input for creating smiley-emoji is the same as creating an emoticon, i.e. typing " :) ", we believe that the studies are relevant, because the association will be similar.

It is in the context of this debate, and its potential implications that this study is to be viewed. The aim of this study is to increase the knowledge about the properties of emoji. This paper makes an attempt to study the underlying mechanism of the effect of emoji on the sentiment of a text. This is done by presenting participants with the same textual content but with different emoji. This study is interested in entertaining the notion, that emoji does influence the sentiment and therefore the meaning of a message. Furthermore, this paper also aims at gaining more knowledge on whether or not the interpretation of emoji vary between genders. On this basis, we hypothesize that:

***H*: Emojis modulate the perceived sentiment of written sentences.**

The interpretation of gender (MW, CC)

This study's focus is two-sided. In addition to the modulation of sentiment, our field of investigation is also that of the interpreted gender of the emoji. In contrast to natural languages, emoji are in most cases pictures, that try to portray items or feelings in a non-arbitrary way. What is meant by non-arbitrary in this context is that emoji refer to their real world referent by use of pictures, that are meant to reflect the physical attributes of the referent, and are therefore not arbitrary. That is to say, emoji resemble faces and objects.

In turn, this means that the interpreted gender of emoji also relates to the interpreted gender of the referent the emoji is referring to. In other words, if we want to study how facial emoji are perceived, we have to look at how their references, the faces, are perceived.

Previous research has shown very precise predictions of gender, based on the metric relations between different parts of the face (Gilani, Rooney, Shafait, Walters, & Mian, 2014). The study tried to make a computational model of the perceived gender of faces. One of the most precise predictors was shown to be the geodesic distance between the eyes (Gilani et al., 2014). As the studied emoji in this paper do not have any noses, this will not be possible measure for our study. Gilani et al. (2014) also proposed other significant predictors between the genders, such as jaw, chin and eye size. The relation between these findings and the interpretation of the gender of emoji are hard to match, as the emoji are simplifications of the human face, and do not follow the physiological norms, most notably perhaps the head having the form of a circle instead of an ellipsis. It is therefore hard to give measures on aspects such as the jaw and the chin of the face.

Instead of the physiological aspects of the emoji, one might instead look at the emotions or ideas they represent. Brebner (2003) did a study on the differences between gender in the experience of emotions. Brebner tested the differences in experience between eight emotions:

Anger, Fear, Joy, Sadness, Contentment, Guilt, Pride and Affection. The results showed a difference in the intensity of the emotions. Males scored higher on Pride, whereas females scored higher on Affection and Sadness (Brebner, 2003).

Although the effect sizes were small, this could serve as a more concrete bias in the interpretation of emoji. From this, one might expect emoji that communicate emotions of Affection and Sadness to be more associated with femininity, and emoji that communicate Pride to be associated to masculinity. This is in line with studies examining the stereotypes of males and females in relation to emotions. Kelly & Hutson-Comeaux (1999) examined the stereotypical emotions of men and women, manipulating the context the emotions were presented in. The findings show, that overreactions to happy and sad events were deemed more characteristic for women in a personal context, but being more characteristic for men in an achievement context. Regardless of context, overreactions in angry events were deemed more characteristic for men (Kelly & Hutson-Comeaux, 1999). This could serve as an indication, that emoji portraying anger would be perceived as being more masculine. This serves as a foundation for our second part of the study and its research question:

***RQ*: Is the interpreted gender of an emoji consistent with existing stereotypes on male and female emotions?**

Methods and materials (MW, CC)

In this study, an experimental and an exploratory approach was used. We recognize an experimental approach as trying to explain a causal relationship between two variables by manipulating the variables in an experiment. In our case, we're trying to explain the sentiment from the emoji used in the sentence. We recognize an exploratory approach as a more open ended research, with results not necessarily statistically describable. In our case, we're trying to describe certain patterns in the interpreted genders of emoji, but without having an initial hypothesis. Most of the existing literature has its emphasis on the usage of emoji, and not its associations. Because of the lack of existing research in the interpreted gender of emoji, the exploratory part of the study is done partly on the basis of research of feelings and gender stereotypes, but with our own search for patterns in the data. The general tendencies of interpretation of the gender of emoji were examined, as well as tendencies for the individual emoji. This examination was then compared to what our expectations would be of the data, if existing stereotypes on gender differences in feelings were analogous to that of emoji.

Participants (MW, CC)

The participants in this study were primarily Danish (93 out of 142), with the second most frequent nationality being American (22 out of 142). Over half of the participants were in the the age group of 19-22 years old (79 out of 142), and ranged from ages of 15 to 69 years old (M = 23.93, SD = 6.48). 75 of the participants were female, 63 were male, and 4 described themselves as "Other". On the question of how often participants used Emoji on a scale from 1-7, where 1 was "not at all" and 7 was "all the time", over half (60%) answered 5 or above.

Materials (MW, CC)

In this study we measured two different things. In the first part of the survey, we measured the sentiment value of a sentence. There was 10 unique sentences, that were shown once each without an emoji, and twice with an either positive or negative emoji. Sentences themselves were grouped into negative and positive sentences.

An example of a positive sentence and a negative sentence:

"I really like you."

"My brother is hurt."

These sentences were then manipulated with either a positive or a negative emoji.

An example of a sentence with a negative and then a positive emoji:

"My brother is hurt 😟"

"My brother is hurt 😀"

Figure 1: Example of the first part of the questionnaire

Participants were then asked to rate the sentiment of the sentences between 1-9, where 1 was very negative and 9 was very positive.

For the second part of the questionnaire, we measured whether participants perceived the previously seen emoji as masculine or feminine. The possible answers did not include a neutral option. However, we allowed participants not to answer questions in the second part of the survey, if they didn't have a particular inclination. (See Appendix for a link to a viewing of the entire survey.)

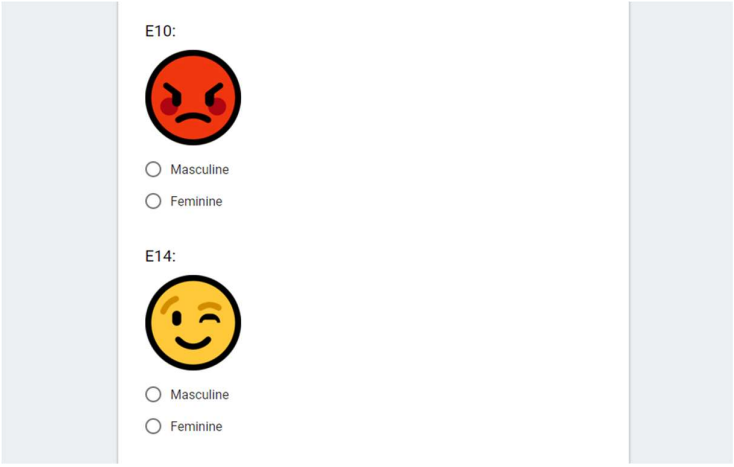

Figure 2: Example of the second part of the questionnaire

Procedure (MW, CC)

Participants were presented with a questionnaire, using the web based Google Forms. First they were asked to fill out the following information: Alias, Age, Gender, Native Language, Country of residence, and how often they used emoji. Participants were gathered primarily with use of sharing the link to the survey with social media such as Reddit and Facebook. The 10 sentences used in the survey were designed to be value-laden, i.e. to have a sentiment in either a positive or a negative direction. Five of the sentences had a negative sentiment and the other five had a positive sentiment. The choice to create sentences with a clearly positive or negative sentiment, was done in order to see how emoji modulate existing sentiment. Additionally, the

clearly defined sentiment allowed us to make a reliable assessment of the sentences, without including a coding procedure.

Participants were asked to rate the sentences on a odd-numbered scale. The odd-numbered scale (1-9) was chosen to allow for a neutral choice of sentiment (5). This is to ensure that participants are not polarized in the interpretation of the sentiment.

20 different emoji were used in the survey, with the aim of having a clear sentiment in either a negative or a positive direction for every emoji. The emoji were taken from Emojipedia, an encyclopedia of emoji ("😀 Smileys & People", n.d.). Only emoji that resembled faces was used. Another consideration that guided our choice of emoji was that of familiarity. We tried to limit the chosen emoji to emoji that were familiar and that we knew were relatively frequent in usage across social media (Rothenberg, 2013). The balance between these criteria, the inclusion of faces, clear sentiment and frequency, were the basis of the chosen emoji.

As sentences were presented three times each, we had to find a way around influencing the answer of our participants. One might imagine, that the order that the questions were presented in, would be a factor to be considered. To overcome this obstacle, all of the questions were randomized. The same was done for the gender-evaluation of emoji. Furthermore, we randomized whether or not "Masculine" or "Feminine" would be the first or the second possible answer. In this way, if the order did have an effect, it would be distributed across the items, and would not play a part in our results. The reasoning behind not including a neutral option for the second part of the study, was done in order to get clear inclinations in either a feminine or masculine direction. As previously mentioned, this could be a question, that participants had reasons for not answering. This could be because of a missing opinion on the matter, or a refusal to put gender labels on the emoji. For this reason, these questions were skippable.

Analysis/coding (MW, CC)

As previously mentioned, emoji are translatable across platforms, but their visual appearance differ. Miller et. al. showed, that this difference in appearance also led to a difference in interpretation (Miller et. al, 2016). On this basis, precautions were taken.

To be sure, that participants received the same stimuli, we made a form of reliability-test with a picture of an emoji, and an identical written emoji. Participants were then asked to say if the two were the same emoji. While the emoji might change according to what device was used, the .jpg is not translated into different languages - it is uniform across platforms and thus we can account for whether to include particular participants data or not. The emoji and picture chosen was that of an emoji with a cowboy hat. This choice was made on the basis of this particular emoji's general low frequency in usage (Rothenberg, 2013). At the time of this survey, the cowboy emoji is not listed on the Emoji Tracker, and has only received a around 400-500 Google searches since its launch (Rothenberg, 2013; Google Trends, n.d.). Furthermore, the cowboy hat emoji is a part of a new generation of emoji, as it was released as an approved part of Unicode 9.0 and Emoji 3.0 in 2016 ("🤠 Cowboy Hat Face", n.d.). Because of these two attributes, its low frequency and novelty, it makes for a good test emoji, as we expect, that if it cowboy hat emoji corresponds with the picture of the emoji as viewed on Microsoft's emoji representation, the chance of the more frequently used and older emoji to render will be relatively high. For this reason, the participants who answered "No" on the question of the picture and the emoji being identical, were filtered out of the survey.

From surveying the data, we could see, that some participants had used the same alias. For instance, 4 participants chose the alias "Cowboy". This is a concern, as we expect some variation in answers to be due to individual differences, and wanted to capture that in our model with the inclusion of alias as a random intercept. Additionally, the inclusion is also essential

for our model to live up to the assumption of independence. To work around this issue, we had to have a unique value for every participant. Google Forms automatically include a timestamp, for when participants finish the survey. As this is unique for every participant, this was used instead, and a manual change of alias was not deemed necessary.

As the data was collected using a scale from 1-9, similar to that of a Likert-scale, the data is technically ordinal (Likert, 1932). In order to generate comparability of the different conditions, we use the mean sentiment of the different conditions to compare how sentiment modulates according to emoji.

This approach has been the topic of some statistical debate for some time, and we want to make clear, that there is still some doubt about the validity of this method. The critique has risen due to the interpretation of the intervals of the scale. This is especially the case for the typical Likert-scales, as they go from ordinal categorical values *(I feel great, I feel good, I feel ok, I feel bad, I feel horrible)*, to integers *(1, 2, 3, 4, 5)*. There is no reason to assume, that there is the same distance between the statement *"I feel ok"* and *"I feel good"*, as there is between *"I feel ok"* and *"I feel bad"*, but that is the effect of the transformation of the values. When transformed to strict integers, and analyzed statistically, the computer doesn't care, it just crunches the numbers (Norman, 2010). It is hard to conclude anything on the basis of these results, as the scales themselves are influencing what the differences between the numbers are. In our case, there are some key resemblance points, but also important differences.

Our survey is similar to that of a Likert scale in that it tries to boil fairly complex human behavior down to that of integers. When that is stated, it is important to note that the previously mentioned transition was not the case for our data. The data was ordinal but also numeric from the get go. As we only used labels on 1 and 9 on the scale, there is little reason to believe, that the intervals between the different numbers isn't equal. This seems to be an advantage that this

type of scale has over that of a Likert-scale. Even though our data is technically ordinal, we propose that there is sufficient reasons to treat this data as intervals. The debate is not without reason however. As the data is not continuous, we don't see the normal distributions that are key to the assumptions of linear models, yet we do see it in our residuals. Furthermore, the variance is not equal, and our data is heteroscedastic. Although this is the case, the linear model and its results were used. This is done in light of the work of Geoff Norman, showing the robustness of parametric tests, such as the ANOVA but also on regression models, when done on data resembling that of a Likert-scale, even in the case of heteroscedastic and non-normally distributed data (Norman, 2010).

Model building (MW, CC)

In terms of emoji's impact on sentiment, the aim was to create a model that predicted sentiment from emoji-usage or the lack thereof. For this reason, the data was given different conditions in relation to emoji accompanying the sentence: 1 for negative, 2 for positive and 3 for no emoji. If a difference in means is observed, in accordance to their valence, this would indicate that they modulate sentiment. As previously mentioned, we expect some variance to be determined by the individual differences in participants. We also expect the items i.e. the sentences presented will also be a factor that has to be taken into account. This is due to valence-difference in the sentences, with or without emoji. This is necessary to account for the variance between the positive and the negative sentences, but also for the variance in valence in the two groups of sentences. For this reason, we use a mixed model, with items and participants as random intercepts, and condition as a fixed effect in predicting sentiment. For this the lmer-function was used (Bates, Maechler, Bolker & Walker, 2017). The model looks like this:

$$Sentiment \sim Condition + (1|Timestamp) + (1|Item)$$

Before applying our model to our data, the underlying assumptions relating to linear models were checked. The assumption of normality of residuals is upheld, as well as independence due to the inclusion of a timestamp as a random intercept. To check for the assumption of homoscedasticity, Levene's test was used, and showed a significant result, indicating that the data did not live up to this assumption. For this reason, a Welch's T-test was performed as it does not assume homoscedasticity. The results were significant, but it was not able to include random intercepts, which are vital to the assumption of independence. Instead, we used the lsmeans package to calculate the adjusted means for each group from the model. Lsmeans also explores post hoc pairwise comparisons between different conditions (Lenth, 2017). Although the robustness of parametric tests have been shown to be strong by Norman (2016), we wanted to increase the validity of our results in light of the debate of the methods used. We used Bonferroni's adjustment of the p-values to be more conservative in our attitude towards the data. Because this study is also interested in the differences in emoji interpretation between gender, an interaction effect between condition and gender is also tested. The model is identical to the one above, with the only exception being the inclusion of the interaction effect:

$$Sentiment \sim Condition * Gender + (1|Timestamp) + (1|Item)$$

For the coding, analysis and visual representation of the data, RStudio v. 3.4.2 was used. To get further analytical data, a number of packages in RStudio was used. The anova()-function from the lmer-package was used to get the F-statistics from the model, using Type III Wald and Satterthwaite's approximations to degrees of freedom. (Bates et al., 2017) The r.squaredGLMM-function from the MuMIn-package was used to calculate the effect size. (Bolker, 2017) The lsmeans-function from the lsmeans package was used to do individual pairwise comparisons using the Bonferroni adjustment, based on the adjusted means of the different conditions. (Lenth, 2017)

Results

Experimental results:

The analysis of the sentiment was done with a mixed effects model with sentiment as the outcome variable, condition as a fixed effect and participant and item as random effects. Condition i.e. emoji significantly changed the sentiment, $F(2, 27) = 11.13$, $p < .001$, $R^2 = .26$. Using the Bonferroni post hoc test, only the difference between the negative and positive emoji showed significant difference ($p < .001$).

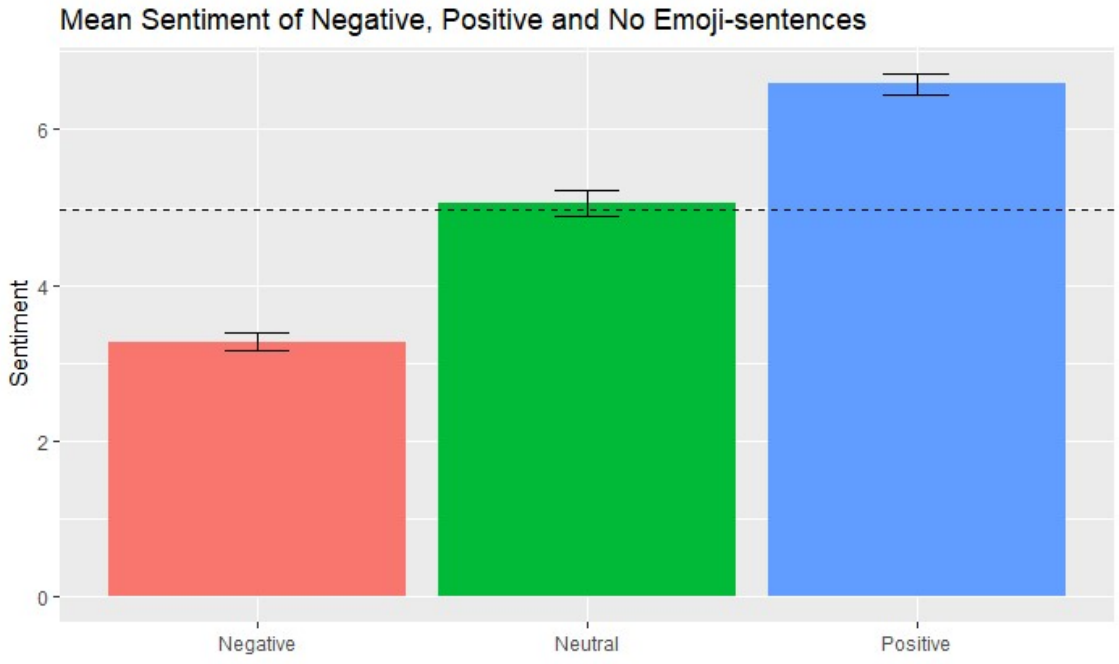

Figure 3: Differences in the mean sentiment of sentences. The different conditions represent the different type of emoji used in the sentence, with neutral being without an emoji. The dashed line represents the global mean.

Furthermore, there was found a significant interaction effect between gender and condition, $F(4, 2548) = 4.14$, $p = .002$, $R^2 = .27$. However, the individual pairwise comparisons with the Bonferroni adjustment did not reveal any significant difference between gender and condition (all ps $> .05$).

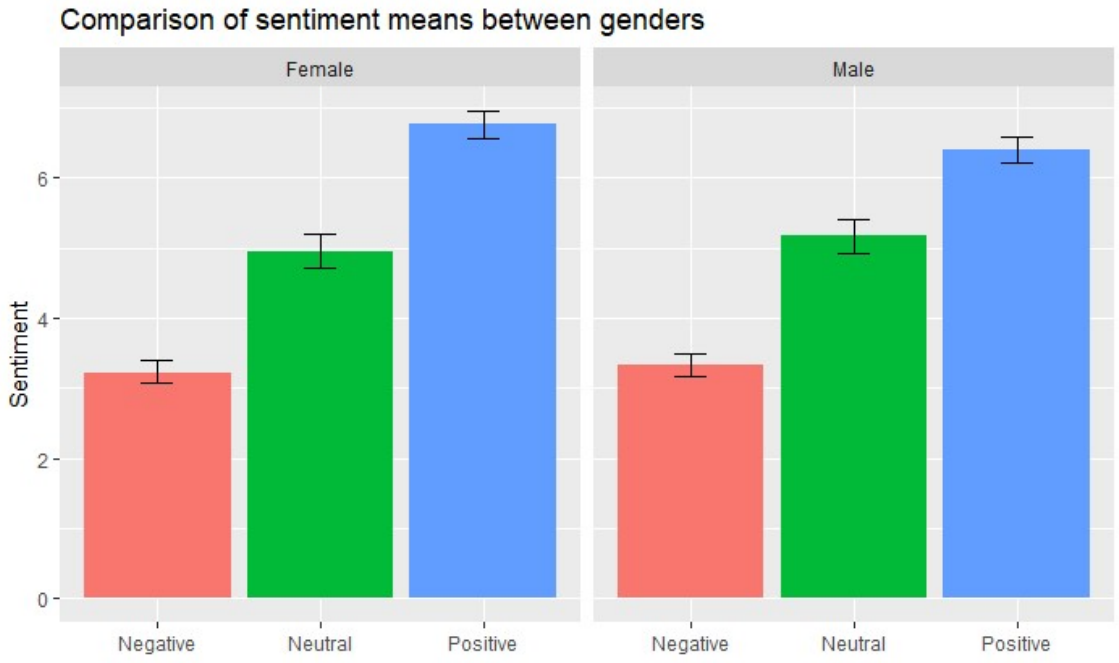

Figure 4: Comparison of sentiment means between genders. The different conditions represent the different type of emoji used in the sentence, with neutral being without an emoji.

Exploratory discoveries:

In terms of the more exploratory examination of the data, the following observations were done. The most pressing of the observations have been reported in shortened format below its graphical representation:

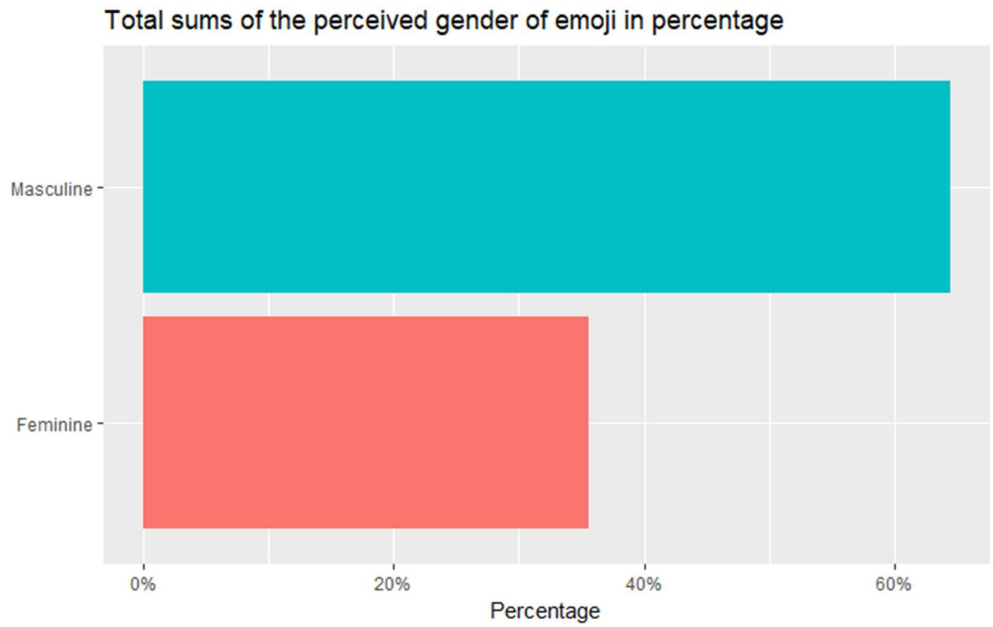

Figure 5: The total percentages of the summed emoji and their interpreted gender.

Emoji are primarily perceived as being male (64%).

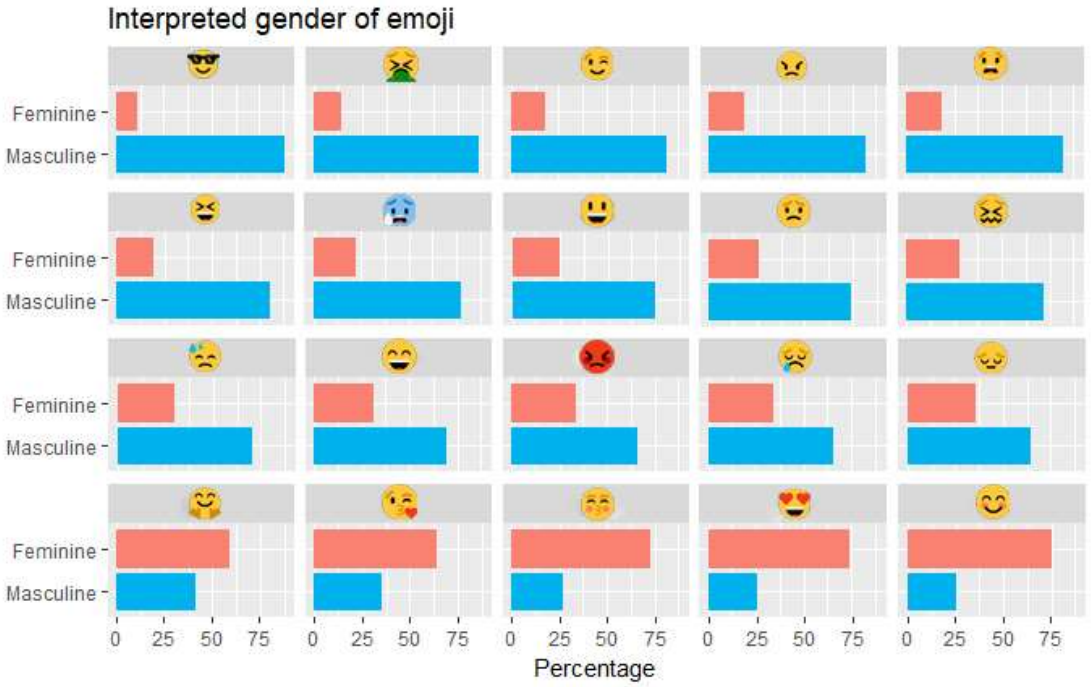

Figure 6: The interpreted gender of emoji, ranked from most masculine (top left) to most feminine (bottom right).

The top row, being the most masculine emoji, contains 2 out of 5 positive emoji, while the bottom row, is the only row with emoji being interpreted as more feminine than masculine, consists of exclusively positive emoji.

Discussion:

Discussion on the Hypothesis:

The results from our experimental study show support for our hypothesis, that emoji modulate the sentiment of a sentence. The difference however was only significant between the positive and the negative emoji. The reason for this could be due to ambiguity in the neutral condition, thus deviating enough to cross with both data from the positive and negative condition. There does seem to be indications in the data, that would suggest the hypothesis to be significant across all conditions. From examining the bar plot (Figure 3), we see a clear difference between the means, and no overlap between the confidence intervals. Another interesting piece of information is the resemblance between the global mean and that of the neutral condition. Furthermore, both of the means are almost exactly 5, the middle of our scale. This suggests, that in the neutral condition, the sentiment of the positive and negative sentences balance each other out. When considering the global mean, the same can be said; there is a balance between the negative and the positive modulation of the emoji. This is also in line with the observation, that the modulation of sentiment deviates relatively equally from the neutral condition in both the positive and the negative condition. The missing significance in the other conditions might be due to the more conservative approach in the post hoc testing. The overall results, indicating a sentiment modulation due to emoji-usage, are in line with that of Derks et al. (2007) and in contrast to that of Walther & D'addario (2001).

The results from the interaction effect is at first glance hard to interpret. This is most likely due to significance level transferring from the previous model. From the post hoc test, there is no indication of a significant interaction effect. The only result resembling a tendency, was that males generally tend to rate the positive condition lower than females. This can also be viewed in the bar plot of the gender comparison (Figure 4). The basis for the hypothesis of an interaction effect was based on the transfer of gender differences in other fields of study. It could simply be, that this transfer doesn't happen in relevance to our study. Additional research is required to conclude whether or not this is the case.

Elaborations on the interpreted gender of emoji

A number of things can be stated about the interpreted gender of emoji. First of all, there seems to be a difference in the positivity of emoji in accordance to their perceived gender. As stated in the observations about the interpreted gender of emoji (Figure 6), there seems to be a skewness in positivity, where feminine emoji are perceived as more positive than masculine emoji. It could be a topic of discussion on whether or not this is a general tendency or only present in our sample of emoji.

If there was a link between general gender stereotypes and the interpretation of the gender of emoji, we would expect more explicit signs of affection and sadness from predominantly feminine emoji, and explicit signs of pride and anger from the predominantly masculine emoji. This is based on the studies by Brebner (2003) and Kelly & Hutson-Comeaux (1999). From the exploratory examination of the data, there does seem to be indications, that the interpreted gender of emoji is in line with general gender stereotypes. From looking at the bottom row of the chart, the most feminine emoji, we do see more explicit signs of affection. This is based on items present such as hearts, hugs and kisses, which are not present in any of the other emoji. Another indication for a link, is in the third row. Looking at the last two emoji in this

row, we see indications of explicit sadness, such as a tear and a frowning face. This could serve as evidence for the more explicit sadness in females. It is important to note, that this is only the case in relation to the other emoji in this study. Both of the sad emoji, are on averaged rated more masculine than feminine, which could serve as a counter argument. The link could also be critiqued, as other emoji, which also show explicit sadness, are deemed more masculine than feminine, even in relation to the rest of the emoji of the study. This is the case for 😢and 😟. One might argue, that these are less extreme in their sadness - both have open eyes, and the water shown on the blue emoji seem more reminiscent of sweat than a tear.

Another case for the similarity is on the masculine end of the spectrum. The most masculine emoji, 😎, is described by Emojipedia as "denoting a sense of cool." ("😎 Smiling Face With Sunglasses", n.d.). This could be seen as an indication in line with the notion, that excessive pride is perceived more masculine than feminine. Another item in line with gender stereotypes, is that of anger. The 4th most masculine smiley, 😠, is described as "(...) anger, grumpiness or annoyance at a situation." ("😠 Angry Face", n.d.) Interestingly, the arguably more extreme version of representation of anger, colored red, 😡, is deemed more feminine than the uncolored version. Although this is still deemed more masculine than feminine, the difference could be in other domains of gender stereotypes. The most feminine emoji, 😊, has what Emojipedia calls rosy cheeks ("😊 Smiling Face With Smiling Eyes", n.d.), which could have similarities with that of makeup. This could also serve as an explanation of why this emoji is deemed the most feminine. Although less obvious, the same red circles can be found on the angry red emoji, which could explain the more feminine tendencies of this emoji. Although there are certainly exceptions to be accounted for, there seems to be some indications that suggest, that the perceived gender of emoji follow the gender stereotypical emotions.

To create a more complete interpretation of the gender of emoji, it would be interesting to include the participant's identification with the emoji - both in terms of their familiarity with the emoji but also in terms of their own usage of the emoji. One might expect, that with high frequency in their own usage of an emoji, participants would be more likely to pick their own gender. This could be done with a measuring either of how often they used particular emoji with a Likert-scale, or ideally with a comprehensive dataset containing participant's emoji-usage over an extended amount of time - although this might create some ethical problems. This would also make room for more individual difference in emoji usage to be included in models.

It could be discussed whether or not the study of the gender of emoji is an actual representation of reality. As there was no neutral option, participants were forced to polarize their answers. If the neutral option was included, emoji would might be viewed as predominantly gender neutral. This could also give interesting insights, in cases where the emoji are interpreted to have an either masculine or feminine identity. If these cases were in line with gender stereotypes, much of the unexplained emoji that do not fall in either stereotypical gender camp could strengthen the idea of predominantly feminine and masculine emoji.

General discussion on the validity of the paper

One might discuss the participants effect on legitimacy of the study. Previous studies have shown variance in emoji perception between geographical areas. As the study contains participants from 19 countries, the geographical difference might have been a factor (Miller et al., 2016). Many of the countries were only represented with one participant, and therefore the geographical and cultural difference was hard to account as the sample size was so small. It

would be interesting to do a similar study with a substantial participant pool from an array of different cultures and different geographical areas.

As participants were found through sharing links on Reddit and Facebook, there is little to no possibility of verifying the data. This means, that the same person could have taken the survey twice and used a different alias, or have lied about the personal information. There might even be people with malicious intent, so called “trolls”, might have taken the test and answered wrongly on purpose. Since we didn’t require login and that the survey was anonymous, it would be easy to do so. We hope that the number of participants will balance any such incidents out.

Another way the study could be refined, is to have a completely unknowing group of participants. A portion of the participants were students, who had been introduced to the hypothesis and underlying assumptions of the study. This could also serve as an explanation a tiny portion of the data, which was in direct opposition to what the hypothesis would expect. For example, when asked to rate the sentence "My brother is hurt. 😟", 2 participants gave the rating 9, which was specified as “Very positive”. The reverse can be viewed on the positive sentences, where a sentence such as "You are my favorite person. 😄", was given the rating 1 by 3 participants, which was specified as “Very negative”. This was not in line with what one would anticipate as a normal response in this scenario. One might debate whether or not this data should have been included. They were chosen to be included on several reasons. Based on our limited number of participants, it would be costly for the validity of the research to exclude any participants without sufficient reasons. Additionally, previous research indicated not only differences in meaning across platforms but also within (Miller et al., 2016). This is also in line with the study conducted on Chinese consumers by Jaeger & Ares (2017), where 33 emoji were rated. Eight of these ended up not having a defined category, as they were interpreted without any strong relation to any of the 39 terms of emotions.

One might debate the functionality of the Cowboy-test of reliability of emoji-convergence in the survey. This is so even though the test did manage to exclude a big portion of the participant pool (36,8% of the participants answered “No” on this question). An email was included in the study to give feedback on the study. Some participants reported, that even though they answered “Yes” on the cowboy-emoji, the test for the emoji actually showing in the manner wanted, some emoji didn’t show in the test. Another problem was the words used in the question. The exact wording was “Does the emoji match the picture?”, and here a more clear distinction could have been used. The intended question was whether or not they were identical. One might imagine, that some participants answered “Yes”, if the emoji wasn’t identical but *resembled* the emoji from the picture. This would in turn lead to more variance between participants in accordance to the results from Miller et. al. (2016), which showed the difference between the sentiment across different platforms i.e. versions of the emoji. This also creates another point of critique, as the sample size decreases substantially. With a larger sample size, one might hope to find statistically significant differences in the post hoc testing for more than just the positive and negative condition. This is reinforced by visually inspecting the results (See Figure 3).

As stated previously, there is an existing discussion in the usage of parametric tests on data of resemblance to the Likert-scale. If the findings could be replicated or improved upon with the use of data, that is not of this character, it could strengthen the belief in the robustness and applicability of these tests on this kind of research.

Considerations for future research:

From the acquired knowledge of this study, a list of reflections on the subject are listed here:

- Previous research has shown the different effects emoji can have on conversation. (Derks et al., 2008) However, this has not been done in relation to the shift in valence that emoji create. It would be interesting for future research to look at the different effects, such as irony, and to study what that particular effect does to the valence of a sentence. In our case, this could have been done partially by including an optional field of text for each question, where participants were asked to fill in the reasoning behind the given answer. The same principle could have been used in the interpreted gender of emoji, which potentially could have lead to a stronger case for the relation between gender stereotypes and interpreted gender of emoji.
- As previously stated, there has been found a difference in emoji-usage and perception between cultures (Lu et al., 2016). It would be interesting to do a study with enough participants from each culture to make statistical inferences about every group. This would be interesting for both the sentiment, but also for the interpreted gender of the emoji to see, if there could be found a difference between cultures.
- Instead of conducting the experiment as a survey, it would be interesting to do it as a face-to-face experiment. If one videotaped one's participants while they took a similar study, one could look at the facial responses, and have participants reason for their answers with direct questions. This could give further insights in the considerations and effect of emoji usage.
- Building on the exploratory findings of this study, it would be interesting to find a link between more masculine emoji being associated more negatively than feminine emoji. Preliminary research in the perceived gender of emoji is perhaps needed first to conduct a complete study of this phenomenon.

Conclusion

- The statistical analysis supports the claim that emoji modulate the sentiment of a sentence. There seems to be very little evidence, that suggests a difference between gender in this regard, although a tendency for men assign lower values for positive emoji was found.
- In the elaborative examination of the data, there was found indications that emoji perception is in line with existing stereotypes concerning gender differences in emotions. More research is needed to increase the certainty of this claim.

References

Websites & Documents:

Bates, Maechler, Bolker & Walker (2017). *Package 'lme4'.* Retrieved January 4, 2018 from The Comprehensive R Archive Network-website: https://cran.r-project.org/web/packages/lme4/lme4.pdf

Bolker (2017). GLMM FAQ, Retrieved January 4, 2018, from Bolker's Github-website: https://bbolker.github.io/mixedmodels-misc/glmmFAQ.html

Google Trends. (n.d.). [*Search data from Google, from 2004 till now. Search term:* 🍜] Retrieved December 29, 2017 from Google Trends: https://trends.google.com/trends/explore?date=all&q=%F0%9F%A4%A0

Messenger (2017). *The Art of Communication: Messages That Matter.* Retrieved from FB Newsroom website: https://fbnewsroomus.files.wordpress.com/2017/11/messagesthat-matter_editorial-2.pdf

Rothenberg, M. (2013) [Counter and tracker of various emoji used on twitter in realtime] *emojitracker: real time emoji use on twitter* Retrieved December 28, 2017 from Emojitracker: http://emojitracker.com/

Schenker, M. (2016, October 11). THE SURPRISING HISTORY OF EMOJIS. Retrieved from: https://www.webdesignerdepot.com/2016/10/the-surprising-history-of-emojis/

Twitch Emotes. (n.d.). Retrieved January 2, 2018 from TwitchEmotes: https://twitchemotes.com/

What is the full list of emoticons? Retrieved January 4, 2018 from Skype Support: https://support.skype.com/en/faq/FA12330/what-is-the-full-list-of-emoticons

🤠 Cowboy Hat Face. (n.d.). Retrieved December 16, 2017 from the Emojipedia: https://emojipedia.org/face-with-cowboy-hat/

😊 Smiling Face With Smiling Eyes. (n.d.). Retrieved December 27, 2017 from the Emojipedia: https://emojipedia.org/smiling-face-with-smiling-eyes/

😎 Smiling Face With Sunglasses. (n.d.). Retrieved December 27, 2017 from the Emojipedia: https://emojipedia.org/smiling-face-with-sunglasses/

😀 Smileys & People. (n.d.). Retrieved January 04, 2017 from the Emojipedia: https://emojipedia.org/people/

Papers and articles

Barbieri, F., Ronzano, F., & Saggion, H. (2016). What does this emoji mean? A vector space skip-gram model for twitter emojis. Paper presented at the *Lrec,*

Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, syntax, and semantics. *Language in Mind: Advances in the Study of Language and Thought,* , 61-79.

Brebner, J. (2003). *Gender and emotions* doi:10.1016/S0191-8869(02)00059-4

Christiansen, M. H., & Kirby, S. (2003). *Language evolution: Consensus and controversies* doi://doi.org/10.1016/S1364-6613(03)00136-0

Derks, D., Bos, A. E. R., & Grumbkow, J. v. (2007). *Emoticons and social interaction on the internet: The importance of social context* doi://doi.org/10.1016/j.chb.2004.11.013

Derks, D., Bos, A. E. R., & von Grumbkow, J. (2008). Emoticons in computer-mediated communication: Social motives and social context. *CyberPsychology & Behavior, 11*(1), 99-101. doi:10.1089/cpb.2007.9926

Dimberg, U., & Lundquist, L. (1990). *Gender differences in facial reactions to facial expressions* doi://doi.org/10.1016/0301-0511(90)90024-Q

Gilani, S. Z., Rooney, K., Shafait, F., Walters, M., & Mian, A. (2014). Geometric facial gender scoring: Objectivity of perception. *Plos One, 9*(6), e99483. Retrieved from https://doi.org/10.1371/journal.pone.0099483

Hamza Alshenqeeti. (2016). Are emojis creating a new or old visual language for new generations? A socio-semiotic study. *Advances in Language and Literary Studies, 7*(6), 56-69. doi:10.7575/aiac.alls.v.7n.6p.56

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). *Human neural systems for face recognition and social communication* doi://doi.org/10.1016/S0006-3223(01)01330-0

Hudson, M. B., Nicolas, S. C., Howser, M. E., Lipsett, K. E., Robinson, I. W., Pope, L. J., . . . Friedman, D. R. (2015). Examining how gender and emoticons influence facebook jealousy. *Cyberpsychology, Behavior, and Social Networking, 18*(2), 87-92.

Jaeger, S. R., & Ares, G. (2017). *Dominant Meanings of Facial Emoji: Insights from Chinese Consumers and Comparison with Meanings from Internet Ressources* doi:12048/10.1016/j.foodqual.2017.04.009

Kelly, J. R., & Hutson-Comeaux, S. (1999). Gender-emotion stereotypes are context specific. *Sex Roles, 40*(1), 107-120. doi:1018834501996

Kralj Novak, P., Smailović, J., Sluban, B., & Mozetič, I. (2015). Sentiment of emojis. *Plos One, 10*(12), e0144296. Retrieved from https://doi.org/10.1371/journal.pone.0144296

Lenth, R. V. (2017). Using lsmeans.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology,*

Lu, X., Ai, W., Liu, X., Li, Q., Wang, N., Huang, G., & Mei, Q. (2016). Learning from the ubiquitous language: An empirical analysis of emoji usage of smartphone users. Paper presented at the *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing,* 770-780.

Marengo, D., Giannotta, F., & Settanni, M. (2017). *Assessing personality using emoji: An exploratory study* doi:12048/10.1016/j.paid.2017.02.037

McIntyre, E. S. (2016). From cave paintings to shakespeare and back again: What are emoji and should I be afraid? (doctoral dissertation). *From Cave Paintings to Shakespeare and Back again: What are Emoji and should i be Afraid?,*

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes, 45*(3), 211-236. doi:10.1080/01638530802073712

Niedenthal, P. M. (2007). Embodying emotion. *Science, 316*(5827), 1002-1005. doi:10.1126/science.1136930

Walther, J. B., & D'Addario, K. P. (2001). The impacts of emoticons on message interpretation in computer-mediated communication. *Social Science Computer Review, 19*(3), 324-347. doi:10.1177/089443930101900307

Whorf, B. L. (1956). Language, thought, and reality: Selected writings of….(edited by john B. carroll.).

Wolf, A. (2000). Emotional expression online: Gender differences in emoticon use. *CyberPsychology & Behavior, 3*(5), 827-833. doi:10.1089/10949310050191809

Yoshimura, R., Rothfus, E., Liu, J., & Zheng, B. (2017). Perceiving sex in text messages: Pepperdine students use screenshots to distinguish individuals' sexes. *Pepperdine Journal of Communication Research, 5*(1), 16.

Appendix

Link to the complete survey:

https://goo.gl/forms/Uu17BRGcPnDvfJ2Z2