

What's behind those smiling eyes: Examining emoji sentiment across vendors

Ashley A. Shurick

Google
Mountain View, CA, USA
ashleyshurick@google.com

Jennifer Daniel

Google
Mountain View, CA, USA
jenniferx@google.com

Abstract

To date, a significant body of literature has examined the unique characteristics involved in communicating emotion. These days, emotional intent is frequently conveyed through emoji in online communications. If we are to design technology that assists in conveying emotional intent, we need to better understand how particular design details can play a role in interpretation - particularly across different vendors, many of which have slightly varying designs of the same emoji. With this research we compared the sentiment associated with 45 emoji across four commonly used vendors (iOS, Android, Samsung, and WhatsApp) to identify areas of misalignment and begin identifying features that influence sentiment interpretation. 6,250 US mobile phone users were asked to select the sentiment that best described 10 randomly-selected emoji. Results indicate a high level of sentiment agreement for emoji expressing positive affect, surprise, and those with external cues (such as a Face with Medical Mask). Lower levels of agreement were assigned to emoji that expressed negative affect, ambiguous expressions, or had closed eyes. Potential reasons for the low levels of agreement include less familiarity with specific emoji and differences between vendors in how they illustrate facial features.

Introduction

Humans are inherently social creatures and frequently use facial expressions to communicate and convey emotion (Darwin 1965). Evidence suggests that six basic emotions (fear, sadness, anger, surprise, happiness, and disgust) may be universally recognized across cultures (Izard 1971; Ekman and Friesen 1975). To better understand the component parts that comprise specific emotional expressions, (Ekman and Friesen 1978) developed the Facial Action Coding System (FACS) which systematically broke down expressions into Action Units (AUs), or movements of individual facial muscles. FACS has since been used in a variety of applications, including to identify emotional expression. For example, (Kohler et al. 2004) used FACS to assess the recognition of happy, sad, angry, and fearful expressions, and determined that each distinct emotion could be recognized by

a specific group of AUs. Sad faces, for example, were associated with lowered eyebrows and raised chin (Kohler et al. 2004).

While specific AU's have been used to identify distinct emotional expression in human faces, less research has highlighted the contribution that individual facial features have on the interpretation of computerized facial expressions, such as emojis. Emoji, defined as "a small digital image or icon used to express an idea, emotion, etc. in electronic communications"¹ are being used with increasing frequency in all forms of computer-mediated communication (CMC). Smiley emoji in particular are commonly used in messaging to convey nuance and intent when plain text is not sufficient.

Prior research has examined the sentiment associated with emoji renderings by analyzing and characterizing the text surrounding the emoji (Liu, Li, and Guo 2012; Novak et al. 2015). Research has also directly examined emoji interpretation by having raters assign a specific sentiment to an emoji rendering (Miller et al. 2016). One important contribution to come out of research on emoji sentiment is the concept of fragmentation whereby a sender may intend to express a specific emotion or sentiment but the receiver may see a different expression which may communicate a different sentiment. Due to different rendering of the same Unicode character across vendors (Google, Samsung, Apple, etc.) and 3rd Party Apps (e.g., WhatsApp) emoji expressions are often fragmented (Miller et al. 2016; Tigwell and Flatla 2016). Research has highlighted a need for emoji design convergence to help alleviate misinterpretation both within and between vendors (Miller et al. 2016), but little research has examined the specific influence that individual facial features have on sentiment interpretation. If eyebrows are tilted inward on one emoji rendering and tilted outward on another emoji rendering, what effect, if any, does that have on interpretation?

Understanding the facial attributes that influence perceived sentiment is essential to creating meaningful convergence of emoji design and reducing fragmentation, and is critical for identifying opportunities for new emoji to resolve interoperability. Furthermore, identifying emojis with high levels of sentiment agreement could improve keyword

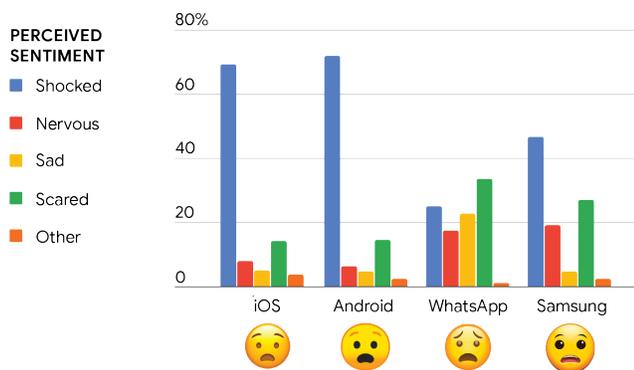


Figure 1: Four designs of “Anguished Face” emoji and associated sentiments

search for emoji predictions.

As a first step in trying to understand whether individual attributes such as eyebrows, mouth curvature, coloring, etc. influence interpretation, we compared the sentiment associated with smiley emoji across iOS, Android, Samsung, and WhatsApp to see whether emoji were consistently interpreted across renderings by these different vendors.

Method

Participants

8,681 US residents were recruited through Qualtrics to participate. Of these, 2,431 were excluded due to missing data resulting in a final sample of 6,250. Participants were roughly split in terms of primary device type with 2,188 iOS users, 2,104 Samsung users, and 1,957 Android users that did not use Samsung. The sample was roughly 40% male (2,472) and 60% female (3,750). Approximately 32% of the sample (2024 participants) were 18-34 years old, 38% of the sample (2358 participants) was 35-54 years old, and 30% of the sample (1868 participants) was above 55 years.

Procedure

All participants completed a survey via Qualtrics, an online survey platform. Participants were presented with the image of one emoji rendering and asked to select the sentiment or descriptor that best described the emoji. Four descriptors were provided to each participant along with an “other” category that allowed for free-form input. Each participant completed the task 10 times. Each emoji was rated by a minimum of 270 participants and the order of emoji was randomly presented to participants.

Selected Emoji

45 emoji were selected from the Unicode consortium. Smiley emoji were selected due to their high frequency of use and variance in anthropomorphic characteristics. Of all possible smiley emoji we selected emoji with complex design and lots of features, for example, emoji with teeth or tongue, hands, variation in eyebrow shape, variation in eye shape, contextual elements such as a handkerchief, etc.

Additionally, we wanted to examine whether frequency of emoji use would have any effect on sentiment interpretation, so we included 10 of the top 30 most frequently used emoji and 10 emoji that ranked in the top 31-80 most frequently used emoji².

Vendor Selection

To compare sentiment across vendors we selected some of the world’s most popular smartphone operating systems: Android, iOS, and Samsung³ as well as a popular messaging app, WhatsApp⁴. We specifically compared iOS version 13.1, Samsung version One UI 1.5, WhatsApp 2.19.352 and Android 10.

Sentiments

Sentiments were sourced from the Common Locale Data Repository Project (CLDR), definitions found from emoji related dictionaries, and emoji shortcodes. Among the types of data that CLDR includes is locale-specific information and keywords associated with emoji. Emoji shortcodes are codes used on messaging apps to swiftly insert without switching to your emoji keyboard. For example typing, :joy:

shows Face With Tears of Joy (😄) on apps that support shortcodes. These codes are not standardized and vary slightly from vendor to vendor.

Data Analysis

In order to compare sentiment between emoji, the joint probability of agreement was calculated by summing the number of raters who assigned a given sentiment and dividing by the total number of raters for the emoji. This calculation was done for each emoji rendering, allowing us to directly compare sentiment between vendors. This measure of inter-rater agreement was chosen over other measures because of the nature of the study design: raters assigned sentiment to 10 emoji chosen at random, meaning that no raters assigned sentiment for the same set of emoji.

In addition to calculating agreement, variation ratio was used to better understand how well the sentiment with the highest level of agreement (or the mode) reflected the dataset as a whole. The variation ratio was calculated as follows:

$$1 - (f/N)$$

where f is the number of cases for the most frequently assigned sentiment, and N is the total number raters.

Variation ratio reflects the proportion of answers which are not in the mode category and was calculated for each emoji rendering to compare between vendors. For example, in Figure 1, “shocked” has the most user selections for iOS, Android, and Samsung emoji renderings and therefore was used as the mode when calculating variation ratio. Alternatively, “scared” had the most user selections for the WhatsApp emoji rendering and therefore was used as the mode

²<https://home.unicode.org/>

³<https://gs.statcounter.com/os-market-share/mobile/worldwide>

⁴<https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/>

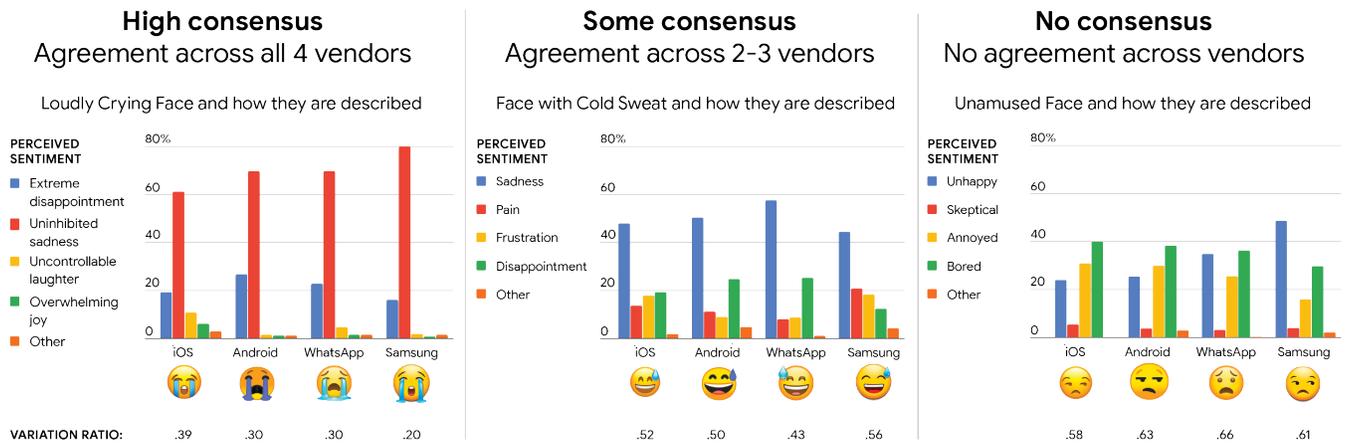


Figure 2: Examples of consensus and variation ratio: (Left) Example of an emoji with high levels of agreement across all 4 vendors. (Center) example of an emoji with some level of agreement across two vendors. (Right) Example of an emoji with no agreement across vendors. Variation levels vary based on level of agreement, with higher values indicating more disagreement and lower values indicating greater consensus. Y-axes represent percent of raters who assigned a given sentiment.

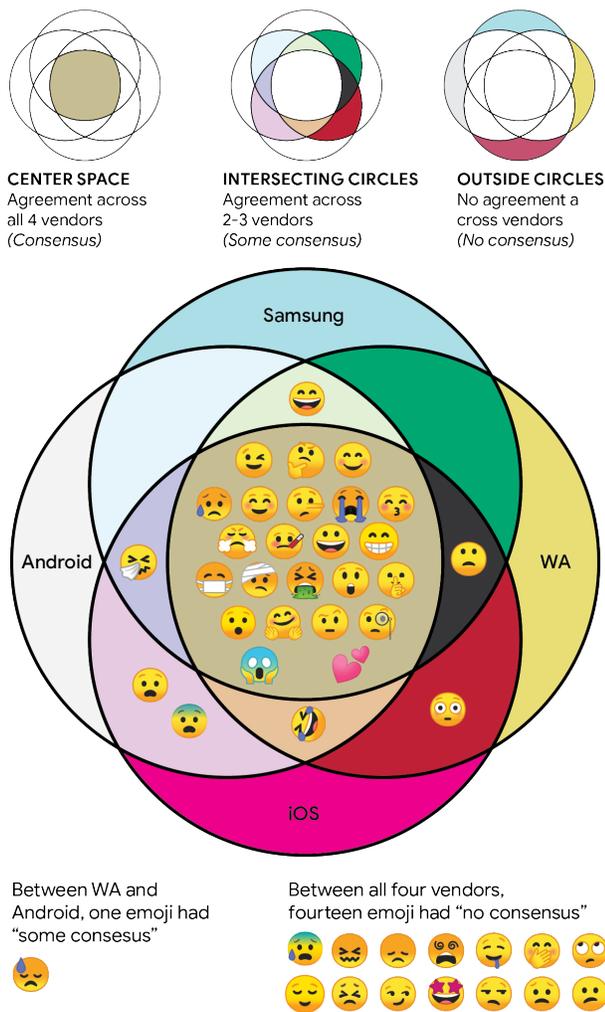


Figure 3: Model of emoji agreement

when calculating variation ratio. Lower numbers reflect less variation and more agreement, and higher numbers reflect more variation and less agreement.

In addition to calculating agreement and variation ratio “other” responses were analyzed to look for patterns - essentially, to discover whether a high proportion of participants agreed on an interpretation that was not provided by the four multiple choice options. If less than 10% of participants assigned an “other” response to a given emoji, the responses were summed into a 5th “other” category that was used in calculating percentage of agreement and variation ratio. Two emoji had >10% participants assign “other” responses for at least one of the vendors - dizzy face (🤪) and drooling face (🤤). These responses were examined to identify overlap. Very few participants were in agreement (<5%) so all “other” responses were aggregated and used as a 5th category.

Our goal was to identify emoji with high and low levels of agreement to look for patterns in design and identify facial attributes that may influence sentiment. Each emoji had 5 possible interpretations (4 assigned sentiments and a 5th “other” category) so there was a 20% chance that a rater selected any given response. For our categorization we deemed that a minimum of 50% of raters needed to agree on a sentiment in order for the emoji to be deemed as having some level of agreement - or consensus. This threshold allowed us to identify the largest areas of agreement and disagreement between sentiments so that we could further explore how individual facial features may influence interpretation.

After determining consensus, sentiment was examined across vendors and grouped according to vendor overlap:

- An emoji was deemed to have consensus when the same sentiment was assigned across all four vendors.
- An emoji was deemed to have some consensus when the same sentiment was assigned across two or three vendors.

Table 1: Variation ratios across vendors

	iOS	Android	WhatsApp	Samsung
CONSENSUS	0.34	0.36	0.35	0.35
2-3 VENDOR CONSENSUS	0.42	0.43	0.44	0.53
NO CONSENSUS	0.59	0.57	0.58	0.59
AVERAGE	0.43	0.44	0.44	0.46

- An emoji was deemed to have no consensus when a different sentiment was assigned across all four vendors, and/or raters did not achieve a level of agreement of at least 50%.

These groupings were used to distinguish between emojis that may benefit from vendor convergence vs emoji that already had convergence (for examples of agreement and variation ratio calculations across these three groupings, please see Figure 2). The one caveat is for two emoji (HUGGING FACE (🤗), and FACE WITH ONE EYEBROW RAISED (😬)). Raters agreed upon two sentiments that were synonyms⁵ (enthusiasm and excitement for HUGGING FACE, and suspicious and skeptical for FACE WITH ONE EYEBROW RAISED) so percentage of agreement was combined across both.

Results

Model of Agreement

Consensus was achieved across all four vendors for 23 of the 45 tested emojis (😱😄😬😁😭😞😏😇😈😎😡😢😣😤😥😦😧😨😩😪😫😬😭😮😯😰😱😲😳😴😵😶😷😸😹😺😻😼😽😾😿😺😻😼😽😾😿). Emoji that have high consensus tend to be positive, surprised, or have clear contextual elements, such as a face mask. Eight emoji (😄🏊😭😞😟😠😡😢) had some consensus across 2-3 vendors and 14 emoji (🤔😬😏😇😈😎😡😢😣😤😥😦😧😨😩😪😫😬😭😮😯😰😱😲😳😴😵😶😷😸😹😺😻😼😽😾😿) had no consensus across vendors. Emoji with some or no consensus tended to have negative or ambiguous expressions or closed eyes. For a model that summarizes overlap between emoji, see Figure3.

Variation Ratio

Average variation ratios are represented in Table 1. Inherently, variation ratios for emojis with no consensus are higher than for emojis with consensus. Overall, vendors did not differ significantly in their average variation ratios.

Frequency of Emoji Use

Of the 10 most frequently used emoji, 7 were deemed to have consensus across all vendors. Of the 10 less frequently used emoji only two were deemed to have consensus across all vendors.

⁵<https://www.merriam-webster.com/>

Discussion

The current research set out to compare the sentiment associated with emojis across four different vendors in order to identify areas of common overlap and opportunities for designs to intentionally align or deviate. Results indicate a high level of agreement for 23 of the 45 emoji tested. Specifically, raters are in general agreement for the sentiment of emoji with positive affect, surprise, and those with external cues. Raters were less aligned for the remaining 22 tested emoji, with 8 emojis having consensus across 2-3 vendors and 14 emojis having no consensus across vendors.

There are a number of reasons that emojis may have different interpretations across vendors. Emoji are frequently used in the context of a conversation - without context, emoji may be interpreted differently by different people. Additionally, some raters may be less familiar with some emoji, which may cause the rater to question the associated sentiment. Results demonstrated that 7 of 10 most frequently used emoji had consensus, while only 2 of 10 less frequently used emoji had consensus. It may be that raters are in general agreement for emoji they see and use on a regular basis. Future research should explore this idea in more detail.

Another possible reason emojis may have different interpretations across vendors is because the facial features associated with each emoji can vary across vendors. For example, in Figure 1, the eyebrows of the Anguished Face emoji for WhatsApp are tilted inwards while the eyebrows of the emoji for iOS, Android and Samsung are tilted outwards. The variation ratio for WhatsApp is much higher than the other three vendors as raters were split between four possible sentiments. The variation ratio for iOS and Android were much lower as raters were in general alignment that the emoji expressed a “shocked” sentiment. While Samsung has a comparable expression to Android and iOS, the emoji has a tongue and teeth which may create a more nuanced interpretation. This interpretation is further supported by a variation ratio that is lower than WhatsApp but higher than iOS and Android.

A second example highlighting how slight adjustments in facial features may influence the resulting interpretation may be seen in the data for “slightly frowning face” (see Figure 4). More than 50% of raters across iOS, WhatsApp and Samsung agreed that “sad” best described the “slightly frowning face” emoji. However, less than 50% of raters assigned “sad” for the Android emoji rendering and there was more variance in the interpretation (the Android rendering had a variation ratio that was $\geq 15\%$ higher compared to other vendors). The Android emoji has a wider mouth with less frown compared to the other vendors which may contribute to the nuanced interpretation. To help create a more

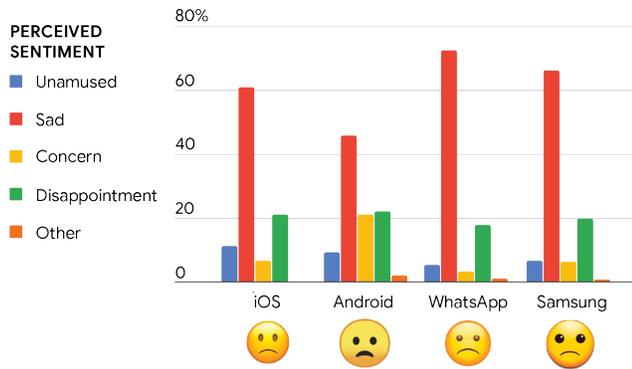


Figure 4: Four variations of “Slightly Frowning Face” emoji and associated sentiments.

unified interpretation of emojis vendors may seek to align the facial features for each emoji for existing codepoints. Vendors with lower levels of agreement (or higher variation ratios) may want to focus on updating the design of their emoji in order to improve alignment. In the future, to prevent further fragmentation, Unicode could consider more explicit design guidelines.

Limitations

Though we studied 45 of the existing smiley emoji, there are currently 3,225 total emoji Unicode characters. Likewise, we studied only four of the most popular emoji vendors, but there are a number of other emoji fonts with their own unique emoji renderings. Future research should continue to clarify the facial attributes associated with emoji renderings across a variety of different vendors.

The findings presented here are further limited by the fact that emoji were examined out of context (i.e., not in the presence of a larger conversation). Additionally, this research did not examine how individual differences may influence the associated interpretation of sentiment - for example, the ways in which gender, age, or culture influence interpretation. Future research should explore these issues.

Conclusion

This research helped shed light on the way in which facial features like eyebrows, mouths, and eyes may influence emoji interpretation. Future research should continue to explore the influence that specific facial features have on emoji design, which could prove highly impactful for the Unicode consortium to consider when introducing new smiley emoji.

Acknowledgements

Special thanks to Cecile Basnage and Christina Ahmadzai for help coding the survey.

Additional Information

For more information on tested emoji, please request access to the following: <https://docs.google.com/document/>

References

Darwin, C. 1965. The expression of the emotions in man and animals (vol. 526). Chicago, IL: University of Chicago press.(Original work published in 1872).

Ekman, P., and Friesen, W. 1975. Unmasking the face: A guide to recognizing emotions from facial clues. Englewood Cliffs, NJ: Prentice Hall. *Elfenbein, HA* 86:203–235.

Ekman, P., and Friesen, W. 1978. Manual of the facial action coding system (facs). Trans. ed. Vol. Consulting Psychologists Press, Palo Alto.

Izard, C. 1971. The face of emotion. *Appleton-Century-Crofts, New York*.

Kohler, C. G.; Turner, T.; Stolar, N. M.; Bilker, W. B.; Brensinger, C. M.; Gur, R. E.; and Gur, R. C. 2004. Differences in facial expressions of four universal emotions. *Psychiatry research* 128(3):235–244.

Liu, K.-L.; Li, W.-J.; and Guo, M. 2012. Emoticon smoothed language models for twitter sentiment analysis. In *In Proceedings of the 26th AAAI Conference on Artificial Intelligence*, 1678–1684. AAAI Press.

Miller, H. J.; Thebault-Spieker, J.; Chang, S.; Johnson, I.; Terveen, L.; and Hecht, B. 2016. "blissfully happy" or "ready to fight": Varying interpretations of emoji. In *In Proceedings of the 10th International AAAI Conference on Web and Social Media*, 259–268. AAAI Press.

Novak, P. K.; Smailović, J.; Sluban, B.; and Mozetič, I. 2015. Sentiment of emojis. *PLoS ONE* 10(12):e0144296.

Tigwell, G. W., and Flatla, D. R. 2016. Oh that’s what you meant! reducing emoji misunderstanding. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, MobileHCI ’16*, 859–866. New York, NY, USA: Association for Computing Machinery.