

# Not Just Iconic: Emoji Interpretation is Shaped by Use

**Brianna O’Boyle**

Dept. of Linguistics  
San Diego State University  
San Diego, CA, USA, 92182  
boboyle@sdsu.edu

**Gabriel Doyle**

Dept. of Linguistics  
San Diego State University  
San Diego, CA, USA, 92182  
gdoyle@sdsu.edu

## Abstract

Where do the meaning of emoji come from? Though it is often assumed that emoji are fully iconic, with meanings derived from their visual forms, we argue that this is only one component of their meaning. We surveyed users and non-users of the Chinese social media platform WeChat for their interpretations of emoji specific to WeChat. We find that some emoji show significant differences in their interpretations between users and non-users, as well as how familiar a person is with the specific emoji’s use. We argue that this reflects a more complex process for building the meaning of emoji on a platform than pure iconicity.

## 1 Introduction

How do you know what an emoji means? The answer may seem obvious: what it looks like is what it means. The Unicode emoji U+1F971, for instance, is a yawning face. It corresponds to a universal human symbol for tiredness. In this sense, emoji are different from words, whose forms generally provide little information about their meaning. *Chat* means very different things in English and French, for instance. As a result, emoji may seem to be a fairly trivial part of sentiment analysis, easier than dealing with the complex meanings of a text.

But this simplified division into semantically-transparent emoji and semantically-opaque text is not the whole story. Some words’ forms do convey information about their meanings. Onomatopoeia, like *quack* or *meow*, are attempts to transcribe a sound, and aim for semantic transparency. While onomatopoeic words seem like very special cases, recent work has increasingly found evidence of sound symbolism even in words whose wordforms seem completely unrelated to their meaning (see Svantesson, 2017 for a review). Wordforms also correlate with the perceived complexity of the ideas they express (Lewis and Frank, 2016). These subtle

links between words’ forms and meanings are even used by children as part of language acquisition (Imai et al., 2008).

These results suggest that words are not as semantically opaque as they seem. In a similar vein, then, we might ask if emoji are as semantically transparent as they seem. We examined people’s interpretation of emoji from the Chinese social media platform WeChat and found significant deviations in meanings for some emoji between people familiar and unfamiliar with the emoji.

This suggests that emoji interpretation is substantially less transparent than it seems. The meanings of emoji get built up through use, like words’ meanings do, and may drift away from their purely iconic interpretations as they amass platform-specific meanings.

Our key finding is that some, but not all, emoji show significantly different interpretations by regular users of a platform than the first impression these emoji convey to novices. These differences are idiosyncratic, and appear to arise from a combination of increasing familiarity with the platform and its cultural context. In some cases, the difference is sufficient to flip the valence of the emoji’s interpretation between novices and experienced users. We also propose some possible approaches to closing this gap—or at least mitigating the impact of such misinterpretations.

## 2 Background

### 2.1 Emoji meaning and interpretation

Emoji were first created in the 1990s for Japanese cellphones, and have become a prominent component of computer-mediated communication since then. Emoji have a variety of communicative uses, including conveying the tone of a message (Lo, 2008; Hu et al., 2017), adding situational meaning (Kaye et al., 2016), and giving a sense of personality from the author (Sugiyama, 2015). They pro-



Figure 1: Some of the WeChat-specific emoji.

vide paralinguistic information which the reader then combines with the linguistic information of a message to determine the sentiment of the message, though the result is not always a straightforward combination of the text and emoji components (Tian et al., 2017).

Most importantly for this research, some have proposed that emoji represent a cross-linguistic, cross-cultural communication avenue, due to their iconicity and semantic transparency (Alshenqeeti, 2016; Lu et al., 2016). However, for this to be the case, we need to establish that emoji are actually interpretable across the divides of language, culture, and even social media platforms.

There is some evidence that emoji do not seamlessly straddle these divides. Tigwell and Flatla (2016), for instance, collected people’s sentiment ratings on Unicode emoji that were rendered differently on Android and iOS devices. They found significant deviations in how people reacted to these emoji due to their different forms on different devices. Similarly, the intended interpretation of an emoji may rely on idiosyncratic interpretations, as with inside jokes or call-backs (Wiseman and Gould, 2018). Thus, while we can develop emoji sentiment lexicons that give approximate sentiments for an emoji (Kralj Novak et al., 2015), there will be a variety of dimensions in which the actual interpretation will vary from that baseline.

In the present study, we examine how platform familiarity can affect emoji interpretation in a more stripped-down setting. Participants all see the same representation of an emoji, so cross-platform rendering differences will not exist. The emoji are evaluated apart from their text so that there will not be complex interactions between the two. Despite this, we find that familiarity with the emoji and the WeChat platform affect the interpretation of an emoji’s sentiment.

## 2.2 WeChat and its emoji

WeChat is a Chinese messaging, social-media, and payment app with more than one billion users. It is extremely prominent within China, and is used both by Chinese nationals and people of other na-

tionalties within China (e.g., Szurawitzki, 2022 studied German speakers’ use of WeChat).

WeChat represents an exciting test case because it maintains its own lexicon of 105 emoji, which range from minor variants of the familiar Unicode emoji to completely unique forms<sup>1</sup>. Some work exists looking at the usage of WeChat emoji by WeChat users. Li and Yang (2018) used a corpus study of WeChat to examine the motivations behind the use of emoji, and found approximately half of the emoji use was to express or intensify an emotion. Liu et al. (2020) studied feelings of embarrassment from sending an emoji that had the opposite meaning of what was intended, reflecting the dangers of emoji misinterpretation. Both of these studies point to the importance of getting accurate understandings of emoji sentiment for understanding the meaning of a message, as well as hinting at the danger of misinterpretation.

## 3 Experiment

### 3.1 Experimental design

**Survey and participants** To determine how well the meaning of emoji were conveyed by their images and how much their use on the platform contributed, we asked 57 participants to rate the sentiment of the 105 WeChat emoji. For each emoji, participants were shown the emoji and asked to place it on a two-dimensional plot with valence on the x-axis and energy on the y-axis. Participants went through a four-round familiarization phase at the start of the survey to make sure they understood the task, placing four words (*happy*, *sad*, *peaceful*, and *angry*) as examples of each quadrant of the plot. Participants also rated their familiarity with the emoji on a five-point scale. At the end of the survey, participants self-reported information about their frequency and length of WeChat use, proficiency in Chinese, length of residency in China, and familiarity with Chinese culture. The survey was reviewed by our institution’s IRB process.

Our participants ranged in age from 20 to 72 years old, with a mean age of 32.3. 36 of the participants were WeChat users (mean age 34.3), and 21 of the participants were non-users (mean age 29.0). Unsurprisingly, WeChat use correlated both with participants’ self-reported familiarity with Chinese culture and language proficiency. The mean Chinese cultural familiarity, on a six-point scale, was

<sup>1</sup>We exclude three legacy emoji imported from QQ, WeChat’s predecessor, from our analysis.

4.7 for WeChat users and 2.0 for non-users. The mean Chinese language proficiency, again on a six-point scale, was 3.7 for users and 1.9 for non-users.

**Sentiment dimensions** The sentiment dimensions of valence and energy (also called “arousal”) were chosen based on previous usefulness in word sentiment (Mohammad, 2018).<sup>2</sup>

**Statistical testing** We analyzed the data in two ways. For our initial analysis, we divided the participants into 21 users and 36 non-users of WeChat. The visualizations in Figure 2 show this division, with cross-hairs showing the means and standard deviations for each emoji’s rating. This two-way division provides a simple representation of the different interpretations of the emoji between users and non-users. We performed Mann-Whitney U-tests (with Holm corrections for multiple comparisons) on these categorical splits.

To increase our statistical power, we followed up the above analysis with Kendall rank correlation tests between sentiment ratings and participants’ familiarity with each emoji (again, with Holm corrections). Since not all users are equally familiar with all emoji, and non-users may have some familiarity with some emoji (such as those that match Unicode emoji), this gives us a more detailed sense of how much the interpretation is shaped by experience rather than the strict user/non-user dichotomy.

### 3.2 Results

**Different interpretations** Of the 105 emoji we tested, eight had significant correlations between familiarity with the emoji and the valence ratings of the emoji after a Holm correction for multiple comparisons. Figure 2 shows the difference between user and non-user means for the emoji with significant valence correlations. The top left emoji (clapping) had a significant energy difference in addition to its valence difference.

One clear pattern is that familiarity with an emoji generally led to more extreme ratings; participants gave more neutral ratings to emoji that they were less familiar with. Interestingly and importantly, we see that non-users occasionally misunderstand the sign of the valence. The most extreme example is the top-right emoji in Figure 2, which shows strong but nearly opposite valence ratings between users and non-users of WeChat. This emoji is used

<sup>2</sup>Sentiment work often includes a third dimension, dominance, but we omitted this factor due to its correlations with energy, and to keep the task visually simple for participants.

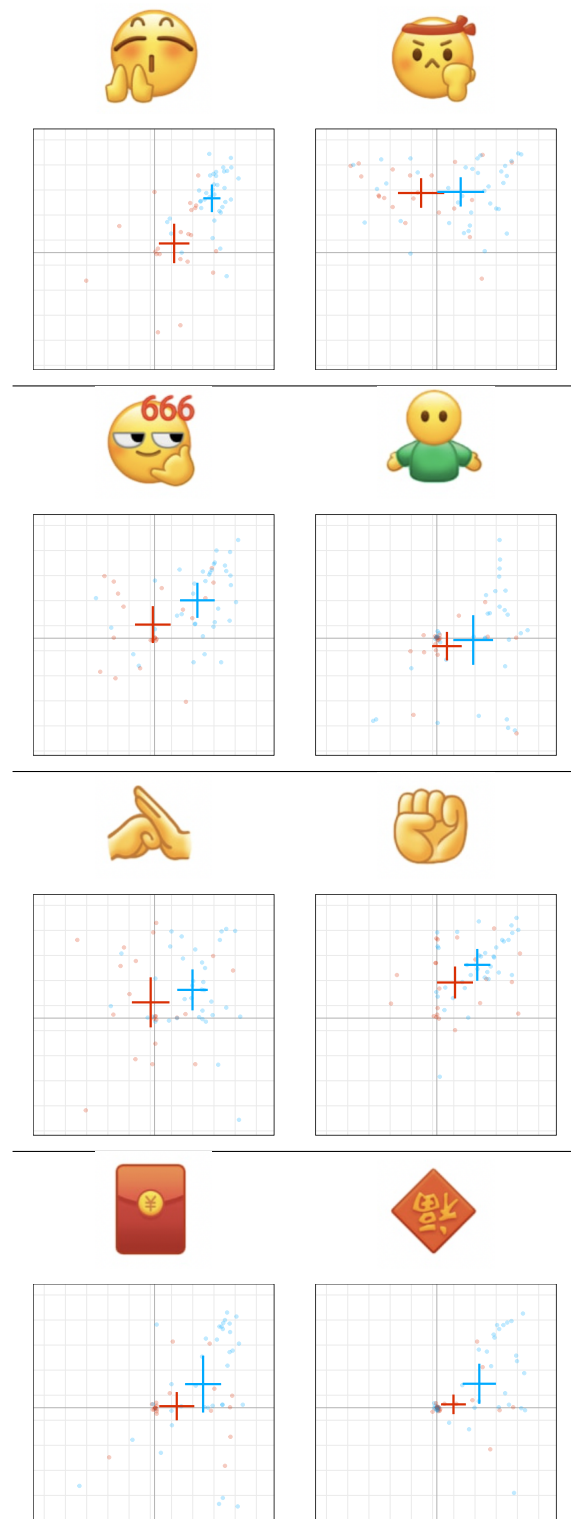


Figure 2: Emoji with significant valence-familiarity correlations. Blue dots are ratings from WeChat users, Red dots from non-users. The crosshairs show the mean and standard deviations for each group’s ratings. The x-axis shows valence ratings; the y-axis shows energy. All emoji had Holm-corrected significances at  $p < .01$  except for the 666-emoji, which had  $p < .05$ .

to express encouragement and likely has relation to the Chinese expression *jīa yóu*, which literally translates to “add oil” but is used as a cheer like in sporting events, or to mark a concerted effort. However, to non-users, the emoji appears to be raising a fist to fight the interlocutor.

Some of these emoji, like *jīa yóu*, require knowledge of Chinese culture to properly assess their meanings. For instance, the “666” emoji in Western cultures is likely to call to mind the Christian devil and negative emotions; however, it also represents Chinese internet slang for “awesome”, explaining some of the interpretive gap.

But not all the emoji misinterpretations can be explained through cultural knowledge. For instance, the green-shirted man is used to signal a hug on WeChat, while non-users find its gesture incomprehensible, regardless of their general Chinese cultural knowledge. Similarly, the clapping face (top-left) and fist (third row, right) lack obvious cultural explanations for the different interpretations; these appear to have platform-specific interpretations that deviate from the information available in their visual forms. Kendall tests found significant correlations between valence and participants’ self-reported Chinese cultural familiarity for only four of these eight emoji (clapping, *jīa yóu*, 666, and fist-in-hand), so cultural knowledge alone is not enough to account for the differences.

**Similar interpretations** The emoji with significant interpretative differences generally show stronger, more positive valence ratings for users than non-users. As a result, one might argue that these differences are nothing more than users being more confident in their ratings than non-users. However, if this were the case, we would expect to see a similar pattern on all the emoji (though not necessarily to statistically-significant levels). To the contrary, we found that users and non-users agree very closely in their interpretation of a number of emoji. This suggests that the deviations in Figure 2 are not just a result of WeChat users growing more confident in their sentiment ratings, but actually developing a richer meaning for some emoji based on their use on WeChat. Figure 3 shows the seven emoji with the smallest distance between the users’ and non-users’ estimates of valence and energy, and illustrates the close agreement between two of them; users and non-users align closely on both mean and standard deviation.

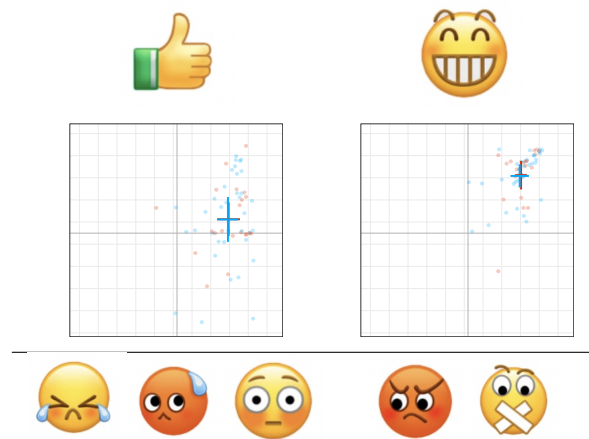


Figure 3: Emoji with the smallest differences between users and non-users, with differences plotted for two of them. The x-axis shows valence ratings; the y-axis shows energy.

## 4 Discussion

Our key finding was that familiarity with an emoji can, in some but not all cases, lead to significant differences in the interpretation of that emoji. Such differences have been found in previous studies of emoji, but the differences were generally tied to changes in the emoji’s form (Tigwell and Flatla, 2016), multiple meanings for the emoji (Miller et al., 2016), or personal experience (Wiseman and Gould, 2018). We found differences even with the same visual forms, and separated from any textual content that might influence the emoji’s meanings.

This suggests that emoji interpretation is not strictly based on form and iconicity, but that emoji develop richer meanings through use on a platform. They are interpreted through a mixture of form and use, similar to words or other linguistic elements, albeit with a much heavier impact of iconic form. As such, the study of emoji interpretation may gain from treating them through a more linguistic lens.

### 4.1 Avoiding misinterpretation

Our results also suggest that the danger of emoji misinterpretation is most acute for new users of a platform, whether because they do not have experience with the range of available emoji, do not know the “dialect” differences between platforms, or simply misinterpret the images the emoji convey.

Social media platforms may wish to acclimate new users to the established meanings of emoji on a platform, to help them avoid misinterpretations. While both Unicode and WeChat emoji have textual



labels, these do not adequately convey nuanced meanings, and alternative methods of conveying richer senses of the emoji sentiments are required.

One possibility is to show users a visualization of the baseline sentiments of an emoji, whether assessed automatically through sentiment analysis of messages containing these emoji or by asking users to participate in a rating experiment like this one.

Another possibility is to provide a “glossary”, with samples of the emoji in use, as dictionaries do with words. This could be curated, with emoji lexicographers selecting sample messages that illustrate the range of meanings, or automatically generated based on recent messages that use the emoji on the platform.

## 4.2 Limitations

The current study was based on primarily American respondents, though approximately half had lived in China for at least a year (including 75% of WeChat users). As a result, our data does not fully reflect the wide range of users or non-users of WeChat. We encountered correlations between our participants’ familiarity with WeChat, the individual WeChat emoji, Chinese culture, and the Chinese language. WeChat use and emoji familiarity had the strongest impacts on emoji interpretation differences, but further work is needed to control for the impacts of these other factors.

Additionally, we limited our analysis to emoji in isolation, as our goal was to assess whether the most basic interpretation of the emoji still relied on experience with the emoji. Emoji are sometimes used by themselves without text, so these results will apply to some real-world usage. But of course, in their general use, emoji tend to appear in richer conversational contexts and are accompanied by other linguistic information. As such, it is not certain that the observed user/non-user sentiment differences will persist for emoji used in conjunction with texts. That said, the assessment of these emoji in isolation can serve as a baseline for future research examining how their sentiment differs in real-world conversations, especially in cases of sarcasm, hyperbole, or irony.

## References

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).

Tianran Hu, Han Guo, Hao Sun, Thuy-vy Nguyen, and Jiebo Luo. 2017. *Spice Up Your Chat: The Intentions and Sentiment Effects of Using Emojis*. *Proceedings of the International AAAI Conference on Web and Social Media*, 11(1):102–111.

Mutsumi Imai, Sotaro Kita, Miho Nagumo, and Hiroyuki Okada. 2008. *Sound symbolism facilitates early verb learning*. *Cognition*, 109(1):54–65.

Linda K. Kaye, Helen J. Wall, and Stephanie A. Malone. 2016. “Turn that frown upside-down”: A contextual account of emoticon usage on different virtual platforms. *Computers in Human Behavior*, 60:463–467.

Petra Kralj Novak, Jasmina Smalović, Borut Sluban, and Igor Mozetič. 2015. *Sentiment of Emojis*. *PLOS ONE*, 10(12):e0144296.

Molly L Lewis and Michael C Frank. 2016. The length of words reflects their conceptual complexity. *Cognition*, 153:182–195.

Li Li and Yue Yang. 2018. *Pragmatic functions of emoji in internet-based communication—a corpus-based study*. *Asian-Pacific Journal of Second and Foreign Language Education*, 3(1):16.

Liyuan Liu, Yen Hsu, Jie Zhang, and Qianling Jiang. 2020. *A study on the embarrassment of senders who missend emojis with opposite meanings on social apps: Taking WeChat as an example*. *Psicologia: Reflexão e Crítica*, 33(1):20.

Shao-Kang Lo. 2008. *The Nonverbal Communication Functions of Emoticons in Computer-Mediated Communication*. *CyberPsychology & Behavior*, 11(5):595–597.

Xuan Lu, Wei Ai, Xuanzhe Liu, Qian Li, Ning Wang, Gang Huang, and Qiaozhu Mei. 2016. *Learning from the ubiquitous language: An empirical analysis of emoji usage of smartphone users*. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 770–780, Heidelberg Germany. ACM.

Hannah Miller, Jacob Thebault-Spieker, Shuo Chang, Isaac Johnson, Loren Terveen, and Brent Hecht. 2016. “Blissfully Happy” or “Ready to Fight”: Varying Interpretations of Emoji. In *Proceedings of the 10th International Conference on Web and Social Media, ICWSM 2016*, pages 259–268.

Saif Mohammad. 2018. *Obtaining Reliable Human Ratings of Valence, Arousal, and Dominance for 20,000 English Words*. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 174–184, Melbourne, Australia. Association for Computational Linguistics.

- Satomi Sugiyama. 2015. [Kawaii meiru and Maroyaka neko: Mobile emoji for relationship maintenance and aesthetic expressions among Japanese teens](#). *First Monday*.
- Jan-Olof Svantesson. 2017. [Sound symbolism: The role of word sound in meaning](#). *WIREs Cognitive Science*, 8(5).
- Michael Szurawitzki. 2022. [The Chinese Messaging Application WeChat as Used by German Speakers in China: Usage Practices, Multilingual Environments, Emojis, and Beyond](#). *Linguistik Online*, 113(1):111–152.
- Ye Tian, Thiago Galery, Giulio Dulcinati, Emilia Molimpakis, and Chao Sun. 2017. [Facebook sentiment: Reactions and emojis](#). In *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media*, pages 11–16, Valencia, Spain. Association for Computational Linguistics.
- Garreth W. Tigwell and David R. Flatla. 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*, pages 859–866, Florence Italy. ACM.
- Sarah Wiseman and Sandy J. J. Gould. 2018. [Repurposing Emoji for Personalised Communication: Why \[Pizza\] means “I love you”](#). In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–10, Montreal QC Canada. ACM.