



OPEN

DATA DESCRIPTOR

Emoji-Dis: A dataset of emojis characterised in 13 discrete emotions

Pilar Ferré¹✉, Juan Haro¹, Miguel Ángel Pérez-Sánchez² & José Antonio Hinojosa^{3,4,5}

Emojis are widely used in digital communication. A key function of these symbols is to convey affective meaning. In order to study emojis scientifically, it is necessary to have normative data characterising them on a number of relevant variables. So far, however, resources in this area are scarce. This article presents subjective norms for 112 emojis in 13 affective dimensions that refer to discrete emotions: anger, disgust, fear, sadness, anxiety, happiness, awe, contentment, amusement, excitement, serenity, relief, and pleasure. This data was collected from 763 Spanish speakers and validated using standard methods applied to existing affective datasets of emojis and words. This is the first normative emoji study to provide data on a large number of positive emotions. Given that most emojis are used to communicate positive emotions, these norms will be of particular value to researchers interested in these graphic icons from a variety of academic fields. The dataset has been made freely available.

Background & Summary

Emojis are graphical icons that represent a wide range of concepts, including facial expressions, objects, living beings, gestures, activities, and symbols (see¹, for a review). They were first introduced in Japan in late 1990s and became popular in 2010 after being standardized in the Unicode version 6.0, and supported by Apple on the iPhone². The original emoji set contained 722 characters. This number has been steadily increasing since then, with 3,790 emojis in the latest Unicode version, released in September 2024 (Full Emoji List, v16.0; <https://unicode.org/emoji/charts/full-emoji-list.html>).

The massive use of emojis in computer-mediated communication (CMC) has greatly stimulated the scientific interest in these graphic icons, which have become a hot topic of research in academic fields as diverse as computer science, communication, marketing, linguistics, and psychology, among others (see¹ and³ for reviews).

Much research in the field of communication has focused on the factors that influence emoji use in CMC, identifying age, gender, psychological traits, mood and culture as relevant variables. Indeed, older users are less likely to use emojis and use fewer emojis compared to younger users⁴. Women have a more positive attitude towards emojis and use them more frequently than men^{5,6}. Furthermore, extroversion⁷ and good mood⁸ are positively associated with the use of emojis. There are also cultural differences, with people from certain cultures tending to use emojis more frequently than others⁹.

Research in semiotics has aimed to identify the semantic, pragmatic and phatic functions of emojis within texts¹⁰. Within this framework, recent theoretical developments have sought to establish the rules that govern the joint construction of meaning by emojis and language¹¹.

Emojis are an important communicative resource used to overcome a clear limitation of CMC compared to face-to-face interactions: The lack of non-verbal cues, such as facial expressions, body language, and prosody. In this sense, emojis have a clear semantic function. Indeed, there is evidence indicating that emojis are semantically integrated with the accompanying text, as the congruence between emojis and the verbal message leads to a facilitated processing^{12–15}. Furthermore, emojis act as semantic cues that help to clarify the meaning of the written message. For example, when communication includes emojis people are better at recognising the

¹Universitat Rovira i Virgili, Department of Psychology and CRAMC, Tarragona, Spain. ²Departamento de Psicología Básica y Metodología, Universidad de Murcia, Murcia, Spain. ³Instituto Pluridisciplinar, Universidad Complutense de Madrid, Madrid, Spain. ⁴Departamento de Psicología Experimental, Procesos Cognitivos y Logopedia, Universidad Complutense de Madrid, Madrid, Spain. ⁵Centro de Investigación Nebrija en Cognición (CINC), Universidad Nebrija, Madrid, Spain. ✉e-mail: mariadelpilar.ferre@urv.cat

conveyed indirect meaning of sentences¹⁶, and at disambiguating ambiguous messages¹⁷. People also show a greater agreement in their ratings of the affective tone of messages when they include emojis².

Another key communicative function of emojis is to convey affective meaning (e.g.^{18–20}). In this regard, emojis facilitate the communication of subtle emotional cues, such as sarcasm²¹, and are used to both intensify and reduce the intensity of messages with an emotional content²². Moreover, these graphic symbols influence the mood of the receiver, with advertisements containing emojis producing a more positive affect in consumers than those that do not contain them²³. This also has social consequences, such that emojis contribute to maintain emotional closeness in CMC²⁴.

The emotional function of emojis has prompted researchers to characterise their affective content, leading to the development of emotional lexicons. These resources describe emojis in terms of their positivity or negativity or in terms of their association with specific emotions (e.g., fear or happiness). In one of the first studies in this line, Novak *et al.*² created the Emoji Sentiment Ranking, in which 751 emojis were classified into positive, negative and neutral. This distribution was obtained by asking a small number of human annotators to label the affective polarity of the tweets in which the emojis appeared. Due to the subjectivity and time-consuming nature of human annotation, other studies have adopted automated methods. For example, Fernández-Gavilanes *et al.*²⁵ used the official definitions in Emojipedia (an online encyclopaedia of emojis) to create an emoji lexicon, while Kimura and Katsuragi²⁶ relied on emotional vectors to automatically assign emotional labels to emojis. These vectors were based on the frequency of co-occurrence between emojis and emotional words (i.e., words that refer to emotions).

As useful as they are, a limitation of these approaches is that they do not consider users' perceptions²⁷. This is a relevant issue, given the existing differences between emoji developers' intended meanings and users' interpretations²⁷; the ambiguity of emoji interpretations, both between people and across platforms²⁸; and the fact that many existing emojis may have more than one interpretation²⁹.

As an alternative, several normative studies have been developed. In these, many users rate the affective properties of large sets of stimuli. This research has mostly adopted a dimensional approach to emotions, which describes emotional states in terms of a few affective dimensions, the most relevant of which are valence (i.e., the hedonic tone of an experience, ranging from highly unpleasant to highly pleasant) and arousal (i.e., the degree of activation elicited by the experience, ranging from highly calming to highly exciting^{30,31}). Within this framework, these two dimensions can be used to describe the affective properties of any stimulus, such as pictures (e.g.^{32,33}) and words (e.g.^{34,35}). This approach has been applied to emojis recently, although only a few normative datasets are available so far. The most comprehensive studies in this regard are those of Ferré *et al.*³⁶, who developed Emoji-SP with Spanish users; Rodrigues *et al.*²⁷, who created the Lisbon Emoji and Emoticon Database (LEED) with Portuguese users; and Scheffler and Nenchev³⁷, who obtained affective ratings from German speakers. All of them collected valence and arousal ratings for a large set of emojis (1031, 253, and 107, respectively). These normative datasets also included other variables such as familiarity, clarity, visual complexity^{27,36,37}, frequency of use^{36,37}, aesthetic appeal and meaningfulness²⁷. Other smaller scale studies collecting valence and arousal ratings have been published in recent years (^{5,19,20,38}), which included a set of 70, 33, 74 and 18 face emojis, respectively).

There is another framework for describing the human affective space, derived from basic (discrete) emotion theories, which posits the existence of a set of biologically determined emotions, with specific neural signatures, cognitive appraisals, facial expressions and behavioural action tendencies (e.g.³⁹). There is variability in the number of emotions proposed, but most authors agree on five discrete emotions: happiness, sadness, anger, fear, and disgust (see⁴⁰). Although much smaller in number than those focusing on valence and arousal, some normative studies have collected data on the association of different types of stimuli, such as words, with these five discrete emotions (e.g.^{41–44}).

As can be seen above, the most widely accepted proposal encompasses four negative emotions (sadness, anger, fear, and disgust) and one positive emotion (happiness). More recent emotion theories, however, emphasise the need for a further differentiation of positive emotions. They conceptualise positive emotions as a set of specific states that facilitate an adaptive response to different types of fitness-relevant resources, and that have their own psychological, behavioural and neurophysiological markers (⁴⁵, see also⁴⁶, for an overview). In light of this, it is important to develop emotional lexicons that cover a wider range of positive emotions. This makes particular sense for emojis. The reason is that most emojis are positive² and are used to communicate positive emotions¹⁷. Two recent studies have moved in this direction. Godard and Holtzman⁴⁷ and Shoeb and de Melo⁴⁸ developed a lexicon of emojis which included surprise and three positive emotions (anticipation, joy and trust) in addition to the typical four negative emotions (anger, disgust, fear and sadness). However, only Godard and Holtzman⁴⁷ provided information on user perception through subjective ratings from a large number of participants. On the other hand, positive emotions are even more varied than those considered in the two aforementioned studies^{45,46}.

We present a normative dataset with subjective ratings of 112 face emojis in 13 discrete emotions: anger, sadness, fear, disgust, anxiety, happiness, awe, contentment, amusement, excitement, serenity, relief, and pleasure. Therefore, the study has been conducted within the framework of basic emotion theories³⁹. Furthermore, in line with recent theoretical proposals⁴⁵, we included eight positive emotions to capture variability in positive emotional experiences. We selected these particular emotions because they have been included in some normative studies focusing on words (e.g.⁴⁶). This could potentially facilitate future research involving both types of stimuli. Of note, Emoji-SP³⁶ contains all these emoji, so there is also data available on valence, arousal, visual complexity, familiarity, frequency of use, and clarity. This resource will therefore be very useful for a more complete description of the affective properties of emojis, combining the dimensional³⁶ and basic (current data) emotion approaches. It will be also useful for research with emojis in general. Some areas of application include sentiment analysis, social media studies, and chatbot design. For example, the inclusion of emojis has been shown to

| Variable | Mean | SD | Min | Max |
|-------------|------|------|------|------|
| Amusement | 2.23 | 1.03 | 1.07 | 4.72 |
| Anger | 1.61 | 0.80 | 1.00 | 4.73 |
| Anxiety | 1.89 | 0.77 | 1.03 | 3.80 |
| Awe | 1.98 | 0.66 | 1.14 | 3.95 |
| Contentment | 2.43 | 0.90 | 1.29 | 4.36 |
| Disgust | 1.55 | 0.59 | 1.00 | 4.54 |
| Excitement | 2.11 | 1.03 | 1.03 | 4.59 |
| Fear | 1.60 | 0.64 | 1.00 | 3.60 |
| Happiness | 2.29 | 1.19 | 1.06 | 4.69 |
| Pleasure | 2.26 | 0.96 | 1.10 | 4.28 |
| Relief | 2.16 | 0.91 | 1.13 | 4.00 |
| Sadness | 1.71 | 0.86 | 1.00 | 4.72 |
| Serenity | 2.33 | 0.90 | 1.15 | 4.32 |

Table 1. Descriptive statistics of the variables.

improve the accuracy of sentiment analysis^{49,50}. In a different field, analysing social media posts can reveal differences in emoji usage relating to gender, age⁵¹ and personality⁵². Notably, emoji use not only provides information about the sender, but also influences how the receiver perceives him/her⁵³, even when the sender is not human. Indeed, some studies have demonstrated that using emojis in human–chatbot interactions can enhance the chatbot’s perceived credibility, competence, social attractiveness⁵⁴ and warmth⁵⁵. The dataset developed in this study can benefit all these fields, as researchers can use it to select emojis that are well characterised in terms of their affective and non-affective properties. This will contribute to the rigour of studies and the validity of conclusions.

Methods

Data were obtained from 763 participants recruited through Prolific and three Spanish universities: Universidad de Murcia (Murcia, Spain), Universitat Rovira i Virgili (Tarragona, Spain), and Universidad Complutense de Madrid (Madrid, Spain). Of these, 504 were female (66% of the sample) and 259 were male (34%). The mean age was 24.07 years ($SD = 9.68$; range = 17–70). Participants were recruited via the official communication channels of the universities’ teachers and the Prolific platform. They received either academic credit (university students) or monetary compensation (Prolific participants) for their participation. They signed an informed consent document prior to the study. The study was approved by the Ethics Committee of the University of Rovira i Virgili (reference CEIPSA-2021-PR-0044).

We selected 112 emojis from the Unicode Emoji 13 version, published by the Unicode Foundation (<http://unicode.org/emoji/charts/full-emoji-list.html>). All emojis represented faces, most of them anthropomorphic, although there were also some faces of animals (cats, monkeys) and of different creatures (e.g., demons, aliens, ghosts). We chose faces that potentially expressed an emotion.

The emojis were presented in their Facebook version. Participants rated them on 13 affective dimensions that refer to discrete emotions: anger, disgust, fear, sadness, anxiety, happiness, awe, contentment, amusement, excitement, serenity, relief, and pleasure. Participants were given the name of the emotion and its definition. These definitions were taken from⁴² and⁴⁶. The instructions and definitions of the emotions can be found in the OSF repository at <https://doi.org/10.17605/OSF.IO/WV7UB>⁵⁶. Participants had to indicate the extent to which the emoji reflected the discrete emotion they were rating. They were provided with a 5-point Likert scale, ranging from “not at all” to “extremely”, to make their ratings. Participants also had the option “I don’t know the emoji” to indicate that they did not know a particular emoji.

Thirteen questionnaires were created, one for each discrete emotion. The 112 emojis appeared randomly in each questionnaire, in PNG format of size 72×72 . The questionnaires were created and administered online using TestMaker⁵⁷. The average time to complete each questionnaire was approximately 10 minutes. Each participant completed one or more questionnaires.

Data cleaning was carried out prior to analysis. One-hundred and three responses were removed. We removed 49 responses whose ratings showed a low correlation with the mean of all ratings on the same questionnaire (i.e., $r < 0.1$). Correlations close to zero are assumed to reflect idiosyncratic response patterns, whereas negative correlations indicate that the participant understood the rating scale in the reverse order. We also removed 54 responses from those participants who completed the same questionnaire twice. There was an average of 79.92 valid responses per questionnaire (range = 69–90, $SD = 6.24$), and each emoji received an average of 68.69 valid ratings (range = 65–90, $SD = 6.23$).

Data Records

The dataset is available in XLSX (“emojis_discrete_emotions_database.xlsx”) and PDF (“emojis_discrete_emotions_database.pdf”) format in the OSF repository: <https://doi.org/10.17605/OSF.IO/WV7UB>⁵⁶. Each row contains data for a specific emoji, and each column corresponds to a specific variable. The **id** column contains a unique identifier code for each emoji in the database. This is followed by the ASCII code (**ascii_code** column). The **emoji** column represents the visual form of the emoji in their standard Unicode form. The **category** column classifies each emoji into a specific category for organisation purposes. The dataset contains ratings for 13

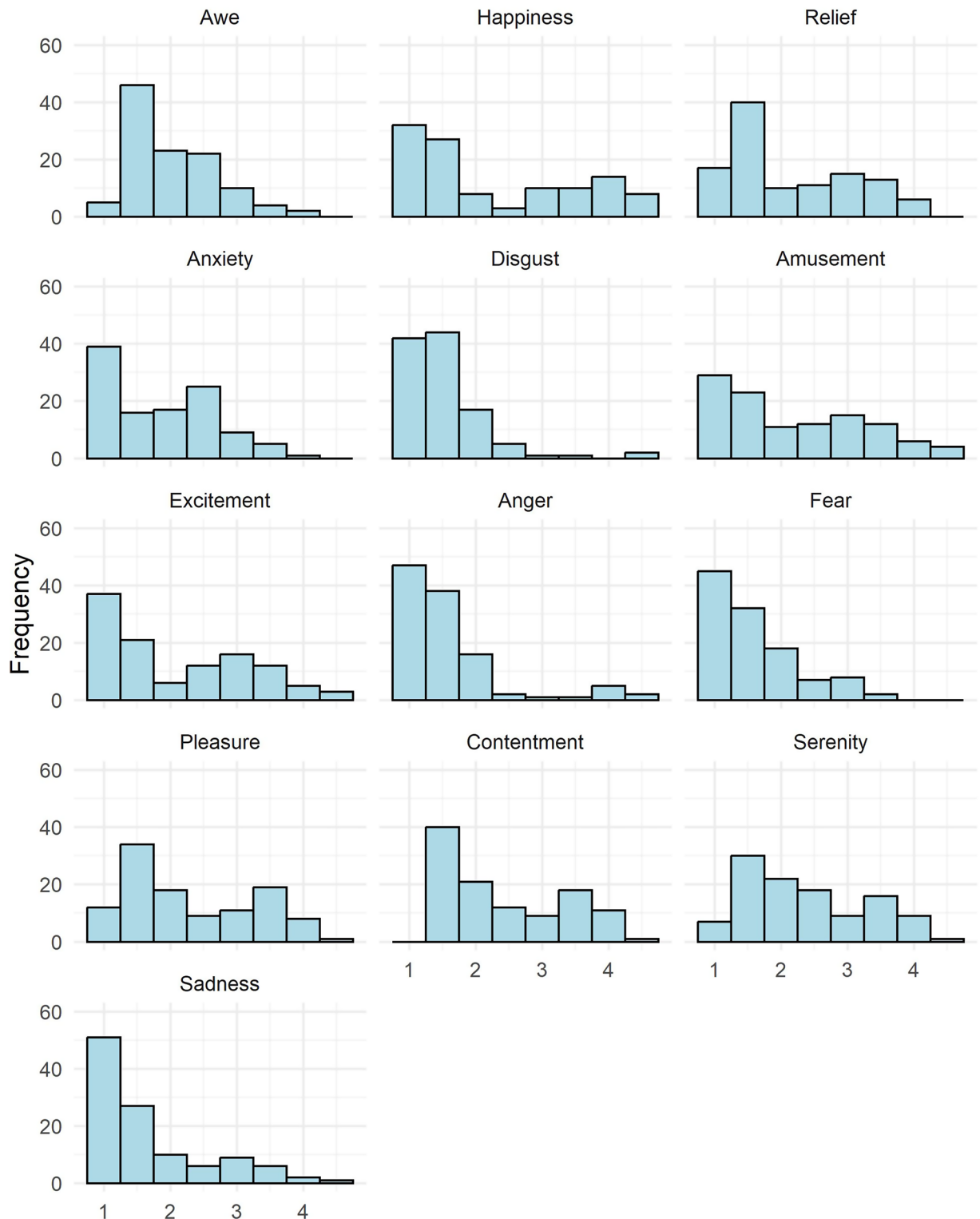


Fig. 1 Histograms with the distribution of the ratings of the variables.

discrete emotions (rated from 1 to 5, see Methods), with the corresponding columns for mean ratings (`[emotion]`), standard deviation (`[emotion]_sd`), number of valid responses (`[emotion]_n`), and number of 'don't know' responses (`[emotion]_dk`). This structure is repeated for the thirteen emotions included in the database: anger, disgust, fear, sadness, anxiety, happiness, awe, contentment, amusement, excitement, serenity, relief, and pleasure.

An XLSX file containing the raw data for the 13 discrete emotions is also available in the OSF repository. The file ("raw_data.xlsx") consists of 13 sheets, each corresponding to a discrete emotion. Each column represents

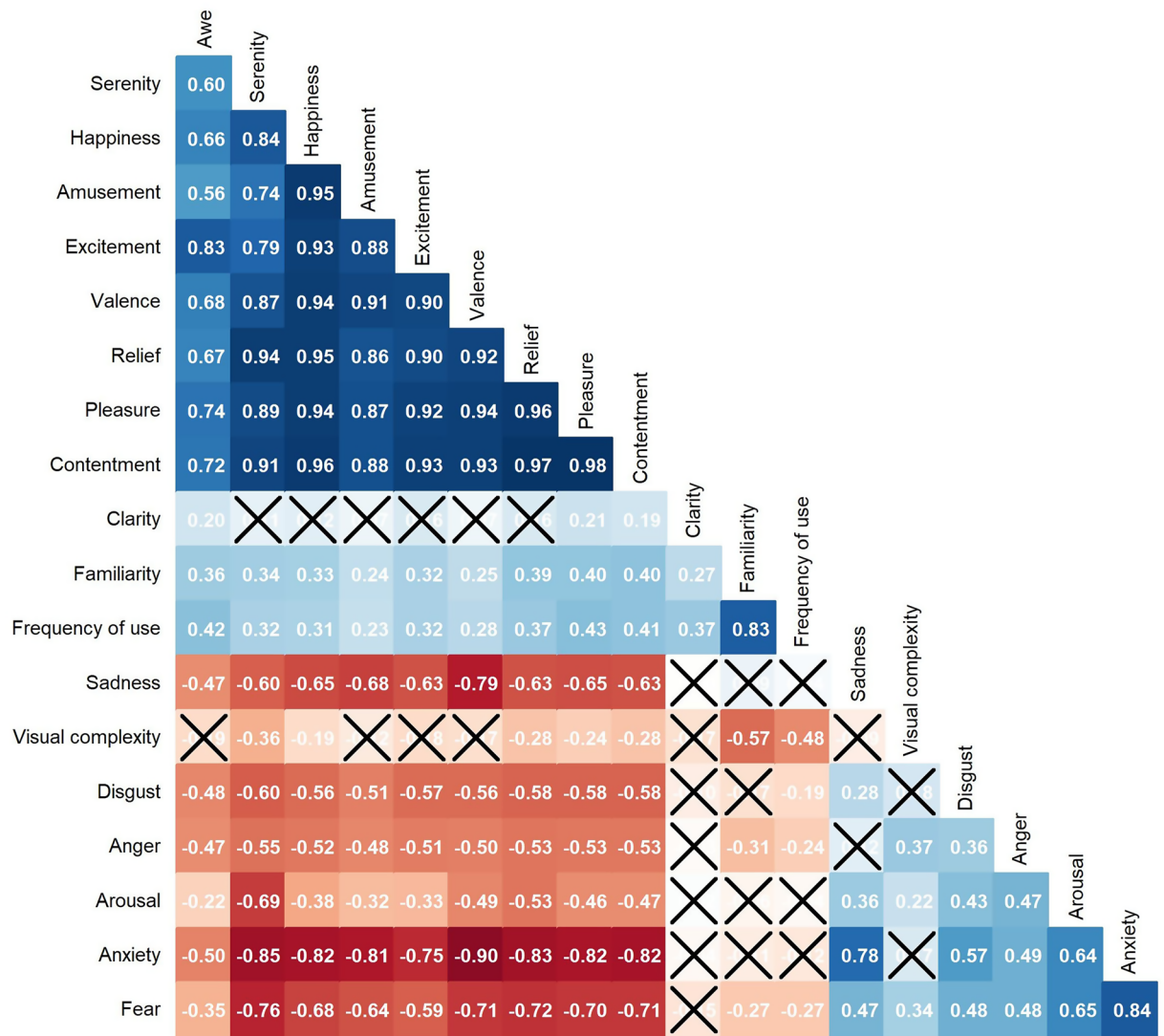


Fig. 2 Correlogram showing bivariate correlations between the variables. Positive correlations are shown in blue, with darker shades indicating stronger relationships, while negative correlations are shown in red. The “X” in the matrix indicates correlations where the p-value is greater than 0.05. All other correlations have p-values less than 0.05.

an emoji (identified by its id), while each row contains the participant’s rating for each emoji on a scale from 1 to 5, with blank cells indicating “don’t know” responses.

Technical Validation

We measured the interrater reliability of each questionnaire by means of the intraclass correlation coefficient (ICC⁵⁸). Two-way random effects based on the absolute agreement of multiple raters (2,k) were used. The ICCs were all statistically significant (all $p_s < .001$; $M = 0.97$, $SD = 0.01$, range = 0.96–0.98), indicating excellent support for the reliability of the data.

Descriptive statistics and distributions of the variables included in the dataset are presented in Table 1 and Fig. 1. All the discrete emotions are negatively skewed, with mean scores below 2.5, which places them below the midpoint of the measurement scale (see Table 1). Nevertheless, each discrete emotion contains emojis with mean scores across the full range of the scale. These results are consistent with those of previous word datasets on discrete emotions^{42,44,46}.

Bivariate correlations between the variables are shown in Fig. 2. We included in these correlations the ratings of visual complexity, familiarity, frequency of use, clarity, emotional valence and emotional arousal from the normative study by Ferré *et al.*³⁶. Discrete negative emotions (anger, fear, disgust, anxiety, sadness) show a strong negative correlation with valence and a moderate positive correlation with arousal. In contrast, discrete positive emotions (amusement, awe, excitement, happiness, pleasure, relief, contentment, serenity) show a strong positive correlation with valence and a moderate negative correlation with arousal. Furthermore, strong positive correlations are observed between variables associated with negative emotions and similarly between

those associated with positive emotions. These results are very similar to those reported by Hinojosa *et al.*⁴⁶, who assessed the same discrete emotions using words. The consistency of these results underscores the robustness of the patterns observed in this emoji dataset.

Limitations and future studies. This is the first large-scale normative dataset of emojis covering several positive emotions. One limitation, however, is that these data have been gathered only from Spanish people and from a larger proportion of women than men. Previous normative studies of emojis using a dimensional approach show a high correlation between valence and arousal ratings across different countries^{36,37}. This suggests that the perception of the emotional properties of emojis is highly consistent. Therefore, a similar degree of correspondence can be expected for basic emotion ratings, such as those included in this dataset. However, individual differences (e.g., gender and age) and cultural differences may exist, and these are worth investigating in further research. Future studies should also cover a wider range of emotions and more complex emotional states that can be associated with emoji use (e.g.⁵⁹).

Code availability

No custom code was used. ICC: Function ICC, from the *psych* package version 2.4.6. Descriptive statistics: Function *summarise_all*, from the *dplyr* package version 1.1.2. Histograms: Function *ggplot*, from the *ggplot2* package version 3.5.1. Correlations: Function *corrplot*, from the *corrplot* package version 0.95. Software: R version 4.3.1 and RStudio version 2023.06.1 + 524were used.

Received: 20 December 2024; Accepted: 24 July 2025;

Published online: 29 July 2025

References

- Bai, Q., Dan, Q., Mu, Z. & Yang, M. A systematic review of emoji: Current research and future perspectives. *Front. Psychol.* **10**, 2221, <https://doi.org/10.3389/fpsyg.2019.02221> (2019).
- Kralj Novak, P., Smailović, J., Sluban, B., Mozetič, I. Sentiment of Emojis. *PLoS One* **10**, <https://doi.org/10.1371/journal.pone.0144296> (2015).
- Tang, Y. & Hew, K. F. Emoticon, Emoji, and Sticker Use in Computer-Mediated Communication: A Review of Theories and Research Findings. *Int. J. Commun.* **13**, 2457–2483 (2019).
- Boutet, I., Goulet-Pelletier, J.-C., Sutera, E. & Meinhardt-Injac, B. Are older adults adapting to new forms of communication? A study on emoji adoption across the adult lifespan. *Comput. Human. Behav.* **13**, 100379 (2024).
- Jones, L. L., Wurm, L. H., Norville, G. A. & Mullins, K. L. Sex differences in emoji use, familiarity, and Valence. *Comput. Human. Behav.* **108**, 106305 (2020).
- Prada, M. *et al.* Motives, frequency and attitudes toward emoji and emoticon use. *Telemat. Inform.* **35**, 1925–1934 (2018).
- Hall, J. A. & Pennington, N. Self-monitoring, honesty, and cue use on Facebook: The relationship with user extraversion and conscientiousness. *Comput. Human. Behav.* **29**, 1556–1564 (2013).
- Konrad, A., Herring, S. C. & Choi, D. Sticker and emoji use in Facebook messenger: Implications for graphicon change. *J. Comput. Mediat. Commun.* **25**, 217–235 (2020).
- Lin, T.-J. & Chen, C.-H. A preliminary study of the form and status of passionate affection emoticons. *Int. J. Des.* **12**, 75–90 (2018).
- Danesi, M. *The Semiotics of Emoji*. (Bloomsbury, London, 2017).
- Logi, L. & Zappavigna, M. A social semiotic perspective on emoji: How emoji and language interact to make meaning in digital messages. *New Media Soc.* **25**, 3222–3246 (2023).
- Barach, E., Feldman, L. B. & Sheridan, H. Are emojis processed like words?: Eye Movements reveal the time course of semantic processing for emoji-fied text. *Psychon. Bull. Rev.* **28**, 978–991 (2021).
- Pfeifer, V. A., Armstrong, E. L. & Lai, V. T. Do all facial emojis communicate emotion? the impact of facial emojis on perceived sender emotion and text processing. *Comput. Human. Behav.* **126**, 107016 (2022).
- Robus, C. M., Hand, C. J., Filik, R. & Pitchford, M. Investigating effects of emoji on neutral narrative text: Evidence from eye movements and perceived emotional valence. *Comput. Human. Behav.* **109**, 106361 (2020).
- Weissman, B., Engelen, J., Baas, E., Cohn, N. The lexicon of emoji? conventionality modulates processing of emoji. *Cogn. Sci.* **47**, <https://doi.org/10.1111/cogs.13275> (2023).
- Holtgraves, T., Robinson, C. Emoji can facilitate recognition of conveyed indirect meaning. *PLoS One* **15**, <https://doi.org/10.1371/journal.pone.0232361> (2020).
- Riordan, M. A. The communicative role of non-face emojis: Affect and disambiguation. *Comput. Human. Behav.* **76**, 75–86 (2017).
- Cherbonnier, A. & Michinov, N. The recognition of emotions beyond facial expressions: Comparing emoticons specifically designed to convey basic emotions with other modes of expression. *Comput. Human. Behav.* **118**, 106689 (2021).
- Jaeger, S. R., Roigard, C. M., Jin, D., Vidal, L. & Ares, G. Valence, arousal and sentiment meanings of 33 facial emoji: Insights for the use of emoji in consumer research. *Food Res. Int.* **119**, 895–907 (2019).
- Kutsuzawa, G., Umemura, H., Eto, K., Kobayashi, Y. Classification of 74 facial emoji's emotional states on the valence-arousal axes. *Sci. Rep.* **12**, <https://doi.org/10.1038/s41598-021-04357-7> (2022).
- García, C., Turcan, A., Howman, H. & Filik, R. Emoji as a tool to aid the comprehension of written sarcasm: Evidence from younger and older adults. *Comput. Human. Behav.* **126**, 106971 (2022).
- Sampietro, A. Emoji and Rapport Management in spanish WhatsApp chats. *J. Pragmat.* **143**, 109–120 (2019).
- Das, G., Wiener, H. J. D. & Kareklas, I. To emoji or not to emoji? examining the influence of emoji on consumer reactions to advertising. *J. Bus. Res.* **96**, 147–156 (2019).
- Gibson, W., Huang, P. & Yu, Q. Emoji and communicative action: The semiotics, sequence and gestural actions of 'face covering hand'. *Discourse Context Media* **26**, 91–99 (2018).
- Fernández-Gavilanes, M., Juncal-Martínez, J., García-Méndez, S., Costa-Montenegro, E. & González-Castaño, F. J. Creating Emoji Lexica from unsupervised sentiment analysis of their descriptions. *Expert Syst. Appl.* **103**, 74–91 (2018).
- Kimura, M., Katsurai, M. Automatic construction of an emoji sentiment lexicon. *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, 1033–1036, <https://doi.org/10.1145/3110025.3110139> (2017).
- Rodrigues, D., Prada, M., Gaspar, R., Garrido, M. V. & Lopes, D. Lisbon emoji and Emoticon Database (LEED): Norms for emoji and emoticons in seven evaluative dimensions. *Behav. Res. Methods* **50**, 392–405 (2018).
- Tigwell, G. W., Flatla, D. R. Oh that's what you meant! *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, 859–866, <https://doi.org/10.1145/2957265.2961844> (2016).

29. Cześtochowska, J. *et al.* On the context-free ambiguity of emoji. *Proceedings of the International AAAI Conference on Web and Social Media*. **16**, 1388–1392, <https://doi.org/10.1609/icwsm.v16i1.19393> (2022).
30. Bradley, M. M. & Lang, P. J. *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings* (Citeseer, 1999).
31. Russell, J. A. Core affect and the psychological construction of emotion. *Psychol. Rev.* **110**, 145–172 (2003).
32. Kurdi, B., Lozano, S. & Banaji, M. R. Introducing the open affective standardized image set (OASIS). *Behav. Res. Methods* **49**, 457–470 (2017).
33. Marchewka, A., Żurawski, K. & Jednoróg, A. Grabowska, The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, Realistic Picture Database. *Behav. Res. Methods* **46**, 596–610 (2014).
34. Ferré, P., Guasch, M., Moldovan, C. & Sánchez-Casas, R. Affective norms for 380 Spanish words belonging to three different semantic categories. *Behav. Res. Methods* **44**, 395–403 (2012).
35. Warriner, A. B., Kuperman, V. & Brysbaert, M. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav. Res. Methods* **45**, 1191–1207 (2013).
36. Ferré, P., Haro, J., Pérez-Sánchez, M. Á., Moreno, I. & Hinojosa, J. A. Emoji-SP, the Spanish emoji database: Visual complexity, familiarity, frequency of use, clarity, and emotional valence and arousal norms for 1031 Emojis. *Behav. Res. Methods* **55**, 1715–1733 (2023).
37. Scheffler, T. & Nenchev, I. Affective, semantic, frequency, and descriptive norms for 107 face emojis. *Behav. Res. Methods* **56**, 8159–8180 (2024).
38. Fischer, B. & Herbert, C. Emoji as affective symbols: Affective judgments of emoji, emoticons, and human faces varying in emotional content. *Front. Psychol.* **12**, 645173, <https://doi.org/10.3389/fpsyg.2021.645173> (2021).
39. Ekman, P. Facial expression and emotion. *Am. Psychol.* **48**, 384–392 (1993).
40. Briesemeister, B. B., Kuchinke, L. & Jacobs, A. M. Discrete emotion norms for nouns: Berlin affective word list (denn-bawl). *Behav. Res. Methods* **43**, 441–448 (2011).
41. Čoso, B., Guasch, M., Bogunović, I., Ferré, P. & Hinojosa, J. A. Crowd-5E: A Croatian psycholinguistic database of affective norms for five discrete emotions. *Behav. Res. Methods* **55**, 4018–4034 (2023).
42. Ferré, P., Guasch, M., Martínez-García, N., Fraga, I. & Hinojosa, J. A. Moved by words: Affective ratings for a set of 2,266 Spanish words in five discrete emotion categories. *Behav. Res. Methods* **49**, 1082–1094 (2017).
43. Hinojosa, J. A. *et al.* Affective norms of 875 Spanish words for five discrete emotional categories and two emotional dimensions. *Behav. Res. Methods* **48**, 272–284 (2016).
44. Stadthagen-González, H., Ferré, P., Pérez-Sánchez, M. A., Imbault, C. & Hinojosa, J. A. Norms for 10,491 Spanish words for five discrete emotions: Happiness, disgust, anger, fear, and sadness. *Behav. Res. Methods* **50**, 1943–1952 (2018).
45. Shiota, M. N. *et al.* Beyond happiness: Building a science of discrete positive emotions. *Am. Psychol.* **72**, 617–643 (2017).
46. Hinojosa, J. A. *et al.* The bright side of words: Norms for 9000 Spanish words in seven discrete positive emotions. *Behav. Res. Methods* **56**, 4909–4929 (2024).
47. Godard, R. & Holtzman, S. The multidimensional lexicon of emojis: A new tool to assess the emotional content of Emojis. *Front. Psychol.* **13**, 921388, <https://doi.org/10.3389/fpsyg.2022.921388> (2022).
48. Shoeb, A. A., de Melo, G. EMOTAG1200: Understanding the association between emojis and emotions. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, <https://doi.org/10.18653/v1/2020.emnlp-main.720> (2020).
49. Liu, C. *et al.* Improving sentiment analysis accuracy with emoji embedding. *J. Saf. Sci. Resil.* **2**, 246–252 (2021).
50. Yoo, B. & Rayz, J. T. Understanding Emojis for Sentiment Analysis. *The International FLAIRS Conference Proceedings* **34**, <https://doi.org/10.32473/flairs.v34i1.128562> (2021).
51. Koch, T. K., Romero, P. & Stachl, C. Age and gender in language, emoji, and emoticon usage in instant messages. *Comput. Human Behav.* **126**, 106990 (2022).
52. Kennison, S. M., Fritz, K., Hurtado Morales, M. A. & Chan-Tin, E. Emoji use in social media posts: relationships with personality traits and word usage. *Front. Psychol.* **15**, 1343022, <https://doi.org/10.3389/fpsyg.2024.1343022> (2024).
53. Cavalheiro, B. P., Prada, M. & Rodrigues, D. L. Examining the effects of reciprocal emoji use on interpersonal and communication outcomes. *J. Soc. Pers. Relatsh.* **41**, 2147–2168 (2024).
54. Beattie, A., Edwards, A. P. & Edwards, C. A Bot and a Smile: Interpersonal Impressions of Chatbots and Humans Using Emoji in Computer-mediated Communication. *Commun. Stud.* **71**, 409–427 (2020).
55. Yu, S. & Zhao, L. Emojifying chatbot interactions: An exploration of emoji utilization in human-chatbot communications. *Telemat. Inform.* **86**, 102071 (2024).
56. *Emoji-Dis: A dataset of emojis characterised in 13 discrete emotions.* <https://doi.org/10.17605/OSF.IO/WV7UB> (2024).
57. Haro, J. *Testmaker: Aplicación para crear cuestionarios online*, (available at <http://jharo.net/dokuwiki/testmaker>) (2012).
58. Koo, T. K. & Li, M. Y. A guideline of selecting and reporting intraclass correlation coefficients for Reliability Research. *J. Chiropr. Med.* **15**, 155–163 (2016).
59. Buono, P., De Carolis, B., D'Errico, E., Macchiarulo, N. & Palestra, G. Assessing student engagement from facial behavior in on-line learning. *Multimed. Tools Appl.* **82**, 12859–12877 (2023).

Acknowledgements

Grants PID2021-125842NB-I00 and PID2023-149606NB-I00, funded by MICIU/AEI/10.13039/501100011033 and FEDER, UE. Grant 101182959 from the Horizon Europe Framework Programme, HORIZON-MSCA-2023-SE-01. Grant 2023PFR-URV-00196, funded by URV.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication. More specifically, P.F.: Conceptualization, project administration, supervision, funding acquisition, writing original draft and writing review; J.H.: Conceptualization, methodology, project administration, data curation, writing original draft; M.A.P.: Conceptualization, methodology, writing review; J.A.H.: Conceptualization, methodology, writing review.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Additional information

Correspondence and requests for materials should be addressed to P.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025