Quantifying Emotional Specificity and Ambiguity in Emojis: An Entropy Based Analysis of Discrete Emotion Ratings

Anonymous Author(s)
Affiliation
Address
`email`

Abstract

Emojis are ubiquitous in digital communication, yet their emotional meanings are often ambiguous. A recently released normative data set provides mean ratings of 112 emojis on 13 discrete emotions in Spanish speakers. Based on studies demonstrating that many emoji do not unambiguously depict a single emotion, we introduce an entropy-based *emotion specificity index* (ESI) to quantify how concentrated the ratings of an emoji in one emotion. After baseline correction, we compute Shannons entropy across the 13 emotion ratings and normalize it by the maximum possible entropy. Low values of ESI indicate ambiguous or neutral emojis, whereas high values reflect a strong association with a single emotion. Our analyzes reveal that negative emojis show greater specificity than positive or neutral ones, that the principal component analysis recovers a valence continuum explaining nearly 73 % of the variance, and that ESI is systematically related to affective valence. We discuss applications of the ESI in marketing, health communication, and mental health monitoring and situate our findings within emerging normative datasets and crosscultural research on emoji interpretation.

1 Introduction

Emojis are pictorial symbols used to accompany written language on social platforms, in messaging applications and email. They enrich text by reinforcing tone, expressing emotions or substituting words, yet their interpretation is far from straightforward. Psychological experiments have demonstrated that many emoji are ambiguous because they do not symbolise a single emotion and instead require contextual cues for disambiguation (2; 7). A crowdsourcing study showed that only about 1 % of emoji are interpreted consistently across users, whereas roughly 4 % are as ambiguous as random words (3). At the same time, the concept of *emodiversity*the variety and relative abundance of emotions experienced by an individualhas been formalised using Shannons entropy (4). While emodiversity quantifies the diversity of an individuals emotional life, it has not been applied to pictorial symbols.

Normative datasets provide critical reference values for research on emotion processing. The *EmojiDis* database reports mean ratings of 112 emojis on thirteen discrete emotions in a large sample of Spanish speakers (1). The ratings exhibit the expected structure: positive emotions correlate strongly with each other and negatively with negative emotions, and principalcomponent analysis reveals a dominant valence dimension (1). Complementary datasets have appeared recently. Scheffler and Nenchev collected affective, semantic and descriptive norms for 107 face emojis in German speakers and replicated the quadratic relationship between valence and arousal and found that subjective familiarity correlates strongly with usage frequency and positively with valence and clarity

(5). The EmojiSP dataset provides norms for 1031 emojis across six dimensions and shows that positively valenced emojis are more familiar and frequently used than negative ones, again replicating a Ushaped valencearousal relationship (6).

Beyond normative data, research has examined how context, culture and individual differences shape emoji interpretation. Aldunate and colleagues found that perceived mood in ambiguous messages tends to be negative regardless of emoticon valence and that negative mood perception is especially pronounced when positive emoticons accompany negative text; response times are slower for incongruent messages, indicating that emoticon valence interacts with message valence during disambiguation (7). Chen et al. investigated how gender, age and culture influence emoji comprehension and reported that United Kingdom participants were more accurate than Chinese participants for most emotions except disgust; cultural differences were only partly mediated by familiarity, and Chinese participants sometimes use the smile emoji sarcastically (8). Dynamic or animated emojis introduce additional dimensions: a recent Frontiers study showed that rhythmic motion increases arousal for all dynamic emojis and that motion effects on valence depend on the emotion category, recommending that rhythm and motion be considered when designing animated emojis (9).

Emoji usage also depends on personality traits and social context. Liu and Sun found that shyness, neuroticism, extraversion and agreeableness correlate with different reasons for using emojis or stickers and that people adjust their frequency of use depending on the audience and conversation type (10). Kennison and colleagues analysed emoji use in Twitter posts and discovered that users who deploy the most emojis have the lowest openness to experience, while emoji use was unrelated to other Big Five traits; frequent emoji users also employed more words relating to family, positive emotion and sadness (11). In marketing, a systematic review concluded that emojis attract attention, stimulate social interaction, enhance consumers experiences and influence purchase decisions (12). And in medicine, the animated emoji scale (AES) for dental anxiety showed strong correlations with established scales and was preferred by 74.5 % of children, underscoring the potential of emojis in health assessment (16). These findings highlight the need for a quantitative measure of emotional specificity to guide the selection and design of emojis across applications.

We therefore ask: *How specific are individual emojis with respect to discrete emotions?* Using the EmojiDis dataset we introduce an entropybased emotion specificity index (ESI) and explore its relationship with affective valence. We hypothesise that emojis depicting clear negative expressions (e.g., disgust or anger) will exhibit high specificity, whereas neutral or skeptical faces will be ambiguous. We also consider how these findings inform realworld applications (Figure 1).

2 Methods

2.1 Dataset

We analysed the publicly available *EmojiDis* database (1). Each row corresponds to a unique emoji and includes its Unicode code point, category label (e.g., *facesmiling*, *faceneutralskeptical*) and mean ratings on thirteen discrete emotions: anger, disgust, fear, sadness, anxiety, happiness, awe, contentment, amusement, excitement, serenity, relief and pleasure. Ratings were obtained from 763 Spanish speakers on a 15 Likert scale. We subtracted 1 from each mean rating to treat the lower bound as a neutral baseline and clipped negative values to zero.

2.2 Emotion specificity index

Let r_i denote the baselineadjusted mean rating of an emoji on emotion i, with $i \in \{1, \ldots, 13\}$. We define the probability $p_i = r_i / \sum_j r_j$ over all nonzero adjusted ratings and compute Shannons entropy $H = -\sum_i p_i \log p_i$. Following emodiversity research (4), the *emotion specificity index* is

$$\mathrm{ESI} = 1 - \frac{H}{\log 13}, \tag{1}$$

where $\log 13$ is the maximum entropy for thirteen equally likely emotions. Low values of ESI indicate that ratings are spread across many emotions (ambiguity), whereas high values indicate concentration in a single emotion.

2.3 Valence index and principalcomponent analysis

To situate each emoji on a positivenegative continuum we defined a *valence index* as the average of the eight positive emotions (amusement, awe, excitement, happiness, pleasure, relief, contentment, serenity) minus the average of the five negative emotions (anger, disgust, fear, anxiety, sadness). Positive values denote positive affect; negative values denote negative affect. We standardised the thirteen emotion variables and performed principalcomponent analysis (PCA) to identify latent dimensions.

2.4 Correlation analysis and software

Pearson correlation coefficients were computed among the thirteen emotion ratings. Analyses were carried out in Python using `pandas`, `numpy` and `scikitlearn`.

3 Results

3.1 Discreteemotion structure and valence dimension

Consistent with earlier reports (1), the correlation matrix exhibited strong positive correlations among positive emotions and strong negative correlations between positive and negative emotions (see Figure 6). PCA revealed that the first principal component explained 72.8 % of the variance and loaded positively on all positive emotions and negatively on all negative emotions. This component captures an affective valence continuum (Figures 2 and 3). The second component loaded heavily on anger and sadness and accounted for 7.4 % of the variance.

3.2 Emotion specificity index distribution

The ESI ranged from roughly 0.015 to 0.233 (mean 0.12). Negative emojis displayed higher specificity than positive or neutral emojis. Table 1 lists the five most specific and five most ambiguous emojis. Face vomiting (🤮) and angry face with horns (👿) exhibited high ESI values and were strongly associated with *disgust* and *anger*, respectively. Conversely, grimacing face (😬) and zippermouth face (🤐) had very low ESI and were associated with *anxiety* or *anger*, indicating high ambiguity.

Table 1: Five most specific and five most ambiguous emojis according to the emotion specificity index (ESI). High specificity values indicate concentration of ratings in a single emotion; low values indicate ambiguity. Dominant emotions correspond to the highest adjusted rating.

Emoji	Category	ESI	Dominant emotion
🤮	faceunwell	0.23	disgust
👿	facenegative	0.23	anger
👺	facecostume	0.22	anger
😣	faceconcerned	0.22	anxiety
😖	faceconcerned	0.22	anxiety
😬	faceconcerned	0.02	anxiety/anger
🤐	faceneutralskeptical	0.03	anxiety/anger
🙈	monkeyface	0.04	anxiety
🙄	faceneutralskeptical	0.05	disbelief
🤔	facehand	0.05	curiosity

3.3 Relationship between ESI and valence

ESI values were negatively correlated with the valence index ($r = -0.46$), indicating that more negative emojis tend to convey specific emotions. A scatter plot of emojis in the PC1PC2 plane coloured by valence index (Figure 2) shows that negative emojis cluster on the left, whereas positive emojis cluster on the right. When the same plot is coloured by ESI (Figure 3), highspecificity emojis appear primarily among the negative cluster, whereas ambiguous emojis span the central and positive regions.

4 Discussion

4.1 Interpreting the emotion specificity index

Our entropybased ESI provides a quantitative measure of how clearly an emoji conveys a discrete emotion. High specificity implies a concentrated emotion profile and low ambiguity. Negative emojis, especially those representing anger and disgust, exhibit high specificity. This pattern may reflect the distinct facial configurations associated with negative emotions and the stronger evolutionary pressures on recognising threats. In contrast, neutral or skeptical faces have dispersed ratings across emotions and thus convey ambiguous feelings.

The valence continuum extracted by PCA aligns with the dimensional emotion theory used in many normative datasets (5; 6). The negative correlation between ESI and valence suggests that positive emojis often serve more generic functions (e.g., signalling friendliness or politeness) rather than conveying a specific discrete emotion. This observation complements work showing that positive emojis are more frequently used and more familiar than negative ones (6).

4.2 Crosscultural differences and individual factors

Our analyses used data from Spanish speakers and may not generalise globally. Research on emoji comprehension across cultures reveals notable differences. Chen et al. found that UK participants were more accurate than Chinese participants in identifying most emotions and that cultural differences were not fully explained by familiarity or platform; Chinese participants often used the smile emoji for sarcasm (8). These findings imply that universal facial emotions do not necessarily translate to universal emoji meanings. Personality traits also influence emoji use. Liu and Sun reported that shyness, neuroticism, extraversion and agreeableness correlate with different reasons for using emojis and stickers, and that people adjust usage depending on conversation partners and context (10). Kennison et al. observed that heavy emoji users scored lower on openness to experience and that emoji use was related to word categories such as family and sadness (11). Such individual differences likely modulate both the perceived specificity of emojis and their selection in communication.

4.3 Applications

Marketing and consumer engagement. Businesses increasingly deploy emoji in advertising and social media to stimulate interaction and influence purchasing decisions. A recent review noted that emojis attract attention, enhance creativity and innovation in marketing messages, but ambiguous emojis may hinder comprehension and must be used judiciously (12). Empirical studies with South African Generation Z consumers showed that emojis elicit emotional responses and increase purchase intention, especially among older members of the cohort (13). Our ESI can guide marketers in selecting highspecificity positive emojis (e.g., hearteyes or party face) to evoke clear positive feelings, while avoiding neutral emojis that may be misinterpreted.

Health communication and patientprovider interaction. Emoji can reduce the cognitive burden of health messages and increase engagement. Lin and Luos informationdesign study emphasised that emojis should be used judiciously alongside text in health materials and noted growing applications in doctorpatient communication and psychological assessment (14). In a crosscountry survey of cancer community app users, most participants reported using emojis to express emotions and believed emojis could improve communication with healthcare providers, yet they warned that variation in emoji appearance and cultural interpretation could lead to miscommunication (15). High ESI emojis may serve as reliable icons in symptom checklists or pain scales. The animated emoji scale for dental anxiety demonstrated that motion emojis are a childfriendly and valid tool for assessing anxiety, with strong correlations to established measures and a clear preference among children (16). Dynamic emojis also elicit higher arousal than static ones, and the effects of rhythm and motion on valence vary by emotion category (9), suggesting design principles for future mHealth tools.

Mentalhealth monitoring and mHealth. Selfhelp apps that prompt users to log mood with emojis are emerging as low burden tools for ecological momentary assessment. Van Buren et al. found that adolescents appreciated emojibased mood tracking but emphasised the need for professional support to interpret entries and avoid misunderstandings (17). Selecting emojis with high specificity could improve the reliability of mood logs; for example, using the face vomiting emoji to denote disgust or the smiling face with hearteyes for affection. Researchers should also consider normative ratings and crosscultural differences when designing such tools.

5 Limitations and future work

Our analysis is constrained by the characteristics of the EmojiDis dataset. Ratings were obtained exclusively from Spanish participants and contained more women than men, limiting generalisability. Future studies should compute ESI values using normative data from diverse populations and explore crosscultural consistency. Although we subtracted a neutral baseline from ratings, alternative transformations could be considered. Furthermore, context and cooccurring text dramatically change emoji interpretation (7); incorporating textual context into specificity measures is an important direction. Finally, dynamic effects and motion should be integrated into future indices to capture the richer emotional expressiveness of animated emojis.

6 Conclusion

We introduced an entropybased emotion specificity index to quantify how strongly an emoji conveys a particular discrete emotion. Applied to the EmojiDis dataset, the ESI revealed that negative emojis are generally more specific than positive or neutral emojis, and that neutral faces are highly ambiguous. Together with principalcomponent and valence analyses, our results provide a quantitative foundation for selecting and designing emojis for research and realworld applications. By integrating normative data, crosscultural findings and insights from marketing, health and mHealth contexts, our study offers guidance for leveraging emoji in communication while acknowledging their limitations.

Figures

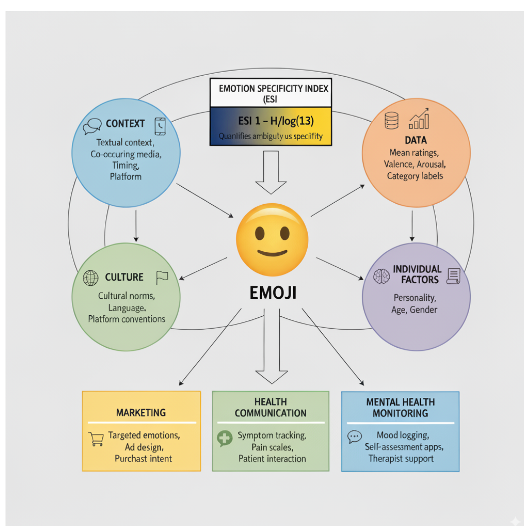

Figure 1: Conceptual model of Emoji sepcificity and its application

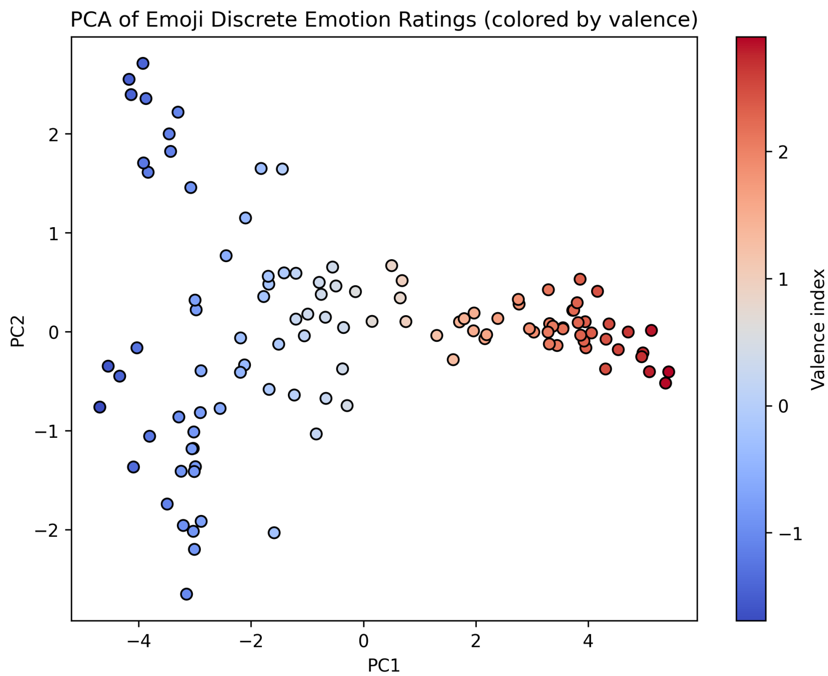

Figure 2: Principalcomponent representation of emoji ratings coloured by the valence index. Warm colours indicate positive valence and cool colours indicate negative valence. The first principal component corresponds to a valence continuum.

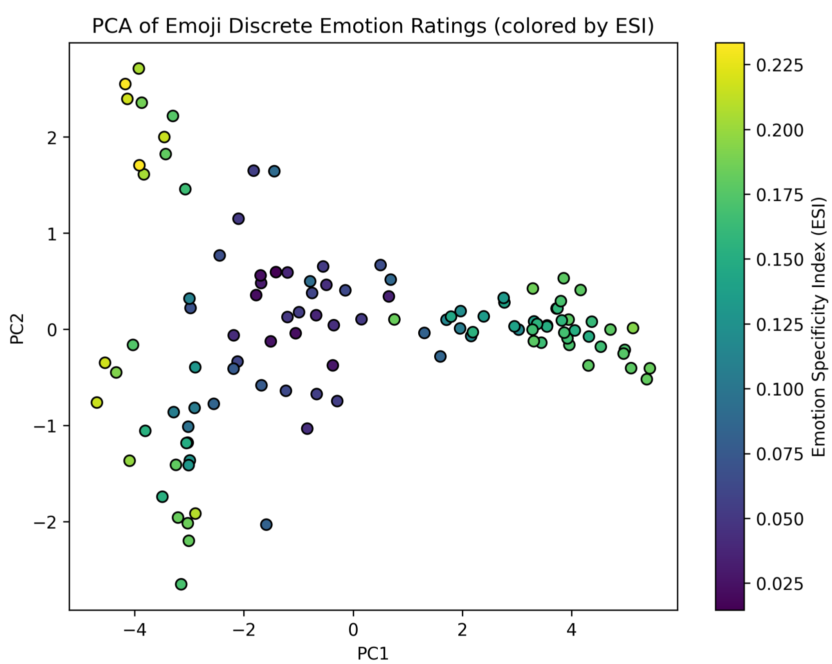

Figure 3: Principalcomponent representation coloured by the emotion specificity index (ESI). High ESI values (yellow) indicate that an emoji is strongly associated with a single emotion, while low values (dark blue) indicate ambiguity.

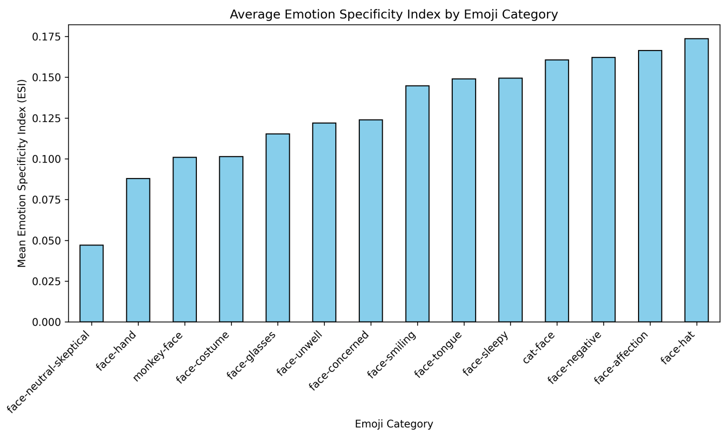

Figure 4: Mean emotion specificity index (ESI) and valence distributions across emoji categories. Categories such as *facehat* and *faceaffection* exhibit high specificity, whereas categories like *faceneutralskeptical* and *facehand* show low specificity.

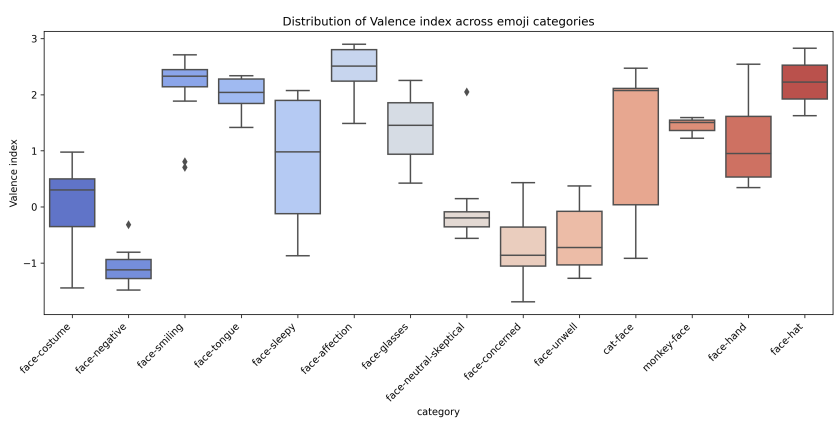

Figure 5: Distribution of the valence index across emoji categories. Positive categories show high valence and narrow spreads, while ambiguous categories show wide distributions.

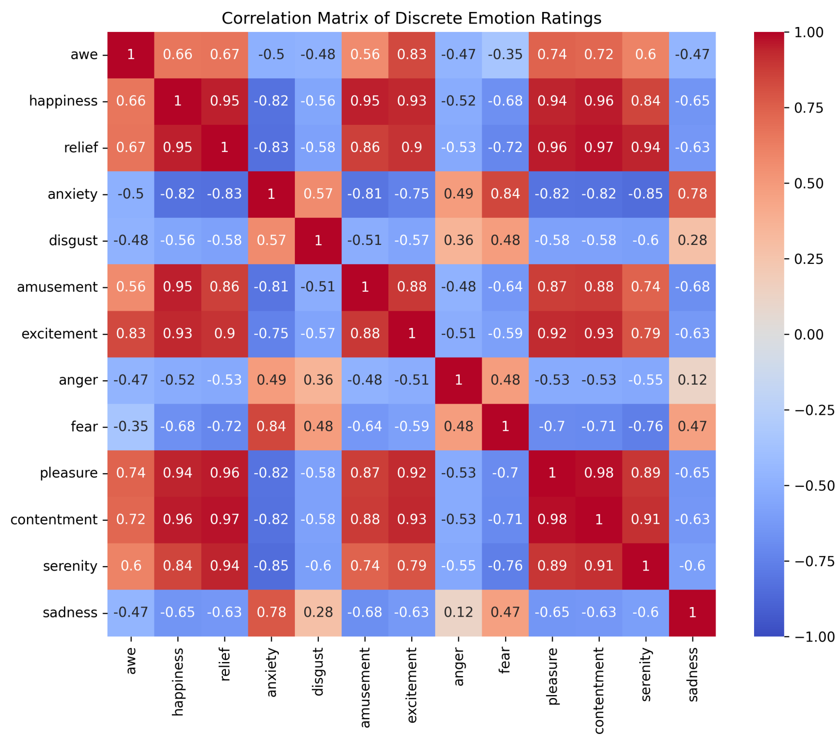

Figure 6: Correlation matrix of the thirteen discrete emotion ratings. Warm colours indicate positive correlations and cool colours indicate negative correlations. Positive emotions correlate strongly with each other and negatively with negative emotions.

References

[1] P. Ferré, P. Haro, M. Á. PérezSánchez, I. Moreno & J. A. Hinojosa. EmojiDis: a dataset of emojis characterised in 13 discrete emotions. *Sci Data* **12**, 1313 (2025).

[2] B. Fischer & C. Herbert. Emoji as affective symbols: affective judgements of emoji, emoticons and human faces varying in emotional content. *Front. Psychol.* **12**, 645173 (2021).

[3] J. Czstochowska, A. Nenchev & T. Scheffler. On the contextfree ambiguity of emoji. In *Proc. ICWSM* **16**, 13881392 (2022).

[4] Emodiversity Project. For researchers. Available at `https://emodiversity.org` (accessed 15 Aug 2025).

[5] T. Scheffler & I. Nenchev. Affective, semantic, frequency and descriptive norms for 107 face emojis. *Behav. Res. Methods* **56**, 81598180 (2024).

[6] P. Ferré, J. Haro, M. Á. PérezSánchez, I. Moreno & J. A. Hinojosa. EmojiSP, the Spanish emoji database: visual complexity, familiarity, frequency of use, clarity, and emotional valence and arousal norms for 1031 emojis. *Behav. Res. Methods* **55**, 17151733 (2023).

[7] N. Aldunate, M. VillenaGonzález, F. RojasThomas, V. López & C. A. Bosman. Mood detection in ambiguous messages: the interaction between text and emoticons. *Front. Psychol.* **9**, 423 (2018).

[8] Y. Chen, X. Yang, H. Howman & R. Filik. Individual differences in emoji comprehension: gender, age and culture. *PLOS ONE* **19**, e0297379 (2024).

[9] S. Zhang & X. Zhao. A study of dynamic emoji emotional responses based on rhythms and motion effects. *Front. Psychol.* **14**, 1247595 (2023).

[10] S. Liu & R. Sun. To express or to end? Personality traits are associated with the reasons and patterns for using emojis and stickers. *Front. Psychol.* **11**, 1076 (2020).

[11] S. M. Kennison, K. Fritz, M. A. Hurtado Morales & E. ChanTin. Emoji use in social media posts: relationships with personality traits and word usage. *Front. Psychol.* **15**, 1343022 (2024).

[12] B. Bak & S. Kyunho. Emoji research: a systematic review of the design, use, function and applications of emoji. *Front. Psychol.* **14**, 1137729 (2023).

[13] M. Maraule, R. Duffett & T. Edu. Modelling emoji online marketing on websites among young consumers: the moderation effect of age. *Future Bus. J.* **11**, 91 (2025).

[14] T. S. Lin & Y. Luo. Visualtextual integration: emoji as a supplement in health information design. *Int. J. Design* **18**, 3758 (2024).

[15] M. N. Essex, J. G. Andrews, K. E. Hockenberry & S. L. Edwards. The potential for emojis to facilitate communication between patients and healthcare professionals. *Front. Commun.* **9**, 1402788 (2024).

[16] J. V. Setty, I. Srinivasan, S. Radhakrishna, A. M. Melwani & M. K. D. R. Use of an animated emoji scale as a novel tool for anxiety assessment in children. *J. Dent. Anesth. Pain Med.* **19**, 227233 (2019).

[17] S. van Buren, J. Kuiper, M. van Roost & T. Vriesema. Can an emoji a day keep the doctor away? An explorative mixedmethods feasibility study to develop a selfhelp app for youth with mental health problems. *Front. Psychiatry* **10**, 593 (2019).

Agents4Science AI Involvement Checklist

1. **Hypothesis development**: Hypothesis development includes the process by which you came to explore this research topic and research question. This can involve the background research performed by either researchers or by AI. This can also involve whether the idea was proposed by researchers or by AI.

   Answer: D

   Explanation: I uploaded the dataset to GPT-5 agent and let it find scientific questions to explore.

2. **Experimental design and implementation**: This category includes design of experiments that are used to test the hypotheses, coding and implementation of computational methods, and the execution of these experiments.

   Answer: D

   Explanation: GPT-5 agent did everything except that I provided a dataset to it.

3. **Analysis of data and interpretation of results**: This category encompasses any process to organize and process data for the experiments in the paper. It also includes interpretations of the results of the study.

   Answer: D

   Explanation: All including analyzing data, plotting figures and writing the article did by GPT-5 agent. Only Figure 1 is generated by nano banana by feeding the article generated by GPT-5 agent to it.

4. **Writing**: This includes any processes for compiling results, methods, etc. into the final paper form. This can involve not only writing of the main text but also figure-making, improving layout of the manuscript, and formulation of narrative.

   Answer: D

   Explanation: GPT-5 agent did all including latex formatting. I only edited very limited formats specifically for emoji symbol to make it rendered correctly.

5. **Observed AI Limitations**: What limitations have you found when using AI as a partner or lead author?

   Description: Just not perfect but definitely the quality of the paper is very similar to master level. There are some spelling errors in the Figure 1 that is generated by nano banana.

Agents4Science Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract clearly shows the goal of this research, the method used, the results achieved and its potential applications.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: In the paper, it has a Section 5 to discuss the limitations of the work

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: The hypothesis has been clearly proposed in the Introduction section. It states "emojis depicting clear negative expressions (e.g., disgust or anger) will exhibit high specificity, whereas neutral or skeptical faces will be ambiguous." Data analysis has been done to test it.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: The dataset used in this study is open access and has been cited. All codes generated by GPT-5 agent and its process can be found on https://chatgpt.com/share/68c352f9-2750-8010-88a6-1980860f6895

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer:

   Justification: The dataset used is open access and it has been cited in reference 1. The coding and data analysis process by GPT-5 can be seen through https://chatgpt.com/share/68c352f9-2750-8010-88a6-1980860f6895

   Guidelines:

   - The answer NA means that paper does not include experiments requiring code.
   - Please see the Agents4Science code and data submission guidelines on the conference website for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so No is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [NA]

Justification: NA

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer:

   Justification: In figure 4, a box plot was presented to show the deviations. Statistical significance tests are not performed.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, or overall run with given experimental conditions).

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [NA]

   Justification: NA

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the Agents4Science Code of Ethics (see conference website)?

   Answer: [NA]

   Justification: NA

   Guidelines:

   - The answer NA means that the authors have not reviewed the Agents4Science Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: In the section 4, it talks about potential applications of this work

    Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations, privacy considerations, and security considerations.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies.