Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-026-06821-z OPEN

Beyond words: emoji patterns in cross-cultural branding

Altug Tanaltay[1], Selcen Ozturkcan[1,2]✉ & Nihat Kasap[1]

This study explores how brands and users communicate through emojis across culturally distinct markets, focusing on Turkish- and English-speaking users on social media platform X. It examines cross-cultural differences in emoji usage patterns, specifically their prevalence, diversity, preferences, and emotional meanings, by applying statistical and information theory methods to large-scale social media data collected over an extended period. Our analysis combines descriptive statistics with two empirical studies. The first examines differences in emoji frequency, richness, and preference between the two language groups. The second explores the emotional semantics of emojis by comparing their co-occurrence with emotion-related words, allowing us to assess cross-cultural variation in their emotional interpretation. The findings indicate that brands use more emojis in Turkish posts, whereas English-speaking users tend to use multiple emojis once they do. Although the most frequently used emojis show increasing convergence over time, the emotional meanings assigned to them vary, particularly for positive and neutral emojis, which display greater divergence than negative ones. These results suggest a shared but culturally distinct visual language of emotion in branding. This study enhances marketing communication theory by demonstrating that emoji usage embodies both universal and culturally specific trends in brand messaging. Unlike prior work that often contrasts Western and East Asian cultures, this research includes Turkey as an intermediary case, both as a moderately individualistic society and as an emerging market. By providing a comprehensive cross-cultural analysis of emoji use, the study offers evidence for the potential emergence of a global emoji language while emphasizing the importance of cultural context in interpretation. These insights are valuable for global brand managers seeking to tailor emoji strategies for culturally adaptive and emotionally compelling messaging.

[1] Sabancı University, Istanbul, Turkey. [2] Linnaeus University, Vaxjo, Sweden. ✉email: selcen.ozturkcan@lnu.se

Introduction

Digital brand communications have recently evolved beyond text, with emojis now playing a critical role in how brands convey tone and emotion. Emojis, which are small pictographs that supplement or replace words, have become widely prevalent in online messaging, including marketing content (Almaguer et al., 2025; Tanaltay et al., 2025). In fact, over 90% of internet users report using emojis regularly in everyday digital communication (Daniel, 2021; Siddiqui, 2020), and brands have followed suit by integrating emojis into their social media campaigns and customer interactions (Casado-Molina et al., 2022; Ko et al., 2022). For marketers, these colorful symbols are more than playful embellishments; they serve as paralinguistic cues (Luangrath et al., 2017) that add emotional tone and personality to an otherwise plain text by making brand posts more relatable and engaging to users (Sciandra and Sharma, 2024).

Prior marketing studies have identified numerous benefits of incorporating emojis into brand communications, from strengthening persuasive appeal in social media posts and promoting positive word-of-mouth (Lee and Hsieh, 2019) to significantly boosting user engagement with branded content (McShane et al., 2021; Tanaltay et al., 2025). For example, experimental research shows that posts accompanied by emojis are often perceived as warmer and more positive, and people receiving a post with emojis tend to feel happier and less stressed than those receiving the same textual content without emojis (Riordan, 2017). Emojis can even increase positive feelings toward advertisements and social media in the target audience, which in turn can foster stronger brand connections. For example, in the Spanish beer industry, emojis in customer service messages elevate user engagement (Casado-Molina et al., 2022), helping brands build long-term relationships with users by adding positive emotional value to messages, while sports-related emojis improve brand identity and visual appeal (Aydogdu and Katırcı, 2021). For accommodation brands, pairing emojis with specific social media texts increases engagement (Wang et al., 2023). In advertising, innovative emoji use captures attention in crowded spaces (Fakhry, 2021), and when combined with social and political hashtags, they enhance post visibility and impact (Alfano et al., 2023). However, excessive emoji use may also detract from post credibility and source trustworthiness (Koch et al., 2023). Effective marketing communication depends on achieving the right tone and emotional resonance with the audience, and emojis have emerged as a powerful tool to help achieve this goal. Therefore, it is crucial to understand the dynamics of emoji prevalence (how often emojis are used), diversity (the range of emojis), preferences (which emojis are chosen most), and emotional meaning for better brand–user interaction in alternative markets such as developed and emerging. Understanding these dynamics will help enhance post clarity and emotional tone, ultimately increasing the level of user engagement, especially in the globally interconnected world with brands simultaneously serving many international markets.

However, despite their global popularity, emojis are *not* a "universal language" that carries identical meaning across cultures (Chen et al., 2024). Communication research shows that cultural context can significantly influence both how people use emojis and how they interpret them (Gao and VanderLaan, 2020; Guntuku, 2019). On one hand, there are striking similarities in the emoji use worldwide. Certain emotional icons have broad, cross-cultural appeal; for instance, the 😂 *"face with tears of joy"* and the ❤️ *"red heart"* consistently rank among the most frequently used emojis in many countries around the world (Daniel, 2021). This convergence suggests that some aspects of digital emotional expression transcend language barriers. On the other hand, there's a lot of variation and subtlety. People from different cultures often employ emojis in distinct ways and assign them different connotations (Gao and VanderLaan, 2020; Guntuku, 2019). An emoji that seems positive or innocuous in one culture might carry a completely different subtext in another. For example, the "thumbs-up" gesture emoji 👍 is viewed as a friendly sign of approval in the United States, but in parts of the Middle East, it is equivalent to a highly offensive insult (essentially "up yours") (Anderson, 2018, June). Similarly, Eastern and Western audiences have been found to interpret facial emojis differently: Western users focus more on mouth expressions, while East Asian users focus on eye cues (Park et al., 2014), leading the same smiling face emoji to be perceived as happier by Westerners but sometimes less so by Easterners (Gao and VanderLaan, 2020). These cultural variations mean that a brand's well-intended, emoji-laden post could be misunderstood or even backfire if the emoji's emotional meaning doesn't align with the local target audience's interpretation. In other words, emojis are not one-size-fits-all emotional signals in internationally connected marketing communication. This presents a critical challenge and opportunity for global brands: understanding the commonalities and disparities in emoji use across cultures is essential to communicating consistently and effectively with international users. Without such understanding, a campaign that succeeds in one market might falter in another due to an emoji miscommunication, undermining the brand's communication goals of clarity, emotional connection, and audience engagement.

Despite the increasing use of emojis in branding, there is a notable gap in global marketing communication research regarding how the cultural background of markets influences emoji-based interactions between brands and users. Much of the existing emoji research has been conducted in single-culture settings (often developed markets in Western, English-speaking contexts) or in general communication studies, leaving it unclear how cross-cultural differences in emoji use might impact branding outcomes, especially in non-Western, non-English-speaking, and emerging economies. As a result, managers of global brands have had little scholarly guidance on whether an emoji strategy that works in one market (i.e., using a particular emoji to seem friendly or funny) will translate the same way elsewhere. This paper addresses this gap by providing a focused, data-driven comparison of emoji use in two culturally and linguistically distinct market environments: the Turkish-speaking (as an example of an emerging and non-Western) market and the English-speaking (as an example of a developed and Western) market. In doing so, we adopt language as a proxy for culture, a practice supported by prior cross-cultural user research (Luangrath et al., 2017; Luna and Peracchio, 2001). Language is known to activate culturally specific mental structures and influence individuals' affective and cognitive responses in marketing contexts (Luna et al., 2008). This property makes it a reliable lens for capturing cultural variation in communication practices. Additionally, lexical and semantic selections in language have been demonstrated to mirror profound cultural norms and temporal transformations (Agovino et al., 2021), thereby further substantiating the methodological integrity of this approach. We collected a large dataset of brand posts and replies from June 2016 to June 2021 from Platform X (formerly Twitter), a global text-based social media platform where emoji use is widespread. With over 500 million tweets sent daily and emojis embedded in a significant share of brand-user interactions, Platform X offers an ideal environment for observing real-world, cross-cultural communication (Statista, 2025). Together with being the third most used social media environment in Turkey (Statcounter, 2025), its

global reach and focus on concise, emotive messaging make it particularly suitable for examining emoji use in multilingual marketing contexts.

Addressing the overarching research question, *"How do emoji use patterns and emotional meanings differ or align across Turkish- and English-speaking markets in branding communication on Platform X?"*, we initially analyze emoji prevalence and diversity through the application of descriptive statistics and measures derived from information theory. Next, we implement two empirical studies to examine emoji preference and emotional semantics in the context of brand communication on social media in Turkish-speaking versus English-speaking countries. We concentrate on Turkish and English contexts because they offer an informative contrast: with its population of over 85 million inhabitants, Turkey represents a scalable emerging non-Western market with a unique cultural background (often characterized as more collectivist and context-driven in communication), whereas English-speaking markets like the US or the UK (developed and Western) are typically more individualistic and low-context in communication style (Hofstede, 2011; Hofstede et al., 2010). By analyzing these side by side, we can uncover which patterns of emoji use are globally convergent and which are culturally divergent in a branding context. In doing so, we directly respond to the need for a more explicit rationale linking emojis, culture, and communication goals in marketing. Our study is grounded in the idea that effective brand communication must not only leverage emojis to enhance emotional expression but also do so in a culturally intelligent way. We approach this investigation with the goals of determining how often brands and users resort to implementing emojis in each market we examine (prevalence), how varied their emoji vocabulary is (diversity), which emojis are most favored or avoided (preferences), and what emotional sentiments those emojis convey in each culture (semantics) we focus on. Understanding these aspects is crucial for brands aiming to build consistent brand personalities and customer relationships across different regions. Ultimately, this cross-cultural focus is essential for achieving core global marketing communication objectives: ensuring the delivery of messages that are characterized by clarity, cultural resonance, and emotional engagement for the target audience within each localized market.

Our findings offer two key contributions to marketing communication theory and practice. First, we discover evidence of a convergence in emoji preferences across cultures. Despite the many cultural differences between Turkish and English-speaking user groups, users in both contexts gravitate toward a similar set of "favorite" emojis in brand-related communications. In other words, there is a common core of popular emojis (such as positive smileys and heart symbols) that transcends the emerging non-Western market and developed Western market duality. This convergence suggests the emergence of a shared digital emotional language, a valuable insight for global brands, as it indicates certain emojis carry universal appeal and can be used to connect with diverse audiences in a broadly consistent way. Marketers may take advantage of this overlap by incorporating universally popular emojis into their cross-market global campaigns to evoke familiarity and positive sentiment. Second, we uncover clear differences in the emotional semantics of emojis between the Turkish and English-speaking contexts, highlighting the influence of culture on emoji meaning. Even when marketers use the same emoji in both markets, the emotion or implication it conveys can significantly vary across cultural contexts. For example, our analysis reveals that Turkish-speaking users associate certain emojis with different feelings or use contexts than English-speaking users do, indicating that an emoji's emotional flavor is shaped by cultural communication norms. These differences manifest in how emojis accompany particular phrases, how they substitute for words, or how strongly they express sentiment in each language. This contribution is particularly important because it goes beyond simply cataloging which emojis are used to understand how they are used and preferred differently. It warns us that a lighthearted emoji deployed by a brand in one culture might be interpreted more sarcastically or negatively in another culture (and vice versa). Such insights can help practitioners avoid cross-cultural miscommunication. These two findings offer a set of convergent emoji preferences coupled with culturally divergent emoji interpretations, setting a more detailed understanding of emoji-based branding. We suggest that while a global visual lexicon for basic emotions is emerging, cultural context still profoundly shapes communication subtleties and user sentiment.

In summary, this research shows that emoji use in global branding should be strategically adapted across cultures. By comparing Turkish- and English-speaking markets, we provide evidence that emoji use reflects both universal patterns and cultural differences. Our findings highlight that emojis act as emotional language in brand messaging, blending shared symbols with culture-specific meanings (Gao and VanderLaan, 2020; Guntuku, 2019). This offers practical guidance for marketers: while global brands can use common emoji trends to maintain consistency, they must also account for cultural differences to ensure their messages are understood as intended.

The paper is organized as follows: *"Theoretical Background"* outlines the study's foundation and contribution to emoji research, focusing on cultural variability in communication, emojis in brand-user communication, and cross-cultural studies. Next, the *"Empirical Studies and Results"* section discusses the quantitative content analysis methods and findings across descriptive analysis and two empirical studies. Each study is presented with its methodology, results, and discussion sections, considering the implications of our findings, relating them to recent cross-cultural emoji studies and brand communication. The conclusion reflects on the study's limitations and implications and proposes future research avenues.

Theoretical background

The increasing prevalence of emojis in digital communication has prompted interdisciplinary studies that investigate their significance in sentiment analysis, perception, and cultural usage (Bai et al., 2019; Evans, 2017; Guntuku, 2019). This study focuses on the cultural nuances of emoji use in English- and Turkish-speaking markets, building on existing work in communication variability and cross-cultural emoji studies to advance business research applications.

Emojis in cross-cultural communication studies and the theoretical framework. Culture plays a key role in shaping communication, with aspects like individualism and collectivism affecting both language and nonverbal cues, including emojis. Gudykunst's cultural variability in communication (CVC) model highlights how national culture determines communication norms and patterns (Gudykunst, 1997; 1996). In parallel, Hofstede's cultural dimensions offer a widely used framework for quantifying cultural traits, making them particularly valuable in international marketing and consumer behavior research (Hofstede et al., 2010; Soares et al., 2007). These dimensions remain applicable even when original country scores are not employed, as researchers often adapt or remeasure them using updated instruments suited to specific contexts.

As a communication tool, when shared with a common interpretation, emojis reduce misunderstandings, improve emotional connection, and foster cross-cultural understanding in communication (Revita et al., 2023), making them a universal

language on social media (Chen et al., 2024). Researchers and practitioners can gain a better understanding of how brands can incorporate tone and personality into digital communication by using their typology and model. Additionally, emojis possess various meanings and applications in distinct cultural contexts (Bai et al., 2019). Prior studies utilizing Gudykunst's CVC model and Hofstede's individualism-collectivism scores investigated the similarities and differences in emoji usage and preferences between collectivist and individualistic cultures, focusing on various geographic regions, languages, countries, or specific groups within a country. Researchers frequently linked individualistic traits to individuals from Western regions (such as Western Europe, North America, and Australia) while characterizing collectivist qualities as representative of personalities from non-Western regions (including Asia, South America, and Africa) (Green et al., 2005). These studies utilize language as an indicator of users' cultural backgrounds; however, they tend to limit their comparisons to the West and East, emphasizing distinct cultural groups like English-speaking nations and Asian countries that speak Chinese, Korean, or Japanese. For instance, Yuki et al. (2007) and Park et al. (2014) examined the USA and Japan, while Sun et al. (2023) focused on the USA and South Korea. Additional research, including Togans et al. (2021), Gao and VanderLaan (2020), Guntuku (2019), and Revita et al. (2023), has similarly contrasted English-speaking countries with East Asian nations such as China, Korea, and Japan.

There are also examples in previous research that have focused on a wider set of countries and languages. Park et al. (2014) analyzed 1.7 billion tweets from 2006 to 2019 by employing Hofstede's national culture scores (Hofstede, 2011) across 78 countries and found that Platform X users from individualistic cultures preferred horizontal and mouth-oriented emoticons like :), whereas those from collectivist cultures favored vertical and eye-oriented emoticons like ^_^. In a similar vein, Gao and VanderLaan (2020) conducted an online quasi-experimental study with participants from Western and Eastern cultures, revealing that the emotional valence ratings (i.e., happiness/sadness) of emojis are more significantly influenced by the mouth for Westerners and by the eyes for Easterners.

Sun et al. (2023) found that cultural orientation towards individualism or collectivism affects emoji use in digital communication, with intercultural competence and emotional intelligence moderating this relationship. Moreover, Guntuku (2019) identified distinct emoji meanings across Eastern and Western cultures, despite the similar frequencies of use. Furthermore, Li et al. (2019) showed that emoji use for expressing emotions correlates with Hofstede's cultural dimensions.

In studies such as Kejriwal et al. (2021) and Li et al. (2019), where the focus was on a wider set of countries, the analysis was limited to comparing only the frequencies of emojis in messages or how they represented positive or negative sentiments in general, undermining their ambiguity in representing more detailed emotions such as happiness, sadness, anger, fear or disgust. Yet, according to Hofstede et al.'s (2010) cultural scores, English-speaking countries like the USA or the UK exhibit markedly more individualistic cultures (89 and 90 points, respectively) compared to China or South Korea (20 and 18 points, respectively). This study chose to focus on the English language to represent the more individualistic Western cultures, building on prior research that recommends analyzing larger datasets with more detailed information about specific languages and countries. On the other hand, the context of emerging markets has also become key to focus on, especially due to their growing numbers and influence on the global economy, with further percussion effects on the involved managerial decisions (Dinçer et al., 2024). In this regard, Turkey, with its moderate individualism score of 37 and its central geographic position relative to China or South Korea, was chosen to represent an emerging non-Western market. This selection facilitated a more subtle comparison and provided an opportunity to test if the impact of individualism vs. collectivism, non-Western vs. Western, or emerging vs. developed market distinctions on emoji use in messages also applied to other intermediary countries that had a similar moderate individualism score (see Fig. 1).

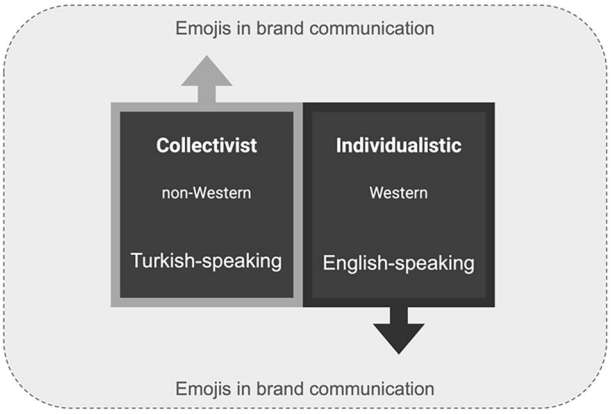

Fig. 1 Comparison of emoji use in brand communication between Turkish (collectivist) and English (individualistic) cultures (source: authors' own creation).

This study builds on earlier research (Li et al., 2019; Park et al., 2014) that examined global brand communication through big data approaches for social media (e.g., Platform X). Like Kejriwal et al. (2021), who examined emoji usage in 4 million tweets from 30 different countries, we look at how emoji use differed by language and, consequently, by market. Kejriwal et al. (2021) previously noted regional variations and universally used emojis, particularly in the Middle East. We expand their research by examining the use of emojis in local and global branding contexts (Vandergriff, 2013), taking into account how meaning is influenced by the surrounding language and text. We explore use patterns in greater detail, looking at the emotions expressed and the differences in time between Turkish (representing emerging markets with less individualistic cultures than those of East Asia) and English-speaking (representing developed markets with Western cultures). As with common constructs used in prior research, our study compares the prevalence, diversity, and preferences of emojis in brand communication, along with semantic emotional interpretation. By doing this, we add to the body of literature by offering a comparative viewpoint that incorporates an intermediate cultural context between individualism in East Asia and the West. This approach enables us to investigate whether emoji use is converging into a common visual language among Turkish- and English-speaking users, two culturally and linguistically distinct markets. We begin by analyzing emoji prevalence and diversity with descriptive analysis, and then examine emoji selection preferences in Study 1, addressing the research question: "RQ1: *What are the variations in emoji selection preferences between Turkish- and English-speaking users on Platform X?*"

Emojis in brand-user communication studies. Emojis are interpreted and understood in varying ways depending on the particular context in which they are employed (Gawne and McCulloch, 2019). Also, in text-based marketing, textual paralanguage, such as emojis, is an essential tool for communicating social cues and emotional tones (Luangrath et al., 2017).

Therefore, to uncover the emoji use in international communication and marketing research, it is essential to understand how users in different markets interpret them, considering the ambiguity of their meaning under different circumstances and contexts (Jaeger et al., 2019). The ability of emojis to capture users' emotions (Farhoudinia et al., 2024) in social media advertising activities renders them a useful tool for evaluating emotions regarding products and services (Bai et al., 2019). While emojis do carry emotional weight, the extent of their impact and the conditions under which they influence cognitive processing may require further investigation to understand their effectiveness in business communication (Kaye et al., 2021).

Experimental studies reveal that negative emojis can diminish the perceived rationality of online reviews (Kim and Gupta, 2012), while positive emojis can enhance purchase intentions (Das et al., 2019) and message expressiveness (Boutet et al., 2021). Positive emojis also increase perceptions of warmth and positivity in workplace communications (Cavalheiro et al., 2024). However, excessive emoji use may also harm message credibility (Koch et al., 2023). Former research that used content analysis found that emojis are also crucial in customer care on Platform X (Casado-Molina et al., 2022), increase X engagement (McShane et al., 2021), and when paired with travel content, they boost social media interaction (Wang et al., 2023).

While previous research has demonstrated the usefulness of emojis across various business domains, recent advances in computational methods and data availability have enabled more detailed evaluations of their impact on digital marketing and brand communication using sentiment analysis (Farhoudinia et al., 2025). For example, Kastrati et al. (2024) demonstrated that large-scale emoji-driven data annotation, combined with deep learning models enables robust sentiment and emotion analysis on X. Along these lines, Tanaltay et al. (2025) used the same method, first calculating the probability distribution of emojis for Ekman (2005)'s six basic emotions *(happiness, sadness, fear, surprise, anger, and disgust)* using co-frequency analysis of emojis and emotional words in Platform X posts. They then generated a preliminary dataset to train deep learning classifiers to predict emotions in brand posts and user replies. Finally, they evaluated the effect of dominant emotions in brand posts on user engagement. A common note of this work is that the emotional meanings of emojis may vary according to culture, context, and time, thus underlining the importance of analyzing their meaning across different contextual settings.

Our research enhances external validity by building on previous studies through the analysis of cross-cultural big data sourced from Platform X, focusing on how emojis express emotions in global branding communication. We specifically examine emoji usage in English- and Turkish-speaking markets to gain insights into cultural perceptions and improve the generalizability of emoji research across emerging Non-Western and developed Western markets. In Study 2, we pose the research question: "RQ2: *What are the differences in emotional semantic meanings associated with commonly used emojis in global and local branding communication across Turkish- and English-speaking markets on Platform X?*" Additionally, to analyze the emotional usage of emojis and align with prior research, we adapted co-frequency analysis of emojis and emotional words based on Ekman's (2005) six basic emotions, utilizing English (Zhu, 2014) and Turkish (Tocuoglu, 2019) word lexicons.

Empirical studies and results

Research design. The research design and the steps followed for analysis are shown in Fig. 2. After collecting Turkish and English brand posts on Platform X, along with associated user responses to these posts, preprocessing and quantitative analysis methods were applied. Emoji use was first examined through a descriptive analysis comparing the prevalence and diversity of emojis across the two languages that we focused on. This was followed by two empirical studies that investigated the differences and similarities in emoji use within social media branding contexts in English and Turkish. Regarding descriptive analysis of the prevalence of emojis, emoji frequencies in posts were counted, and using alternative regression models, the difference in mean number of emojis in posts and replies was assessed per language. Then, the number of unique emojis was counted per post, and by calculating the Type Token Ratio metrics, the diversity of emojis preferred across the English and Turkish datasets was evaluated.

Emojis in all messages were ranked by their aggregate frequencies by language and then segmented into three popularity phases in *Study 1: high, moderate*, and *low* popularity. The similarity of emojis in each popularity phase across the two languages was measured and compared using Jaccard Distance.

Finally, in *Study 2*, co-occurrence matrices of emotional words and emojis were processed semantically to attain probability distributions by six basic emotions of Ekman: happiness, sadness, anger, disgust, fear, and surprise (Ekman, 2005). Then, the

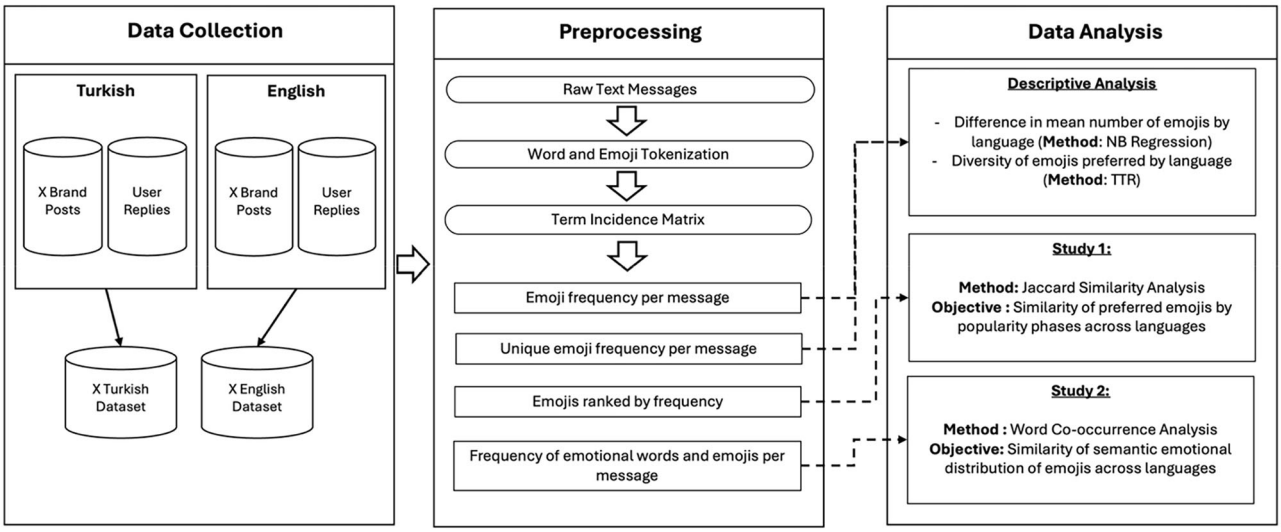

Fig. 2 Research Design (Pre-processing and data analysis steps are applied for English and Turkish datasets separately) (source: authors' own creation).

similarity of these distributions was compared across languages using Jensen-Shannon Divergence.

Data collection and pre-processing. We collected data from Platform X (formerly Twitter), where emojis were widely used in brand-user interactions. Considering its high daily tweet volume over the globe and popularity in Turkey (Statcounter, 2025; Statista, 2025), it provided a suitable context for studying emoji use in multilingual marketing. A substantial dataset from Platform X was acquired through its Search API through an academically verified research account access, which enables the collection of posts in textual format by setting specific keywords and a defined time period. The API also offered several filtering options, including the language of posts or the post type as 'retweet' or 'reply.' This study purposefully included all Fortune 500 companies that maintained Platform X accounts in both English and Turkish languages as of 2021, focusing on global brand communication for a cross-cultural analysis. The final sample contained 33 brands of fast-moving consumer goods (FMCG), fast food, technology, automotive, apparel, retail, finance, and logistics, as also listed in Table 1. (*Please see* Table A1 *in Supplementary Information for the number of posts downloaded by industry and categorized as posts and replies*.)

By targeting specific brand names and accounts as keywords, all Platform X posts from these brands, along with the respective Platform X user responses that were associated, were retrieved from June 2016 to June 2021. A custom-built Java library facilitated the downloading and storage of the posts in an Elasticsearch index to expedite searches. The dataset initially comprised 8,101,034 English and 735,459 Turkish Platform X posts. Among these, English posts included 2,708,874 brand posts and 5,392,160 user replies. In comparison, the Turkish ones featured 245,422 brand posts and 490,037 user replies (*please see* Table A1 *in Supplementary Information for the number of posts downloaded by industry and categorized as posts and replies*).

The plain text content of the posts was stored separately in raw text format with their timestamp information, as the collected Platform X data contained excess metadata that was unnecessary for our analysis. Using Python, raw text was tokenized into word and emoji sequences for each post in the dataset. From emoji and word token sequences, word incidence matrices were formed by posts represented by a vector of frequency of distinct words and emojis. A three-dimensional sparse matrix format was used to store the data in memory and on disk. In this format, the first dimension represented the post rows, the second dimension contained indices for non-empty entries, and the third dimension held the count values for each non-empty index, which corresponded to previously counted words and emojis. This approach significantly reduced the volume of disk use and processing time compared to repeatedly querying raw text data. Subsequent analyses were performed using Python's Numpy and SciPy modules after generating the sparse matrices. Extracted from these matrices, in descriptive analysis, the sum of the count of all emojis by post was used to compare the mean number of emojis, and the sum of distinct emojis by post was used to measure the diversity of emojis used by English- and Turkish-speaking Platform X users. In Study 1, emojis were ranked and further analyzed based on their popularity by summing all the appearances of each emoji across the data sets. For Study 2, co-occurrence matrices of emojis and emotional words were further extracted from the word incidence matrices, which were accomplished by taking the dot product of the word incidence matrix and its transposed version.

Descriptive analysis of emoji use across languages

Prevalence of emojis. To analyze the prevalence of emojis across languages, the difference in the mean number of emojis in posts between English and Turkish datasets was statistically evaluated. The data was segmented into conversational components, such as brand posts and user replies, to differentiate between posts from brands and users. As indicated in Table 2 and Fig. 3, the distribution of emoji numbers in brand posts and user replies is predominantly right-skewed, with the first quartile, median, and third quartile values being zero. Most brand posts do not contain emojis (Posts: 72% for Turkish, 80% for English; Replies: 84.8% for Turkish, 85.4% for English), and among those that do, most include only one emoji (Posts: 24% for Turkish, 14% for English; Replies: 9.8% for Turkish, 8.6% for English). Turkish posts are

Table 1 Brands selected for platform X data collection.

FMCG	Fast Food	Technology	Apparel	Automotive	Finance	Logistics	Retail
Gillette, Coca-Cola, L'Oreal, Algida/Walls, Orkid/ Alldays, Pepsi, Flormar, Oriflame, Yves Rocher	Burger King, KFC, McDonald's	Vodafone, Samsung, Siemens, Sony, Microsoft, HP, Huawei, Netflix	Levi's, Marks and Spencer H&M	Hyundai, Toyota, BMW	Allianz, Axa, HSBC	UPS, DHL	Watsons, Carrefour

Table 2 Summary statistics.

	English			Turkish		
	Posts	Replies	Overall	Posts	Replies	Overall
Mean	**0.28**	**0.29**	**0.29**	**0.36**	**0.27**	**0.30**
St. Dev.	0.69	1.25	1.09	0.73	1.24	1.10
Q1, Q2, Q3	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0	0, 0, 0
Maximum	100	136	136	51	136	136
No Emoji[a]	80.8%	85.4%	84%	71.5%	84.9%	80%
Has Emoji[b]	19.2%	14.6%	16%	28.5%	15.1%	20%

Q1, Q2, and Q3 correspond to the first quantile, the median, and the third quantile.
[a]Proportion of brand posts and user replies that contain no emoji.
[b]Proportion of brand posts and user replies containing at least one emoji.

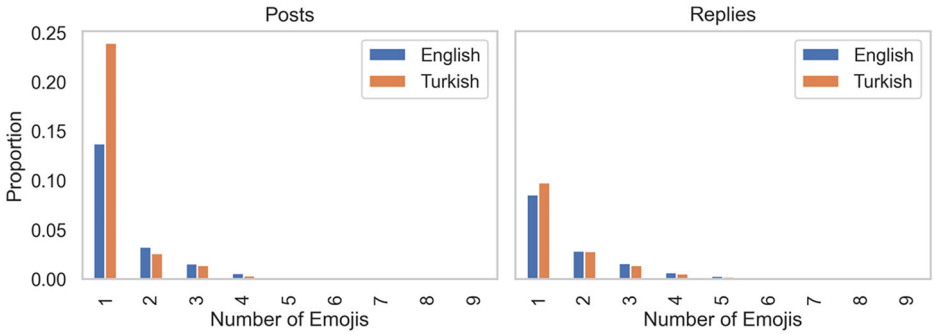

Fig. 3 Emoji distribution frequency: brand posts and user replies. (source: authors' own creation).

Table 3 Parameter estimates of ZINB regression.

	Posts			Replies		
	Estimate	**Std. Err.**	***p***	**Estimate**	**Std. Err.**	***p***
English (base, odds)	−9.8977	2.224	0.0*	−8.0123	4.779	0.094
Turkish (odds)	−0.5118	0.014	0.0*	−1.9131	5.077	0.706
English (base, rate)	−0.8191	0.005	0.0*	−1.2363	0.002	0.0*
Turkish (rate)	−0.2049	0.007	0.0*	−0.0775	0.005	0.0*
Dispersion	0.9860	0.011	0.0*	7.1146	0.018	0.0*

*$p < 0.01$.

more likely to contain a single emoji than English posts, but English posts slightly more commonly include multiple emojis.

To test the significance of the difference in the mean number of emojis used in posts across languages, a Zero Inflated Negative Binomial Regression (ZINB) model was employed due to the skewness in distribution and the excess number of zeros in emoji frequencies. A binary language variable was established, with English set as the base category. The number of emojis in each post served as the dependent variable, as shown in Eq. 1.

$$y = \beta_0 + \beta_1 \times Turkish \quad (1)$$

where y represents the number of emojis in a post, β_0 is the intercept, which is the expected number of emojis for posts in English, and β_1 is the slope, representing the change in the number of emojis when the post is in Turkish.

The results of the ZINB model are shown in Table 3. The ZINB model has two main components. The first one is the "Odds" component, where the likelihood that an observation belongs to the excess-zero group is reported. A more negative coefficient means lower odds of having zero emojis. The second component is the "Rate" component, where expected count of the variable under concern is reported. A more negative coefficient indicates fewer emojis used on average. The ZINB model allows the variance to be greater than the mean, and the "dispersion" parameter quantifies this excess variance.

Regarding brand posts, the odds of having zero emojis are significantly higher in English than Turkish, as indicated by the highly significant negative estimate for English (−9.90, $p < 0.001$) and a smaller, yet also significant, negative coefficient for Turkish (−0.51, $p < 0.001$). This suggests that English posts are much more likely to contain no emojis than Turkish posts. In replies, however, while the English base estimate remains large (−8.01), it is not statistically significant ($p = 0.094$), and the Turkish estimate is not significant either ($p = 0.706$). This suggests that differences in the odds of zero emoji use in replies are less pronounced and not statistically reliable.

Both English and Turkish show highly significant negative coefficients for the rate component in posts and replies (all $p < 0.001$), indicating that emoji use is generally low in both languages, even when emojis are used. However, the rate is lower in English than in Turkish. For example, in posts, Turkish has a less negative rate estimate (−0.20) compared to English (−0.82), implying that brands posting in Turkish used emojis more frequently when they did use them. A similar pattern holds in replies, though the difference is smaller (−0.08 for Turkish vs. −1.24 for English).

Overall, the findings suggest that, on average, brands tended to use more emojis in Turkish posts compared to English. Users show no significant difference in the likelihood of using emojis between the two languages; however, English speakers are more likely to use multiple emojis after using one, while Turkish speakers tend to stick with a single emoji.

Diversity of emojis. The objective of this next analysis is to compare the richness/diversity of emojis favored between English and Turkish datasets. A standard technique for assessing the lexical diversity of a dataset or text involves computing the Type-Token Ratio (TTR). TTR is defined as the ratio of different word types to the total number of word tokens (Baayen, 2001). Since the goal was to ascertain the diversity of emojis used rather than words, TTR was modified to represent the number of distinct emoji types divided by the total number of emoji occurrences shown in Eq. 2.

$$TTR_i = \frac{V(C_i)}{C_i} \quad (2)$$

where i refers to language, $V(C_i)$ is the distinct count of emojis, and C_i is the total count of emojis in corpus i.

However, when considering the richness of words, it is recognized that the size of the dataset significantly influences this measure (Tweedie and Baayen, 1998). Thus, TTR was normalized by dividing the dataset into chunks of 1000 posts, and mean TTR was calculated over all chunks of data for each

language. To assess the significance of the difference in TTR between the English and Turkish datasets, we performed both a t-test and a Monte Carlo simulation. The Monte Carlo simulation was conducted by combining all posts into a single dataset, disregarding language differences, and calculating the normalized TTR on randomly sampled data to generate a probability distribution of TTR. This distribution was then utilized to calculate a p-value, which indicates the probability of observing TTR values as extreme as, or more extreme than, those observed in the actual data. If the *p*-value is below the designated significance level, it could be interpreted as evidence against the null hypothesis, which suggests that the observed difference occurred by chance. The procedure was also repeated on a monthly basis from 2016 to 2021. This time, monthly diversity measures ($TTRM_i$) are computed for English and Turkish and plotted as a function of time to further analyze whether the variety of preferred emojis amongst languages is also changing through time *(please see* Appendix B *of Supplementary Information for the detailed procedure).*

The counts of distinct emojis used in each language are summarized in Table 4. The number of distinct emojis counted in English (2598) was higher than the number counted in Turkish (1840), likely due to the larger number of observations in English. However, there was a high degree of overlap in the emojis used in both languages, with 1798 emojis preferred in English also being preferred in Turkish. This figure corresponds to a Jaccard similarity of 69% and 98% for English and Turkish, respectively, suggesting that nearly all emojis preferred in Turkish were also preferred in English.

To further investigate the similarity in emoji use between the two languages, we compared the richness, or diversity, of emojis preferred using the TTR normalized by corpus size. The results, presented in Table 5, show that the overall mean diversity of emojis used in Turkish (0.04) was slightly lower than that in English, based on both the *t*-test and permutation test ($p < 0.001$). However, when the monthly mean diversity was analyzed, we found that Turkish was 1% above English, with similar means and lower variance. The observed monthly difference was not statistically significant (*p*-values of 0.481 for the *t*-test and 0.410 for the permutation test), suggesting that it may have occurred by chance.

The trend in emoji diversity over time was also analyzed from 2016 to mid-2021, and similar findings were observed with regard to the presence of differences between the languages. Figure 4 presents the distribution of overall and monthly diversity statistics for English and Turkish, as well as the monthly change over time. Both languages showed a similar trend, starting with values above 0.4 and converging toward 0.25 in a decreasing fashion.

Our results show that a more diverse set of emojis is preferred in English. However, the mean monthly diversity of emojis decreases for both languages from 2016 to 2021, meaning that the diversity of the emojis preferred to degrade similarly. A second question arises if the preferred set of emojis becomes more similar following the decrease in richness, which will be summarized in the next section with Study 1.

Discussion. The findings regarding the frequency of emoji use across different cultural contexts revealed notable discrepancies in existing research. Guntuku (2019), for example, observed that Westerners used more emojis compared to their Japanese and Chinese counterparts, suggesting a greater inclination toward emoji use in less context-dependent cultures. Unlike Guntuku (2019), research by Kejriwal et al. (2021) and Togans et al. (2021) suggested that East Asian users use emojis more, likely because their communication styles often need extra context, which emojis help provide and make conversations friendlier.

Our results complement these studies, adding a broader perspective by analyzing branding social media data and emoji use in Turkey compared to Western practices. The findings showed that brands in their Turkish posts used more emojis than those in English. Among users, however, the difference in emoji use between Turkish and English speakers was minimal, with a preference for using a single emoji per post in Turkish versus multiple emojis in English-speaking contexts. Our results revealed a comparable trend in emoji usage among individualistic English speakers and collectivist Turkish speakers, differing from earlier studies that examined more varied cultural contexts. Managerially, brands operating in collectivist cultures can benefit from using emojis to foster a sense of community, while those targeting individualistic cultures can use them to enhance personal connections and emotional expression. This differentiation in emoji use ensures that posts resonate more deeply with target audiences and their cultural contexts.

Regarding the diversity of emojis preferred by users, Feldman et al. (2017) argued that language proficiency affects how emojis are selected, especially when people communicate in a language other than their native one. Later, Togans et al. (2021) experimented on participants from the USA and East Asia and showed that the latter group used a more diverse set of emojis. According to their results, diversity and preference for emojis did not appear to be a function of English proficiency but were dependent on cultural traits such as collectivism and interdependent self-construal. In contrast, our results showed that, while more collectivist in cultural nature, Turkish speakers used a less diverse set of emojis. This finding aligned with Kejriwal et al. (2021)'s observation that emoji diversity is directly affected by the country, and it was observed less in Turkey compared to Europe and the US. Despite these differences, our study showed that emoji diversity decreases and converges over time for both

Table 4 Count of distinct emojis in brands dataset.

	English	Turkish
Distinct Emojis	2598	1840
Difference[a]	800	42
Jaccard Similarity	69%	98%

[a]Number of different emojis preferred compared to the other language.

Table 5 Diversity Statistics.

	English	Turkish	Difference	*t*-stat	**p*-val	***p* perm.*
Overall Diversity	0.33	0.29	0.04	4.982	<0.001	0.00
Standard Deviation	0.16	0.18				
Monthly Diversity	0.34	0.35	−0.01	−0.706	0.481	0.410
Standard Deviation	0.06	0.09				

**p*-value for *t*-test ***p*-value of permutation test, no samples were observed with difference of 0.04 following randomized experiments.

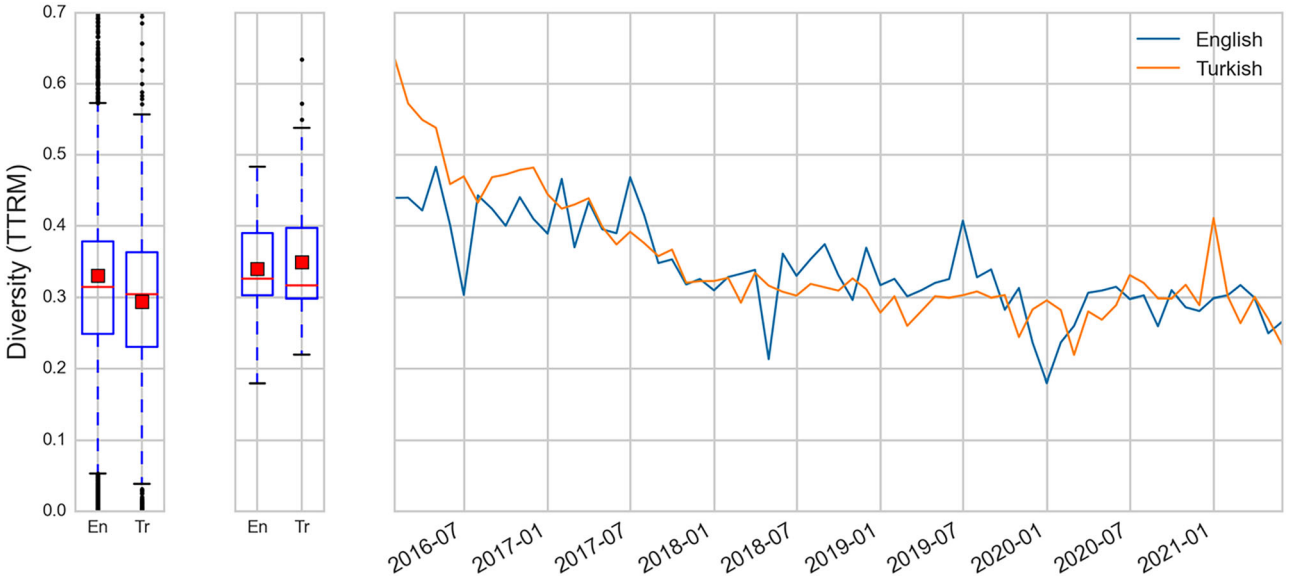

Fig. 4 Diversity analysis. *From left to right: Overall diversity, monthly diversity, and diversity by months.* (source: authors' own creation).

English and Turkish speakers, hinting at a globalization of digital communication practices where dominant social media platforms are leading to more uniform emoji use patterns across Turkish- and English-speaking groups.

These findings carry important implications for marketing scholarship and practice, particularly in the realm of international brand communication. As emojis become increasingly central to digital branding, their function transcends mere stylistic embellishment; they serve as affective anchors that shape emotional tone, relational signaling, and brand personality in culturally specific ways. For marketers, understanding the semiotic dynamics of emoji use enables the crafting of more culturally congruent content strategies. In collectivist cultures, where social harmony and implicit emotional cues are paramount, emojis can be tactically deployed to foster inclusion, soften commercial messaging, and signal communal values. Conversely, in more individualist contexts, emojis may be leveraged to personalize brand voice, convey humor, or enhance authenticity in one-to-one interactions.

Moreover, the observed convergence in emoji diversity over time suggests an emergent global semiotic economy in which dominant digital platforms subtly standardize affective expression. This raises critical questions for marketing strategy: To what extent should brands localize their symbolic repertoire to align with cultural expectations, and when might it be more effective to adopt a cosmopolitan, cross-cultural emoji vernacular that resonates across linguistic boundaries? The tension between global consistency and local sensitivity remains central to digital brand positioning, and emojis, though small in form, are powerful levers in navigating this balance. Future marketing research should continue to explore how visual language elements like emojis interact with broader constructs such as brand anthropomorphism, emotional branding, and cultural congruity, especially in algorithmically curated communication environments.

Study 1: Similarity of preferred emojis across languages. To address RQ1: *"What are the differences in emoji selection preferences between Turkish- and English-speaking users on Platform X?"* This study examined the similarities of emojis utilized in posts during various popularity phases. Furthermore, we analyzed how these similarities changed over time across these phases.

Methodology. To explore variations in emoji selection preferences between Turkish and English-speaking users on Platform X, the study first analyzed the frequency of each emoji's use across the two datasets. The analysis revealed a Pareto-like distribution in both languages, where a small proportion of emojis accounted for most of their use; approximately 10% of emojis represented 90% of all appearances in the Turkish dataset. About 9% did so in the English dataset, suggesting that emoji preferences might follow a power-law distribution (please see Fig. A1 in Supplementary Information).

A standard method to identify power-law distributions involves log-transforming both ranks and frequencies of data and seeking a linear pattern, as outlined by Clauset et al. (2009). However, this study found that a consistent power-law distribution across all ranks did not best fit the data for this particular dataset (Fig. A2 in Supplementary Information). Instead, following the methodology proposed by Kejriwal et al. (2021), the emojis were counted, ranked according to frequency, and then classified into three popularity phases: high, moderate, and low. Each phase was found to fit a power-law distribution with distinct parameters, providing a better view of emoji popularity and use dynamics forming emoji sets for comparability across languages (please see Fig. A3 in Supplementary Information for details on popularity phases).

To quantify the similarity of popular emojis between English and Turkish, we calculated the Jaccard Distance (JD) for each emoji popularity phase:

$$Jaccard\ Similarity_i(N_i^{en}, N_i^{tr}) = \frac{|N_i^{en} \cap N_i^{tr}|}{|N_i^{en} \cup N_i^{tr}|} \quad (3)$$

$$JD_i\left(N_i^{en}, N_i^{tr}\right) = 1 - Jaccard\ Similarity_i(N_i^{en}, N_i^{tr})$$

where i refers to the popularity phase, N_i^{en} and N_i^{tr} denote the distinct set of emojis in phase i for the English and Turkish languages, respectively.

Jaccard Similarity is a measure of similarity between two sets of data, ranging from 0 to 100%, which quantifies the proportion of distinct items that are common in both sets. On the other hand, JD is a measure of dissimilarity between the two sets, which is calculated by subtracting the Jaccard similarity from 100%. In our case, JD was a suitable statistic that was not affected by the imbalanced data size of each group and the slight differences in

the ranking of the items in the sets. To determine whether the JD statistic was significant or due to chance, we also applied a Monte Carlo permutation test, but replacing the fourth step with the JD_i statistic. The procedure was also repeated monthly from 2016 to 2021. During this time, monthly similarity statistics (JD_i) were calculated for English and Turkish and plotted as a function of time, helping to examine whether the trend changed over time.

Results. As presented in Table 6, JD values increased across the high, moderate, and low popularity levels of 0.42, 0.49, and 0.65, respectively, to indicate that emojis associated with higher popularity were more consistently used across languages, while less popular emojis showed greater divergence in use patterns. The *p*-values for the three popularity levels calculated by Monte Carlo simulations, from highest to lowest, were 0.006, 0.028, and 0.0. These results suggest that the JD values for the highest and lowest popularity levels were statistically significant, but the JD value for the moderate level was observed more than 1% of the time in the random dataset, suggesting that the computed statistic may have occurred by chance. Additionally, JD increased significantly to 0.65 in the low popularity level. In contrast, there was a significant level of similarity in the most popular emojis between the two cultures, with a decrease in similarity for the lower popularity levels. The infrequent use of these emojis may render them as statistical noise without a distinct pattern.

We also measured similarity as a function of time and observed some noticeable patterns. In Fig. 5, the monthly JD statistics between English and Turkish emoji preferences are plotted. For the two lower popularity levels, the preferred emojis were mostly dissimilar (JD 70% on average). Both curves followed a similar pattern, converging at around 74% with slopes close to zero. However, the most popular 200 emojis showed a trend toward increasing similarity over time, starting at 90% JD.

Table 6 Similarity distance in JD across popularity levels.

Popularity	JD[a]	*p*-value**
High	0.42	0.006 < 0.01
Moderate	0.49	0.028
Low	0.65	0.000 < 0.01

[a]Jaccard Distance. **Probability of observing measured JD statistic in Monte Carlo permutation test.

When considering the decreasing diversity trends in the previous section together with the increasing similarity of popularly preferred emojis, it appears that the preferences in both communities are converging towards a smaller and more similar subset of emojis.

Discussion. We adapted Kejriwal et al.'s (2021) method to categorize the similarity of preferred emojis based on their popularity phases. They ranked emojis according to their frequencies and segmented them into three popularity phases: high, moderate, and low. We observed that their distribution followed a power-law pattern, with each phase displaying distinct attributes. According to Kejriwal et al. (2021) and Li et al. (2019), the similarity among the most favored emojis diminishes as you move from the USA (West) to Japan (East Asia). However, our findings suggested a different pattern when a less collectivist and more individualistic-oriented Turkey is considered, suggesting that while the proportion of common emojis preferred decreased from high to low popularity phases, the similarity of the highly popular group among English and Turkish Platform X users increased over time. When considering the combined implications of Studies 2 and 3, it appeared that while there may initially be cultural variations in emoji use, the trend over time was towards a reduction in diversity and an increase in the commonality of certain emojis. Also aligned with Revita et al. (2023)'s findings, this pattern suggested that digital communication, particularly the use of emojis, is an area where global cultural convergence is observable. This trend may be influenced by the widespread appeal of emojis that represent fundamental human emotions and expressions, transcending cultural boundaries.

Managerially, the findings highlight the delicate balance between leveraging universally popular emojis for global consistency and using culturally specific ones to engage localized markets effectively. For instance, global brands can maintain a consistent brand identity while strategically adapting emoji use to local preferences for heightened engagement. Brands in English-speaking markets should consider using a broader variety of emojis to address diverse communication preferences, while a more focused use may better suit markets like Turkey, where less diversity in emoji use is observed. Moreover, the trend of decreasing similarity in emoji use across lower popularity levels underscores how shared cultural conventions may emerge around highly popular symbols, while less frequently used emojis leave greater room for localized or culturally specific interpretation.

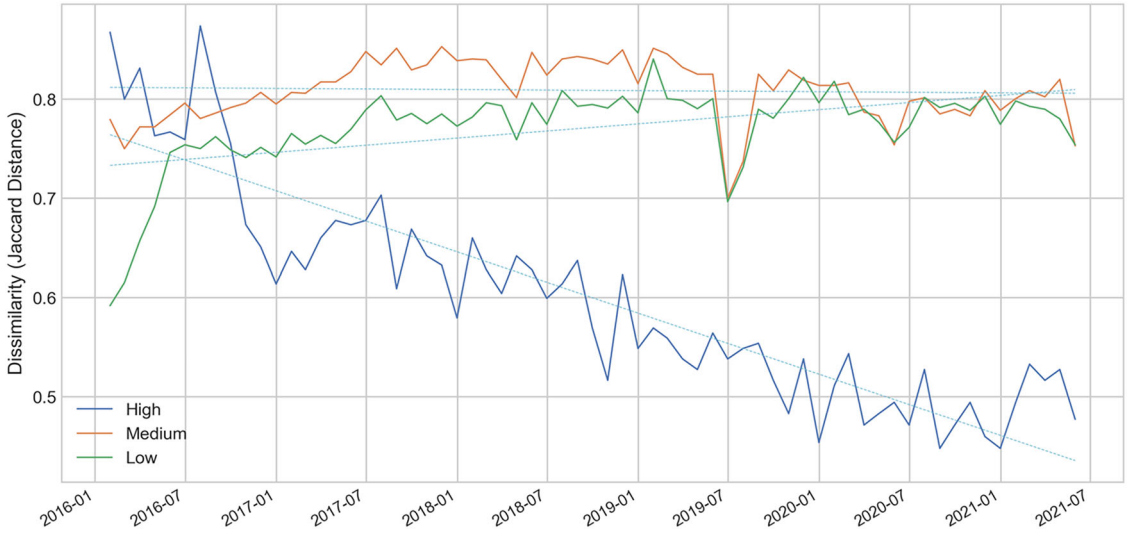

Fig. 5 Similarity distance of emojis by popularity levels and time. (source: authors' own creation).

Understanding this divergence is important in global brand communication, as it highlights the need for careful adaptation of less common visual elements to avoid misalignment with cultural expectations.

From a marketing perspective, the evolving convergence in high-frequency emoji use points to the emergence of a shared symbolic repertoire, which might be termed a form of "global emoji grammar" that offers brands an increasingly stable set of affective cues for cross-market communication. Yet this symbolic stability should not be mistaken for uniformity in interpretation. While highly popular emojis may afford a baseline of global recognizability, their contextual meanings remain mediated by cultural sensibilities, usage norms, and audience expectations. As such, marketers must navigate a semiotic double bind: on one hand, leveraging these ubiquitous symbols to craft cohesive brand narratives across borders, and on the other, attuning to the subtle divergences in interpretation that persist beneath the surface of apparent convergence.

This tension is particularly pronounced when deploying less popular or emerging emojis, which tend to carry more localized or experimental meanings. These symbols, though potentially rich in resonance, also carry higher semiotic risk; they may be interpreted differently or fail to register across cultures. Here, brands have an opportunity to differentiate by cultivating emoji literacy that is both culturally informed and contextually agile. Incorporating real-time cultural monitoring, audience feedback, and localized A/B testing into social media strategy can help brands fine-tune their use of symbolic elements in ways that maximize emotional relevance while minimizing cultural misalignment. Ultimately, emojis are not only tools of expression but also instruments of relational branding, and their strategic use in digital communication offers a unique lens through which to study the dynamic interplay of culture, technology, and marketing in a globalized media landscape.

Study 2: Semantic emotional representation of emojis. This study aims to address RQ2: "*What are the differences in emotional semantic meanings associated with commonly used emojis in branding communication across Turkish- and English-speaking markets on Platform X?*" To do so, we derived probability distributions of six basic emotions for each emoji in each language based on their co-occurrence with emotional words. These distributions were then compared across languages to assess the cross-cultural similarity in the emotional interpretation of emojis.

Methodology. A single emoji might represent a combination of emotions at different levels, making it difficult to support the idea that an emoji is used for a single emotion by different users. Co-occurrence analysis is a common method for exploring underlying patterns in text data (Harris, 1981) and was shown to be useful in user-emoji preference studies (Kaye et al., 2021; Liu et al., 2023; Wang et al., 2023). Riloff (1996), Tang et al. (2013), Song et al. (2015), and Chen et al. (2021) provide examples of how emojis can be analyzed through their co-occurrence with emotional seeds to calculate the discrete probability distribution of emojis across six basic emotions: happiness, surprise, fear, sadness, anger, and disgust (Ekman, 2005).

To understand how emojis represent basic emotions, we implemented an analysis of their semantic use as a similar approach to address the research question: "*What are the differences in the semantic representations of six basic emotions (happiness, surprise, sadness, anger, fear, and disgust) associated with each emoji between Turkish- and English-speaking users on Platform X?*" The top 90 words scored for each basic emotion as emotional seed words from the Turkish Emotion Lexicon for Turkish (Tocuoglu, 2019) and the NRC Emotion Lexicon (Zhu, 2014) for English were selected as seed words. Then, the relevance of an emoji to a set of emotional words is determined using the co-occurrence frequencies of emojis with the selected emotional words. These results allowed us to obtain a discrete probability distribution for each emoji, showing their relevance to the six basic emotions. For a particular emoji, the probability of representing one of the six emotions was calculated as follows:

$$P_i(e_j) = \frac{\sum_k co\_frequency(e_j,\ ew_{ik})}{\sum_k \sum_i co\_frequency(e_j,\ ew_{ik})} \quad (4)$$

where i refers to basic emotions as happiness, surprise, fear, sadness, anger, and disgust. *ej* denotes the emoji at index j. $cof\_requency(e_j, ew_{ik})$ refers to the requency that that emoji *ej* and the emotional word ew_{ik} at index k in the i'th emotion category co-occur within a post in the dataset. The denominator is the sum of all co-occurrences with all the emotional words.

The emotion probability distributions of emojis in English and Turkish were contrasted using Jensen-Shannon Divergence (JS), a symmetrical distance metric for discrete probability distributions (Endres and Schindelin, 2003). JS is computed as the weighted sum of Kullback-Leibler Divergence (MacKay, 2003), weighted by the probabilities of the events under consideration, which in this case are the probabilities of observing six emotions for each emoji. To obtain an overall understanding of emotional patterns in semantic use, common emojis for each language were manually categorized based on the sentiment they convey, namely positive, negative, or neutral, using emoji sentiment scores provided by Novak et al. (2015b). For instance, a 🙂 represents a positive sentiment and is labeled as "Positive," while 🙁 represents a negative sentiment and is labeled as "Negative." Subsequently, two binary variables were created to indicate the negativity and positivity of an emoji, with neutrality serving as the reference point. The JD distribution followed a bell-shaped curve; thus, an OLS regression model was employed to assess the significance of differences based on sentiment. This methodology allows for distinguishing how similarly emojis were utilized across different sentiments and identifying emojis that are emotionally comparable or distinct in their use.

Results. Among the 200 most popular emojis selected for comparison, 128 were common in both English and Turkish posts. These common emojis accounted for more than 80% of all emoji use in both datasets. As previously described, we calculated the probability distributions of happiness, surprise, anger, disgust, sadness, and fear for each emoji based on their co-occurrences with emotional words. We then quantified the differences in the emotional use of emojis between the two languages using the Jensen-Shannon divergence. Then, emojis were coded by sentiment labels from Novak's ranking scores (Novak et al., 2015a) as positive, neutral, or negative.

The distribution of distances by positive, neutral, and negative sentiments is shown in Fig. 6. The most emotionally diverse lexical use was observed for neutral symbols (mean: 0.369, std: 0.161). On the other hand, emojis with negative sentiment were more similarly used than those with positive sentiment, with mean distances of 0.208 (std: 0.04) and 0.288 (std: 0.06), respectively.

To test whether the observed differences in distances were statistically significant, we fit an OLS regression model as described earlier. Before the model fitting process, we created two dummy variables to represent the negativity and positivity of an emoji, with neutrality as the baseline. The output of the OLS model is shown in Table 7, where all estimates are significant,

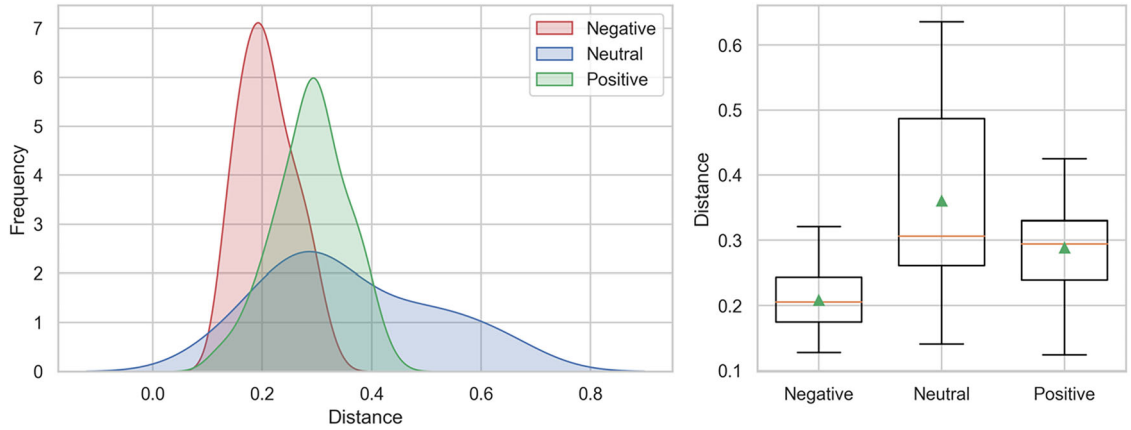

Fig. 6 Distribution of distances by sentiments. Left: Distance densities by sentiments; Right: Box plots of distances by sentiments. (source: authors' own creation).

Table 7 OLS estimates of sentiments and mean distances.

	Estimate	Std. Err.	*t*	*p*-value
Neutral (constant)	0.3699	0.021	18.005	0.00*
Negative	−0.0815	0.022	−3.626	0.00*
Positive	−0.1614	0.024	−6.806	0.00*

*$p < 0.01$.

suggesting that negative emojis are more semantically similar in use across cultures in terms of the combination of emotions they represent. In contrast, positive emojis are used less similarly in brand posts and replies across languages.

Figure 7 provides a more detailed view of emotion distributions for common emojis across two languages. The first two plots show how these emojis align with six basic emotions: happiness, surprise, anger, disgust, sadness, and fear, while the bottom plot reports the Jensen-Shannon divergence scores that capture cross-language differences. In English, these emojis appear most frequently with happy words. In Turkish, however, positive emotions such as happiness and surprise are distributed more evenly alongside negative emotions like anger, disgust, sadness, and fear. Notably, although certain emojis may seem tied to a single emotion, in practice, they are often used interchangeably with words expressing different emotions.

Drawn from the co-occurrence-based emotional profiles shown in Fig. 7, a notable example is the ❤️ red heart emoji, which appears frequently in both Turkish and English posts. In the English dataset, its emotional profile is dominated by happiness and surprise, reflecting affectionate or joyful interpretations. In contrast, its use in Turkish shows a broader emotional range, with notable co-occurrence not only with happiness but also with sadness and disgust, possibly reflecting more ambivalent or bittersweet emotional contexts in which the emoji is used (i.e., expressions of support during difficult times or emotionally intense situations).

Another shared emoji, 😍 (smiling face with heart-eyes), similarly exhibits cultural divergence. While English-speaking users primarily associate it with happiness, Turkish posts show a more blended emotional profile, with co-occurrences indicating surprise and sadness as secondary associations, suggesting a possible extension of use toward admiration mixed with nostalgia or longing.

Conversely, emojis with negative sentiment, such as 😢 (crying face) or 😠 (angry face), tend to show more stable emotional meanings across languages. As shown in the lower panel of Fig. 7, the Jensen-Shannon divergence scores for these emojis are consistently low, confirming that users in both cultures assign them similar emotional interpretations, likely due to the more universally recognizable expressions of negative emotions.

By contrast, neutral or ambiguous emojis, such as 🙃 (upside-down face) or 🤔 (thinking face), exhibit the highest divergence scores, indicating substantial variation in their emotional meanings across cultures. These emojis may serve more context-dependent or ironic functions that differ in interpretation between Turkish and English speakers.

Taken together, these examples support our finding that while emoji preferences may converge across cultures, the emotional meanings attached to the same emojis often diverge, especially for positive and neutral sentiment emojis. This highlights the importance of understanding emojis' cultural interpretation in global branding and their semiotic flexibility in digital messaging.

Discussion. In this study, each emoji was represented in a 6-dimensional vector space based on its relevancy to the six basic emotions of happiness, surprise, sadness, fear, disgust, and anger. The probabilistic distance between English and Turkish datasets for each emoji was then measured, after which the emojis were grouped according to their general sentiment (positive, negative, or neutral), and the distribution of these distances was evaluated.

Li et al. (2019) identified differences in emoji use that correspond with Hofstede's Cultural Dimensions Model, emphasizing notable disparities in how people use emojis across cultural lines. They found that users from more individualistic cultures (e.g., the USA and the UK) are more likely to use emojis with positive sentiment, while users from collectivist cultures are more inclined toward negative sentiment emojis. Similarly, Gao and VanderLaan (2020) noted that individualistic cultures, favoring direct expression, tend to interpret positive emojis more literally and diversely, leading to a broader range of positive emoji interpretations and use. Collectivist cultures, on the other hand, rely on shared understanding and context to interpret emotions in a different and more consistent way, especially when it comes to negative emotions. Kaye et al. (2021), studying how emotional valence affects emoji processing in terms of response speed and accuracy, showed that while positive emojis are processed more quickly, they lack the universal interpretation of negative emojis. Positive emotions are culturally dependent, resulting in varied use, while negative emojis are processed relatively consistently because negative emotions are universally recognized. Our findings indicate that negative emojis have a more consistent

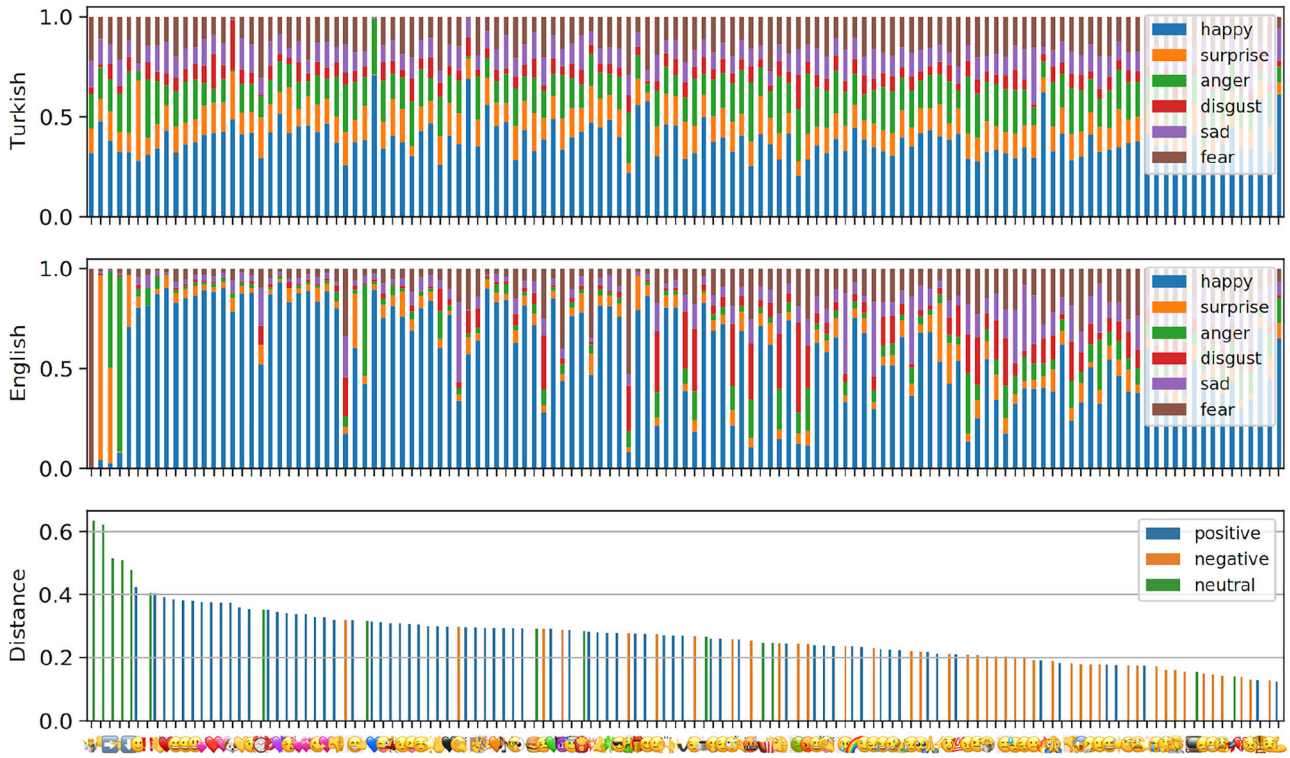

Fig. 7 Emotional distribution of emojis and the distance of their distributions between languages. (source: authors' own creation). *Top: Emotional probability distribution of emojis in Turkish; Center: Emotional probability distribution of emojis in English; Bottom: Distance in divergence colored by sentiments.

use pattern, reflecting their emotional semantics, than positive emojis between individualistic English-speaking and collectivist Turkish-speaking markets in branding communications.

Study 1 revealed that, despite cultural differences between Turkish- and English-speaking user groups, both groups tend to favor similar emojis in brand-related communications. However, Study 2's findings suggest that there are clear differences in the emotional semantics of emojis regarding their usage. Thus, understanding the emotional connotations of different emojis and their cultural interpretations is crucial for emotionally intelligent communication strategies. This approach can prevent miscommunication and increase emotional clarity, particularly in international marketing campaigns. A significant takeaway is the crucial impact of strategic emoji selection on brand perception and customer interaction. Prior research has shown that emojis can influence users' emotional responses and interaction behaviors, especially when aligned with cultural and contextual expectations (Koch et al., 2023; McShane et al., 2021). Marketers should recognize that the choice of emojis can directly influence post credibility and user engagement metrics, such as likes and shares, as demonstrated by Das (2019) and Ko et al. (2022). The careful selection of emojis can enhance perceived playfulness, boosting engagement with brand posts on platforms like X (McShane et al., 2021). However, overuse or inappropriate use of emojis can decrease post credibility and trustworthiness (Koch et al., 2023).

As shown, the emotional use of emojis differs significantly between English and Turkish, with emojis conveying negative sentiment being used more similarly than those conveying positive sentiment. Brands should tailor their campaigns to align with the emotional expectations and norms of each cultural group. For instance, a brand might choose to use different sets of emojis in their campaigns targeting Turkish- and English-speaking markets to align more closely with the emotional expectations and cultural norms of each group. Additionally, our findings reveal that the emotional use of emojis differs across cultural contexts, aligning with earlier research that highlights the roles of communication context, timing, and social media platforms in shaping emoji usage (Franco and Fugate, 2020). Maintaining uniform emoji use across various cultures, times, and platforms supports the credibility and integrity of brand communications. Moreover, reciprocal emoji use can positively impact interpersonal relationships, promoting stronger communication outcomes and building a meaningful bond between brands and users (Cavalheiro et al., 2024).

Overview of key constructs and empirical insights. To bring together the main parts of our analysis, Table 8 offers a clear summary of the key constructs examined. These constructs are emoji prevalence and diversity (descriptive analysis), selection preferences (Study 1), and differences in emotional semantics (Study 2). For each, we outline how it was defined, the theory behind it, the specific research question it relates to, the methods we used, and what we found. Each of the constructs in the analysis reflects a different aspect of how emojis function in brand communication and how their use varies across Turkish- and English-speaking users, two linguistically and culturally distinct groups.

Conclusions, limitations, and future research directions

Conclusions. This study set out to investigate how cultural contexts shape the use of emojis in digital brand communication, specifically through a comparative analysis of English-speaking and Turkish-speaking user interactions on Platform X. Through an empirical analysis of emoji use and preferences in a dataset naturally generated by users in a marketing context, we focused on cross-cultural differences between English-speaking and Turkish-speaking communities, which exhibit varying levels of social collectiveness. By drawing on large-scale, longitudinal data and employing information theory-based methods, the research

Table 8 Overview of constructs, theoretical anchors, methods, and findings.

Construct	Definition	Theoretical Anchor	Related RQ	Models	Measures	Key Results
Prevalence	Frequency of emoji use in brand posts and user replies across Turkish and English datasets	Paralinguistic cue theory (Luangrath et al., 2017); Hofstede's cultural dimensions (2010), (Gudykunst, 1997) CVC Theory	Descriptive, supporting RQ1	Zero-Inflated Negative Binomial Regression (ZINB)	Emoji count per post; binary language variable	Emoji usage is generally low in both languages. Turkish brand posts contain more emojis than English ones; difference is less pronounced among user replies. English users are more likely to use multiple emojis when they do use any.
Diversity	Richness of emoji vocabulary used in posts	Lexical richness in communication; cultural convergence/ divergence theories (Revita et al., 2023)	Descriptive, supporting RQ1	Type-Token Ratio (TTR); Monte Carlo permutation testing; *t*-tests	TTR normalized by 1,000 posts (overall and monthly)	English users employ a more diverse set of emojis. However, emoji diversity in both languages converges downward over time (2016-2021), indicating possible globalization of emoji vocabulary.
Similarity of Preferred Emojis (Study 1)	Overlap in top-used emojis between Turkish and English, segmented by popularity level	Cultural variability in emoji preference; power-law dynamics in emoji use (Kejriwal et al., 2021)	**RQ1:** What are the variations in emoji selection preferences between Turkish- and English-speaking users?	Power-law segmentation (high/ moderate/low popularity); Jaccard Distance; Monte Carlo simulations	Emoji frequency ranks; Jaccard Distance across popularity levels	Popular emoji sets are increasingly similar across languages, especially in the high-frequency group. Less popular emojis diverge more. Over time, the most-used emoji sets converge, supporting a trend toward a shared emoji repertoire.
Semantic Emotional Difference (Study 2)	Differences in usages along with emotional words of Ekman's 6 basic emotions.	Emotion semantics; Hofstede's (2010) dimensions; Gudykunst (1997) CVC Theory	**RQ2:** What are the differences in emotional semantic meanings associated with commonly used emojis?	Co-occurrence analysis; Jensen-Shannon Divergence; OLS regression with sentiment category dummies	Probability distribution across 6 basic emotions; semantic sentiment labels (positive/neutral/ negative)	Negative emojis are used more similarly across cultures; positive emojis show greater divergence. Emotional meanings of emojis vary by culture, but some convergence is visible in negative/ emotionally basic forms.

contributes a nuanced perspective to the literature on cross-cultural digital communication. Prior research on cultural differences in emoji preferences has frequently indicated correlations between national context and emoji usage. However, these studies often relied on data collected over short timeframes and conducted analyses at aggregate levels, thereby overlooking the temporal dynamics and evolving nature of social behavior. By contrast, our approach integrated information theory-based methods and accounted for temporal variation, thereby responding to prior calls for more detailed and context-sensitive analyses using alternative variables and more differentiated cross-country comparisons (Gao and VanderLaan, 2020; Guntuku, 2019). The research is conceptually grounded in Gudykunst's cultural variability in communication (CVC) framework (Gudykunst, 2005; Gudykunst, 1996) and Hofstede's cultural dimensions (Hofstede, 2011; Hofstede et al., 2010), which together offered a robust interpretive lens to examine how non-verbal cues such as emojis function within cultures characterized by different communication styles. The CVC framework elucidated the emotional layering typical of high-context cultures such as Turkey, while Hofstede's individualism–collectivism dimension revealed the more explicit, low-context tendencies of emoji use in English-speaking environments. As noted by Sun et al. (2023) and Togans et al. (2021), the individualism–collectivism dimension has proven particularly powerful in revealing how cultural orientations shape both the semantics and frequency of emoji use (Gao and VanderLaan, 2020; Sun et al., 2023; Togans et al., 2021).

Our findings revealed a dual pattern of divergence and convergence. On one hand, we observed culturally specific deployments and interpretations of emojis, particularly in how emotional nuances were conveyed and received. On the other hand, a gradual convergence was also evident, with both linguistic groups increasingly employing a shared repertoire of high-frequency emojis. This trend suggests the emergence of a global visual lexicon, though with notable asymmetries. Specifically, negative emojis tended to be interpreted more consistently across cultural contexts, while positive emojis displayed greater variability, supporting earlier findings by Kaye et al. (2021) and Li et al. (2019).

Our research underscores the value of examining emoji use not merely as a stylistic device but as a culturally embedded mode of emotional expression. It advances communication marketing theories by demonstrating that emojis function as more than decorative elements in brand communication; they are semiotic tools that reflect, reproduce, and potentially reshape cultural values in digital environments. For practitioners, these insights highlight the strategic importance of culturally intelligent emoji use, balancing universal recognizability with local resonance. As digital communication continues to evolve, brands that navigate these cultural subtleties with care and precision will be better equipped to foster meaningful engagement across diverse audiences.

Limitations. While this study offers substantive theoretical and empirical contributions to the understanding of cross-cultural emoji use in digital brand communication, it remains bounded by several limitations that delineate the scope of its insights and illuminate promising pathways for future inquiry.

One salient limitation lies in the platform specificity of the analysis. Conducted exclusively on Platform X, the study reflects communicative behaviors shaped by its character-constrained, rapid-exchange environment. Such a format inherently privileges brevity, intensifies reliance on symbolic compression, and may thus amplify the use of emojis as affective shorthand. Yet this context is but one manifestation of digital interaction. Platforms that encourage richer multimedia expression or sustained dialogue, such as Instagram, WhatsApp, or TikTok, may foster different paralinguistic practices. As such, extending this research across diverse digital ecosystems would offer a more textured understanding of how technological affordances mediate cultural patterns of expression.

Moreover, the study's use of language as a proxy for culture, while grounded in established cross-cultural frameworks, necessarily flattens the intricate heterogeneity of cultural identity. Turkish and English, though useful entry points into differing cultural paradigms, do not exhaust the complexity of users' lived experiences or communicative repertoires. Cultural orientations are shaped not only by language but also by intersecting factors such as age, education, gender, geography, socio-political context, and transnational exposure. A finer-grained analysis that integrates such dimensions could reveal subtler layers of variation and shed light on emergent hybridities in global digital cultures.

The brand-centered context of the dataset also introduces conceptual boundaries. While brand-user interactions are highly relevant for studying emotive expression in commercial discourse, they represent only a slice of the broader digital communicative landscape. The emotional scripts governing engagement may differ markedly from those animating private conversations, activist messaging, or community-driven content. Future work might therefore consider how emoji use operates across varying discursive domains, each with its own affective textures, relational norms, and cultural registers.

Methodologically, the reliance on information-theoretic measures and co-occurrence with curated emotional seed words provides analytic clarity, but at the cost of overlooking certain expressive subtleties. Emojis are semiotically rich and context-sensitive; they often convey irony, ambivalence, or layered meaning that eludes computational models calibrated to detect explicit sentiment. While the chosen approach affords replicability and large-scale applicability, it could be fruitfully complemented by interpretive or participatory methods, ranging from discourse analysis and digital ethnography to user-centered experiments that further investigate how individuals experience and ascribe meaning to these visual symbols.

Finally, although the observed pattern of convergence in emoji use gestures toward a shared digital lexicon, such convergence should not be mistaken for cultural uniformity. Similar patterns of usage may mask divergent interpretations, and surface-level regularities can obscure enduring differences in emotional encoding and communicative intent. Emoji use, like language itself, is continuously evolving, negotiated within communities, and shaped by global flows and local constraints. Continued longitudinal research is needed to trace how meanings stabilize, fragment, or shift over time within and across cultural boundaries.

Future research directions. Building on the theoretical scaffolding and empirical findings of this study, several promising directions for future research emerge, each capable of advancing our understanding of emoji use as a culturally embedded communicative practice within digital brand discourse.An immediate extension would involve broadening the platform ecology under investigation. While Platform X provides a compelling case of paralinguistic expression under structural constraint, future studies might explore whether similar patterns hold across more visually immersive or conversational environments. Investigating how interface design, multimodal integration, and temporality shape emoji usage across platforms such as Instagram, WhatsApp, or TikTok would yield a more nuanced understanding of how technological affordances mediate cultural expression.

Table 9 Suggested open-ended research questions for advancing the study of emoji use in cross-cultural digital communication.

No	Suggested question for future research
1	How do platform-specific affordances influence the cultural encoding and decoding of emojis in digital communication?
2	In what ways do demographic variables (e.g., age, gender, or education) interact with cultural norms to shape emoji use across digital environments?
3	How do diasporic and multilingual communities navigate emoji semantics across linguistic and cultural boundaries?
4	To what extent do algorithms and predictive emoji suggestions influence cultural convergence or divergence in emoji preferences?
5	How are emojis used to signal identity, alignment, or resistance in subcultural and activist digital spaces?
6	What role do emojis play in the performance and perception of emotional labor in professional or branded online communication?
7	How does the meaning of commonly used emojis evolve over time within specific linguistic and cultural communities?
8	Can emojis function as a form of visual lingua franca, and if so, what are the limits of their universality?
9	How do different cultural contexts interpret emoji sequences or combinations, especially those intended to convey complex emotions or narratives?
10	What is the role of irony, satire, or ambiguity in emoji use, and how is such nuance culturally situated?

Equally important is the expansion of cross-linguistic and cross-cultural comparisons beyond the English–Turkish dyad examined here. Incorporating a wider range of linguistic and cultural communities, including those informed by Confucian, Arab, Latin American, or African worldviews, would enrich the typology of global emoji usage. These investigations could also engage with linguistic relativity, examining how grammar, writing systems, and culturally bound metaphors influence the visual-emotive palette of digital users.

A deeper investigation of demographic and psychographic variables would likewise yield valuable insights. Factors such as age cohort, gender identity, professional orientation, and digital fluency may modulate emoji interpretation and deployment, intersecting with cultural values in complex ways. Disaggregating these dimensions would allow researchers to detect subcultural dynamics, generational divergence, and identity-driven emoji appropriation that are obscured at the macro level.

From a methodological standpoint, future research could benefit from multi-modal and mixed-method approaches that move beyond probabilistic co-occurrence models. Ethnographic inquiry, user diaries, affective mapping, or controlled experimental designs could illuminate the affective, strategic, and sometimes ironic deployment of emojis in everyday digital life, particularly in ways that defy standard sentiment classification.

Finally, theoretical development remains an open frontier. Concepts such as "platformed affect," "emotive capital," or "visual vernaculars" deserve closer attention as scholars seek to account for how emoji use intersects with symbolic labor, emotional curation, and algorithmic recommendation systems. Of particular interest is the entanglement of personalization algorithms with emoji semantics: How might such systems reinforce normative emotional repertoires or inadvertently standardize symbolic expression across cultural contexts?

To support these trajectories, a curated set of open-ended research questions is presented in Table 9, each designed to catalyze deeper and more diverse scholarly engagement with emoji use across digital cultures.

Data availability

Compliance with data protection laws and platform terms of service We comply with our countries' data protection laws and the relevant platform's terms of service. The data supporting this study's findings were derived from publicly available posts and replies on Platform X (formerly Twitter) between 2016 and 2021. Data collection, processing, and dissemination were conducted in compliance with applicable data protection legislation in Turkey (Republic of Türkiye - Personal Data Protection Law No. 6698, KVKK) and Sweden (Kingdom of Sweden - Ethical Review Act 2003:460), as well as with Platform X's terms of service and developer policies. The study did not involve the processing of special categories of personal data as defined under the applicable legal frameworks. As per the platform's requirements, raw textual content and any reversible textual representations are not redistributed. Moreover, to minimize the risk of user re-identification, all user-identifying information was removed prior to analysis, and only tweet identifiers and irreversible, derived features are part of the open access dataset. Anonymization and minimization of re-identification risk. We confirm that we anonymized all personal identifiers. All personal identifiers, including usernames, user IDs, and profile metadata, were removed prior to analysis. No attempts were made to identify, profile, or contact individual users. To further minimize the risk of re-identification, raw textual content and any reversible textual representations are not redistributed. The shared data consists solely of tweet identifiers and aggregated, irreversible derived features, such as emoji and seed-word count incidence matrices. Data sharing and access conditions. We confirm that we used secure, controlled-access repositories or data use agreements for sharing. The publicly released dataset includes tweet identifiers (IDs), emoji, and seed-word count incidence matrices, and associated metadata necessary to reproduce the analyses. These materials are deposited in Zenodo, a secure and widely used research data repository, and are publicly accessible at: https://zenodo.org/records/18093119. Access to tweet text, where permitted by Platform X and subject to content availability, requires rehydration using the provided tweet identifiers in accordance with Platform X's terms of service. No additional data use agreements or controlled-access arrangements apply beyond compliance with the platform's access conditions.

Code availability

All code used for data preprocessing, feature extraction, and analysis based on the processed datasets is openly available in a public GitHub repository: https://github.com/atanaltay/beyond_words_emoji_patterns.git. The repository contains scripts required to reproduce the analytical workflow using the shared derived data. No scripts for collecting or redistributing raw tweet content are provided.

Received: 18 November 2024; Accepted: 18 February 2026;
Published online: 10 March 2026

References

Agovino M, Bevilacqua M, Cerciello M (2021) Language as a proxy for cultural change. A contrastive analysis for French and Italian lexicon on male homosexuality. Qual Quant 1–24. https://doi.org/10.1007/s11135-021-01121-x

Alfano M, Reimann R, Quintana IO, Chan A, Cheong M, Klein C (2023) The affiliative use of emoji and hashtags in the black lives matter movement in Twitter. Soc Sci Comput Rev 41(6):2145–2160. https://doi.org/10.1177/08944393221131928

Almaguer J, Felix R, Harmeling CM (2025) Emoji marketing: Toward a theory of brand paralinguistics. Int J Res Mark 42(1):95–112. https://doi.org/10.1016/j.ijresmar.2024.06.002

Anderson D, Stuart M, Abadi M, Gal S (2018) Common hand gestures in the US that are offensive in other countries. Retrieved from https://www.businessinsider.com/hand-gestures-offensive-different-countries-2018-6#:~:text=The%20thumbs,of%20approval%20in%20most%20countries

Aydogdu V, Katırcı H (2021) Common language of new era in sport clubs: emojis. J Messenger 13(1):63–80

Baayen RH (2001) Word frequency distributions (Vol. 18): Springer Science & Business Media

Bai Q, Dan Q, Mu Z, Yang M (2019) A systematic review of emoji: current research and future perspectives. Front Psychol 10:2221. https://doi.org/10.3389/fpsyg.2019.02221

Boutet I, LeBlanc M, Chamberland JA, Collin CA (2021) Emojis influence emotional communication, social attributions, and information processing. Comput Hum Behav 119: 106722. https://doi.org/10.1016/j.chb.2021.106722

Casado-Molina AM, Rojas-de Gracia MM, Alarcón-Urbistondo P, Romero-Charneco M (2022) Exploring the opportunities of the emojis in brand communication: the case of the beer industry. Int J Bus Commun 59(3):315–333. https://doi.org/10.1177/2329488419832964

Cavalheiro BP, Prada M, Rodrigues DL (2024) Examining the effects of reciprocal emoji use on interpersonal and communication outcomes. J Soc Pers Relationsh https://doi.org/10.1177/02654075231219032

Chen Y, Yang X, Howman H, Filik R (2024) Individual differences in emoji comprehension: gender, age, and culture. PloS one 19(2):e0297379. https://doi.org/10.1371/journal.pone.0297379

Chen Z, Cao Y, Yao H, Lu X, Peng X, Mei H, Liu X (2021) Emoji-powered sentiment and emotion detection from software developers' communication data. ACM Trans Softw Eng Methodol (TOSEM) 30(2):1–48. https://doi.org/10.1145/3424308

Clauset A, Shalizi CR, Newman MEJ (2009) Power-law distributions in empirical data. SIAM Rev 51(4):661–703. https://doi.org/10.1137/070710111

Daniel J (2021) The most frequently used emoji of 2021. THE UNICODE EMOJI MIRROR PROJECT. Retrieved from https://home.unicode.org/emoji/emoji-frequency/?utm_source=chatgpt.com

Das G, Wiener HJ, Kareklas I (2019) To emoji or not to emoji? Examining the influence of emoji on consumer reactions to advertising. J Bus Res 96:147–156. https://doi.org/10.1016/j.jbusres.2018.11.007

Dinçer MAM, Yıldırım M, Dil E (2024) As an emerging market Turkish culture's quest to be positioned on Meyer's cultural map. Rev Int Bus Strat 34(1):126–151. https://doi.org/10.1108/RIBS-03-2023-0023

Ekman P (2005) Basic emotions. In: Handbook of cognition and emotion. pp 45–60

Endres DM, Schindelin JE (2003) A new metric for probability distributions. IEEE Trans Inf Theory 49(7):1858–1860. https://doi.org/10.1109/TIT.2003.813506

Evans V (2017) The Emoji Code: how smiley faces, love hearts and thumbs up are changing the way we communicate: Michael O'Mara Books

Fakhry NAM (2021) Employment of emoji in social awareness advertising campaigns. Int Des J 11(5):49–60. https://doi.org/10.21608/idj.2021.191678

Farhoudinia B, Ozturkcan S, Kasap N (2025) Fake news in business and management literature: a systematic review of definitions, theories, methods, and implications. Aslib J Inf Manag 77(2):306–329. https://doi.org/10.1108/AJIM-09-2022-0418

Farhoudinia B, Ozturkcan S, Kasap N (2024) Emotions unveiled: detecting COVID-19 fake news on social media. Human Soc Sci Commun 11(1). https://doi.org/10.1057/s41599-024-03083-5

Feldman LB, Aragon CR, Nan-Chen C, Kroll JF (2017) Emoticons in text may function like gestures in spoken or signed communication. Behav Brain Sci 40. https://doi.org/10.1017/s0140525x15002903

Franco CL, Fugate JM (2020) Emoji face renderings: exploring the role emoji platform differences have on emotional interpretation. J Nonverbal Behav 44(2):301–328. https://doi.org/10.1007/s10919-019-00330-1

Gao B, VanderLaan DP (2020) Cultural influences on perceptions of emotions depicted in emojis. Cyberpsychol Behav Soc Netw 23(8):567–570. https://doi.org/10.1089/cyber.2020.0024

Gawne L, McCulloch G (2019) Emoji as digital gestures. Language@ Internet 17(2). https://scholarworks.iu.edu/journals/index.php/li/article/view/37786

Green EG, Deschamps JC, Paez D (2005) Variation of individualism and collectivism within and between 20 countries: a typological analysis. J Cross Cultural Psychol 36(3):321–339. https://doi.org/10.1177/0022022104273654

Gudykunst WB (1997) Cultural variability in communication: an introduction. Commun Res 24(4):327–348. https://doi.org/10.1177/009365097024004001

Gudykunst WB (2005) Theorizing about intercultural communication. Sage, 2005

Gudykunst WB, Matsumoto Y, Ting-Toomey S, Nishida T, Kim K, Heyman S (1996) The influence of cultural individualism-collectivism, self construals, and individual values on communication styles across cultures. Hum Commun Res 22(4):510–543. https://doi.org/10.1111/j.1468-2958.1996.tb00377.x

Guntuku SC, Li M, Tay L, Ungar LH (2019) Studying cultural differences in emoji usage across the east and the west. Paper presented at the International AAAI Conference on Web and Social Media. https://doi.org/10.1609/icwsm.v13i01.3224

Harris ZS (1981) Co-occurrence and transformation in linguistic structure. In: Papers on syntax. Dordrecht: Springer, pp 143–210 https://doi.org/10.1007/978-94-009-8467-7_8

Hofstede G (2011) Dimensionalizing cultures: the Hofstede model in context. Online readings in psychology and culture 2(1). https://doi.org/10.9707/2307-0919.1014

Hofstede G, Hofstede GJ, Minkov M (2010) Cultures and organizations: software of the mind (Rev. 3rd ed. New York: McGraw-Hill

Jaeger SR, Roigard CM, Jin D, Vidal L, Ares G (2019) Valence, arousal and sentiment meanings of 33 facial emoji: Insights for the use of emoji in consumer research. Food Res Int 119:895–907. https://doi.org/10.1016/j.foodres.2018.10.074

Kastrati M, Kastrati Z, Shariq Imran A, Biba M (2024) Leveraging distant supervision and deep learning for Twitter sentiment and emotion classification. J Intell Inf Syst 62(4):1045–1070. https://doi.org/10.1007/s10844-024-00845-0

Kaye LK, Rodriguez-Cuadrado S, Malone SA, Wall HJ, Gaunt E, Mulvey AL, Graham C (2021) How emotional are emoji?: exploring the effect of emotional valence on the processing of emoji stimuli. Comput Hum Behav 116: 106648. https://doi.org/10.1016/j.chb.2020.106648

Kejriwal M, Wang Q, Li H, Wang L (2021) An empirical study of emoji usage on Twitter in linguistic and national contexts. Online Soc Netw Media 24. https://doi.org/10.1016/j.osnem.2021.100149

Kim J, Gupta P (2012) Emotional expressions in online user reviews: How they influence consumers' product evaluations. J Bus Res 65(7):985–992. https://doi.org/10.1016/j.osnem.2021.100149

Ko EE, Kim D, Kim G (2022) Influence of emojis on user engagement in brand-related user generated content. Comput Hum Behav 136: 107387. https://doi.org/10.1016/j.chb.2022.107387

Koch T, Denner N, Crispin M, Hohagen T (2023) Funny but not Credible? Why using (many) emojis decreases message credibility and source trustworthiness. Soc Media+ Soc 9(3):20563051231194584. https://doi.org/10.1177/20563051231194584

Lee CT, Hsieh SH (2019) Engaging consumers in mobile instant messaging: the role of cute branded emoticons. J Prod Brand Manag 28(7):849–863. https://doi.org/10.1108/JPBM-08-2018-1997

Li M, Chng E, Chong AYL, See S (2019) An empirical analysis of emoji usage on Twitter. Ind Manag Data Syst 119(8):1748–1763. https://doi.org/10.1108/IMDS-01-2019-0001

Liu T, Shi K, Hu L, Liu Y, Liu Y (2023) A new instrument for measuring customers' perceptions of service warmth: a big data and machine learning approach. SAGE Open 13(4):21582440231218803. https://doi.org/10.1177/21582440231218803

Luangrath AW, Peck J, Barger VA (2017) Textual paralanguage and its implications for marketing communications. J Consum Psychol 27(1):98–107. https://doi.org/10.1016/j.jcps.2016.05.002

Luna D, Peracchio LA (2001) Moderators of language effects in advertising to bilinguals: a psycholinguistic approach. J Consum Res 28(2):284–295. https://doi.org/10.1086/322903

Luna D, Ringberg T, Peracchio LA (2008) One individual, two identities: frame switching among biculturals. J Consum Res 35(2):279–293. https://doi.org/10.1086/586914

MacKay DJC (2003) Information theory, inference and learning algorithms. Cambridge University Press. https://doi.org/10.1086/586914

McShane L, Pancer E, Poole M, Deng Q (2021) Emoji, playfulness, and brand engagement on Twitter. J Interact Mark 53(1):96–110. https://doi.org/10.1016/j.intmar.2020.06.002

Novak K, Smailović J, Sluban B, Mozetič I (2015a) Emoji sentiment ranking. In: url:http://kt.ijs.si/data/Emoji_sentiment_ranking

Novak K, Smailović J, Sluban B, Mozetič I (2015b) Sentiment of emojis. PLoS ONE 10(12:):e0144296. https://doi.org/10.1371/journal.pone.0144296

Park J, Baek YM, Cha M (2014) Cross-cultural comparison of nonverbal cues in emoticons on Twitter: evidence from big data analysis. J Commun 64(2):333–354. https://doi.org/10.1111/jcom.12086

Revita I, Zalfikhe FA, Boholano H, Tuan NT, Huszka B, Datuk ZDM (2023) Emoticons unveiled: a multifaceted analysis of their linguistic impact. J Arbitrer 10(3):260–274. https://doi.org/10.1111/jcom.12086

Riloff E (1996) Automatically generating extraction patterns from untagged text. In: Paper presented at the Proceedings of the National Conference on Artificial Intelligence. https://cdn.aaai.org/AAAI/1996/AAAI96-155.pdf

Riordan MA (2017) Emojis as tools for emotion work: communicating affect in text messages. J Lang Soc Psychol 36(5):549–567. https://doi.org/10.1177/0261927X17704238
Sciandra M, Sharma N (2024) Picture perfect or symbolic backfire? assessing the detrimental effects of brand emoji use on consumers' brand attitudes. In: The 2024 Association of Marketing Theory and Practice Proceedings, 35. https://doi.org/10.20429/amtp.2024.29
Siddiqui I (2020) Linguistic impact on emojis in online communication. Int'l JL Mgmt Hum 3:1689. https://ijlmh.com/linguistic-impact-of-emojis-in-online-communication/
Soares AM, Farhangmehr M, Shoham A (2007) Hofstede's dimensions of culture in international marketing studies. J Bus Res 60(3):277–284. https://doi.org/10.1016/j.jbusres.2006.10.018
Song K, Feng S, Gao W, Wang D, Chen L, Zhang C (2015) Build emotion lexicon from microblogs by combining effects of seed words and emoticons in a heterogeneous graph. In: Paper presented at the Proceedings of the 26th ACM conference on hypertext & social media. https://doi.org/10.1145/2700171.2791035
Statcounter(2025) Social Media Stats Turkey. Retrieved from https://gs.statcounter.com/social-media-stats/all/turkey
Statista (2025) Most popular social networks worldwide as of February 2025, by number of monthly active users. Retrieved from https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/
Sun J, Lasser S, Lee SK (2023) Understanding emojis: cultural influences in interpretation and choice of emojis. J Int Intercult Commun 16(3):242–261. https://doi.org/10.1080/17513057.2022.2036790
Tanaltay A, Ozturkcan S, Kasap N (2025) Decoding digital engagement: a comparative analysis of English and Turkish brand post popularity dynamics on platform X. J Res Interact Mark 19(2):221–245. https://doi.org/10.1108/JRIM-10-2023-0368
Tang D, Qin B, Liu T, Li Z (2013) Learning sentence representation for emotion classification on microblogs. In: Paper presented at the CCF International Conference on Natural Language Processing and Chinese Computing. https://doi.org/10.1007/978-3-642-41644-6_20
Tocuoglu MA, Alpkocak A (2019) Lexicon-based emotion analysis in Turkish. Turkish J Electr Eng Comput Sci 27(2):1213–1227. https://doi.org/10.3906/elk-1807-41
Togans LJ, Holtgraves T, Kwon G, Zelaya TEM (2021) Digitally saving face: An experimental investigation of cross-cultural differences in the use of emoticons and emoji. J Pragmat 186:277–288. https://doi.org/10.1016/j.pragma.2021.09.016
Tweedie FJ, Baayen RH (1998) How variable may a constant be? Measures of lexical richness in perspective. Comput Humanit 32(5):323–352. https://doi.org/10.1023/A:1001749303137
Vandergriff I (2013) Emotive communication online: a contextual analysis of computer-mediated communication (CMC) cues. J Pragmat 51(1):1–12. https://doi.org/10.1016/j.pragma.2013.02.008
Wang X, Cheng M, Zhu J, Jiang R (2023) When texts meet emoji: a multi-stage study of tourism brands. J Travel Res 63(8):2062–2077. https://doi.org/10.1177/00472875231203396
Yuki M, Maddux WW, Masuda T (2007) Are the windows to the soul the same in the East and West? cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. J Exp Soc Psychol 43(2):303–311. https://doi.org/10.1016/j.jesp.2006.02.004
Zhu X, Kiritchenko, S, Mohammad SM (2014) Nrc-canada-2014: recent improvements in the sentiment analysis of tweets. In: Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014). pp 443–447 https://doi.org/10.3115/v1/S14-2077

Acknowledgements

We gratefully acknowledge the editor and anonymous reviewers for their insightful and constructive feedback, which has been invaluable in improving the clarity, rigor, and overall quality of our manuscript.

Author contributions

The first author conducted the research, retrieved the data, conducted the analysis, and drafted the manuscript as part of the PhD thesis successfully completed in 2023. The second author and PhD co-advisor provided extensive guidance throughout the research process, co-wrote sections of the manuscript, and offered critical feedback on the manuscript. The third author and PhD main advisor oversaw the overall project and provided valuable feedback on the manuscript.

Funding

Open access funding provided by Linnaeus University.

Competing interests

The second author (S.O.) was a member of the Editorial Board of this journal at the time of acceptance for publication. The manuscript was assessed in line with the journal's standard editorial processes, including its policy on competing interests. The other authors declare no competing interests.

Ethical approval

We followed ethics approval and informed consent where required and applicable. This article does not contain any studies with human participants performed by any of the authors. More precisely, the research involved no direct interaction, recruitment, intervention, or manipulation of individuals and was limited to non-interventional analysis of publicly accessible user-generated content. Under the applicable regulatory frameworks in both the Republic of Türkiye and the Kingdom of Sweden, formal ethics committee approval is required only for research involving interventions, identifiable biological material, or the processing of sensitive personal data. Because none of these conditions applied, ethics committee approval was not required or applicable to seek. Since the study relied exclusively on publicly available data and involved no direct engagement with individuals, individual informed consent was not required. This study analyzed publicly available content posted on Platform X (formerly Twitter) between 2016 and 2021. The research involved no direct interaction with individuals, no recruitment, and no physical or psychological intervention; it consisted solely of non-interventional analysis of publicly accessible user-generated content. The first and last authors are affiliated with Sabancı University, Turkey, and the corresponding author holds a secondary affiliation at the same institution. According to the applicable law in Turkey (Republic of Türkiye - Personal Data Protection Law No. 6698, KVKK, https://www.kvkk.gov.tr/Icerik/6649/Personal-Data-Protection-Law), ethics committee approval is only required where personal data are processed in ways subject to data protection regulation, particularly in relation to special categories of personal data. On the other hand, research based exclusively on publicly available sources, without the processing of sensitive personal data and without intervention, is exempt from ethics committee review. This interpretation is consistent with institutional research ethics guidelines established in Turkey as per the TR Dizin Ethical Principles flowchart (https://trdizin.gov.tr/wp-content/uploads/2022/04/TRDizin_etik_ilkeleri_akis_semasi.pdf) which governs research and publication practices. As the present study relied solely on publicly accessible social media content and did not process sensitive personal data as defined under applicable legislation, ethics committee approval was not required under the applicable Turkish regulatory framework. The corresponding author's primary affiliation is Linnaeus University, Sweden. According to the applicable law in Sweden (Kingdom of Sweden - Ethical Review Act 2003:460, https://etikprovningsmyndigheten.se/en/what-the-act-says), ethical review is required only for research involving physical interventions on individuals, methods intended to affect participants physically or psychologically, identifiable biological material, or the processing of sensitive personal data. Because the present study analyzed publicly available material, involved no intervention, and did not process sensitive personal data as defined under Swedish legislation, formal ethical approval was not required under Swedish law. All procedures were conducted in accordance with applicable national regulations and institutional guidelines in both jurisdictions and with internationally recognized ethical principles for research involving human-related data, including the World Medical Association Declaration of Helsinki (https://www.wma.net/policies-post/wma-declaration-of-helsinki) and the Belmont Report (https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html).

Informed consent

This article does not contain any studies with human participants performed by any of the authors. The study did not involve direct contact, recruitment, participation, or intervention with individuals but relied exclusively on publicly available social media posts and replies. In accordance with internet-mediated research ethics and data protection principles, no attempts were made to identify, profile, or contact users. Usernames, profile metadata, and any potentially identifying information were removed prior to analysis. Only aggregated and irreversible derived features (e.g., emoji frequency counts and co-occurrence statistics) were used in the analyses, and no identifiable personal data are reported in the manuscript. Because the study relied exclusively on publicly available content and involved no direct interaction with individuals, individual informed consent was not required under the applicable regulatory and ethical frameworks in Turkey and Sweden.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-026-06821-z.

Correspondence and requests for materials should be addressed to Selcen Ozturkcan.