Opico: A Study of Emoji-first Communication in a Mobile Social App

Sujay Khandekar
University of Illinois at Urbana-Champaign
khandek2@illinois.edu

Joseph Higgs
University of Illinois at Urbana-Champaign
joebh5@gmail.com

Yuanzhe Bian
University of Illinois at Urbana-Champaign
yb3@illinois.edu

Chae Won Ryu
University of Illinois at Urbana-Champaign
cwryu2@illinois.edu

Jerry O. Talton
Carta, Inc.
jerry.talton@carta.com

Ranjitha Kumar
University of Illinois at Urbana-Champaign
ranjitha@illinois.edu

ABSTRACT

In the last two decades, Emoji have become a mainstay of digital communication, allowing ordinary people to convey ideas, concepts, and emotions with just a few Unicode characters. While emoji are most often used to supplement text in digital communication, they comprise a powerful and expressive vocabulary in their own right. In this paper, we study the affordances of "emoji-first" communication, in which sequences of emoji are used to describe concepts without textual accompaniment.

To investigate the properties of emoji-first communication, we built and released Opico, a social media mobile app that allows users to create *reactions* — sequences of between one and five emoji — and share them with a network of friends. We then leveraged Opico to collect a repository of more than 3700 emoji reactions from more than 1000 registered users, each tied to one of 2441 physical places.

We describe the design and architecture of the Opico app, present a qualitative and quantitative analysis of Opico's reaction dataset, and discuss the implications of Emoji-first communication for future social platforms.

CCS CONCEPTS

• **Human-centered computing** → **Collaborative and social computing**.

KEYWORDS

emoji; social media; mobile app

ACM Reference Format:
Sujay Khandekar, Joseph Higgs, Yuanzhe Bian, Chae Won Ryu, Jerry O. Talton, and Ranjitha Kumar. 2019. Opico: A Study of Emoji-first Communication in a Mobile Social App. In *Companion Proceedings of the 2019 World Wide Web Conference (WWW '19 Companion), May 13–17, 2019, San Francisco, CA, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3308560.3316547

WWW '19 Companion, May 13–17, 2019, San Francisco, CA, USA

ACM ISBN 978-1-4503-6675-5/19/05.
https://doi.org/10.1145/3308560.3316547

1 INTRODUCTION

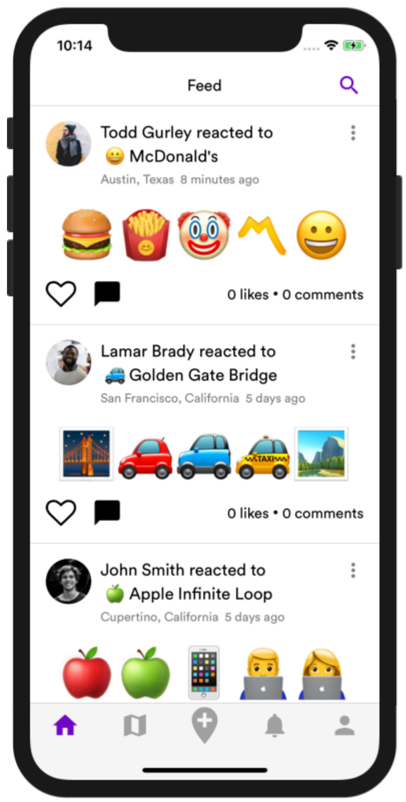

Figure 1: To study the properties of emoji-first communication, we built and deployed Opico, a social media mobile app that allows users to create *reactions* — sequences of one to five emoji — and share them with a network of friends.

Since they were first proposed in 1999 [21], emoji have exploded in popularity, allowing internet users to quickly convey ideas, concepts, and emotions with just a few Unicode characters. These pictorial — often playful — representations offer universal appeal: as of June 2018, the Unicode 11 standard includes 2,823 emoji in its specification [3], and Facebook reports that 92% of users between the ages of 13 and 18 (and 77% of users between the ages of 56 and 64) employ them in communications [2].

Researchers have studied the role emoji play in digital communication, focusing particularly on the way emoji are commonly used to supplement text by emphasizing key concepts or replacing certain words [18]. Some examples exist, however, of people composing messages *entirely* in emoji: Emoji Dick, for instance, is a fanciful sentence-by-sentence translation of Herman Melville's Moby Dick [23].

In this paper, we study the affordances of this kind of "emoji-first" communication, in which sequences of emoji are used to convey meaning without textual accompaniment. Rather than attempt to co-opt existing platforms to conduct such a study at scale in the wild, we built and released Opico, a social media mobile app that allows users to create emoji *reactions* — sequences of between one and five emoji characters — share them with a network of friends, and discuss their meanings (Figure 1).

We released Opico on the Apple App Store and Google Play in early 2018. As of the time of this writing, the app has gained more than 1000 registered users, and generated more than 3700 emoji reactions to a set of more than 2400 physical places. This paper describes the design and architecture of the Opico app, presents a qualitative and quantitative analysis of the resultant dataset, and reports the results of a focused study conducted with 23 Opico users.

Our results suggest that Emoji-first communication represents a powerful and promising mechanism for certain types of digital messages. We demonstrate how the constraints of an emoji-only vocabulary encourage users to succinctly express complex concepts, make jokes, and tell stories. We discuss the implications of the Opico experiment for future social communications platforms, and make a number of data-driven observations about potential ways to evolve the emoji standard.

2 RELATED WORK

A number of researchers have studied emoji communication on various social media and messaging platforms [5, 8, 10, 18, 24, 26, 30]. On these platforms, emoji mainly supplement text, providing affective context, reinforcing concepts found in the text, and/or replacing specific words [5]. Most *emoji sequences* present in this communication comprise repeating emoji, used to emphasize a singular concept [10, 18].

Other studies describe the unexpected affordances of emoji. On Venmo, a peer-to-peer payments platform, people use emoji as shorthand to describe payments. Caraway et al. examined how this emergent emoji use case added a social layer to the app and encouraged users to compose creative sequences of emoji to engage friends in their social network [8]. Zhou et al. studied non-textual communication on WeChat in China, and observed that people use emoji and stickers to converse about sensitive topics and those that are tedious to express in text [30]. Finally, Vidal et al. found that a high percentage of food-related tweets contain emoji, suggesting that they can be an effective representation for describing culinary experiences [26]. This paper examines whether the social, affective, and efficiency affordances of emoji can be amplified in an emoji-first setting, in which users initiate communication with sequences of emoji without textual accompaniment.

While emoji-based communication has many desirable properties, it is more ambiguous than text-based expression [19]. The meaning of emoji is often context-dependent, and tied to the sender's intent [10, 14]. Textual context, however, does not always mitigate miscommunication [20]. Since the artwork for a particular emoji symbol may vary substantially between the major mobile vendors, people's interpretations of emoji may vary across platforms [25]. Emoji usage and interpretation varies across age, gender, geographical location, and social status [6, 9, 16, 29]. This paper is the first to study the interpretability and ambiguity of standalone emoji sequences.

Despite the ambiguity inherent in emoji-based communication, researchers have trained useful models that capture the meaning and sentiment of emoji, suggesting that widely-accepted interpretations of emoji do exist [7, 12, 15, 27, 28]. Many of these models are emoji-word embeddings trained on social media and messaging data, which predict emoji meaning based on textual context [7, 12, 28]. Other models are trained explicitly on crowdsourced annotations for emoji to support retrieval tasks (e.g., emoji keyboard search) [22]. This paper provides evidence that these types of models could be extended to emoji sequences in the future.

3 THE DESIGN OF OPICO

To test the affordances of emoji-first communication, we built Opico, a mobile social app released on the iOS App Store and Google Play. Opico is built to allow users to author and share emoji *reactions* to physical places. Each reaction comprises a sequence of emoji that range between one and five Unicode characters. The five-character maximum provides a Twitter-like constraint, encouraging authors to be succinct and creative, and making reactions easy to parse and consume.

Each reaction is tied to a specific physical location retrieved from the Google Places API [4]. Users can leave a reaction to describe the location itself, a particular experience they've had, or their sentiment about the place. Reactions may be as straightforward as simple one-emoji 👍 (similar to a Facebook "like"), or tell more complicated stories, for instance using 🚘🚦🛬🏃😡 to illustrate the experience of being late for a flight due to traffic. Accordingly, Opico serves as a review app, lifelogging platform, and a way to check-in with friends.

3.1 Login and Onboarding

When a user first downloads Opico, they are prompted to authenticate with Google or Facebook account credentials. If the user is new to the platform, the app leads them through a short onboarding sequence that explains Opico's general concept, walks them through the process of creating their first reaction, and helps them find other users to follow. After onboarding, users are taken to the feed, which shows a timeline of recent reactions.

3.2 Posting Reactions

3.2.1 Locations. To create a new reaction, a user must first select a location. Opico displays a map view of nearby businesses and points of interest, and allows users to search by address and by keywords.

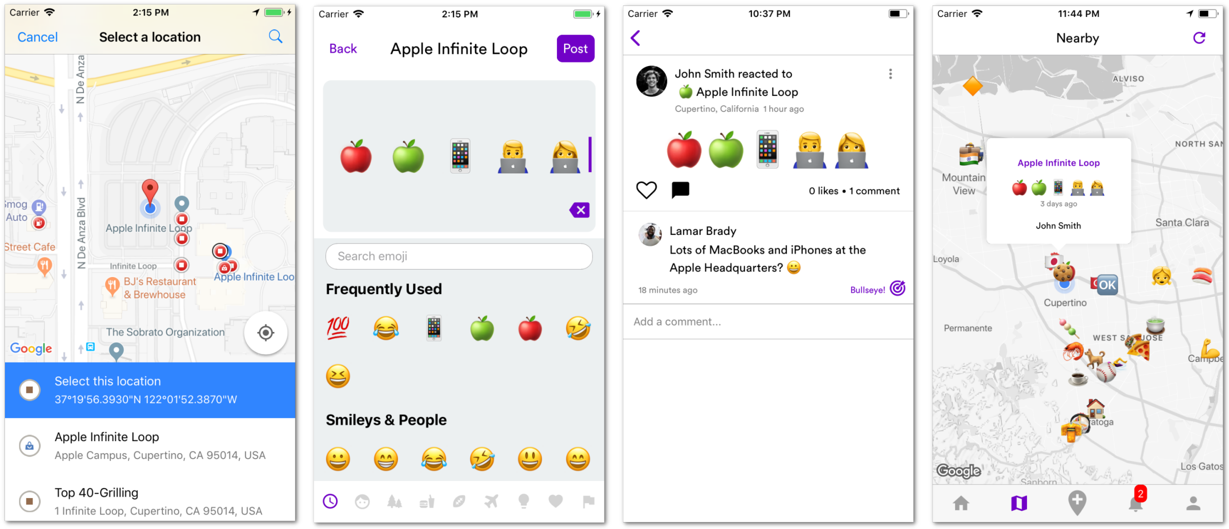

Figure 2: To create a new emoji reaction, a user picks a location and selects up to five emoji. Other users can comment on the reaction in words and earn a "bullseye" 🎯 if the author deems their description correct. In addition to a feed view, Opico provides a map view for viewing reactions.

3.2.2 Keyboard. After a location has been selected, the user is presented with the Opico emoji keyboard. The keyboard presents the entire set of Unicode emoji in a large, scrollable list, divided into a set of curated categories. The bottom row of the keyboard displays quick-jump icons for each category, and the top of the list shows a personalized subset of frequently-used emoji for easy recall.

The keyboard also exposes search functionality so that users can retrieve emoji via keyword. While the Unicode standard provides official names and aliases for each emoji, Opico exposes a broader set of keywords, bootstraped from synonym databases and augmented by logging unsuccessful queries. When a user runs an unsuccessful emoji search, the app logs the query string so that Opico's developers can consider which emoji should have been returned.

3.3 Social Affordances

3.3.1 The Feed. The feed is Opico's primary social view. It shows a list of reactions — sorted in reverse chronological order — from the set of users the logged-in user follows. To ensure a sufficient quantity of fresh content and encourage engagement and discovery, the feed also includes reactions that are "one-hop" away in the network, or from friends of friends. Users can search for new people to follow in the feed's accompanying tab.

3.3.2 Profile & Location. Opico exposes a profile page for each user — which displays a chronological list of their reactions — and a similar profile for each location in Opico's dataset. These views provide a way for a user to track their own travels and recent check-ins, as well as to see a summary of the reactions for a particular place. In addition, user profiles provide links to lists of followers and followees.

3.3.3 Likes, Comments, and Bullseyes. Opico provides a simple ❤️ "Like" mechanism for users to engage with one another's posts, as well as freeform comment functionality for more substantive discussions. Comments in Opico are text fields that are *not* restricted to emoji, and can be used to reply to the content of a reaction, ask questions, or — in the case of ambiguous or difficult-to-interpret posts — guess a reaction's meaning. To encourage users to engage in this latter way, Opico also provides a 🎯 "bullseye" indicator, which the author of a reaction may assign to a single "on-target" comment.

3.3.4 Emoji Maps. Since Opico reactions are location-based, the app also provides a map view of nearby reactions using a Google maps backend. Each location is represented by a single emoji in this view, chosen by picking the most frequently-used emoji character across all user reactions at the location. Tapping on the emoji reveals the most recent reaction for the location, and tapping on the callout takes the user to the reaction itself. The map makes it easy for users to discover locations that have been visited by their network; an alternate, personalized view shows only the user's own reactions.

3.3.5 External Sharing. Opico also allows users to crosspost reactions to their Facebook profile by toggling a button in the posting flow. Crossposts generate a URL with an OpenGraph image displaying the emoji reaction on a map with a small description underneath. This sharing flow allows users to advertise their reactions to friends who have not yet joined the Opico network.

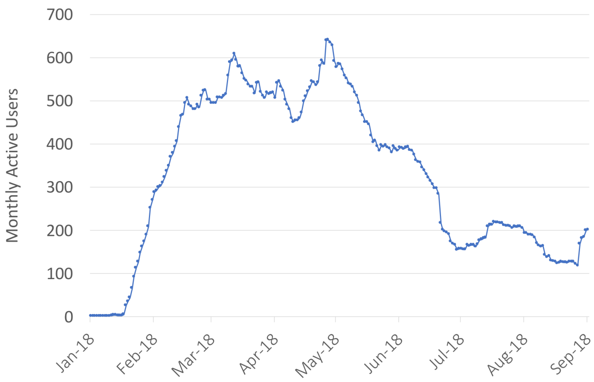

Figure 3: Monthly active Opico users from 1/1/18 to 9/1/18.

4 USERS AND ENGAGEMENT

Opico was released in January 2018 on the Apple App Store and Google Play, and marketed on the campus of the University of Illinois at Urbana-Champaign through posters and word-of-mouth. Since its release, Opico has gained more than 1000 registered users, and generated more than 3700 emoji reactions to more than 2400 physical places.

The platform's registered users are 71.5% male and 28.5% female; 32.98% use Android devices, 67.02% use iOS devices; and 90.73% are from the United States. Most of the users (51.13%) are between the ages of 18 and 24, 33.19% of them are between 25 and 34, 11.39% are between 35 and 44, and 4.3% are between 45 and 54.

Figure 3 shows monthly active users from launch to September 2018, with peak engagement — more than 500 users — between the months of March and May, and usage tapering off during summer break. As of September 2018, users have produced 3769 reactions, given out 6888 likes, left 802 comments, and awarded 126 bullseyes.

On average, each user follows 14.28 ($\sigma = 22.88$) users and is followed by 11.89 ($\sigma = 20.94$) users; 49.74% of the 5826 following relationships are reciprocal.

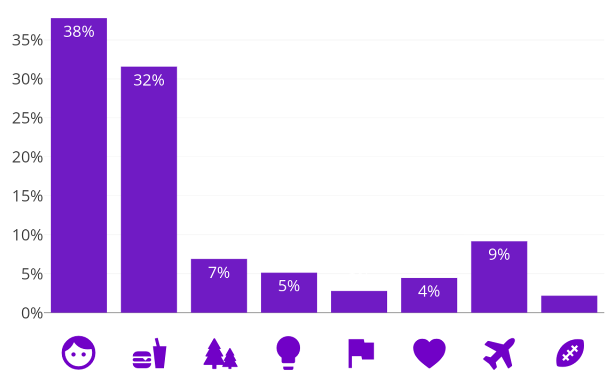

Figure 4: The breakdown of emoji usage by emoji keyboard categories.

Figure 5: Opico emoji reactions often describe activities such as working, traveling, or waiting.

5 USAGE AND REACTION PATTERNS

The 3769 reactions comprise 4.82 emoji on average ($\sigma = 2.17$), which are distributed in the following way across emoji keyboard categories: 38% of the emoji come from the "Smileys & People" category, 32% from "Food & Drink," 9% from "Travel & Places," 7% from "Animals & Nature," 5% from the "Objects," 4% from "Symbols," 3% from "Flags," and 2% from "Activities" (Figure 4).

A significant fraction of emoji used in reactions — "Smileys & People" — describe how people felt at or about a location. "Food and Drink" is the next most popular category of emoji, since most reactions were left for restaurants (58%). The distribution over emoji categories closely matches the distribution over place types: overall, 72% of the reactions were left for food-related places which include restaurants, bars, cafés, and bakeries; and 5% of the reactions were left for airports and lodging. Emoji from the "Animals & Nature" category are often used to describe food, but are also used to describe abstract concepts. For example, the snail 🐌 and turtle 🐢 emoji are commonly used to describe "slow" situations, and the mouse 🐭 is used to describe things that are "small."

Many reactions read like short reviews: since most reactions are associated with food-related establishments, they describe what users ate or drank, and how pleasing the experience was. Other reactions describe activities such as working, traveling, and waiting. Work-related reactions often involve emoji with laptops — 💻, 👩‍💻, 👨‍💻 — and coffee ☕. Travel-related emoji often describe where people are leaving from 🛫, where they are landing 🛬, and "going home soon" 🔜🏠. Finally, waiting-related reactions are often characterized by sequences of people — 👫👫👬 — to indicate lines, and sequences of hourglasses — ⌛⏳ — and clocks — 🕐🕑🕒 — to indicate the passage of time. Figure 5 illustrates many of the common patterns associated with these activities.

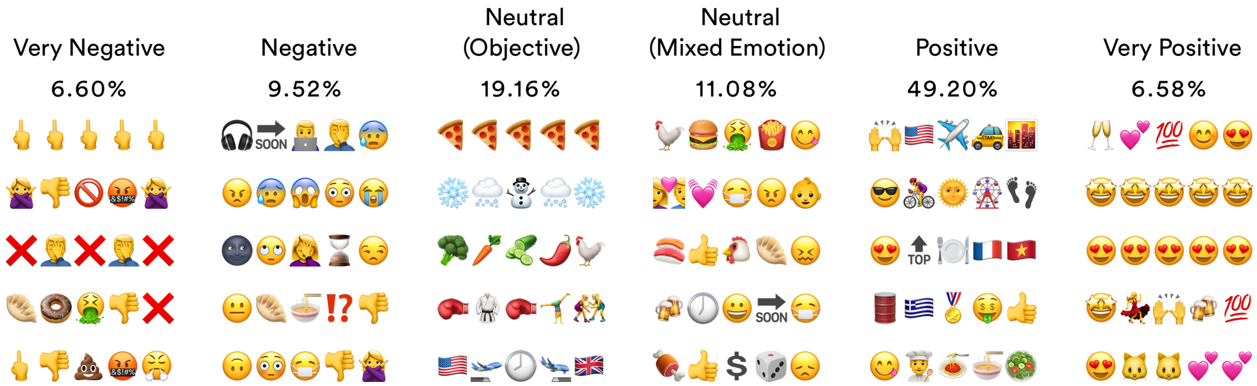

Figure 6: A selection of Opico reactions, categorized by sentiment based on the Emoji Sentiment Ranking [15]. Reactions can be neutral in two ways: they comprise only non-emotive emoji, or they comprise mixed emotions that cancel each other out.

5.1 Emoji Collocation Analysis

To identify patterns and activities in Opico's reaction data, we extracted frequently-occurring emoji bigrams and trigrams, and ranked them by their point-wise mutual information (PMI) [17]. Unlike in the English language, we found very little ordering consistency within consecutive sequences of emoji, so we formed bigrams and trigrams from unordered pairs and triplets, respectively. Since PMI is biased toward infrequent events, we consider only bigrams present in at least five reactions and trigrams present in at least three; additionally, both types of n-grams must have been used by at least two different Opico users to be counted.

Our collocation analysis yielded 248 emoji bigrams with PMIs greater than zero. Many of the bigrams and trigrams are pairs of the same emoji, which is a common pattern in normal emoji usage [18]. Emoji repetition such as 🌮🌮 ($n = 12$) may be used to communicate the plural form of the emoji, or an explicit count. Other times repetition is used for emphasis: 🌶️🌶️ ($n = 7$) and 😀😀 ($n = 8$), may indicate levels of spiciness and happiness, respectively.

Other bigrams pair similar types of emoji, such as those for beer 🍺🍻 ($n = 28$), working on a laptop 👨‍💻👩‍💻 ($n = 19$), and seafood 🦀🦐 ($n = 7$). Or, they bring together commonly-associated items and concepts, such as beer and pool 🍺🎱 ($n = 5$), burgers and fries 🍔🍟 ($n = 57$), coffee and work 💻☕ ($n = 20$), or taking off and landing 🛫🛬 ($n = 8$). Only a few bigrams pair an emotive emoji with a non-emotive one. Most of these bigrams express positive sentiments toward a food item such as pancakes 🍦😋 ($n = 8$), except for one: ⌛😡 ($n = 5$) signifying anger over a long wait.

Perhaps, some of the most interesting bigrams in the dataset are ones that qualify food. When food emoji are paired together, one emoji often describes the main dish, while the other describes the flavor or an ingredient found in that dish. For example, reactions such as 🍫🍰 ($n = 9$) and 🍦🍫 ($n = 7$) signify chocolate cake and ice cream, and ☕🍫 ($n = 11$) represents a 'mocha' or 'hot chocolate.' Similarly, combinations such as 🍔🐔 ($n = 8$) and 🍞🥑 ($n = 8$) can capture dishes such as chicken burger and avocado toast.

Many bigrams pair international flags with food items to describe food from certain countries or cultures: for instance, 🇲🇽🌮 ($n = 10$) and 🇮🇹🍝 ($n = 8$). Many of these bigrams are unsurprising since tacos and pasta are well-known Mexican and Italian dishes. However, some combinations are more unusual, such as 🇮🇳🥞 ($n = 8$). From comments that received bullseyes, we know that this bigram is used in the reaction dataset to communicate the savory Indian pancake-like dish dosa. Since there is no dosa emoji, users repurposed existing emoji to express this new idea.

Overall, we identified 118 trigrams with PMIs greater than zero. Trigrams exhibit patterns similar to bigrams. Many trigrams feature the same emoji repeated three times either to express plurality or emphasis. Reactions such as 😍😍😍 ($n = 21$) or 😠😠😠 ($n = 3$) describe an extra emphasis on an emotion, whereas reactions such as 👫👫👫 ($n = 3$), or 👪👪👪 ($n = 4$) describe "waiting in line" or crowds.

Similar to bigrams, trigrams also bring together similar types of emoji: alcoholic beverages 🍷🍹🥃 ($n = 3$), breakfast foods 🍞🍳🥓 ($n = 3$), and fitness activities 🏃‍♀️🏋️‍♀️💪 ($n = 3$). Trigrams also capture common food dishes, such as a chicken avocado salad 🐔🥑🥗 ($n = 4$), and spicy Indian curry 🇮🇳🌶️🍛 ($n = 3$).

6 SENTIMENT ANALYSIS

Since many emoji explicitly describe a feeling, a reaction's sentiment can be analyzed in a direct manner. We leverage the Emoji Sentiment Ranking [15] lexicon to compute a sentiment score for each reaction in our dataset. The ranking lexicon assigns a sentiment score to each emoji, where positive and negative values correspond to positive and negative sentiment, respectively. Accordingly, emoji with neutral sentiment are assigned a score of 0.

To compute an overall sentiment score for a reaction, we sum up the sentiment scores of its constituent emoji. Based on the sentiment score, we characterize each reaction as having either a strong positive, positive, neutral, negative, or strong negative sentiment. Reactions can be neutral in two ways: either they comprise only non-emotive emoji, or they comprise both positive and negative emoji that cancel each other out. Figure 6 shows representative reactions from each of the six categories.

The majority of Opico's reactions contain positive sentiments (49.02%). This result is unsurprising given the fact that two most frequently used emoji in the reaction dataset are 😍 or 👍, and suggests that users are often incentivized to react as a way to log and remember their favorite places.

Negative emoji reactions were less common (9.52%), and often described the reason someone was upset: the restaurant was too expensive, or a specific dish was bad, etc. Strong positive (6.58%) and negative (6.60%) reactions comprise repetitive positive and negative emoji, respectively. In these extreme instances, it appears users care more about emphasizing the intensity of their emotion than relating a precise story about why they felt that way.

Users create objective neutral reactions (19.16%) to describe a location without expressing an opinion; such users may use the app to track what they did more than how they felt. 11.05% of the reactions in the dataset capture both positive and negative emotions. For example, the reaction 🐓🍔🤮🍟😋 describes how a user did not like the chicken sandwich at a restaurant, but thought the fries were good.

Our approach for measuring sentiment is not foolproof. For example, some reactions expressed feelings of happiness by using emoji to describe certain idiomatic phrases in the English language such as "so happy I could die" or "feeling like I'm in heaven," with reactions such as 🍣🍜😍💀😇 and 😍🍩➡️⚰️😇. These are strong positive reactions, but our scoring approach categorizes both as neutral. Another failure case is irony. For example, the "weary face" emoji 😩 used in the context of describing food can mean "to die for." In the future, we hope that the Opico dataset can be used to build more nuanced, context-sensitive models of emoji sentiment.

7 EFFICIENCY AND AMBIGUITY

One of the primary motivations behind building Opico was to understand whether people could use an emoji-only communication medium to succinctly express complex ideas and stories. A reaction's text translation can help us measure both *efficiency* — how much information can a reaction encode — and *ambiguity* — whether multiple people interpret the message correctly and in the same way. We collected translations through the Opico's in-app comments and bullseye mechanism, and conducted a study with Opico users.

7.1 Bullseyes

Opico's "bullseye" feature allows us to organically collect text-based translations of emoji-only reactions, since bullseyes 🎯 are meant to be awarded to comments that correctly guessed the meaning of a reaction. As of September 2018, users awarded bullseyes to 126 comments. The average length of comments with bullseyes is 26.41 characters ($\sigma = 24.68$), or 4.98 words ($\sigma = 4.67$).

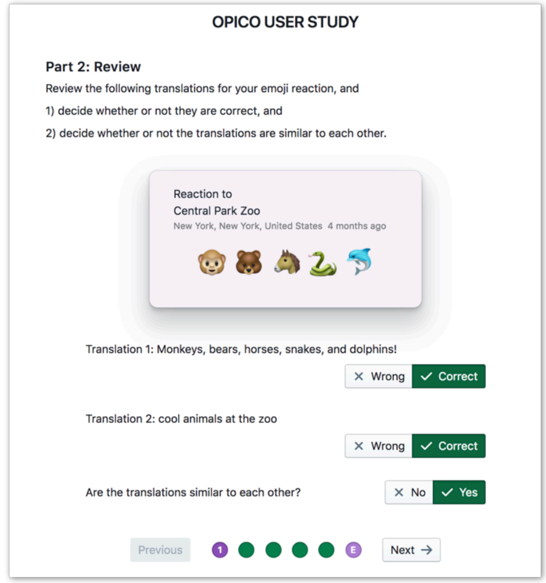

Figure 7: After translating a set of reactions, study participants were asked to review their own reactions and verify whether the translations produced for them were correct.

Creative reactions often elicit comments where other users try to guess the meaning. For example, users were able to guess that the reactions 🐍🎤👸🐍😁 and 🎤🎫🐍💋✨, which were left at stadiums, described Taylor Swift concerts. These reactions allude to the snake emoji's special meaning to Taylor Swift and her fan base. Similarly, users were able to guess that ⚫🐆👍👍👍 and 🐆💎🎮🍿😊, which were left at movie theaters, correspond to the movies "Blank Panther" and "Jumanji." These examples demonstrate how users leverage contextual clues based on location and timeframe to decipher the meaning of reactions.

Many comments that earned bullseyes did not describe the entire reaction, but just the part the author wanted verified. For example, the reaction 🍻📺🏀👏 elicited the comment 'Warriors!', which earned a bullseye despite failing to mention anything about "beer" or "watching the game on TV." Accordingly, we deem the in-app bullseye feature a useful — but incomplete — translation mechanism.

7.2 Study Methodology

To explicitly measure the efficiency and ambiguity of location-based emoji reactions, we recruited participants for a two-part study by emailing the 129 Opico users who had left five or more reactions. The email solicited participation in a user study to better understand how people interpret emoji-based reactions in exchange for emoji swag.

The first part of the study asked users to translate ten emoji reactions written by others into text. We anonymized each reaction, and provided its location and the time it was posted. If a reaction

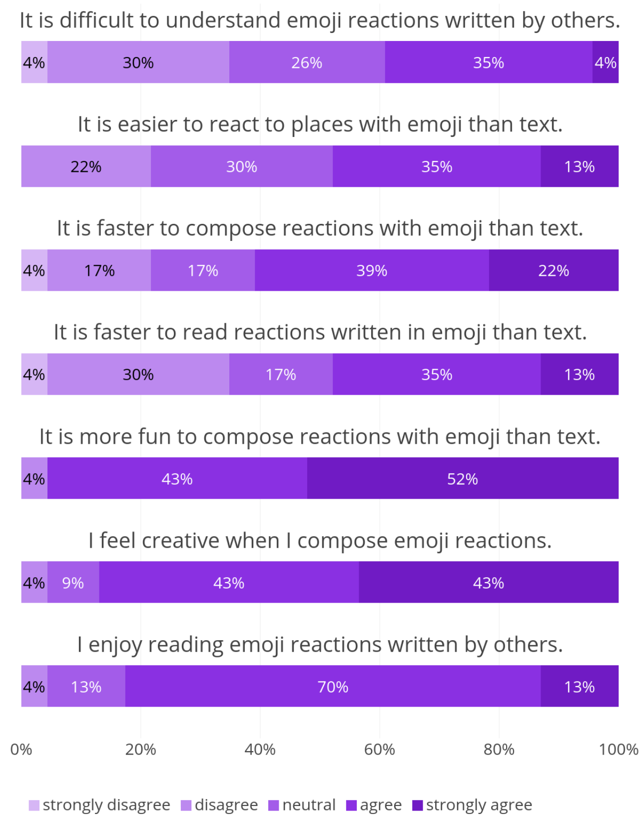

Figure 8: 23 participants were asked to assess statements about emoji reactions on a five-point scale.

was inscrutable, users could select an "I don't know" response instead of composing a translation. To create a representative set of reactions to translate, we randomly sampled five reactions from each participant, and collected two independent translations for each reaction.

After participants completed the first part of the study, we asked them to review their own reactions, and verify whether or not the translations produced by the other participants were correct (Figure 7). In addition, for each reaction with two translations, we asked users to evaluate whether or not the translations were similar to one another. After verifying the translations for their reactions, users completed a survey which asked them to reflect on their experiences composing and consuming location-based emoji reactions. The survey comprised seven questions, and elicited responses on a 5-point Likert-scale, ranging from "strongly disagree" to "strongly agree."

7.3 Study Results

Through our email campaign, we recruited 23 participants. By sampling five reactions from each participants, we created a pool of 115 reactions for participants to translate.

The study's first part produced 219 text translations, and 11 "I don't know" responses. Participants generated two translations for 105 reactions, one translation for nine reactions, and zero translations for one reaction.

In the second part of the study, participants determined that 144 text translations (65.75%) correctly described their own reactions. Out of the 105 reactions with two translations, both translations were correct for 49 reactions (46.67%), and 93 (88.57%) had at least one correct translation.

Additionally, 62 reactions (59.05%) had two similar translations: 45 (72.58%) where both were correct, 6 (9.68%) where both were wrong, and 11 (17.74%) where one translation was right and the other wrong. Some participants indicated that it was possible for reactions to be similar in content and not correctness: for the reaction "🕐🕔🕗😴😩" left at the O'Hare International Airport, the translation "Lots of time waiting at Ohare - sad and sleepy" was marked correct, while the translation "slept in and almost missed flight" was marked wrong.

The average length of the reactions used in the study was 4.10 emoji ($\sigma = 0.99$) and the average length of the resultant text translations was 6.27 words ($\sigma = 3.98$) or 35.15 characters ($\sigma = 21.24$). These results raise interesting questions about whether emoji are more "efficient" than text for conveying information, and in what circumstances.

Figure 8 presents the results of the user survey, which suggests that not all users felt comfortable reading and writing with emoji: 39% of the participants agreed or strongly agreed that "it is difficult to understand emoji reactions," while 22% of the participants disagreed that "it is easier to react to places with emoji than text." Similarly, 21% of the participants disagreed or strongly disagreed that "it is faster to compose reactions with emoji than text," and 34% that "it is faster to read reactions written in emoji than text." Although some users found it difficult to compose and understand emoji reactions, the majority of participants reported that they felt creative (86%) and had fun (95%) while writing them, and that they enjoyed reading them (83%).

8 DISCUSSION AND FUTURE WORK

The results of our study indicate that most location-based emoji reactions are interpretable (since 89% of the reactions received at least one correct translation) and that these interpretations are fairly consistent (59% of reactions were interpreted the same way).

One interesting avenue for future work is studying how ties within one's social network affect one's ability to understand reactions. Users leverage contextual clues based on location and timeframe to decipher the meaning of reactions; to what extent do they also use external knowledge based on personal relationships? Since emoji are often used in personalized ways [29], we wonder how effective emoji-first communication is as a vehicle for broadcasting messages whose true meaning is only understood by those "in the know."

The study also reveals a wide range of comfort with the emoji medium: around half of participants indicated that it is faster to read and write in emoji than in text! As emoji usage increases, we hypothesize that this percentage will also grow.

Below, we sketch a few promising directions for improving emoji-based communication based on Opico: developing better emoji input modalities; informing the emoji specification; and building new, useful emoji interactions.

8.1 Improving Emoji Input Modalities

While Opico's emoji keyboard provides category-based browsing and search, it can still be difficult for someone unfamiliar with the emoji lexicon to select the right emoji. Data collected through informal user interviews suggests that some Opico users did not finish posting reactions because they experienced choice paralysis or were afraid of appearing foolish if they used the "wrong" emoji.

New emoji keyboards can better scaffold emoji novices by suggesting emoji *autocompletions*, or offering templates for longer sequences of emoji based on context [22]. Many messaging apps already suggest an emoji replacement based on the last typed word: this idea could be extended to include longer sequences. On Opico, the keyboard could suggest entire reaction templates based on the type of location, learned from data.

8.2 Informing the Emoji Specification

Although the Unicode lexicon now includes more than 2800 emoji, it still comprises far fewer concepts than the English language. Our informal interviews revealed that some users felt frustrated when they could not find an exact emoji for the concept they wished to convey, while others found this constraint made them more creative. Observing how people creatively repurpose existing emoji to represent new concepts can inform future emoji design.

While most emoji are represented with a single Unicode code point, others are combinations of two or more code points. For example, the 👩‍❤️‍👩 emoji is a combination of the 👩, ❤️, and 👩 code points. The character that joins these code points together is called a Zero Width Joiner, or ZWJ, and, using ZWJ characters with emoji creates Emoji ZWJ Sequences [1]. One benefit to ZWJ sequences is that vendors can introduce their own artwork for any combinations of emoji without first requiring the emoji to be formally ratified by the Unicode consortium.

Opico's dataset — especially the results of our collocation analysis — can help inform future ZWJ sequences that vendors should supply. For example, two frequently used emoji, such as 🥑 and 🍞, can be combined into a single artwork for "avocado toast." The following reactions 🍵🥛⚫, 🌹🍵➕⚫, 🌺🍵⚫😍, ☕🍵⚫⬛, all describe "boba" or "bubble tea," which is a popular drink in the dataset. These reactions suggest that vendors could implement a ZWJ for bubble tea by combining the code points 🍵, 🥛, and ⚫. In the future, apps could also automatically detect semantically meaningful sequences, and leverage visual blending techniques to dynamically generate emoji that express new concepts [11].

8.3 Realizing the Social Affordances

Although the current implementation of Opico allows us to study how users compose and read sequences of emoji, it does not take full advantage of its own representation format. With emoji representations, traditionally complex NLP operations such as summarization and sentiment analysis are greatly simplified. Opico can leverage the emoji representation to easily produce aggregations and visualizations that can reveal interesting patterns about users and places.

For users who utilize the app for life-logging, Opico could provide personalized visualizations that reveal how they spend their time, what kind of food they eat, and so on. Emoji-based search could allow users to quickly run complex queries: a user could find "coffee places to work at" by querying with 💻☕. For users who are looking for a place to eat, Opico could augment its current map visualization to show how nearby restaurants are trending based on sentiment analysis.

Opico reactions could also be aggregated and used to power ranking and recommender systems. The data generated by Opico suggests that an emoji-based rating platform could help overcome the J-shaped distribution that often plagues five-star systems [13]. In contrast to reviews on highly-trafficked e-commerce sites (which are invariably bimodal), sentiments expressed in Opico — where users are more inclined to describe mediocre experiences — follow a more normal distribution. In addition to encouraging users to rate more frequently, emoji-based ratings could provide more descriptive and nuanced feedback for goods, services, and experiences.

REFERENCES

[1] 2016. Emoji ZWJ Sequences: Three Letters, Many Possibilities. https://blog.emojipedia.org/emoji-zwj-sequences-three-letters-many-possibilities/.
[2] 2017. 77% of 56-64 Year Olds Use Emojis On Messenger. https://blog.emojipedia.org/77-of-56-64-year-olds-use-emojis-on-messenger/.
[3] 2018. 157 New Emojis in the 2018 Emoji List. https://blog.emojipedia.org/157-new-emojis-in-the-final-2018-emoji-list/.
[4] 2019. Google Places API. https://developers.google.com/places/web-service/intro
[5] Wei Ai, Xuan Lu, Xuanzhe Liu, Ning Wang, Gang Huang, and Qiaozhu Mei. 2017. Untangling Emoji Popularity Through Semantic Embeddings. In *Proceedings of the 11th International Conference on Web and Social Media*. AAAI Press.
[6] Francesco Barbieri, German Kruszewski, Francesco Ronzano, and Horacio Saggion. 2016. How Cosmopolitan Are Emojis?: Exploring Emojis Usage and Meaning over Different Languages with Distributional Semantics. In *Proceedings of the ACM Conference on Multimedia*. 531–535.
[7] Francesco Barbieri, Francesco Ronzano, and Horacio Saggion. 2016. What does this Emoji Mean? A Vector Space Skip-Gram Model for Twitter Emojis. In *Proceedings of the Language Resources and Evaluation Conference*.
[8] Monica Caraway, Daniel A. Epstein, and Sean A. Munson. 2017. Friends Don't Need Receipts: The Curious Case of Social Awareness Streams in the Mobile Payment App Venmo. *Proceedings of the ACM Conference on Computer-Supported Cooperative Work and Social Computing* (2017).
[9] Zhenpeng Chen, Xuan Lu, Wei Ai, Huoran Li, Qiaozhu Mei, and Xuanzhe Liu. 2018. Through a Gender Lens: Learning Usage Patterns of Emojis from Large-Scale Android Users. In *Proceedings of the 2018 World Wide Web Conference*.
[10] Henriette Cramer, Paloma de Juan, and Joel Tetreault. 2016. Sender-intended Functions of Emojis in US Messaging. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM.
[11] João Miguel Cunha, Pedro Martins, and Penousal Machado. 2018. How Shell and Horn make a Unicorn: Experimenting with Visual Blending in Emoji. In *Proceedings of the 9th International Conference on Computational Creativity*.
[12] Bjarke Felbo, Alan Mislove, Anders Søgaard, Iyad Rahwan, and Sune Lehmann. 2017. Using Millions of Emoji Occurrences to Learn Any-domain Representations for Detecting Sentiment, Emotion and Sarcasm. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.
[13] Nan Hu, Paul A. Pavlou, and Jie Zhang. 2009. Overcoming the J-Shaped Distribution of Product Reviews. *Commun. ACM* 52, 10 (2009).
[14] Tianran Hu, Han Guo, Hao Sun, Thuy-vy Thi Nguyen, and Jiebo Luo. 2017. Spice Up Your Chat: The Intentions and Sentiment Effects of Using Emojis. In *Proceedings of the 11th International AAAI Conference on Web and Social Media*.
[15] Petra Kralj Novak, Jasmina Smailović, Borut Sluban, and Igor Mozetič. 2015. Sentiment of Emojis. *PLOS ONE* 10, 12 (12 2015), 1–22.
[16] Xuan Lu, Wei Ai, Xuanzhe Liu, Qian Li, Ning Wang, Gang Huang, and Qiaozhu Mei. 2016. Learning from the Ubiquitous Language: An Empirical Analysis of Emoji Usage of Smartphone Users. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*.

[17] Christopher D. Manning and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*. MIT press.

[18] Gretchen McCulloch and Lauren Gawne. 2018. Emoji Grammar as Beat Gestures. In *Proceedings of the 1st International Workshop on Emoji Understanding and Applications in Social Media*.

[19] Hannah Miller, Jacob Thebault-Spieker, Shuo Chang, Isaac Johnson, Loren Terveen, and Brent Hecht. 2016. "Blissfully happy" or "Ready to fight": Varying Interpretations of Emoji. In *Proceedings of the 10th International Conference on Web and Social Media*.

[20] Hannah Jean Miller, Daniel Kluver, Jacob Thebault-Spieker, Loren G. Terveen, and Brent J. Hecht. 2017. Understanding Emoji Ambiguity in Context: The Role of Text in Emoji-Related Miscommunication. In *Proceedings of the 11th International Conference on Web and Social Media*.

[21] Arielle Pardes. 2019. The Complete History of Emoji. https://www.wired.com/story/guide-emoji/

[22] Henning Pohl, Christian Domin, and Michael Rohs. 2017. Beyond Just Text: Semantic Emoji Similarity Modeling to Support Expressive Communication. *ACM Transactions on Computer-Human Interaction* 24, 1, Article 6 (2017).

[23] Will Radford, Andrew Chisholm, Ben Hachey, and Bo Han. 2016. 📞👨⛵🐳👌; or "Call me Ishmael" - How do you Translate Emoji?. In *Proceedings of the Australasian Language Technology Association Workshop*.

[24] Channary Tauch and Eiman Kanjo. 2016. The Roles of Emojis in Mobile Phone Notifications. In *Adjunct Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM.

[25] Garreth W. Tigwell and David R. Flatla. 2016. Oh That's What You Meant!: Reducing Emoji Misunderstanding. In *Adjunct Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM.

[26] Leticia Vidal, Gastón Ares, and Sara R. Jaeger. 2016. Use of Emoticon and Emoji in Tweets for Food-related Emotional Expression. *Food Quality and Preference* 49 (2016), 119 – 128.

[27] Sanjaya Wijeratne, Lakshika Balasuriya, Amit Sheth, and Derek Doran. 2016. EmojiNet: Building a Machine Readable Sense Inventory for Emoji. In *Social Informatics*. Springer International.

[28] Sanjaya Wijeratne, Lakshika Balasuriya, Amit P. Sheth, and Derek Doran. 2017. A Semantics-Based Measure of Emoji Similarity. In *Proceedings of the International Conference on Web Intelligence*.

[29] Sarah Wiseman and Sandy J. J. Gould. 2018. Repurposing Emoji for Personalised Communication: Why 🍕 Means "I Love You". In *Proceedings of the ACM Conference on Human Factors in Computing Systems*.

[30] Rui Zhou, Jasmine Hentschel, and Neha Kumar. 2017. Goodbye Text, Hello Emoji: Mobile Communication on WeChat in China. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*.