# ReactionBot: Exploring the Effects of Expression-Triggered Emoji in Text Messages

MIKI LIU, University of Washington, USA
AUSTIN WONG, University of Washington, USA
RUHI PUDIPEDDI, University of Washington, USA
BETTY HOU, University of Washington, USA
DAVID WANG, University of Washington, USA
GARY HSIEH, University of Washington, USA

In this paper we present ReactionBot, a system that attaches emoji based on users' facial expressions to text messages on Slack. Through a study of 16 dyads, we found that ReactionBot was able to help communicate participants' affect, reducing the need for participants to self-react with emoji during conversations. However, contrary to our hypothesis, ReactionBot reduced social presence (behavioral interdependence) between dyads. Post study interviews suggest that the emotion feedback through ReactionBot indeed provided valuable nonverbal cues: offered more genuine feedback, and participants were more aware of their own emotions. However, this can come at the cost of increasing anxiety from concerns about negative emotion leakage. Further, the more active role of the system in facilitating the conversation can also result in unwanted distractions and may have attributed to the reduced sense of behavioral interdependence. We discuss implications for utilizing this type of cues in text-based communication.

CCS Concepts: • **Human-centered Computing** → Computer supported cooperative work

**KEYWORDS**

Eomji; Computer-Mediated Communication; Affect

**ACM Reference format:**

Miki Liu, Austin Wong, Ruhi Pudipeddi, Betty Hou, David Wang, Gary Hsieh 2018. ReactionBot: Exploring the Effects of Expression-Triggered Emoji in Text Messages. In *Proceedings of the ACM on Human-Computer Interaction,* Vol. 2, CSCW, Article 110 (Nov. 2018). ACM, New York, NY. 16 pages. https://doi.org/10.1145/3274379



## 1 INTRODUCTION

According to a 2013 survey by Swyft Media, 74% of people in the U.S. regularly use emoticons, stickers or emoji in their online communication [37]. Instagram also found that 50% of all captions and comments on Instagram have an emoji or two, with faces accounting for six of the top ten emoji used [30]. Since the first documented use of emoticons in 1982 by Scott Fahlman [12], these pictorial representations of facial expressions have become an integral part of everyday text-based communications.


2573-0142/2018/11 - ART110 $15.00

https://doi.org/10.1145/3274379

Much prior research has explored the role of emoticon and emoji in communication. Research in computer-mediated communication (CMC) has found that these graphical representations can act as substitutes for the nonverbal cues that are present in face-to-face communication, but often not present in the computer mediated contexts [46]. Their use can help communicate facial affect, enable better understanding between communicators, facilitate relationship development, and even influence recipient's emotions through emotional contagion [7,26,48].

However, as explained through the Social Information Processing (SIP) Theory, while communicators are able to use emojis (and other cues available through the CMC) to convey social information, "the rate of exchange is slower online, not only because both instrumental and relational information must be conveyed in a limited bandwidth, but because typing and reading are slower than speaking, looking, and listening" [47]. And that "time becomes a critical predictive variable" in the amount of social information conveyed through CMC. Through their experiments, Walther and colleagues have shown that while those using text-based communication can arrive at a fully developed impressions of their chat partners, those in FtF groups arrive at that point sooner [42].

Thus, given the limitations associated with using emoji to communicate valuable social cues during text-based communication, what if instead of relying on communicators themselves to insert emoji into conversations, we use the communicators' facial expressions to trigger the emoji? For example, suppose you received a funny message. Instead of having to enter in a smiley face to indicate your reaction to the message, the system monitors your facial expressions and upon detecting that you are smiling, it automatically reacts to the message with a 😀 (:grinning:). How might this type of expression-triggered emoji differ from the intentional emoji users would currently express using current systems? Would this reduce the communicators' need to use these affect cues? In what ways can this improve or detract from communication?

To explore these issues, we designed and built ReactionBot. ReactionBot is implemented as a chatbot in Slack, a team messaging app. ReactionBot uses the users' computer webcam to monitor users' facial expressions and uses the emotion detected to insert emoji into the text-based communication. To examine ReactionBot's role and impact on communication, we conducted a controlled experiment with 32 participants, where 14 participants (7 dyads) communicated using Slack with ReactionBot enabled, while 18 others (9 dyads) communicated with ReactionBot disabled (control). We found that while ReactionBot supplemented more accurate emotional expressions in conversations that could facilitate user interaction and engagement, it also increased anxiety and led to undesired distraction.

Results from this study offer both practical and theoretical contributions. We present the first implementation and evaluation of an expression-triggered emoji system for text-based communication – ReactionBot. Our study offers insights both on how to design similar systems to support the communication of affect as well as contexts when they may be most helpful. Concurrently, our work also improves our understanding of facial-expression communication through computer-mediated communication. The study of ReactionBot allows us to examine the role of intentionality in communicating nonverbal cues.

## 2 RELATED WORK

Nonverbal communication, such as body language, appearance, and voice, play an important role in traditional Face-to-Face (FtF) communication. Among their many benefits, they support interpretation, increase understanding, and help convey empathy (e.g., [9]). These cues can be even more critical than language content for "emotive, relational, or attributional outcomes" [4].

Facial expressions, specifically, can have stronger effects than vocal or spatial cues in FtF communication (p. 142).

Researchers have extended this line of research on nonverbal cues in communication to computer mediated communication (CMC). Initial research focused on the cue-filtered out theories, such as media richness theory and social presence theory [47]. These theories argue that the lack of nonverbal cues in many CMC channels such as chat and email is detrimental to relationship development. The less rich media prevents valuable social cues and gestures from being communicated, reducing the equivocality of a message, making them less efficient [6]. More recent theories, such as the social information processing theory (SIP), argue that it is not that the intimacy seen in face-to-face relationships cannot be established through CMC, is just that it takes longer [44]. Communicators are motivated to develop interpersonal impressions, and they adapt the media to communicate nonverbal cues to their communicating partners.

An example of such adaptive use is the use of emoticons or emoji in text-based communication to communicate affect. Emoticons are pictorial icons, using punctuation marks, letters, and numbers to display emotion or sentiment (e.g., :-)). Emoji, emoticons' successor, are pictographs of faces, objects, and symbols (e.g., 😀). Both emoticons and emoji enable users to convey a wide range of facial expressions, such as joy, anger, sadness, fear, disgust, and surprise. The introduction of emoji keyboard on iOS and Android platforms, along with the Unicode 6.0 encoding specification which included hundreds of emoji characters as part of the specification, has led to the wide adoption of emoji.

Research studying the use emoticons and emoji have examined their uses and potential benefits. Earlier studies of virtual-worlds (i.e., MUDs and MOOs) showed that its users are able to use emoticons to overcome the lack of nonverbal cues [33,40]. People increase their usage of emoticons overtime, and the emoticon usage predicted the development of online friendship [40]. While an earlier study of the effect of emoticons on message interpretation suggests emoticons' contribution is outweighed by the message content [46], more recent studies suggest that emoticons do influence online message interpretation [7,15,25]. Emoticons can help strengthen the intensity of a message [7]. Studies of emoticon use in the Instant Messaging (IM) context showed that emoticon use predicted enjoyment, which increases personal interaction, and then information richness [19]. In his chapter on Computer-Mediated Communication, Whittaker also note that communication media that afford the sharing of facial expressions can support attention, understanding and agreement [48]. Researchers have also found through field experiments that emoticon usage can influence receivers' emotions through emotional contagion [26]. More recent studies of emoji usage have also supported these prior findings, that people rely on these non-textual elements to convey nonverbal cues and helps them provide "emotion indicators" [24,50].

Despite their use to serve similar nonverbal communication functions, one important difference between emoticons/emoji and facial expressions is intentionality. As noted by Walther and D'Addario in their work on emoticons, while "one may unconsciously smile FtF, … it is hard to imagine someone typing a :-) with less awareness than of the words he or she is selecting" [46]. In other words, instead of a more natural interaction where one's facial expressions are being shown directly and immediately to the conversation partner, communicators of CMC have to deliberately insert in emoji in the conversations. This limits the efficiency of the channel in communicating valuable social cues. Researchers have indeed found that "CMC operates at a rate different from face-to-face communication in terms of users' ability to achieve levels of impression and relational definition equivalent to face-to-face interaction"[45]. And that more time is needed for relationship development via CMC than FtF.

| Emotion | Emoji | Emoji Name |
|---|---|---|
| Anger | 😠 | :angry: |
| Fear | 😧 | :anguished: |
| Disgust | 😖 | :confounded: |
| Happiness | 😀 | :grinning: |
| Surprise | 😮 | :open_mouth: |
| Sadness | 🙁 | :slightly_frowning_face: |
| Contempt | 😒 | :unamused: |

Table 1. Seven basic emotions detected using Microsoft's Emotion API with their corresponding emoji used by ReactionBot.

Another implication of this is the potential truthfulness of these nonverbal cues. As discussed by Marvin [28], "a participant might frown at the keyboard and but [sic] strategically decide to type a smile." Adopting the concept of signal theory in this context [8], one could argue that while emoticons or emoji can be conventional signals for people's affect, they are self-descriptions that can be easily faked. Whereas while facial expressions are one of the most controllable of nonverbal cues [9], people are not always aware of their facial expressions and are often direct and involuntary representations of internal states [23]. Therefore, facial expressions are more truthful signals of people's underlying affect.

Text-based communication has long been considered a less information rich media compared to FtF, but advances in technology are now enabling us to augment leaner media with social cues. Research has begun to explore this design space in CMC. BodyChat [41], examined the inclusion of nonverbal cues from communicators' body language when communicating via avatars. AffectIM communicates affect through avatars though it uses natural language analyses to determine affect [31]. More recent research has also explored the detection of facial expressions through images and communicated the detected emotions in texts to communicators, e.g., "Jane Smiles" [13]. In our work, we seek to explore how to design a system that uses facial-expression triggered emoji in chat, and to study the potential benefits of this type of enriched text-based communication channel. Would it reduce the need for communicators to provide social cues signals themselves? How would people perceive and use these expression-triggered cues?

## 3 REACTIONBOT

Nonverbal communication, such as body language, appearance, and voice, play an important role in traditional Face-to-Face (FtF) communication. Among their many benefits, they support interpretation, increase understanding, and help convey empathy (e.g., [9]). These cues can be even more critical than language content for "emotive, relational, or attributional outcomes" [4]. Facial expressions, specifically, can have stronger effects than vocal or spatial cues in FtF communication (p. 142).

To explore the idea of expression-triggered emoji, we designed and built ReactionBot. ReactionBot is a chatbot built on the Slack platform. It identifies facial expression via the user's webcam and automatically attaches emoji to user messages based on facial expressions. We chose

Slack as the platform for the bot as it is one of the increasingly popular instant messaging systems, with more than nine million weekly active users [35]. It also has a number of emoji features, such as allowing users to react to messages via emoji.

There are two key components to ReactionBot. One is a piece of code that captures users' facial expressions and then assess users' emotions. Our system runs on OSX; this code is implemented in Swift. When an image is captured, it is then sent to Microsoft's Emotion API, which assesses 7 basic emotions of the detected faces: anger, fear, disgust, happiness, surprise, sadness, and contempt [11]. The call to Emotion API returns confidence scores for each of the emotions, ranging from 0 to 1. We should also note that with our context of use, where the user is typing on their laptops, the captured images are generally full-face views in front of the camera without any occlusion.

The other piece of code is the Slack bot coded in Node.js, using Slack's API. This bot monitors users' conversations in the pre-specified slack channels. When an appropriate event has been triggered (i.e., a new message arrived or user start typing a new message), it calls the Swift code to capture users' images. When the confidence scores are returned from the Emotion API calls, the slack bot then decides which emoji to attach, if any, to which messages. The corresponding emoji to the seven basic emotions are presented in Table 1. The Node and Swift codes interact via a websocket.

Below, we will describe the two scenarios (receiving and sending messages) for which we have implemented ReactionBot and discuss some of our design rationale.

## 3.1 Scenarios of Use

*3.1.1 Receiving Scenario.* In the receiving scenario, the user has just received a message from another user. ReactionBot uses the webcam to capture four images immediately after, at 0, 1, 2, and 3 seconds. It then examines the returned confidence scores. We chose the threshold of confidence to be 0.5 based on our testing of the Emotion API. In other words, any of the seven basic emotions have a confidence of 0.5 or higher, the corresponding emoji will be attached to the received message as "reactions." This helps the receiver to communicate their emotional response. If none of the emotions are detected, no emoji are attached.

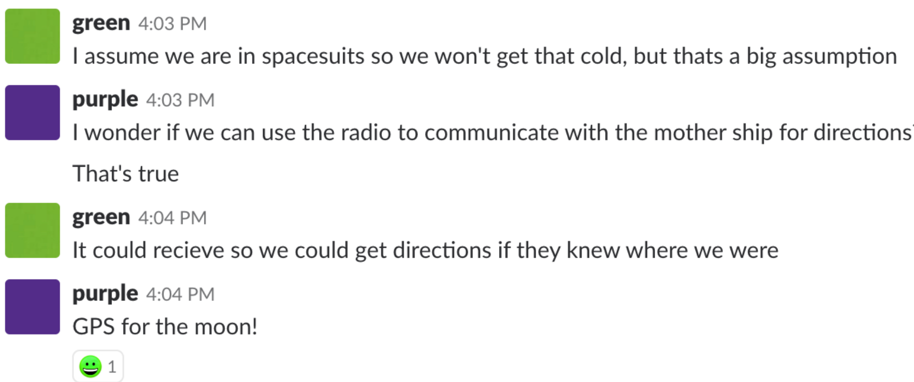


Figure 1. ReactionBot's receiving scenario. Here, green's ReactionBot reacted on their behalf to the received message "GPS for the moon!"

In the example depicted in Figure 1, after the "purple" user had sent a message, the "green" user (in this case the recipient of the message) had smile. ReactionBot observed the smile, and thus attached a ":grinning:" emoji to "purple's" message.

*3.1.2 Sending Scenario.* The second scenario is for outgoing messages. The idea here is to also enable users to communicate their affect for outgoing messages. Whenever the user starts typing in the Slack channel, ReactionBot will start capturing the users' images. It will capture five user images over five seconds, at 1 second intervals, starting when the typing event is triggered through the Slack API. Then, when the user actually sends the message, ReactionBot will both stop the call for webcam images if it was before the 5 seconds are up, and it will attach any emotion detected during those 5 seconds to the outgoing message.

In the scenario depicted in Figure 2, after the "purple" user had smiled when they were composing the message. The system then attached a ":grinning:" emoji to the message when it is sent.

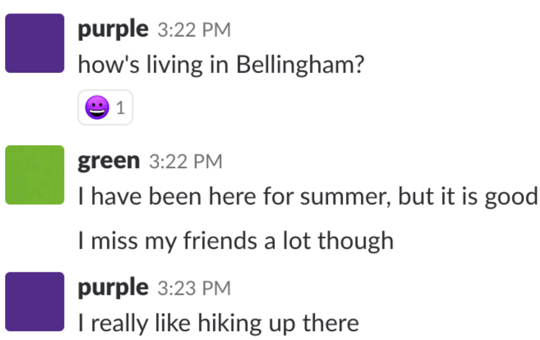


Figure 2. ReactionBot's sending scenario: Here, purple's ReactionBot reacted on their behalf to the sent message "how's living in Bellingham?"

## 4 HYPOTHESES

With the ReactionBot system, we pose two hypotheses about the use of expression-triggered emoji during conversations. First, we hypothesize that ReactionBot offers a richer communication channel than the existing chat systems. The use of facial-expression triggered emoji enables the communication of social cues that are currently not directly visible in chat, and hence reduces the need for users to explicitly communicate their affect. Instead of the users sending or reacting using emoji themselves, the system would facilitate that. This would therefore lead to a decrease in emoji usage by participants when communicating using ReactionBot.

**H1. Participants using ReactionBot will use fewer emoji.**

Another potential outcome from a medium that provides more social cues is that communicating through it could result in more attention paid to the participants in the communication [43]. This would affect the sense of social presence between the communicators, where social presence is defined as "the degree to which a person is perceived as a 'real person' in mediated communication." [17] Social presence is important in supporting communication effectiveness [21], facilitating task-performance and social relationships [51]. Prior work has indeed found that richer communication channels facilitate a stronger social presence. Biocca et al., for example, conducted a study showing that dyads who worked on a decision-making task, a desert survival task, reported higher social presence when interacting FtF than when using a teleconferencing system [2]. Yoo and Alavi have also found that triads who worked on a decision-making task reported higher social presence when interacting via video conferencing than audio conferencing [49]. Thus, we also hypothesize that:

**H2. Participants using ReactionBot will report higher social presence with their partners**.

## 5 EXPERIMENT

To explore the use of ReactionBot, we conducted a controlled lab study. During the between subjects study, pairs of participants interacted over Slack, running on two Macbook Airs. Half of the participants chatted on Slack with ReactionBot (ReactionBot), whereas the other half chatted on Slack without the ReactionBot (control). In both conditions, participants were told that they were to work on a decision-making task with a partner (task described below). They were given no explicit instructions on how they should use Slack (and emoji). This means they were not restricted to using the seven emoji offered through the ReactionBot system. Participants in both conditions could type any emoji in-line as well as attach any (reaction) emoji to messages. Participants in the ReactionBot condition were additionally told about the ReactionBot that will monitor their facial expressions and attach reactions on their behalf. But they were given no explicit instructions on how they should use the system.

During the study, participants were prompted to spend five minutes to chat and introduce themselves. They were then given 25 minutes to complete a collaborative ranking task but were allowed to finish before the given time limit if they completed it early. After the task, participants completed a survey on their task experience, and were asked additional questions from the facilitator regarding the emoji usage and their overall experiences. The entire study lasted for an hour. Participants saw each other during the study introduction but did not see each other again until the end of the study. Dyads both sat in the same room with their backs to each other. Participants wore ear plugs with noise-cancelling headphones so they would not hear each others' reactions auditorily.

The collaborative task they worked on was the Moon Landing task [18], a version of the decision-making tasks often used CMC studies (like the desert survival task) [2,5,36]. The basic premise of the task was that they were asked to pretend to be space crew on the Moon, and due to mechanical difficulties, need to survive and travel 200 miles to reach the rendezvous point. They were able to salvage 15 items intact and undamaged from their crash landing, but can only take 5 out of the 15 items with them. They are asked to choose and rank the top 5 items from the 15 they would take. Participants were given 5 minutes to work on the problem by themselves, then they interacted through Slack to discuss their solutions. They then were asked to individually select and rank the top 5 items again. While there is an established optimal decision to this task (ranking given by NASA) so that we can assess task performance, the task is designed to be open-ended and ambiguous. In other words, it is a type of task that can benefit from richer communication media [6].

### 5.1 Participants

Participants were recruited through on campus flyers at a large U.S university on the West Coast, as well as posts on the local Craigslist. Participants were compensated with a $15 gift card for an hour of their time.

In total, 32 individuals participated in our study (14 participants in the experimental condition, and 18 in the control condition). 20 of the participants were female and 12 were male. Ages of participants ranged from 18 years to 50 years, with a primary concentration of 18 to 30 year olds due to the nature of our recruitment process at a university. Majority of the participants do not use Slack regularly (17 participants). However, they are all frequent users of messaging/chat apps (25 participants use these apps 20 times or more a week). The two conditions had equivalent groups in terms of general demographics and in terms of their use of Slack and messaging/chat apps.

### 5.2 Measures

A number of different measures were examined in our study. First, we collected and examined conversational data. Specifically, we compared the amount of emoji usage between conditions. For our analysis, we counted both the in-line emoji used as well as the reactions to the messages. We also included the two instances of emoticon use that did not get displayed as emoji (O_O, and ¯_(ツ)_/¯). This count provided a measure for the total number of social cues expressed through emoji.

To explore the effect of ReactionBot on social presence, we asked a number of post-study questions using a scale developed by Biocca, Harm, and Burgoon [1]. Specifically, social presence is broken down into numerous sub constructs: co-presence (isolation/inclusion: 2 items; mutual awareness: 6 items), psychological involvement (mutual attention: 8 items; empathy: 6 items; mutual understanding: 6 items), and behavioral engagement (behavioral interdependence: 6 items; mutual assistance: 4 items; dependent action: 2 items).

We also explore decision-making task efficiency and performance between conditions. For efficiency, we recorded the amount of time the dyads used to arrive at a decision. For performance, Moon landing task performance was analyzed for each individual's original response, as well as the pairs' response after discussing. For each response, the score was calculated by taking each of the top six rankings from the participant(s), finding the positive difference between it and the ranking given by NASA for that item, and calculating the sum of the differences.

Finally, each pair of participants was interviewed together after completing the moon-landing task. The interview was recorded and included questions about the system, closeness with partner, and perception of program. For analyses, three of the authors first worked together to identify thematic codes. They then individually coded the transcripts. Finally, they convened again to discuss the insights gained.

### 5.3 Results

The average amount of time participants used in the decision-making task was 15 minutes and 8 seconds (minimum around 7 minutes and maximum just slightly over 25 minutes). Those using ReactionBot spent less time to reach a decision (M=799s, SD=313s) than those in the control (M=994s, SD=372s); however, the difference was not statistically significant (t(14)=-1.11, p=0.30).

In terms of task performance, there was a statistical difference in pre-discussion decision scores between the control and ReactionBot conditions, where participants in the Control condition made better decisions as individuals before the discussion (Mcontrol = 14.9, MReactionBot = 20.4, F(1,14)=5.75, p=0.03). Recall that higher score means worse performance (a larger deviation from the expert ratings). To control for this baseline difference, we conducted a regression analyses using the difference between the average pre- and post-discussion scores as the dependent variable, controlling for the pre-scores. The difference was not statistically significant between the two conditions (LSMcontrol = 8.59, LSMReactionBot = 7.38, p=0.72).

To test our first hypothesis about emoji use, we used Poisson regressions due to the emoji usage being a non-normal count data. We found that participants used fewer emoji during the Moon Landing Task in the ReactionBot condition compared to the Control condition (Mcontrol = 5.44, MReactionBot = 1.43), while there is no significant difference between conditions for the total number of emoji used (combining both system-triggered and self-triggered; Mcontrol = 5.44, MReactionBot = 5.86). This suggest that while the overall emotion being communicated during the task is the same in both conditions, the ReactionBot has replaced some of the affect sharing that the participants previously had to do by themselves (H1 supported, Figure 3). We should also note that there was no significant difference in emoticon use between conditions during the conversational part of the study (prior to the Moon Landing Task; Mcontrol = 1.56, MReactionBot = 1.29). This suggests that (1) participants may not have felt inhibited to send emoji by themselves initially, and (2) participants may need a little bit of time (5 minutes in our study) to adapt their emoji use to the system.

To test our second hypothesis, we compared the averaged social presence score between conditions. We used a mixed-effects regression with the 8 sub-constructs of social presence score as the DVs, the condition as the IV, and session id modeled as a random effect. We found that while all but one of the sub-constructs were rated higher by participants in the Control condition, only one of them was statistically significant – behavioral interdependence (LSMcontrol = 5.51, LSMReactionBot = 4.63, p=0.02). As we will explain later, and supported by our interview findings, the lowered sense of behavioral interdependence may be attributed to the salience of the system in mediating the communication.

### 5.4 Interview Findings

Through post-study interviews, we sought to explore the potential benefits and drawbacks of using ReactionBot in text-based communication. We group our findings into these two categories.

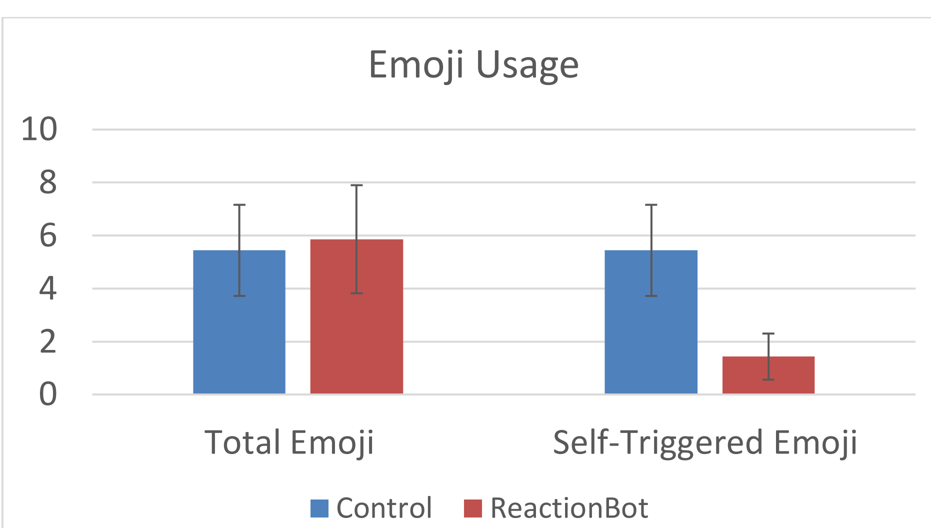


Figure 3. Emoticon usage during the experiment. Where total emoticons use includes both the self-triggered emoticons and the system-triggered emoticons.

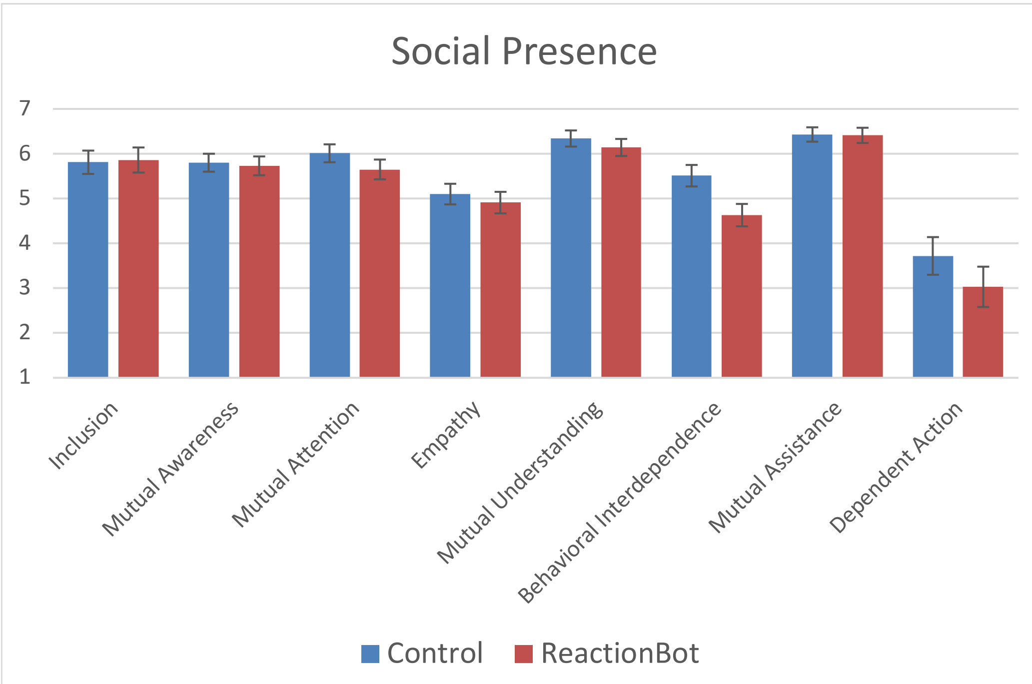


Figure 4. Ratings for the sub-constructs of social Presence between conditions.

*ReactionBot Benefits: Facilitated Valuable Affective Feedback and Increased Awareness of Own Emotions*

Post-study interviews suggested that the emotion feedback through ReactionBot provided valuable nonverbal cues, as the "immediate response" allowed participants to "know how your chat partner is feeling at that moment and how they are responding to your…conversation or your statements" (P32). When asked about the system-triggered emoji, one participant mentioned how it was "exciting" to see the emotional feedback because he knew that he and his partner "were on the same ground" (P5). Another said that, "because it kept on sending the happy [emoji], it just felt more comfortable" and that this was "validating" (P6). He noted that, by feeling more comfortable, "that probably helped with the part where we were actually discussing about the moon" (P6).

Further, many participants suggested that the use of ReactionBot provided a more accurate reading of their partner's emotions in comparison to manually typing an emoji. As stated by the participant, "I usually use emoji or type 'hahaha' but I'm not actually like laughing, but I find it funny I guess. But this is more accurate" (P4). Another stated that the ReactionBot was useful because "you can see how truly they think" but "if they send the emoji, does it actually mean they think that way?" (P3) This could be useful for detecting sarcasm, as "you can see the tone they are using" (P29) with the ReactionBot and be able to tell if the sender's emotion was not aligning with the emotion th ey were expressing through their text. One participant suggested that an application for ReactionBot could even be for doctors or therapists, in the event that one could not meet in person, because the emoji sent by the patient could provide valuable information for the doctor or therapist (P5).

Interesting, we also found that using the system increased participants' awareness of their own emotions. In one example, the system observed a surprised expression and reacted with the corresponding emoji "😮" to the message. Participant then realized they had that expression: "yeah and I was like woah. I didn't know how I did that either" (P30). The system was able to detect and send the surprise emoji without the participant being aware that they were expressing that emotion. Another participant mentioned how the system "made [her] more aware when [she] was smiling at a conversation" (P30). Thus, the system not only increased participant's awareness of their partner's emotions, but also of their own emotions.

*ReactionBot Drawbacks: Increased Anxiety from Potential (Negative) Emotional Leakage, and Distractions*

The inclusion of this facial expressions requires more self-presentation management. This is especially important for negative emotions than positive emotions. Multiple participants mentioned how they tried to convey more positive emotions or smile more, and that potential users could try to hide negative emotions and instead show an emotion that they do not actually feel: "If they're not happy, they might like try and smile more, bigger emotions, so that they're positive, because usually people are more attracted towards happy people. Or you could manipulate the conversation by acting like you're sad but you're not actually sad, so like people pity you" (P3). One participant said that this made it "a bit worrisome that now I need to worry about my facial expressions along with what I'm saying because before when you're talking online you don't really care what face you make" (P11). Another says: ""I feel like there is a certain amount of advantage you can get from not displaying your emotions in chat vs the ReactionBot it sounds like you wouldn't be able to hide how you actually felt" (P13).

Participants also expressed that the system created distraction from the task. There are multiple aspects to this point. The system-triggered emoji themselves can be distracting, especially in a more task-based context. One said that in a "business" context, "the emoticons were...irrelevant" (P12). Another thought the system was "fun" and wanted to get an emotion out of their partner: "it wasn't helping me get my work done quicker" (P29). Finally, the emoji can also be distracting when the system is wrong. "Sometimes I get distracted by the little emoji, because I didn't think at that moment I was smiling but it was showing it as smiling" (P6).

## 6 DISCUSSION

In this work, we designed and studied ReactionBot, a system to help convey facial-expressions through emoji in text-based communication. Based on computer-mediated communication (CMC) research and theories, we hypothesized that enabling this type of non-verbal cues would enrich the communication channel and enhance social presence. This is because the system-mediation would reduce the need for users to communicate this type of social cues themselves. Further, because using facial-expressions as the basis for the emoji is more intentional and more natural compared to the self-reacted emoji, these cues could be a more accurate communication of affect and be perceived as such.

Through our study, we found that while the overall number of emoji used during the task was the same with or without the ReactionBot system, when ReactionBot is in place, participants used fewer emoji themselves. This suggests that the system is performing some of the nonverbal cues sharing work that communicators had to do themselves. An interesting future research question is whether this type of expression-triggered cues can accelerate the formation of relationship. Facial expression is but one type of nonverbal cues; this follow-up study may help us better tease out the specific role of sharing facial-expression in facilitating relationship formation.

Through our interviews, we also found that this emotional feedback, though not perfectly accurate, is generally perceived as more genuine than self-triggered emoji. The facial-expressions triggered reactions help “validate” conversations and facilitate the interpretation of the messages as one can more easily understand the “tone” of communication. These important cues may help minimize miscommunication between partners, so that chat partners can gain a more accurate sense of how their messages are affecting their partners. Relatedly, recent research has highlighted the problem of emoji misinterpretation, where it is shown that people can and do interpret both sentiment and semantic meaning of emoji differently [29]. The use of expression-triggered emoji can offer an unequivocal mapping of emoji and underlying facial-expressions, which can help minimize the discrepancies in interpretation between senders and receivers.

A third type of benefit that we did not hypothesize, but also came up in the interviews, is that the ReactionBot enabled people to become more aware of their own facial expressions and affect. As we have found, and supported in prior work, people are not always aware of their emotions and moods [39]. The expression-triggered emoji can then provide feedback to users and enhance their own emotional awareness. This type of awareness is a critical part of emotional intelligence [34], which can have important applications in contexts such as mental health [27] and job performance [32].

Contrary to our hypothesis, the system did not lead to an increase in social presence between the communicating partners. In fact, the behavioral interdependence measure was found to be higher in the control condition, specifically in the sub-construct on behavioral interdependence [2]. Behavioral interdependence measures the strengthen of mutual impact “e.g., my behavior was in direct response to the other’s behavior.” One interpretation of our result is that including facial expression information in text communication undermined social presence. However, this would seem at odds with much prior work showing that richer channels result in higher social presence [2,49]. Instead, using interview findings, we believe our finding is likely due to the increased salience of the technology in mediating the communication. With the system handling the communication of emotion, it might seem like the communicators are no longer reacting to another directly. This feeling may be further exacerbated by the design decision to use a chat bot to react on behalf of the users. Having this intermediary in between communicators results in participants feeling “disrupted” whenever the emoji are triggered by the bot and reduce the sense of interdependence between the communication partners. As we will discuss in the next section on design implications, there may be ways to improve our designs to make the facial expression signals seem more direct.

Another drawback we noted was the increase in anxiety about emotional leakage when facial-expressions are communicated. This tradeoff relates to research findings in media spaces about the tension between awareness and privacy [3]. But it perhaps more even more closely relates to discussions about truthful signals and self-presentation in CMC [10]. Emoji based on facial expressions is the reliable, assessment signals, whereas the self-triggered emoji represent the unreliable conventional signals. Most current CMC interaction enable impression management through the conventional signals of self-descriptions. Recent studies have also shown that aside from communicating emotions, people are also using emoji for relationship management [22,38]. The use of facial-expression backed signals, while offers more reliable information, reduces people’s ability manage their self-presentation. What is important to note from our findings, is that participants were specifically concerned about negative emotional leakage rather than positive emotional leakage. This suggests that in our study, potential costs associated with positive emotion leakage is less than the costs from negative emotions leakage. Further, while the facial-expression based signals are more truthful, very few signals are impossible to fake. Future work need to explore how people adopt this type of signals in longitudinal use.

### 6.1 Practical Implications

From a technology design perspective, ReactionBot represents a new type of communication channel that is being enabled by advances in emotion recognition. It is situated in between video chat and text chat. It preserves the text-based affordances (e.g., near-synchronous) while enabling the communication of facial expressions. At the same time, the use of emoji instead of transmitting actual images/videos to communicating patterns limit the type of privacy concerns that exists in video [20].

While we found that this type of a system has the potential to enhance communication of emotions to facilitate a more engaged interaction between communication partners, there are also some limitations with the current design that can be improved. One of which is that the current design seems to weaken the sense of interdependence between communicators. Our interview findings suggest that one way to address this problem may be to better integrate the facial-expression signals into the communication channels, instead of relying a separate chatbot. If the reactions are shown to be from the users themselves, instead of a ReactionBot, the user may feel that the reactions are from them (and their partner perceive the reactions to be from them). But at the same time, what is also needed is a better feedback-loop to highlight the fact that the reactions are ultimately triggered by users themselves. This may be achieved with more frequent sampling and analyses of the facial expressions. As soon as a facial expression is shown, it is reacted in the chat window. Users can then make the mental connection between their expressions and what is being shown.

Another question inherent to this work is under what contexts would this system be useful. In general, our participants thought that this type of system might be more appropriate for social uses than task-oriented uses. This corroborates with recent findings emoticons and emoji can adversely affect perceptions of competence when used in work emails [16]. However, we do believe that there may be task-oriented scenarios where this system can be useful. For example, in text-based remote help tasks, facial expressions may not only improve workers' understanding of helpers' instructions, but also help helpers ascertain whether workers understand the instructions. A similar scenario as suggested by our participants, is for interactions with doctors' or physicians. In these scenarios, one may be willing to share the facial expression information in exchange for potentially a higher quality service.

## 7 LIMITATIONS

We used a controlled experiment to test the ReactionBot system and to explore how said system affect communication between dyads. While this study uses a stylized task, it is one that has been commonly used in research in CMC, specifically those that have studied how CMC affect communication behaviors and social presences. Another potential limitation is the sample size. But small sample size is often associated with power (type 2 rather than type 1 error), so that should not undermine the internal validity of our results. Nonetheless, a longitudinal study with more diverse participants is needed to test the external validity of the findings as well as explore norms that develop when the expression-triggered emoji are used in situ, and in contexts with multiple chat partners.

## 8 CONCLUSIONS

In this work, we present ReactionBot, a system that uses users' facial expressions to attach emoji to text messages. Our experimental study demonstrates that such a system could support the communication of facial expressions, an important type of nonverbal cues, in text-based communication. However, this can come at the cost of increasing anxiety due to concerns about emotion leakage and can create undesired distraction, especially in a task-based context. Additional research is needed to both improve on the design of ReactionBot as well as further study its longitudinal use and adoption across different contexts.

## ACKNOWLEDGMENTS

We would like to thank our participants, reviewers, and members of the Prosocial Computing Lab for providing valuable feedback throughout this work.

## REFERENCES

[1] Frank Biocca, Chad Harms, and Judy K. Burgoon. (2003). Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators & virtual environments*, *12*(5), 456-480.

[2] Frank Biocca, Chad Harms, and Jenn Gregg. (2001, May). The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. In *4th annual international workshop on presence*, Philadelphia, PA (pp. 1-9).

[3] Sara A. Bly, Steven Harrison, and Susan Irwin. (1993). Media spaces: bringing people together in a video, audio, and computing environment. *Communications of the ACM*, *36*(1), 28-46.

[4] Judy K. Bugoon, David D. Buller, and W. Gill Woodall. (1996). *Nonverbal communication: The unspoken dialogue.*

[5] V. Orengo Castellá, A.M. Zornoza Abad, F. Prieto Alonso, and J.M. Peiro Silla. (2000). The influence of familiarity among group members, group atmosphere and assertiveness on uninhibited behavior through three different communication media. *Computers in Human Behavior*, *16*(2), 141-159.

[6] Richard L. Daft, and Robert H. Lengel. (1986). Organizational information requirements, media richness and structural design. *Management science*, *32*(5), 554-571.

[7] Daantje Derks, Arjan ER Bos, and Jasper Von Grumbkow. (2007). Emoticons and social interaction on the Internet: the importance of social context. *Computers in human behavior*, *23*(1), 842-849.

[8] Judith Donath. (2007). Signals in social supernets. *Journal of Computer-Mediated Communication*, *13*(1), 231-251.

[9] Paul Ekman, and Wallace V. Friesen. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *semiotica*, *1*(1), 49-98.

[10] Nicole Ellison, Rebecca Heino, and Jennifer Gibbs. (2006). Managing impressions online: Self-presentation processes in the online dating environment. *Journal of computer-mediated communication*, *11*(2), 415-441.

[11] Emotion API Preview. Microsoft Azure.

[12] Scott E. Fahlman. (n.d.). Smiley Lore :-).

[13] Jackson Feijó Filho, Thiago Valle, and Wilson Prata. (2014, September). Non-verbal communications in mobile text chat: emotion-enhanced mobile chat. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services* (pp. 443-446). ACM.

[14] Mark G. Frank, Paul Ekman, and Wallace V. Friesen. (1993). Behavioral markers and recognizability of the smile of enjoyment. *Journal of personality and social psychology*, *64*(1), 83.

[15] Tina Ganster, Sabrina C. Eimler, and Nicole C. Krämer. (2012). Same same but different!? The differential influence of smilies and emoticons on person perception. *Cyberpsychology, Behavior, and Social Networking*, *15*(4), 226-230.

[16] Ella Glikson, Arik Cheshin, and Gerben A. van Kleef. (2017). The dark side of a smiley: Effects of smiling emoticons on virtual first impressions. *Social Psychological and Personality Science*, *9*(5), 614-625.

[17] Charlotte N. Gunawardena, and Frank J. Zittle. (1997). Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. *American journal of distance education*, *11*(3), 8-26.

[18] Jay Hall, and Wilfred Harvey Watson. (1970). The effects of a normative intervention on group decision-making performance. *Human relations*, *23*(4), 299-317.

[19] Albert H. Huang, David C. Yen, and Xiaoni Zhang. (2008). Exploring the potential effects of emoticons. *Information & Management*, *45*(7), 466-473.

[20] Scott E. Hudson, and Ian Smith. (1996, November). Techniques for addressing fundamental privacy and disruption tradeoffs in awareness support systems. In *Proceedings of the 1996 ACM conference on Computer*

*supported cooperative work* (pp. 248-257). ACM.
[21] Benjamin Kehrwald. (2008). Understanding social presence in text-based online learning environments. *Distance Education*, *29*(1), 89-106.
[22] Ryan Kelly, and Leon Watts. (2015). Characterising the inventive appropriation of emoji as relationally meaningful in mediated close personal relationships. *Experiences of Technology Appropriation: Unanticipated Users, Usage, Circumstances, and Design.*
[23] Adam Kendon. (1987). Speaking and signing simultaneously in Warlpiri sign language users. *Multilingua-Journal of Cross-Cultural and Interlanguage Communication*, *6*(1), 25-68.
[24] Joon Young Lee, Nahi Hong, Soomin Kim, Jonghwan Oh, and Joonhwan Lee. (2016, September). Smiley face: why we use emoticon stickers in mobile messaging. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (pp. 760-766). ACM.
[25] Shao-Kang Lo. (2008). The nonverbal communication functions of emoticons in computer-mediated communication. *CyberPsychology & Behavior*, *11*(5), 595-597.
[26] Katja Lohmann, Sebastian Stefan Pyka, and Cornelia Zanger. (2017). The effects of smileys on receivers' emotions. *Journal of Consumer Marketing*, *34*(6), 489-495
[27] Alexandra Martins, Nelson Ramalho, and Estelle Morin. (2010). A comprehensive meta-analysis of the relationship between emotional intelligence and health. *Personality and individual differences*, *49*(6), 554-564.
[28] Lee-Ellen Marvin, L. (1995). Spoof, spam, lurk, and lag: The aesthetics of text-based virtual realities. *Journal of Computer-Mediated Communication*, *1*(2), JCMC122.
[29] 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 ICWSM*, 2016.
[30] Mariella Moon. (2016). Instagram takes a serious look at how people use emojis. https://www.engadget.com/2015/05/02/instagram-emoji-study/
[31] Alena Neviarouskaya, Helmut Prendinger, and Mitsuru Ishizuka. (2010). User study on AffectIM, an avatar-based Instant Messaging system employing rule-based affect sensing from text. *International Journal of Human-Computer Studies, 68*(7), 432-450.
[32] Kostantinos V. Petrides, Norah Frederickson, and Adrian Furnham. (2004). The role of trait emotional intelligence in academic performance and deviant behavior at school. *Personality and individual differences*, *36*(2), 277-293.
[33] Lynne D. Roberts, Leigh M. Smith, and Clare Pollock. (1996). *Exploring virtuality: Telepresence in text-based virtual environments.*
[34] Peter Salovey, and John D. Mayer. (1990). Emotional intelligence. *Imagination, cognition and personality*, *9*(3), 185-211.
[35] Craig Smith. (2018, March 31). 50 interesting Slack Statistics. https://expandedramblings.com/index.php/slack-statistics/
[36] Leslie D. Setlock, Susan R. Fussell, and Christine Neuwirth. (2004, November). Taking it out of context: collaborating within and across cultures in face-to-face settings and via instant messaging. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (pp. 604-613). ACM.
[37] Adam Stembergh, A. (2014). Smile, You're Speaking Emoji: The Rapid Evolution of a Wordless Tongue.
[38] Satomi Sugiyama. (2015). Kawaii meiru and Maroyakaneko: Mobile emoji for relationship maintenance and aesthetic expressions among Japanese teens. *First Monday 20*, 10 (2015).
[39] Alan Swinkels, and Traci A. Giuliano. (1995). The measurement and conceptualization of mood awareness: Monitoring and labeling one's mood states. *Personality and social psychology bulletin*, *21*(9), 934-949.
[40] Sonja Utz. (2000). Social information processing in MUDs: The development of friendships in virtual worlds. *Journal of online behavior.*
[41] Hannes Högni Vilhjálmsson, & Justine Cassell. (1998, May). Bodychat: Autonomous communicative behaviors in avatars. *In Proceedings of the second international conference on Autonomous agents* (pp. 269-276). ACM.
[42] Joseph B. Walther. (1993). Impression development in computer-mediated interaction. *Western Journal of Communication (includes Communication Reports)*, *57*(4), 381-398.
[43] Joseph B. Walther. Relational aspects of computer-mediated communication: Experimental observations over time. *Organization Science 6*: 189–203, 1995.
[44] Joseph B. Walther. (2008). Social information processing theory. *Engaging theories in interpersonal communication: Multiple perspectives*, 391-404.
[45] Joseph B. Walther. (2011). Theories of computer-mediated communication and interpersonal relations. *The handbook of interpersonal communication*, 4, 443-479.
[46] Joseph B. Walther, and Kyle P. D'Addario. (2001). The impacts of emoticons on message interpretation in computer-mediated communication. *Social science computer review*, *19*(3), 324-347.
[47] Joseph B. Walther, and Malcolm R. Parks. (2002). Cues filtered out, cues filtered in. *Handbook of interpersonal*

*communication*, 3, 529-563.

[48] Steve Whittaker. (2003). Theories and Methods in Mediated Communication: Steve Whittaker. In *Handbook of discourse processes* (pp. 246-289). Routledge.

[49] Youngjin Yoo, and Maryam Alavi. (2001). Media and group cohesion: Relative influences on social presence, task participation, and group consensus. *MIS quarterly*, 371-390.

[50] Rui Zhou, Jasmine Hentschel, and Neha Kumar. (2017, May). Goodbye text, hello emoji: mobile communication on wechat in China. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 748-759). ACM.

[51] Robert W. Zmud, Mary R. Lind, and Forrest W. Young. (1990). An attribute space for organizational communication channels. *Information systems research*, *1*(4), 440-457.

Received April 2018; revised July 2018; accepted September 2018.