

# Emoji-Text Mismatches: Stirring the Pot of Online Conversations

Vanessa Vanzan<sup>1</sup>, Amy Han Qiu<sup>1</sup>, Fahima Ayub Khan<sup>1</sup>,  
Chara Soupiona<sup>2</sup>, and Christine Howes<sup>1</sup>

<sup>1</sup>Department of Philosophy, Linguistics and Theory of Science,  
Faculty of Humanities, University of Gothenburg

<sup>2</sup>Department of Philology, Division of Linguistics, University of Crete

## Abstract

We present an ongoing experimental study of how people respond to commonly observed text-emoji patterns in online text-based chats. Using the Dialogue Experimental Toolkit (DiET; Healey et al. 2003), experiments are conducted to compare responses to the same message followed by emojis of opposite emotional valences.

## 1 Background

Digital communication relies on more than just words. In the absence of face-to-face cues, emojis and emoticons take on this role and are widely used, not only among social media users but also by chatbots for different purposes. Including emojis in a digital marketing strategy offers several advantages, such as enhancing engagement, creating a sense of friendliness, and providing a positive personalised experience.

Emojis may play a crucial role in preserving the users' and others self-image, allowing them to add nuances of politeness, humour, or empathy that might be ambiguous in text-only messages. Based on a pilot study of online discussions of a moral dilemma (Soupiona et al., 2024), participants use emojis with positive and negative valences on an alternative basis when presenting their decisions, (e.g., I tend to kill number one 😊), which aligns with Politeness Theory (Brown and Levinson, 1987; Vlasyan et al., 2018). This pattern of alternating valences helps to balance the emotional tone of the conversation, making it more engaging and reducing the potential for misunderstanding (Derks et al., 2008).

This study will investigate whether emojis that deviate from expected politeness norms can influence interlocutors' responses. This will be realised by inserting emojis with opposite emotional valences (i.e., positive 😊 and

negative 😞) in spontaneous dialogues. The current paper describes the methods and our hypotheses. We anticipate presenting preliminary results at the conference.

## 2 Experimental Design

This study is part of ERC project DivCon: Divergence and convergence in dialogue: The dynamic management of mismatches (Starting Grant 101927977). It has ethical approval from the Swedish Ethical Review Authority (Etikprövningsmyndigheten: 2024-00446-01). Data will be collected with written consent from the participants.

Experiments will be conducted using the Dialogue Experimental Toolkit (DiET; Healey et al. 2003), a text-based chat tool designed for introducing word and turn-level interventions in spontaneous dialogues. The latest mobile version of DiET will be run on the Telegram app. The experiments will be conducted on mobile phones provided to the participants, with Telegram pre-installed. The keyboard will be set to English, and the emoji component will be activated. All messages sent by the participants, together with the sender and sending time, will be saved to the server.

Participants who are unfamiliar with each other will be assigned to triadic conversations and led to separate rooms. They will be instructed to discuss the balloon task, an ethical dilemma in which one of four hot air balloon passengers should jump out to their certain death in order to save the others.

The experiment will insert positive and negative emojis when a participant used a decision-related word (e.g., “kill”, “kick”, “save”, or “keep”). The trigger word list is compiled based on face-to-face conversation data collected for previous studies using the same task (Lavelle

et al., 2012; Howes and Lavelle, 2023)

Emojis used for the interventions are selected from the Emoji Sentiment Ranking (Kralj Novak et al., 2015). Only face emojis were chosen to ensure their comparability. Those that do not fit with the balloon task, such as 🎈, 🎉, and 🙄, were excluded. The emojis are shown in Figure 1:

Positive Emojis	Negative Emojis
😄 Grinning Face with Smiling Eyes	😞 Pensive Face
😊 Smiling Face with Smiling Eyes	😫 Tired Face
😋 Face Savoring Food	😱 Fearful Face
😜 Winking Face with Tongue	😞 Disappointed Face
😇 Smiling Face with Halo	😭 Loudly Crying Face
😂 Face with Tears of Joy	😓 Downcast Face with Sweat
😱 Grinning Face with Big Eyes	😓 Anguished Face
😎 Smiling Face with Sunglasses	😓 Anxious Face with Sweat

Figure 1: Emojis used in interventions

The interventions, i.e., text involving decision-related words with added emoji from the sender (Participant A), are sent to the other two participants (B and C). The participants will receive the same emoji, and the emotional valence is randomized. For example, if participant A sends “Passenger A should be saved”, participant B and C will receive “Passenger A should be saved” followed by either 🙄 or 😊. The intervention will not be applied when the sender included emojis in the message. The three participants are randomly assigned as the sender or receiver of the intervention message. Based on the occurrence rate of emojis observed in the pilot study, the intervention will be inserted every 10 to 15 turns.

To control for the discussion time across conversation groups, participants are instructed to chat for approximately 20 minutes but may extend the conversation if necessary. When the discussion ends, the participants will be asked to fill in a questionnaire about their digital habits before receiving a briefing about the study’s purpose and methods.

### 3 Data and Methods

Introducing emojis with opposite emotional valences after decision-related arguments will cause the text and the emoji to have either congruent or incongruent tones, which may elicit responses with different pragmatic functions. In this study, text-emoji pairings with congruent emotional tones are referred to as

“matched”, and those with incongruent tones “mismatched”. The pairings and elicited responses will be annotated by two annotators, and inter-rater reliability will be measured.

The participants’ responses will also be compared in terms of 1) response times, 2) word count, and 3) the number of emojis used by recipients in later chat. The patterns will also be analysed based on the participants’ demographic features and digital habits.

## 4 Hypotheses

Deviations from expected politeness norms would challenge the participants’ expectations of maintaining face, thus requiring the participant to deal with the perceived impoliteness. Therefore, we hypothesise that:

1. Participants who see mismatched emoji-text pairings, compared to those who see matched pairings, will respond with fewer hedgings (e.g., “perhaps,” “kind of”), as mismatched emoji-text pairings, which may be perceived as impolite or confusing, can prompt participants to adapt their discourse accordingly.
2. Participants will spend a longer time responding to mismatched emoji-text pairings, compared to those who see matched pairings.
3. Participants will write more words responding to mismatched emoji-text pairings compared to those who see matched pairings.
4. Participants who see a mismatched emoji-text combo are more likely to use emojis in their responses, compared to those in the matched condition, as a strategy to restore politeness and manage face-threatening situations.

## 5 Implications

This study examines the role of emojis in shape interactional dynamics in online chats. By analysing response patterns and the use of emojis, the method can be used to study user responses in various digital scenarios and provide more insights for the design and optimization of AI-driven chatbots. This study also reveals the influence of cultural and situational factors on emoji use and response patterns.

## References

- Penelope Brown and Stephen C Levinson. 1987. *Politeness: Some universals in language usage*. 4. Cambridge university press.

- Daantje Derks, Arjan ER Bos, and Jasper Von Grumbkow. 2008. Emoticons and online message interpretation. *Social Science Computer Review*, 26(3):379–388.
- Patrick GT Healey, Matthew Purver, James King, Jonathan Ginzburg, and Greg J Mills. 2003. Experimenting with clarification in dialogue. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 25.
- Christine Howes and Mary Lavelle. 2023. Quirky conversations: How people with a diagnosis of schizophrenia do dialogue differently. *Philosophical Transactions of the Royal Society B*. In press.
- Petra Kralj Novak, Jasmina Smailović, Borut Sluban, and Igor Mozetič. 2015. Sentiment of emojis. *PloS one*, 10(12):e0144296.
- Mary Lavelle, Patrick G. T. Healey, and Rose McCabe. 2012. Is nonverbal communication disrupted in interactions involving patients with schizophrenia? *Schizophrenia Bulletin*.
- Chara Soupiona, Vanessa Vanzan, Amy Han Qiu, Fahima Ayub Khan, and Christine Howes. 2024. The balloon dilemma: Emoji-ing moral decisions. Manuscript submitted for publication.
- Gayane R Vlasyan et al. 2018. Linguistic hedging in the light of politeness theory. *European Proceedings of Social and Behavioural Sciences*.