

Impact of Contradicting Subtle Emotion Cues on Large Language Models with Various Prompting Techniques

Noor Ul Huda¹, Sanam Fayaz Sahito^{2*}, Abdul Rehman Gilal³, Ahsanullah Abro⁴,
Abdullah Alshanqiti⁵, Aeshah Alsughayyir⁶, Abdul Sattar Palli⁷

Department of Computer Science, Sukkur Institute of Business Administration, Sukkur, Pakistan^{1,2,4}
Knight Foundation School of Computing and Information Sciences, Florida International University, United States³
Faculty of Computer and Information Systems, Islamic University (IU), Kingdom of Saudi Arabia⁵
College of Computer Science and Engineering, Taibah University, Kingdom of Saudi Arabia⁶
Anti-Narcotics Force, Ministry of Narcotics Control, Islamabad, Pakistan⁷

Abstract—The landscape of human-machine interaction is undergoing a transformation with the integration of conversational technologies. In various domains, Large Language Model (LLM) based chatbots are progressively taking on roles traditionally handled by human agents, such as task execution, answering queries, offering guidance, and delivering social and emotional assistance. Consequently, enhancing user satisfaction with these technologies is crucial for their effective incorporation. Emotions indeed play an effective role in responses generated by reinforcement-learning-based chatbots. In text-based prompts, emotions can be signaled by visual (emojis, emoticons) and linguistic (misspellings, tone of voice, word choice, sentence length, similes) aspects. Therefore, researchers are harnessing the power of Artificial Intelligence (AI) and Natural Language Processing techniques to imbue chatbots with emotional intelligence capabilities. This research aims to explore the impact of feeding contradicting emotional cues to the LLMs through different prompting techniques. The evaluation is based on specified instructions versus provided emotional signals. Each prompting technique is scrutinized by inducing a variety of emotions on widely used LLMs, ChatGPT 3.5 and Gemini. Instead of automating the prompting process, the prompts are given by exerting cognitive load to be more realistic regarding Human-Computer Interaction (HCI). The responses are evaluated using human-provided qualitative insights. The results indicate that simile-based cues have the highest impact in both ChatGPT and Gemini. However, results also conclude that the Gemini is more sensitive towards emotional cues. The finding of this research can benefit multiple fields: HCI, AI Development, Natural Language Processing, Prompt Engineering, Psychology, and Emotion analysis.

Keywords—Emotion cues; prompt; Large Language Model (LLM); Human Computer Interactions (HCI)

I. INTRODUCTION

The proliferation of conversational technologies has resulted in a significant rise in the incorporation of chatbots across various sectors. A chatbot, defined as a dialogue system engaging with humans through natural language via text, voice, or as an embodied agent with multimodal communication, has become increasingly prevalent [1]. Organizations favor chatbots because of their ability to offer

proactive service, immediate assistance, and cost-cutting benefits [2]. They are extensively employed to automate tasks like tracking deliveries, making reservations, obtaining flight information, and placing orders. The round-the-clock availability and swift response to general queries make them an attractive solution for businesses. In recent times, chatbots have extended their utility to provide social and emotional support in healthcare and personal contexts [3].

Chatbots stand out as the fastest-growing communication channel globally, spanning various domains [4]. The substantial advantages associated with integrating chatbots in service and social areas prompt organizations to make significant investments in this technology. Despite this, research suggests that users still harbor reservations about chatbot interactions and express a preference for human agents [2]. Additionally, a review on chatbot usability and user acceptance indicates that people lean towards natural communication as opposed to machine-like interactions, believing that a human can better comprehend them [5]. The study underscores the importance of user satisfaction in successfully integrating and adopting chatbots. Consequently, enhancing user engagement and satisfaction with chatbot interactions has become paramount to delivering an improved experience and encouraging widespread adoption of the technology [6].

The interaction with the chatbots is achieved through input called as Prompt. The prompt in general, is an input aimed to induce a particular response [7]. It can be anything i.e. written statement, voice, image, action, or physical gesture. Since the development of advanced machine learning models, the term "Prompt" has become a referring point for explicit instructions that are provided to these models to get the output. Prompt engineering involves specifying the prompt with the goal of obtaining the desired output. While navigating different social conditions, humans are programmed to express a wide range of emotions. Emotional expressions have a communication purpose. They transmit information about intentions, feelings, needs for action, and situational assessments. They ease the

*Corresponding Author

coordination of social interactions [8] [9]. Prompts were previously more dependent on certain patterns with specific machine learning models, but nowadays, with the development of massive language models, prompting has become easier. These models enable us to feed prompts in any style or tone of human speech. Emotions are naturally included in human speech. People frequently communicate their emotions in writing prompts for large language models (LLMs), both consciously and unconsciously. These feelings can be expressed in a variety of ways, from subtle to more overt expressions [8].

When an LLM detects an emotion, it examines the dataset on which it was trained to determine the meaning of that emotion [9]. As Fig. 1 shows that, if an emotion has a single and well-defined meaning, the model associates that meaning with the emotion. However, if the emotion has numerous meanings, the model considers the prompt's surrounding context. The model makes a sensible assumption about the emotion's intended meaning based on the context. Once the interpretation is made, the model generates a response that corresponds to the emotion's interpreted meaning.

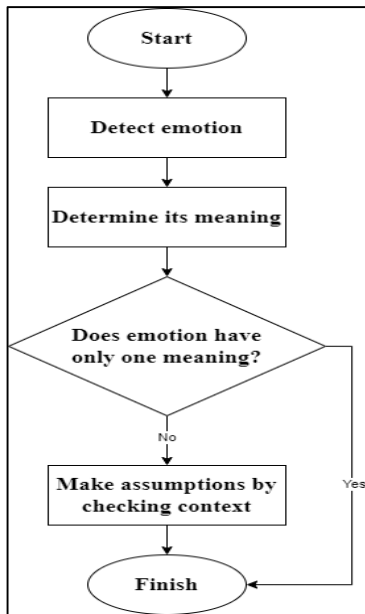


Fig. 1. A flowchart defining LLM's behavior in emotion detection.

It is a common knowledge that LLMs have good understanding of emojis and emoticons [9] [11]; however, these are quite good in interpreting different means of linguistic cues. According to prior researches, LLMs are smart enough to detect emotions when they are explicitly instructed. An example of this can be seen where GPT 3.5 powered SnapAI [12] detected emotions in a lengthy prompt (containing lots of linguistic cues) when instructed explicitly:

The first question to come up is whether these LLMs detect emotions with same perfection evenssss when they are not explicitly instructed, and the second question is what will be the impact of these emotions when these are in contrast to given task. Since zero-shot prompting is the primary method

used by most chatbot users [13], most research on emotion in prompting focuses on this technique. Other advanced prompting methods, on the other hand, are more concerned with specific instructions than emotions. Even if the effect of conflicting emotions on these sophisticated prompting tactics might be small, our aim is to explore it. Our goal in examining this effect is to learn more about the ways in which emotions may affect these prompts.

Prompt: "Okay, so from this prompt of mine you may evaluate the emotions of me. What do you think, and you may write the whole repeat the whole prompt of me by mentioning in bracket that which emotion you find were part of my prompt. So, you may do it, and I know you can do it. So, you may be detailed more detailed more and more detailed, and try to find the emotions in my messages. Now, after that, you can write a conclusion paragraph in describing overall emotions of my message. And also, you may, what can I say, you may identify my tone. What is my prompting tone?"

SnapAI's Response: "Sure! Let's evaluate the emotions in your prompt. Here's the repeat of your prompt with emotions identified:

"Okay, so from this prompt of mine (curiosity) you may evaluate the emotions of me (anticipation, excitement). What do you think, and you may write the whole repeat the whole prompt of me (enthusiasm) by mentioning in bracket that which emotion you find were part of my prompt (eagerness). So, you may do it (encouragement), and I know you can do it (confidence). So, you may be detailed (expectation) more detailed (emphasis) more and more detailed (enthusiasm) and try to find the emotions in my messages (curiosity). Now, after that, you can write a conclusion paragraph (anticipation) in describing overall emotions of my message (expectation). And also, you may, what can I say, you may identify my tone (curiosity). What is my prompting tone (interest)?"

Overall, your prompt reflects a sense of curiosity, anticipation, excitement, enthusiasm, encouragement, confidence, expectation, eagerness, and interest. Your tone throughout the prompt is engaging and shows a genuine desire to understand and explore emotions. Keep up the great work! 😊"

II. LITERATURE REVIEW

Analysis of emotions has always been part of psychological research. In Human-Computer Interaction (HCI), emotions help in enhancing user experience by improving communications and decision-making to adapt to user needs [14].

This literature review's focus is on emotion cues and the relationship between emotions and LLMs. An in-depth search was conducted using several databases, including IEEE Xplore, Google Scholar, ScienceDirect, and ACM Digital Library.

According to linguistics research, emotions are conveyed through texts and images including various ways such as choice of visual elements (facial expressions, emojis, or emoticons), choice of words, formatting, or color themes [15] [16] [17]. Prior to Computer-Mediated Communication (CMC), people used to convey emotions with linguistic cues. They also used to explicitly declare their emotional state in a bracket and adjust it with the text [18]. In 1982, the first emoticon was invented by an American Professor Scott Fahlman using ASCII characters. These emoticons later

transformed into emojis in 1997 by Softbank, and in 1999 by Docomo, using pixel art [19] [20]. There is no proper record of modification in linguistic cues after invention of emoticons and emojis.

The use cases of emoticons and emojis are written in various literature [20] [21] [22]. Emojis have greater impact in cross-culture communications [23]. An individual's personality and behaviour can be determined by linguistic cues from text written by them [24] because these are often used unintentionally; however, as visual cues are always selected with intention, they often create contradiction with text, and create misguidance [25] [26] [27]. Emotion cues works better when they are combination of visual and linguistic cues [23] [28] [29].

The introduction of LLMs gave research a new angle. The majority of retrieved research regarding association of emotion cues with LLMs is about emotion detection. Some of them are about text-based emotion detection [9-11] [30], while others are related to visual emotion analysis [31] [32] [33]. Certain sensitive fields use emotion detection, such as the analysis of suicide notes' emotions [30] and the identification of emotions in autistic children [32]. Emotion detection improves AI systems' understanding and response to users' emotions, assisting in the diagnosis of mental health disorders [34] [35], the customization of marketing campaigns [36][37], and the monitoring of emotional well-being. Most notable applications include mental health chatbots [29] [38- 40], and emotion-aware wearable gadgets [41]. Another study direction is to inject emotions [42] into LLMs and have them act in any given specified role [43] [44]. Emotion prompting is also responsible for the creation of the famous jailbreak "Do Anything Now" [45].

The existing literature doesn't provide a clear definition of how emotions have an impact on different prompting techniques. Additionally, it lacks an explanation of the use cases for incorporating emotions in images. It would be interesting to explore further into these areas in order to better understand the impact of emotions in diverse circumstances.

III. METHODOLOGY

This research analyzed the responses of ChatGPT 3.5 and Gemini on prompts of different techniques, with each having different types of subtle emotions including visual and linguistic cues. Each prompt is constructed manually by exerting cognitive load to be more realistic in context of Human-Computer Interaction (HCI). Emotion cues and prompting techniques are chosen on the basis of their wide recognition and common usage.

A. Establishment of Prompt

Each prompt contains an instruction, and a contradicting emotion cue. The instruction contains an action verb and a context of job, and then an opposite emotion is subtly inserted into the prompt. In visual cues, emojis, or emoticons are directly inserted, while linguistic emotion cues are given by addition or manipulation of words.

The Fig. 2 provides a visual representation of our interaction with LLM-based chatbot. The two inward-pointing arrows are depicting inputs, and the one outward-pointing arrow is representing the LLM's response as output.

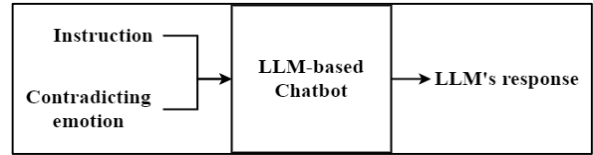


Fig. 2. An illustration of LLM's input and output for this research.

B. Chosen Emotion Cues

The research is based on two sorts of emotion cues—visual and linguistic—to be incorporated into text-based prompts. We have employed emojis [46] and emoticons [47] as visual clues. These are necessary part of digital communication which can be inserted easily. Emojis and emoticons are typed with the typer's willingness in mind, and these are more frequently used than ASCII art and kaomojis [48, 49]. Additionally, as linguistic context, we have selected five emotion cues: 1. Misspellings [29], 2. Tone of voice [50], 3. Word choice [51], 4. Sentence length [52][53], and 5. Similes [54]. Intentional misspelled words cause informality in digital communication [55], while unintentional misspellings i.e. substitution errors, omission errors, and homophone errors tend to occur when text is typed in a rush [56][57]. Tone of voice, and word choice are influenced by specific situation or circumstances in which communication takes place. The receiver assumes an impression of typer based on compiled tone and word choice [58]. A study shows that in excitement, people provide extra details and explanation, and make sentences lengthy, but unrelated context leads to decrement of LLM's accuracy in solving problems. [53]. By expressing distinct and individualized emotions through seeking commonalities between two different things, both similes and metaphors makes statement clearer [59]; however, simile is easier to recognize and interpret, and is more commonly used by people in their daily conversation [60] [61]. These emotion cues are labeled in Table I.

TABLE I. EMOTIONS

Label	Emotion Type
E ₁	Emojis
E ₂	Emoticons
E ₃	Misspellings
E ₄	Tone of voice
E ₅	Word choice
E ₆	Sentence length
E ₇	Simile

C. Chosen Prompting Techniques

Six different prompting techniques are explored in this research which given in Table II. These techniques include

zero-shot prompting which is the most basic kind of prompt with no example [62] [63], sequential thinking prompting by beginning each zero-shot prompt with the phrase “Think step-by-step” [64] [13] [65], few-shot prompting [66] with three examples added to zero-shot prompt, role-playing by assigning a role or persona to chatbot [67] [68] by giving first prompt with phrase “Act as a [role/persona]” and a clear description of the role which is later followed by initial zero-shot prompt as second prompt. The last two techniques are refinement of response of zero-shot prompt through a *feedback prompt* [69], and *template filling prompt* [70] in which we began with an incomplete statement, and then used square brackets to indicate where LLM should write its response.

TABLE II. PROMPT TYPES AND EXAMPLES

Label	Prompt Type	Prompt Text
P ₁	Zero-shot prompt with contradictory emojis	Give only one line response. Recommend me an outfit for summer 🧣❄️👔☐
P ₂	Feedback on response to Zero-shot prompt	don't make it cool and stylish. 🙅🙅🙅
P ₃	Sequential Thinking Prompt	Think step-by-step, and give only one line response. Recommend me an outfit for summer 🧣❄️👔☐
P ₄	3-shot prompt	User: Recommend me a book. AI: "The summer I turned pretty" by Jenny Han User: Recommend me a place AI: Beach User: Recommend me an activity. AI: Swimming User: Recommend me an outfit 🧣❄️👔☐ AI:
P ₅	Role-playing Prompt	Act as an AI who always response in context of summer season; Recommend me an outfit 🧣❄️👔☐
P ₆	Template prompt	Don't write complete paragraph. Just fill this template: 🧣❄️👔☐ One suitable outfit for summer is [name of an outfit].

D. Experimental Mechanism

Each sort of seven emotion cue (E₁ to E₇) is injected in each of the six prompting techniques (P₁ to P₆). The first step in the process is to select an emotion cue E₁ and apply it to prompt P₁. This input is then fed into ChatGPT and Gemini. Both models generate responses, which are evaluated through human feedback. In a separate chat, the same emotion cue E₁ is then applied to the next prompting technique i.e. P₂, and this process continues until all move on to emotion cue E₂, and repeat the process for all prompting techniques. This iterative process continues until all emotion cues and prompting techniques are implemented. This can be mathematically represented using Eq. (1).

$$P_{i\phi} = M_i (\Pi_\phi, E_k) \tag{1}$$

Where, R_{ij} represents the response generated by M_i, the i-th model (ChatGPT or Gemini) using the P_j, the j-th prompting technique (P₁ to P₆) and emotion cue E_k which represents the k-th emotion cue (E₁ to E₇). The whole process is further illustrated in Fig. 3.

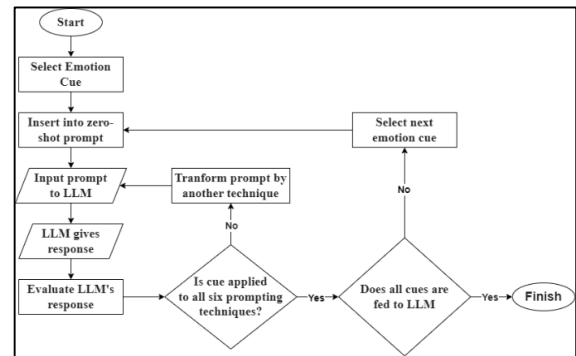


Fig. 3. A flowchart defining prompt feeding process.

Concerning our initial iteration, we started with E₁ that is emojis. The first zero-shot prompt was “Give only one line response. Recommend me an outfit for summer” because it is a common use-case of LLM [71]. For P₁, we inserted “🧣❄️👔☐” as these are contradictory to the prompt, and represent winter and cold weather; however, LLM interprets these emojis in sense of coolness and style. Next, we fed feedback prompt P₂ to not add elements of coolness and style in response. We included the keyword “Think step-by-step” to P₃. Then, moving on to P₄, we added three concise and direct shots, and then inserted the same P₁ right after these shots. In P₅, we designate the LLM to always respond in the context of summer season. Then, we ask it to recommend an outfit, but with contradicting emojis “🧣❄️👔☐”. Lastly, in P₆, LLM was instructed to fill in the template by providing the name of an outfit suitable for summer, but with same counter emojis.

IV. EVALUATION

The evaluation process includes determining if the instructions and emotions were taken into account for each area of prompt engineering. An evaluation based on human judgment was used to carry out this assessment, weighing the overall importance of the prompts, the influence of emotions, and the weight of the instructions. The scales from 1 to 6 are defined to allow consistency and transparency in evaluation process. The detail of the scale is given in Table III below:

TABLE III. EMOTION SCALE DESCRIPTION

Emotion Impact	Scale	Description
Low	1	Emotion is completely neglected, but instruction is considered
	2	Emotion is considered, but it has no impact on LLM's response
Medium	3	Emotion is considered, and it has low impact on response
	4	Both emotion and instruction have 50-50 weightage
High	5	Impact of emotion is higher than impact of instruction
	6	Instruction is completely neglected, only emotion is considered

A. Comparative Analysis

The following tables, labeled as Table IV present the comparative analysis for each emotion:

TABLE IV. EVALUATION OF EMOTION CUES ON EACH PROMPTING TECHNIQUE

Emotion	Prompting technique	LLM	
		ChatGPT 3.5	Gemini
E ₁	P ₁	4	3
	P ₂	1	1
	P ₃	3	6
	P ₄	6	6
	P ₅	4	3
	P ₆	3	1
E ₂	P ₁	1	3
	P ₂	1	1
	P ₃	1	2
	P ₄	1	6
	P ₅	1	4
	P ₆	5	5
E ₃	P ₁	1	4
	P ₂	1	2
	P ₃	5	1
	P ₄	1	1
	P ₅	1	1
	P ₆	1	1
E ₄	P ₁	4	3
	P ₂	3	3
	P ₃	3	3
	P ₄	4	3
	P ₅	3	3
	P ₆	5	1
E ₅	P ₁	1	4
	P ₂	1	4
	P ₃	1	4
	P ₄	1	6
	P ₅	1	5
	P ₆	6	5
E ₆	P ₁	3	5

	P ₂	1	1
	P ₃	1	5
	P ₄	1	1
	P ₅	3	4
	P ₆	1	5
	E ₇	P ₁	5
P ₂		3	4
P ₃		5	5
P ₄		4	4
P ₅		4	5
P ₆		3	5

The evaluation reveals that in ChatGPT, Simile has high impact on each of the prompting techniques, and the emoticons, misspelling, and word choice have neglected impact on five of the six prompting techniques, while in Gemini, simile has highest impact on four of the six prompting techniques, and misspelling has lowest impact on each of the prompting techniques.

B. Evaluation by Experts

To enhance the credibility of our evaluation, we utilized the expertise of four professionals mentioned in the Table V, from relevant fields. Each expert has been assigned a variable for better evaluation:

TABLE V. EXPERTS

Label	Expertise
X ₁	Psychologist
X ₂	Linguist
X ₃	Human-Computer Interaction Expert
X ₄	Prompt Engineer

The selection process of the experts is on basis of their qualifications, experience, and expertise. Since these four experts come from diverse fields, their evaluation and assigned scales differ. Table VI records each expert's evaluation for each prompt using a pre-defined numerical scale from 1 to 6.

TABLE VI. EVALUATION BY EXPERTS

Prompt	X1		X2		X3		X4	
	ChatGPT	Gemini	ChatGPT	Gemini	ChatGPT	Gemini	ChatGPT	Gemini
P ₁ & E ₁	1	1	1	1	3	2	3	3
P ₂ & E ₁	1	4	1	2	2	4	1	1
P ₃ & E ₁	6	6	5	6	6	5	3	6
P ₄ & E ₁	3	3	3	3	3	3	6	6
P ₅ & E ₁	1	1	1	1	1	1	3	3
P ₆ & E ₁	2	1	2	1	1	2	3	1

Each expert brought their unique expertise and perspective to the table. The psychologist focused on how these models affect user emotions [72], while linguistic paid attention to linguistic nuances and biases [73] presented in the generated response. The HCI expert evaluates user experience by considering how well the LLM's response aligns with the intended design of the prompt, and with a friendly contradiction [74, 75], the prompt engineer checks the

response completely on basis of design and creation of the prompt.

V. RESULTS

The impact of contradicting emotion cues is categorized in three means: *High* when the scale is 5 or 6, *Medium* when the scale is 3 or 4, and *Low* when the scale is 1 or 2. The

following table gives a quantitative summary of both models on basis of each prompting techniques.

Across all three impact levels, ChatGPT 3.5 performed best in the sequential thinking and template filling tasks. Its strongest overall performance was in template filling where it achieved high impact three times. Gemini achieved its highest scores on the sequential thinking, three-shots, and template filling tasks, with scores of 3 and 4 in the high impact level. Its strongest performance was in template filling, where it received a score of 4. Looking at the total scores, ChatGPT achieved a total of 7 for high impact, 14 for medium impact, and 21 for low impact. Gemini had totals of 14 for high impact, 16 for medium impact, and 12 for low impact.

Table VII concludes that Gemini is highly sensitive towards emotion cues, while most of the time, ChatGPT prefers the direction of instruction if it is in contradiction to emotion signal. It is an interesting finding that both LLMs caught emotion cue more frequently in template filling.

TABLE VII. IMPACT QUANTITATIVE SUMMARY ON BASIS OF PROMPTING TECHNIQUES

		Impact		
		High	Medium	Low
ChatGPT 3.5	P ₁	1	3	3
	P ₂	0	2	5
	P ₃	2	2	3
	P ₄	1	2	4
	P ₅	0	4	3
	P ₆	3	1	3
	Total	7	14	21
Gemini	P ₁	2	5	0
	P ₂	0	3	4
	P ₃	3	2	2
	P ₄	3	2	2
	P ₅	2	4	1
	P ₆	4	0	3
	Total	14	16	12

For emotion cue E₁, both models showed over 50% impact. E₂ had a larger difference, with Gemini significantly higher at 58.33% versus ChatGPT's 27.77%. E₃ again impacted both models around 27-28%. For E₄, ChatGPT outperformed with 61.11% impact versus Gemini's 44.44%. The largest variation was in E₅, where Gemini achieved a very high impact of 77.77% compared to ChatGPT's more moderate 30.55%. In terms of E₆, Gemini continues its lead with 58.33% impact versus ChatGPT's 27.77%. Finally, for E₇, both models achieved over 65% impact, with Gemini again slightly ahead at 77.77% versus ChatGPT's 66.66%. Fig. 4 highlights that while both models showed varying responses to different emotion cues, Gemini tended to surpass ChatGPT 3.5 in terms of percentage impact, particularly for cues E₂, E₅, and E₇.

The following bar chart presents the percentage of impact for each emotion cue (E₁-E₇) on a scale from 0 to 90%:

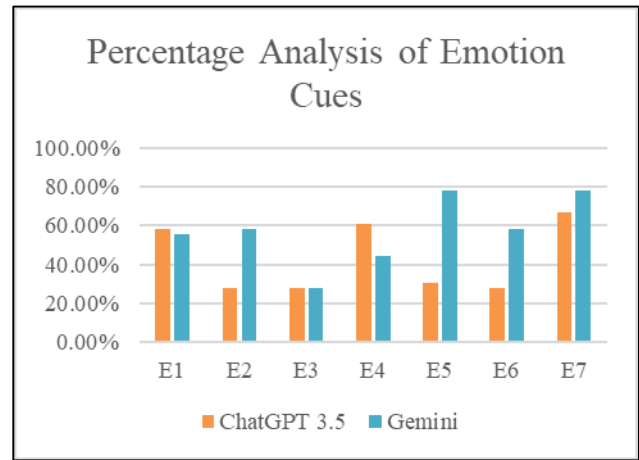


Fig. 4. Percentage analysis of Table VII on basis of emotion cues.

VI. CONCLUSION AND FUTURE WORK

In conclusion, this research aimed to analyze the impact of emotion cues in six different commonly used prompting techniques. Instead of relying on an automated process, we fed prompts manually to the ChatGPT and Gemini, and evaluated it with human assessment. This study is limited to manual feeding of prompts and no any automated model is used. The research obtained significant findings that Gemini is highly sensitive towards emotions even if they are contradictory, and the prompting by template filling also catches emotions while neglecting its instructions most of the time. The research findings have implications for several fields, for instance, understanding how emotional cues impact responses from LLMs in chatbots can contribute to the development of more effective and user-friendly interfaces. This knowledge can inform the design of systems that better cater to users' emotional needs and expectations. Moreover, understanding how LLMs respond to contradicting emotional cues provides valuable insights into the psychological aspects of human-computer interaction. This knowledge can contribute to a deeper understanding of how users perceive and interact with emotionally intelligent systems. In our future research, we aim to explore the nuanced interplay of emotion cues within prompts that incorporate both images and voices. This exploration seeks to deepen our understanding of how visual and auditory elements, in conjunction with textual instructions, influence the responses of LLMs in chatbot interactions.

REFERENCES

- [1] M. Allouch, A. Azaria, and R. Azoulay, "Conversational agents: goals, technologies, vision and challenges," *Sensors*, vol. 21, no. 24, p. 8448, 2021.
- [2] M. Adam, M. Wessel, and A. Benlian, "AI-based chatbots in customer service and their effects on user compliance," *Electronic Markets*, vol. 31, no. 2, pp. 427-445, 2021.
- [3] M. Milne-ives, C. Cock, E. Lim et al., "The effectiveness of artificial intelligence conversational agents in health care: systematic review," *Journal of Medical Internet Research*, vol. 22, no. 10, article e20346, 2020.

- [4] M. Moran, "25+ top chatbot statistics for 2022: usage, demographics, trends," *Tech. Rep., Startup Bonsai*, 2022, September 2022, <https://startupbonsai.com/chatbot-statistics>.
- [5] A. Rapp, L. Curti, and A. Boldi, "The human side of human-chatbot interaction: a systematic literature review of ten years of research on text-based chatbots," *International Journal of Human Computer Studies*, vol. 151, article 102630, 2021.
- [6] J. S. Chen, T. T. Y. Le, and D. Florence, "Usability and responsiveness of artificial intelligence chatbot on online customer experience in e-retailing," *International Journal of Retail and Distribution Management*, vol. 49, no. 11, pp. 1512–1531, 2021.
- [7] "Prompt," *Oxford Learner's Dictionaries, American English*. [Online]. Available: https://www.oxfordlearnersdictionaries.com/definition/american_english/prompt_2#:~:text=1%5Btransitive%5D%20to%20make%20someone,a%20man%20in%20the%20crowd (accessed Feb. 2, 2024)
- [8] E. Kennedy-Moore and J. C. Watson, "Expressing emotion: Myths, realities, and therapeutic strategies," Guilford Press, 2001.
- [9] F. A. Acheampong, C. Wenyu, and H. Nunoo-Mensah, "Text-based emotion detection: Advances, challenges, and opportunities," *Engineering Reports*, vol. 2, no. 7, p. e12189, 2020.
- [10] Gilal, Abdul Rehman, Jafreezal Jaafar, Mazni Omar, Shuib Basri, and Ahmad Waqas. "A rule-based model for software development team composition: Team leader role with personality types and gender classification." *Information and Software Technology* 74 (2016): 105-113.
- [11] S. Zad, M. Heidari, H. James Jr, and O. Uzuner, "Emotion detection of textual data: An interdisciplinary survey," in *2021 IEEE World AI IoT Congress (AIIoT)*, May 2021, pp. 0255-0261. IEEE.
- [12] Help Center - What is My AI on Snapchat and how do I use it? Snapchat, help.snapchat.com/hc/en-gb/articles/13266788358932-What-is-My-AI-on-Snapchat-and-how-do-I-use-it. (accessed Jan. 15, 2024)
- [13] T. Kojima, S. S. Gu, M. Reid, Y. Matsuo, and Y. Iwasawa, "Large language models are zero-shot reasoners," in *Advances in neural information processing systems*, vol. 35, pp. 22199-22213, 2022.
- [14] A. Sears, J.A. Jacko, and J.A. Jacko (Eds.), "The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, Second Edition," 2nd ed., CRC Press, 2007. [Online]. Available: <https://doi.org/10.1201/9781410615862>
- [15] J. T. Hancock, C. Landrigan, and C. Silver, "Expressing emotion in text-based communication," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, April 2007, pp. 929-932.
- [16] S. Aman and S. Szpakowicz, "Identifying expressions of emotion in text," in *International Conference on Text, Speech and Dialogue*, September 2007, pp. 196-205. Berlin, Heidelberg: Springer Berlin Heidelberg.
- [17] S. Zhao, Y. Gao, X. Jiang, H. Yao, T. S. Chua, and X. Sun, "Exploring principles-of-art features for image emotion recognition," in *Proceedings of the 22nd ACM international conference on Multimedia*, November 2014, pp. 47-56.
- [18] V. Evans, "The emoji code: The linguistics behind smiley faces and scaredy cats," Picador, 2017.
- [19] L. Gawne and J. Daniel, "The past and future of hand emoji," in *Proceedings of the 4th International Workshop on Emoji Understanding and Applications in Social Media*, Jul. 2021.
- [20] J. Burge, "Correcting the record on the first emoji set," *Emojipedia Blog*, Mar. 8, 2019. [Online]. Available: <https://blog.emojipedia.org/correcting-the-record-on-the-first-emoji-set/>. (accessed Dec. 18, 2023)
- [21] Y. Tang and K. F. Hew, "Emoticon, emoji, and sticker use in computer-mediated communication: A review of theories and research findings," *Int. J. Commun.*, vol. 13, pp. 27, 2019.
- [22] T. W. Park, S. J. Kim, and G. Lee, "A study of emoticon use in instant messaging from smartphone," in *Human-Computer Interaction. Applications and Services: 16th International Conference, HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part III* 16, Springer International Publishing, 2014, pp. 155-165.
- [23] H. Alshenqeeti, "Are Emojis Creating a New or Old Visual Language for New Generations? A Socio-semiotic Study," *Adv. Lang. Lit. Stud.*, vol. 7, pp. 56-69, 2016. doi: 10.7575/aiac.all
- [24] F. Mairesse, M. A. Walker, M. R. Mehl, and R. K. Moore, "Using linguistic cues for the automatic recognition of personality in conversation and text," *J. Artif. Intell. Res.*, vol. 30, pp. 457-500, 2007.
- [25] A. Wagner, S. Marusek, and W. Yu, "Emojis and Law: contextualized flexibility of meaning in cyber communication," *Social Semiotics*, vol. 30, pp. 396-414, 2020. [Online]. Available: <https://doi.org/10.1080/10350330.2020.1731198>.
- [26] H. Miller et al., "'Blissfully Happy' or 'Ready to Fight': Varying Interpretations of Emoji," in *Proceedings of the International Conference on Web and Social Media*, 2016, pp. 259-268. [Online]. Available: <https://doi.org/10.1609/icwsm.v10i1.14757>.
- [27] H. Miller et al., "Understanding Emoji Ambiguity in Context: The Role of Text in Emoji-Related Miscommunication," in *Proceedings of the International Conference on Web and Social Media*, 2017, pp. 152-161. [Online]. Available: <https://doi.org/10.1609/icwsm.v11i1.14901>.
- [28] N. Cohn, J. Engelen, and J. Schilperoord, "The grammar of emoji? Constraints on communicative pictorial sequencing," *Cognitive Research: Principles and Implications*, vol. 4, 2019. [Online]. Available: <https://doi.org/10.1186/s41235-019-0177-0>.
- [29] "The Emoji Code," *Science*, vol. 357, p. 763, 2017. [Online]. Available: <https://doi.org/10.1126/science.aao5728>.
- [30] B. Desmet and V. Hoste, "Emotion detection in suicide notes," *Expert Systems with Applications*, vol. 40, no. 16, pp. 6351-6358, 2013.
- [31] V. Gajarla and A. Gupta, "Emotion detection and sentiment analysis of images," *Georgia Institute of Technology*, vol. 1, pp. 1-4, 2015.
- [32] P. Rani, "Emotion detection of autistic children using image processing," in *2019 Fifth International Conference on Image Information Processing (ICIIP)*, November 2019, pp. 532-535. IEEE.
- [33] Y. Lu, C. Guo, X. Dai, and F. Y. Wang, "Generating Emotion Descriptions for Fine Art Paintings via Multiple Painting Representations," *IEEE Intelligent Systems*, 2023.
- [34] K. Dheeraj and T. Ramakrishnu, "Negative emotions detection on online mental-health related patients texts using the deep learning with MHA-BCNN model," *Expert Systems with Applications*, vol. 182, p. 115265, 2021.
- [35] Palli, Abdul Sattar, Jafreezal Jaafar, Abdul Rehman Gilal, Aeshah Alsughayyir, Heitor Murilo Gomes, Abdullah Alshantiti, and Mazni Omar. "Online Machine Learning from Non-stationary Data Streams in the Presence of Concept Drift and Class Imbalance: A Systematic Review." *Journal of Information and Communication Technology* 23, no. 1 (2024): 105-139. S. Wang and B. Chen, "Customer emotion analysis using deep learning: Advancements, challenges, and future directions," in *Proceedings of the International Conference on Modern Scientific Research*, 2023, pp. 21-24.
- [36] N. Nguyen, T. H. Nguyen, Y. N. Nguyen, D. Doan, M. Nguyen, and V. H. Nguyen, "Machine learning-based model for customer emotion detection in hotel booking services," *Journal of Hospitality and Tourism Insights*, 2023.
- [37] P. Rathnayaka, N. Mills, D. Burnett, D. De Silva, D. Alahakoon, and R. Gray, "A mental health chatbot with cognitive skills for personalized behavioral activation and remote health monitoring," *Sensors*, vol. 22, no. 10, p. 3653, 2022.
- [38] Gila, Abdul Rehman, Jafreezal Jaafa, Mazni Omar, and Muhammad Zahid Tunio. "Impact of personality and gender diversity on software development teams' performance." In *2014 International Conference on Computer, Communications, and Control Technology (I4CT)*, pp. 261-265. IEEE, 2014.
- [39] F. Booth, C. Potts, R. Bond, M. Mulvenna, C. Kostenius, I. Dhanapala, et al., "A Mental Health and Well-Being Chatbot: User Event Log Analysis," *JMIR mHealth and uHealth*, vol. 11, p. e43052, 2023.
- [40] Z. Yan, Y. Wu, Y. Zhang, and X. A. Chen, "EmoGlass: An end-to-end AI-enabled wearable platform for enhancing self-awareness of emotional health," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1-19.

- [41] S. Muranaka, T. Fukatsu, Y. Takebayashi, M. Kunugi, S. Nakajima, and R. So, "Emotion-injecting prompt for large language model chatbot," 2023.
- [42] "Character.ai," character.ai, [Online]. Available: <https://beta.character.ai/> (accessed Dec. 8, 2023).
- [43] "Social Profiles for Meta's AI Characters," Facebook, September 2023. [Online]. Available: <https://about.fb.com/news/2023/09/social-profiles-for-metas-ai-characters/> (accessed Dec. 8, 2023).
- [44] X. Shen, Z. Chen, M. Backes, Y. Shen, and Y. Zhang, "'do anything now': Characterizing and evaluating in-the-wild jailbreak prompts on large language models," arXiv preprint arXiv:2308.03825, 2023.
- [45] Z. Al-Halah et al., "Smile, be happy:) emoji embedding for visual sentiment analysis," in Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2019.
- [46] B. O'Neill, "ASCII affect: A comparison of emoticons and facial expressions in affective priming," 2019.
- [47] E. Giannoulis et al., "Emoticons, Kaomoji, and Emoji," Emoticons, Kaomoji, and Emoji, 2019. [Online]. Available: <https://doi.org/10.4324/9780429491757-1>.
- [48] T. D. Marten, "The Usage of Abbreviation and Misspelling Occurrence in DOTA 2 Player in Game Chat Log," LANGUAGE HORIZON, vol. 9, no. 3, pp. 1-12, 2021.
- [49] M. Christanti, P. Mardani, and K. Fadhila, "Analysing The Meaning Of Tone Indicators By Neurodivergent Community in Twitter," International Journal of Social Science Research and Review, vol. 5, no. 1, 2022. [Online]. Available: <https://doi.org/10.47814/ijssr.v5i1.118>.
- [50] G. Feng, "On artistic techniques of choice of words in English translation and writing," Journal of Hefei University of Technology.
- [51] D. Das and S. Bandyopadhyay, "Sentence-level emotion and valence tagging," Cognitive Computation, vol. 4, pp. 420-435, 2012.
- [52] F. Shi et al., "Large Language Models Can Be Easily Distracted by Irrelevant Context," ArXiv, abs/2302.00093, 2023. [Online]. Available: <https://arxiv.org/abs/2302.00093>.
- [53] P. Riddell, "Metaphor, simile, analogy and the brain," Changing English, vol. 23, no. 4, pp. 363-374, 2016.
- [54] M. Y. Shipilov et al., "Misspelling of a Lexical Unit as a Marker of Its Semantic Specialization (Based on Erratives Zhosky and Zhosko in Informal Online Communication)," Review of Omsk State Pedagogical University. Humanitarian research, 2022. [Online]. Available: [https://vestnik-omgpu.ru/volume/2022-3-36/vestnik_3\(36\)2022_130-137.pdf](https://vestnik-omgpu.ru/volume/2022-3-36/vestnik_3(36)2022_130-137.pdf).
- [55] G. Logan et al., "Cognitive Illusions of Authorship Reveal Hierarchical Error Detection in Skilled Typists," Science, vol. 330, pp. 683-686, 2010. [Online]. Available: <https://doi.org/10.1126/science.1190483>.
- [56] F. Yiannas, "Food Safety @ the Speed of Thought," Springer, pp. 67-69, 2015. [Online]. Available: https://link.springer.com/chapter/10.1007/978-1-4939-2489-9_19.
- [57] J. Walther et al., "Interpersonal Effects in Computer-Mediated Interaction," Communication Research, vol. 19, no. 1, pp. 52-90, 1992. [Online]. Available: <https://doi.org/10.1177/009365092019001003>.
- [58] Y. Zheng et al., "Acceptable and Reasonable Mistake—On the Birth of Simile and Metaphor and its Aesthetic Value," Journal of Wenshan Teachers' College, 2003.
- [59] L. Ya-ping et al., "On Other Patterns of Simile and Its Usage," Journal of Shangluo University, 2008.
- [60] Y. Song et al., "Simile and Metaphor Interpretation in Children," English Language Teaching, vol. 13, no. 4, pp. 91-103, 2020. [Online]. Available: <https://doi.org/10.5539/elt.v13n4p91>.
- [61] C. Zhou et al., "Prompt consistency for zero-shot task generalization," arXiv preprint arXiv:2205.00049, 2022.
- [62] X. Zhao et al., "Pre-trained Language Models can be Fully Zero-Shot Learners," arXiv, abs/2212.06950, 2022. [Online]. Available: <https://doi.org/10.48550/arXiv.2212.06950>.
- [63] Z. Zhang et al., "Automatic Chain of Thought Prompting in Large Language Models," arXiv, abs/2210.03493, 2022. [Online]. Available: <https://doi.org/10.48550/arXiv.2210.03493>.
- [64] L. Wang et al., "Plan-and-solve prompting: Improving zero-shot chain-of-thought reasoning by large language models," arXiv preprint arXiv:2305.04091, 2023.
- [65] R. L. Logan IV, et al., "Cutting down on prompts and parameters: Simple few-shot learning with language models," arXiv preprint arXiv:2106.13353, 2021.
- [66] Z. Wang et al., "Rolelm: Benchmarking, eliciting, and enhancing role-playing abilities of large language models," arXiv preprint arXiv:2310.00746, 2023.
- [67] Y. Shao, et al., "Character-llm: A trainable agent for role-playing," arXiv preprint arXiv:2310.10158, 2023.
- [68] L. Ouyang et al., "Training language models to follow instructions with human feedback," arXiv (Cornell University), Mar. 2022, doi: <https://doi.org/10.48550/arxiv.2203.02155>.
- [69] D. Rajagopal et al., "Cross-Domain Reasoning via Template Filling," ArXiv, abs/2111.00539, 2021.
- [70] Palli, Abdul Sattar, Jafreezal Jaafar, Manzoor Ahmed Hashmani, Heitor Murilo Gomes, Aeshah Alsughayyir, and Abdul Rehman Gilal. "Combined Effect of Concept Drift and Class Imbalance on Model Performance During Stream Classification." CMC-COMPUTERS MATERIALS & CONTINUA 75, no. 1 (2023): 1827-1845.
- [71] K. O'Byrne et al., "Client Assessment by Novice and Expert Psychologists: A Comparison of Strategies," Educational Psychology Review, vol. 9, no. 3, pp. 267-278, 1997. [Online]. Available: <https://doi.org/10.1023/A:1024739325390>.
- [72] J. Culbertson et al., "Cognitive Biases, Linguistic Universals, and Constraint-Based Grammar Learning," Topics in cognitive science, vol. 5, no. 3, pp. 392-424, 2013. [Online]. Available: <https://doi.org/10.1111/tops.12027>.
- [73] J. M. Carroll et al., "Creating a Design Science of Human-Computer Interaction," pp. 205-215, 1992. [Online]. Available: [https://doi.org/10.1016/0953-5438\(93\)90022-L](https://doi.org/10.1016/0953-5438(93)90022-L).
- [74] A. Patel et al., "A Systematic Approach to Evaluating Design Prompts in Supporting Experimental Design Research," Proceedings of the Design Society: International Conference on Engineering Design, 2019. [Online]. Available: <https://doi.org/10.1017/DSI.2019.282>.