DreamOn: Text-Based Chatbot for Empathic Responses to Dream Descriptions

Rebecca Chiara Q. De Veyra
College of Computer Studies
De La Salle University
Manila, Philippines

Reuben Seth G. Jovellana
College of Computer Studies
De La Salle University
Manila, Philippines

Alejandro Jose E. Morales
College of Computer Studies
De La Salle University
Manila, Philippines

Joaquin Andres D. Rodriguez
College of Computer Studies
De La Salle University
Manila, Philippines

Andrei Dominic O. Viguilla
College of Computer Studies
De La Salle University
Manila, Philippines

***Abstract*—This study investigates interface feedback mechanisms that users perceive as empathetic in dream journaling chatbots. Through a mixed-methods approach with 50 participants interacting with *DreamOn* - our text-based chatbot prototype - we analyzed how message style, visual cues, and interactivity features influence perceived empathy.**

***Index Terms*—empathic computing, dream journaling, chatbot design, human-computer interaction, emotional intelligence**

I. Introduction

II. Related Work

III. Methodology

This section shows the three-phase methodological approach we used in this study. The process involved participant recruitment, the iterative design of our DreamOn prototype, and a mixed-methods data collection strategy to evaluate the system and user perceptions of it.

A. Participant Recruitment

To understand user needs and expectations for the design of the DreamOn system, 21 respondents were recruited via convenience sampling to participate in a user perception survey. Due to this, most of the respondents are fellow college students.

B. Prototype Design

The main focus of this research is DreamOn, a prototype AI chatbot developed to investigate the role of affective feedback in the context of digital dream journaling. The development followed an iterative process. An initial version of the prototype was developed as a baseline, featuring the core chatbot functionality but without the adaptive color and music changes. This allowed for a focused evaluation of the aesthetic feedback in the current iteration. The system was designed to move beyond purely textual interaction and explore how non-verbal, aesthetic cues can shape the user's experience, specifically colors and music.

1) System Architecture and Interaction Flow: The DreamOn system is modeled around a straightforward flow designed to respond to the user's perceived emotional state. The process is initiated when a user inputs a textual description of their dream into the chatbot's interface. This text is then passed to a back-end emotion detection module for analysis. The module processes the narrative and classifies the user's latent emotion into one of six predefined categories. Upon successful classification, the system presents feedback to the application's user interface, dynamically altering the background color palette and playing a music track thematically aligned with the detected emotion. It is important to note that the current prototype's adaptive response mechanism is intentionally focused on these two modalities—visual aesthetics and auditory cues—to isolate and study their specific impact on user perception.

2) Emotion Detection and Classification: The core of the DreamOn system's intelligence lies in its text-oriented emotion recognition module. The initial model was developed using the Google GoEmotions dataset, a large, human-tagged collection of Reddit comments labeled into 27 emotion categories. The model was then fine-tuned to classify textual input into one of six basic emotions, as originally theorized by Ekman: Angry, Disgust, Fear, Joy, Sadness, and Surprise. To ensure applicability to a diverse user base, the model was created and evaluated using dreams written in both English and Tagalog, mirroring the bilingual environment of the targeted audience. This functionality enables the system to handle a broader variety of natural user inputs without limitations of language.

3) Design of Adaptive Aesthetic Feedback: The main objective of this research is to explore the potential of non-verbal cues to convey empathy, an area often less frequently considered than dialogue in HCI research. The design of DreamOn's adaptive feedback was based on concepts of color psychology and emotional connections. Each one of the six detectable emotions is mapped to a distinct interface theme, comprising a specific color palette and a matching musical genre. The mapping used is meant to validate, soothe, or otherwise appropriately respond to the user's emotional state.

TABLE I
MAPPINGS OF EMOTIONS TO INTERFACE THEMES

Detected Emotion	**Theme Name**	**Color Palette (Primary, Secondary, Accent)**	**Psychological Rationale**
Angry	Uplifting	#FF6B6B, #FFD166, #F7F9F7	The use of a red-family color acknowledges the strong association between red and anger found in both linguistic metaphors and empirical studies [1]. This is balanced with warm yellow, a color consistently linked to happiness and optimism, aiming to transition the user from a high-arousal negative state to a positive one [2].
Disgust	Ambient	#88D498, #C7F9CC, #22577A	While some shades of green can be associated with sickness and disgust [3], this theme uses lighter, fresher greens (seafoam, mint). These are psychologically linked to nature and freshness, aiming to evoke a sense of cleansing to counter the feeling of disgust, a concept supported by research showing that cleanliness can reduce the severity of moral judgments tied to disgust [4].
Fear	Emotional	#9B5DE5, #F15BB5, #00BBF9	Blue is widely documented in color psychology studies as being associated with tranquility, stability, and security, which can have a calming effect [5]. Research also indicates that lavender, a light purple, can reduce anxiety and promote feelings of security and comfort, directly countering the arousal of fear.
Joy	Uplifting	#FF9E6D, #FFEA00, #A0E7E5	Yellow is one of the most universally recognized colors for joy and optimism in color-mood association studies [2], [6]. The addition of peach introduces warmth and comfort, reinforcing the positive state.
Sadness	Emotional	#5B8FB9, #B6E2D3, #FADCD9	The choice of blue acknowledges the widespread cultural metaphor of "feeling blue," an association confirmed in perception studies [7]. However, blue is also associated with compassion and sincerity, which can serve to validate the user's emotion. The inclusion of soft seaglass (a seafoam green) introduces elements of healing and hope, drawing on the association of green with nature and renewal.
Surprise	Fantasy	#D77FA1, #BAABDA, #F6F5F5	Purple hues like mauve and lavender are often associated with mystery, magic, and imagination, which aligns with the sense of wonder that can accompany surprise. The use of pure white as a base provides a clean canvas, representing a new or unexpected situation, drawing on white's association with simplicity and clarity [6].

Table I provides a comprehensive explanation for these mappings of emotions to interfaces.

The choice of the music tracks was influenced by the same psychological principles. For example, the "Uplifting Theme" evoked by anger is paired with cheerful and hopeful music, whereas the "Emotional Theme" for sadness includes a more melancholic and soft soundscape. This combined method seeks to establish a comprehensive and engaging emotional atmosphere for the user.

C. Data Collection

The evaluation of the DreamOn concept was designed to address the research from two angles: how reliable it is in detecting the user's emotion and the user's perception of if they felt the application was empathetic towards them.

1) Research Question and Design: This study was guided by this research question: To what extent do adaptive auditory and visual feedback mechanisms, driven by automated emotion detection, influence the perception of empathy in a digital dream journaling chatbot? To answer this question, we approached it with these two stages:

- **Technical Validation:** A quantitative assessment of the emotion detection model's accuracy in classification utilizing a regulated dataset.
- **User Perception Survey:** A survey incorporating both quantitative (Likert scale, multiple-choice) and qualitative (open-ended) questions to evaluate user attitudes, preferences, and expectations regarding empathetic technology, prior to any interaction with the DreamOn prototype. This provides a crucial baseline of user needs.

2) Protocol for Technical Validation: The technical performance of the model was evaluated to establish its reliability. A tailored dataset of 30 short dreams was created specifically for this validation. The dataset was evenly distributed, featuring precisely five unique scenarios for each of the six target emotions (Anger, Disgust, Fear, Joy, Sadness, Surprise). To also cater to Filipino users, the narratives were composed in both English and Tagalog (Three English and two Tagalog narratives per emotion). The validation process involved feeding each of the 30 narratives into the DreamOn chatbot. The model's predicted emotion for each narrative was systematically recorded. The output was then compared against the actual emotion. This was done to get the model's classification accuracy, which was calculated for each of the six emotions individually and as an overall system average.

3) Protocol for User Perception Survey: An online survey was conducted to gather data on user perspectives before introducing them to the DreamOn prototype. The survey was organized into multiple sections:

a) Part 1: Experience and Importance: This section featured inquiries regarding the respondents' previous engagement with digital journaling or mood-tracking applications. It also employed a 5-point Likert scale to assess how important users felt an app was in providing emotional support when talking about personal subjects such as dreams.

b) Part 2: Feature Preferences: This section asked participants to select or rank their preferred feedback styles (e.g., "Calm and validating," "Friendly and casual") along with the UI components they thought would best represent their emotional state (e.g., "Background hue," "Mood music,"

TABLE II
CONFUSION MATRIX AND PER-EMOTION ACCURACY

True Emotion	Predicted						Accuracy
	Angry	Disgust	Fear	Joy	Sadness	Surprise	
Angry	5	0	0	0	0	0	100%
Disgust	0	3	1	0	1	0	60%
Fear	0	0	4	0	1	0	80%
Joy	0	0	0	5	0	0	100%
Sadness	0	0	0	0	5	0	100%
Surprise	0	0	0	1	0	4	80%

"Tone/style of system notifications"). Questions also directly evaluated the perceived influence of visuals and the attractiveness of adaptive features.

c) Part 3: Qualitative Probes: This section featured open-ended questions designed to gather deeper, more intricate information. Participants were requested to explain what contributes to a digital interaction feeling "robotic or dismissive" and to provide details on their positive or negative experiences with current self-reflection applications. Quantitative data from the survey were analyzed through descriptive statistics to see common characteristics. Qualitative data from the open-ended questions underwent thematic analysis, in which responses were coded and classified to reveal common patterns, themes, and important user insights.

IV. Results

This section showcases the results from the evaluation protocol consisting of two parts: the technical performance of the DreamOn system's emotion detection model, followed by the quantitative and qualitative results from the user perception survey.

A. Technical Validation: Emotion Detection Performance

The evaluation of the emotion detection model yielded a distinct understanding of its strengths and weaknesses, delivering essential context for its possible use in a user-facing application.

1) Overall System Accuracy: The emotion detection model attained an overall classification accuracy of 86.67% across the entire set of 30 dream narratives. This figure is derived from 26 accurate predictions made out of a total of 30 scenarios. This outcome suggests that the model is typically effective and operates consistently in most scenarios, offering a robust basis for a flexible system.

2) Per-Emotion Accuracy and Error Analysis: Although the overall accuracy is high, a more detailed examination shows notable differences in performance among the six emotion categories. The model's per-emotion accuracy and specific error patterns are detailed in the confusion matrix in Table II.

The analysis highlights several key findings:

- **High Performance:** The model demonstrated perfect accuracy (100%) for the emotions of Joy (5/5 correct), Sadness (5/5 correct), and Angry (5/5 correct). This suggests a high degree of reliability in identifying these basic emotional states.
- **Moderate Performance:** The model performed well, but not perfectly, for Fear (4/5, 80% accuracy) and Surprise (4/5, 80% accuracy).
- **Poor Performance:** The model struggled most with Disgust, achieving only 60% accuracy (3/5 correct). This category represents the model's primary weakness.

An in-depth analysis of the misclassifications reveals important insights into the model's distinct error patterns:

- One scenario intended to evoke Disgust ("*I dreamt that I was eating my favorite meal when I discovered a dead cockroach...*") was misclassified as Fear.
- Another Disgust scenario, written in Tagalog ("*Tumapak ako sa may kanal na puno ng basura...*"), was misclassified as Sadness.
- One Fear scenario ("*May humila sa akin pababa habang naliligo sa ilog...*") was misclassified as Sadness.
- One Surprise scenario ("*I dreamt that I opened my front door and found a long-lost friend...*") was misclassified as Joy.

These specific errors, particularly the confusion between semantically adjacent but experientially distinct emotions like Disgust/Sadness and Surprise/Joy, are critical for understanding the potential user experience implications.

TABLE III
QUANTITATIVE USER PREFERENCE SURVEY RESULTS (N=21)

Survey Question / Topic	Response Distribution (Count & %)
How important is emotionally supportive feedback?	Very important (5): 3 (14.3%) Important (4): 6 (28.6%) Moderately important (3): 4 (19.0%) Slightly important (2): 4 (19.0%) Not important (1): 4 (19.0%)
Do visuals affect how emotionally supported you feel?	Very much (5): 8 (38.1%) Quite a bit (4): 6 (28.6%) Somewhat (3): 4 (19.0%) Slightly (2): 3 (14.3%) Not at all (1): 0 (0%)
Would you prefer if the app included calming music?	Yes: 16 (76.2%) No: 5 (23.8%)
How likely are you to use an app that adjusts its mood?	Very likely (5): 6 (28.6%) Likely (4): 7 (33.3%) Moderately likely (3): 5 (23.8%) Slightly likely (2): 1 (4.8%) Not likely at all (1): 2 (9.5%)
Which feedback style feels most supportive? (Multiple selections allowed)	Calm and validating: 14 Friendly and casual: 11 Professional and neutral: 6 I don't want feedback: 2

B. User Perception Survey Findings

The survey of 21 potential users provided a rich dataset of attitudes and preferences, establishing a clear user-centered foundation for the design of empathetic digital tools.

1) Quantitative Analysis of User Preferences: The quantitative data from the survey reveal a strong user appetite for emotionally aware and adaptive features in applications for personal reflection. A summary of key responses is presented in Table III.

The trends seen from this data were:

- **Importance of Support:** A notable share of users (45.4%) categorized emotional support as "Important" or "Very important." Although there was a range throughout the spectrum, the desire for support stands out as a significant inclination.
- **Impact of Visuals:** A significant majority of users (68.2%) reported that visuals influence their feelings of emotional support either "Quite a bit" or "Very much." No participant believed that visuals had no impact whatsoever.
- **Preference for Music:** A vast majority (77.3%) indicated that they would like a dream-sharing app to feature soothing music or sounds.
- **Desire for Adaptivity:** A majority of users are open to an adaptable interface, as 63.7% report being "Likely" or "Very likely" to engage with an app that modifies its atmosphere according to their dream journaling.
- **Preferred Style:** When discussing feedback styles, "Calm and validating" was the most commonly chosen option, with "Friendly and casual" as a close second, showing a preference for supportive and affirming communication rather than neutral or formal approaches.

2) Qualitative Themes of Digital Empathy and Dismissiveness: The thematic analysis of open-ended responses offered greater insight into the quantitative results, highlighting how users differentiate between empathetic and robotic elements in digital interactions.

a) Theme 1: Hallmarks of a "Robotic" Interaction: When those who had previously considered an app's response to be dismissive were asked about the reasons for this sentiment, a distinct and consistent theme appeared, focused on a perceived absence of authenticity and personalization. The subtopics consist of:

- **Generic and Canned Responses:** This was the chief subtheme. Users highlighted "clearly pre-prepared/canned replies," a "clear structure/pattern that AI typically uses," and communications you "would immediately recognize... were crafted to convey that." This implies that predictability and the absence of novelty are key indicators of a robotic interaction.
- **Tonal Incongruence:** A significant failure point highlighted by one user occurred when the "Tone is inconsistent with the present mood." This underscores the significance of not only grasping the content but also aligning with the emotional tone of the interaction. A response that doesn't match the tone, even if sincere, can seem shocking and dismissive.
- **Perceived Lack of Understanding:** Responses were characterized as "official and devoid of understanding" and exhibited a "fairly neutral" vocal tone. In an emotional setting, this neutrality was viewed not as impartial but as a deficiency in involvement and compassion.

b) Theme 2: Desired Features for Supportive Reflection: When asked about their likes and dislikes regarding current journaling and mindfulness applications, participants' answers aligned with themes of interactivity, design, and effectiveness:

- **Dislike of Passivity:** A major source of dissatisfaction was the "absence of feedback" from applications that acted solely as inactive storage for ideas. Users indicated a wish for increased interaction beyond mere data input.
- **Value of Aesthetics:** The sensory experience was greatly appreciated. A user felt discouraged by the "limited design options or its simplicity," suggesting that visual appeal influences ongoing participation. In contrast, a different user commended a meditation application for the "narrator's voice being calm and soothing," emphasizing the benefits of quality audio feedback.
- **Tension in Simplicity:** A significant tension was observed between the wish for efficiency and the requirement for captivating design. One user appreciated a "minimalistic" design that allows for "quick journaling without excessive features," while another considered the "basic-ness" to be discouraging. This indicates that a delicate equilibrium needs to be achieved between an efficient user experience and one that is visually and emotionally appealing.

V. Discussion

VI. Conclusion and Future Work

References

[1] A. K. Fetterman, M. D. Robinson, and B. P. Meier, "Anger as 'seeing red': evidence for a perceptual association," *Cognit. Emot.*, vol. 26, no. 8, pp. 1480–1489, 2012.

[2] L. B. Wexner, "The degree to which colors (hues) are associated with mood-tones," *J. Appl. Psychol.*, vol. 38, no. 6, p. 432, 1954.

[3] M. Berthold, and R. Ammann, "How Colours Affect Us," in *Knowledge and Space*, vol. 17, Springer, Cham, 2022, pp. 247-253.

[4] S. Schnall, J. Benton, and S. Harvey, "With a clean conscience: Cleanliness reduces the severity of moral judgments," *Psychol. Sci.*, vol. 19, no. 12, pp. 1219–1222, 2008.

[5] R. M. Gerard, "Differential effects of colored lights on psychophysiological functions," Ph.D. dissertation, Univ. of California, Los Angeles, CA, USA, 1958.

[6] N. Kaya, and H. H. Epps, "Relationship between color and emotion: A study of college students," in *Proc. Centennial Conf. Center Human-Environ.*, 2004, pp. 1–7.

[7] C. A. Thorstenson, A. D. Pazda, and A. J. Elliot, "Sadness and synesthesia: The impact of emotion on color-grapheme pairings," *Conscious. Cogn.*, vol. 33, pp. 384–391, 2015.