2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

Blinking Into Emotion: How Context and LED Frequency Shape Non-Humanoid Robots' Emotional Transparency

Patrizio Scarpato[1], Luca Raggioli[2], Raffaella Esposito[3], Silvia Rossi[1,3]

Abstract— **In social contexts, correctly interpreting communicative signals is essential for understanding an interlocutor's reactions. Emotional cues provide important context, enhancing mutual understanding and enabling more natural, adaptive interactions. For non-humanoid robots, which typically have more limited interaction capabilities, it could be necessary to combine multiple non-verbal signals and to consider the context in which they are used. In this work, we present the design of non-verbal behaviors that enable a non-humanoid robot to communicate its intended emotional state effectively. Our strategy focuses on the use of LEDs as non-verbal cues along with facial expressions for conveying emotions. We conducted a user study to evaluate the effect of these channels with and without context. Our results show that the robot conveys emotions more transparently when context is included, and that blinking LEDs can be an effective channel for communicating emotion. Results suggest that blinking alone is a minimal but functional cue, with richer models performing better without context. When short contextual sentences and more spaced blink frequencies are added, the blinking-only condition performs on par with, or even better than, the full multimodal model for some emotions and, in particular, with respect to participants' perceived arousal. This indicates that carefully designed, simple visual cues can be an effective affect channel for non-humanoid robots.**

I. Introduction

Transparent display of emotions is an essential capability for robots to engage effectively with humans [1], in both cooperative and competitive human-robot interactions (HRI) [2]. In fact, social robots capable of non-verbal emotional expression tend to be more readily accepted by users, facilitating stronger bonds and promoting more natural, transparent, and efficient interactions.

The design of emotional behaviors in robots often draws upon patterns observed in human emotional expression and recognition. Among the various communication modalities investigated in the literature, non-verbal signals such as facial expressions, vocal tone and gestures, have been shown to play a more prominent role in emotional perception than verbal signals, due to their directness and intuitiveness [3], [4]. However, robots often face limitations in facial expressiveness and verbal capabilities, which can limit their ability to convey emotions clearly and effectively [5]. This is the case of robots such as NAO and Pepper, characterized by a rigid plastic face, and limited speech recognition and text-to-speech vocabulary. In the case of non-humanoid robots, such limitations can be further exacerbated by the complete lack of a face, therefore making non-verbal behaviors even more crucial for enabling the recognition and interpretation of others' emotions and intentions [6]. Regardless of the expressiveness of the robot's features, it is also important to consider the impact of contextual information on emotion recognition. In fact, it has been shown how the robot's social context and interaction goals are crucial in the design of effective emotional display.

In this work, we build upon findings in the literature, investigating the use of emotional expressiveness for humanoid robots (e.g., based on LED colors, movements, and facial expressions [7]), to address the display of emotions in non-humanoid robots, characterized by fewer degrees of freedom and without facial features, to propose an effective emotional communication strategy for this class of robots. Specifically, supported by evidence that interactive or blinking lighting systems can be used to elicit emotional and physiological responses in users [8]–[13], we investigated how endowing a non-humanoid robot with colored LEDs and manipulating the blinking frequency can affect the participants' perception of the robot's intended emotion. Moreover, we further assessed the effects of this communication channel by combining it with contextual information that would help disambiguate the internal state of the robot. To this end, we conducted two user studies focusing on the following research questions:

RQ1: Is LED blinking an effective way to convey emotions in non-humanoid robots?

RQ2: Does the context significantly affect the user's perception of the emotions conveyed by the robot?

II. Background and Related Work

Emotional expression can be very useful to enhance social acceptance in HRI [14]. The implementation of behaviors that express emotions in robots promotes higher trust in them [15], and can also reduce stress and stimulate participants' engagement (e.g., interacting or cooperating while playing games [16]). Different communication modalities can be used by robots to express emotional state cues, both verbal (explicit) and non-verbal (implicit), with varying degrees of accuracy [17]. For example, robots with a wide range of motion and the ability to use their voice to communicate can effectively express happiness and fear [17]. However, due to the physical and conversational limitations of robots, many of the cues used by humans cannot be exploited [18]. Characterizing the robots' emotional behavior to combine different cues requires careful design [19], especially in cases where the robot's facial expressivity is limited (e.g., with

[1]Department of Electrical Engineering and Information Technology, University of Naples Federico II, Naples, Italy; [2]Department of Neurosciences and Reproductive and Odontostomatological Sciences; [3]ICAROS center, University of Naples Federico II, Naples, Italy.

Fig. 1. Temi Robot

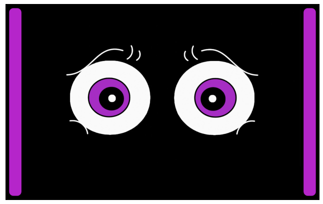

Fig. 2. Fear expression view

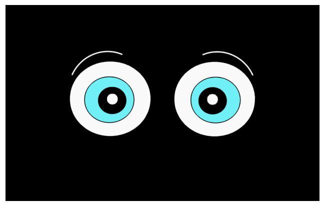

Fig. 3. Neutral expression view

NAO and Pepper) and its body language becomes one of the most immediate channels of communication [14]. For this reason, numerous studies aim to identify the set of non-verbal behaviors that promote more natural and effective human-robot interaction [20], including eyes gaze [21], proxemics, posture, gestures [22], head position [23], and para-verbal sounds [24]. The use of body movements and cues that do not rely on facial expressions is further important in the case of non-humanoid robots, which can be used to enhance emotional expressivity (e.g., wiggling the robot's ears [25]). In the case of non-humanoids, colors, and sounds can also be useful signals to communicate emotions [26]. A non-verbal cue that was observed to accurately convey emotions is the blinking of a visual element, which can facilitate the interpretation of the intensity of the emotion [13], [27].

While robot communication cues are widely considered to be fundamental for humans to understand their emotional state, the context can sometimes be even more important for effective emotion recognition and attitude towards robots [28], [29]. In this work, we will investigate different combinations of the studied communication signals, together with the introduction of a context, to assess how they affect participants' perception of the robot's emotions.

III. Non-Verbal Behaviors Model

In this work, we used the non-humanoid robot Temi[1] (Figure 1). It is 100 cm tall and 45 by 35 cm wide. It has a 10.1-inch HD display, which serves as the main interface for visual interaction, and can be tilted to adjust its angle between 55 and -25 degrees.

We primarily focused on humans' self-reported level of arousal, which we intend to study with respect to the blinking frequency of colored LEDs. In this work, arousal refers exclusively to participants' *self-reported perceived arousal*, and not to directly measured physiological activation. Since Temi does not have any physical LEDs, in the layout of the application, we used two lateral components (see Figure 2), which we treat as LEDs. In order to more thoroughly assess the effect of this communication channel, we also implemented facial expressions, in the form of animated eyes and eyebrows (see Figure 5), and complemented them with paraverbal sounds and movements.

The emotions chosen are the ones described by Ekman's [30], while we adopted the *valence-arousal* model of emotions described by Mehrabian et al. [31]. This model is particularly relevant for our works, as we intend to investigate the effect of the blinking frequency on the participants' subjective self-reported arousal, driven by supporting evidence in the literature on the effect of blinking lights [8], [12], [13]. Moreover, it has also been shown that the presence of a robot's expressions, often based on the use of simulated eyes (e.g., cartoon-like eyes [32]), can be crucial to convey intentions and affects the user's perception, trust and acceptance of robots [10], [11], [33]. For this reason, we designed facial expressions to complement the blinking strategy and to investigate our model more thoroughly.

Below, we describe the two main communication strategies implemented: blinking and facial expressions.

A. Blinking

The robot's tablet is used to display the blinking LEDs. Specifically, the colors used for the LEDs are based on Plutchik's theory of emotions [34]: yellow for happiness, red for anger, green for disgust, purple for fear, dark blue for sadness, and white for surprise (see Table I). Several studies have highlighted the difficulty of finding effective solutions for representing the arousal associated with a particular emotional state. For this reason, we decided to investigate the direct relationship between blink frequency and the arousal associated with a particular emotion, based on observations from the literature [13]. We use this communication channel to promote greater transparency in the emotional display by establishing a direct proportional relationship between the arousal value for the specific emotion and the associated blink frequency. The frequencies were chosen empirically based on trends in the literature, but also with the objective of not being uncomfortable or harmful to the participants [35]. Moreover, we also observed that frequencies higher than 30Hz would result in a less smooth or noticeable flickering of the visual elements (it is to be considered that we are using visual elements on the layout of the Android application, not physical LEDs). We chose, therefore, to use frequencies that were all lower than 10Hz, reported in Table I.

B. Facial Expressions

The blinking LEDs were complemented with facial expressions, which were also displayed on the robot's tablet. Both the appearance and the animation of the facial expressions were designed from the ground up in Blender[2] (Figures 4 and 5), and were based on the Omate Yumi robot and the Classmate robot, as they are perceived as more friendly and likable [36]. The specific facial features were chosen based

[1]https://www.robotemi.com/product/temi/

[2]https://www.blender.org

TABLE I

SIGNALS USED TO CHARACTERIZE THE NON-VERBAL BEHAVIORS OF NON-HUMANOID ROBOTS.

	Happiness	Sadness	Anger	Disgust	Fear	Surprise
Blinking Frequencies	2Hz	1Hz	2.5Hz	0.83Hz	3.33Hz	2Hz
Color	Yellow	Blue	Red	Green	Purple	White
Tilt Angle	50°	-50°	-10°	45°	-53°	20°
Base Angle	0°	0°	0°	30°	-10°	0°
Para-Verbal Sound	"Yeah!"	*Sigh*	"Grr!", *Growl*	"Bleah", "Yuck"	"Ah!"	"Wow"

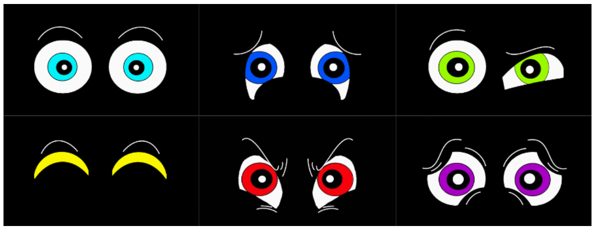

Fig. 4. Facial expression of for the emotions. From left to right, top to bottom: surprise, sadness, disgust, happiness, anger, and fear

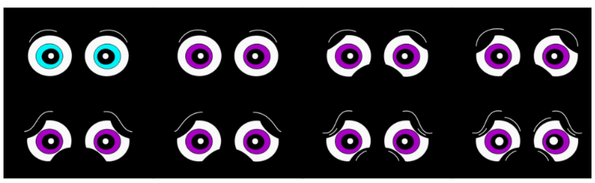

Fig. 5. Progression of expressions of the fear emotion

on Kalegina et al. [37], which analyzed 157 rendered robot faces to understand how different facial features affect human perception. Moreover, eyes and animations were drawn based on established principles of anime and manga character design [38]. Finally, when the robot is not showing any emotion (Figure 3), its eyes blink periodically, to make its face appear more lifelike.

C. Full model of Emotions

Simple movements and para-verbal sounds were combined with the blinking and facial expression channels, to rely on a wider set of signals that have been shown to effectively convey emotion in non-verbal settings [7]. In particular, the movements were implemented by empirically manipulating the tilt angle of the robot's tablet (used to display the eyes) and rotating the robot around its vertical axis using the wheels in its base (see Table I). This characterization is based on findings in the literature, showing that open and fast movements generally convey more positive emotions, while closed and slow movements are perceived as unpleasant [39], and on observations of dogs' behaviors [40]. The para-verbal sounds are also implemented based on the behaviors observed in animals [40], and based on the para-verbal sounds available in the Pepper robot's library (see Table I). Each emotion is expressed through a *holistic bundle* of coordinated cues, including blinking frequency, color, facial animation, para-verbal sounds, and motion. The present study evaluates these bundles as integrated emotional displays, rather than isolating the contribution of individual channels.

IV. EXPERIMENTAL PROCEDURE

We investigated the effectiveness of the model of emotional behaviors described in Section III through two online user studies using questionnaires implemented in Google Forms. Initially, participants provided demographic information (gender, age, and nationality). The participants were shown 6 videos of the robot expressing emotions selected for each condition. For each, the participants were asked to evaluate the perceived emotion, using 5-point Likert scales Self-Assessment Manikin (SAM), measuring valence (from 1="Unpleasant" to 5="Pleasant") and arousal (from 1="Calm" to 5="Excited"). Participants were not provided with explicit emotion labels at any point during the study, while the order of emotional stimuli was randomized. It is important to note that the second User Study builds upon the insights of the first. Specifically, limitations observed regarding the blinking frequencies exposed the need for more pronounced differences between the different emotions. For this work, approval for experimental procedures and protocols was granted by the Ethical Review Committee of the University of Naples Federico II.

V. RESULTS

In this work, emotional transparency is defined as the degree to which participants' perceived emotion matches the robot's intended emotion, operationalized as the distance between participants' ratings and target coordinates in valence–arousal space.

A. User Study I

In this user study, we aim to assess the effectiveness of the blinking LEDs in conveying the arousal associated with each of the chosen emotions. Moreover, we seek to investigate how blinking combined with facial expressions and with a wider set of non-verbal stimuli affects the participants' perception of the robot's emotion. The experimental conditions are obtained based on the communication channels used: **C1**) only blinking; **C2**) blinking and facial expressions; **C3**) full model (blinking, facial expressions, para-verbal sounds, and movements).

We recruited 68 participants (24 females, 44 males), aged between 19 and 65 (mean: 28, std: 7). Participants were evenly distributed across the three experimental conditions (C1–C3), with each participant assigned to a single condition. Most of the participants were Italians (58), the rest of the

TABLE II
VALENCE (V) AND AROUSAL (A) AVERAGE RATINGS ACROSS EMOTIONS AND MODEL CONFIGURATIONS.

	Happiness		**Sadness**		**Anger**		**Fear**		**Disgust**		**Surprise**	
	V	A	V	A	V	A	V	A	V	A	V	A
Baseline	4.52	2.92	1.74	2.08	2.14	3.68	1.72	3.4	1.8	2.4	3.8	3.68
Blinking	3.11	2.31	3.26	2.23	2.37	3.03	2.94	3.0	3.69	2.0	2.86	2.57
Blinking & Face	3.63	2.51	2.29	2.43	2.2	3.23	2.6	2.91	3.03	2.71	2.91	2.69
Full model	3.73	3.09	2.42	2.33	2.61	3.27	2.55	2.91	2.64	2.67	3.39	2.79

nationalities were African-American, Greek, Indian, Russian, Portuguese, and Spanish.

1) Distribution of Emotions in the Cartesian Space: In the dimensional model of emotions (valence-arousal), each emotion can be mapped to a specific area of the Cartesian space [41]–[43]. The values identified by Russell and Mehrabian [41] are used to obtain baseline coordinates for this study. In particular, since the valence-arousal values recorded were both coded with the ranges 1-5 for simplicity, here, we mapped the valence and arousal ranges, [-1, 1] and [0, 1] respectively, to the range [1, 5]. This linear mapping preserves relative distances between emotions while matching the questionnaire scale. Although alternative nonlinear mappings are possible, this choice does not affect relative comparisons between models. As we can see in Figure 6 and Table II, across the six emotions, evaluations of the robot's emotional state using only the blinking channel (blue dots) are in general the most far away from the expected estimates of the baseline (red diamonds). The estimates obtained with full multimodal model (green), on the other hand, almost always lies closest to the baseline, improving the consistency of the evaluation of the arousal for joy and surprise, expressing the intended negativity for sadness and disgust, and tempering the intensity of anger and fear. Overall, with the progressive introduction of communication channels, the predictions are nudged towards the expected baseline target. These observation are made by considering the Euclidean distance between the baseline and each of the different models. However, if we consider the valence-arousal, dimensions individually, this is not always the case. In fact, we can see how for Sadness and Fear, in the arousal dimension, the blinking coordinate is actually closer to the baseline compared to the others. Similarly, for the valence dimension, for Anger, the blinking is closer to the baseline compared to the full model. This shows that, while the Euclidean distances provide a general view on how close the overall perception of the emotions were in each model, by considering the distributions relatively to the dimensions individually, we can better appreciate how the subjective arousal perceived with the blinking model does not dramatically deteriorate compared to the richer models.

Observing the results of the Kruskal–Wallis ANOVA, in Table III, which compares distances between participants' ratings and the baseline values for the three models, we see no significant differences emerging between Blinking, Blinking & Face, and Full model for happiness, sadness, fear, surprise (all $p > .05$), and anger is the only emotion which exhibits a result that is close to statistical significance ($\chi^2 = 4.59, p = .054$). The only exception is disgust, where there is significance ($\chi^2 = 10.13, p = .006$), and post-hoc tests identify a significant difference in the "Blinking versus Full model" pair ($W = 4.39, p = .005$). In the absence of context, blinking is a minimal but functional signal: it is not systematically worse than richer modalities for most emotions (there is no evidence of differences), but it tends to dampen perceived intensity and shows a slight loss of accuracy, especially on disgust, where a difference is observed compared to the Full model. These results are mostly in line with the observations on the distribution of the emotions in the cartesian space, showing that while less accurate as whole, the blinking model would still be localized very close to the other models for Fear, Surprise and Anger, and behaving generally better than the others for the arousal dimension for Sadness and Fear.

2) Evaluation of the spatial relationship between coordinates: In order to evaluate in an intuitive and readable way the participants' perception the robot's emotion, we introduce another metric to measure how well they rated each emotion compared to their expected emotional profile. This metric is based on the decaying neighborhood kernel (see Equation 1), which is typically used in Self-Organizing Maps to control how strongly neighboring neurons are updated, and provides a measure of the activation response of each node [44]. While Euclidean distance between average ratings and target coordinates provides a global measure of emotional transparency, it does not capture how consistently individual participants perceive the intended emotion. Two models may yield similar mean distances while differing substantially in the dispersion of individual responses. To account for this, we introduce a decaying neighborhood kernel that assigns higher activation to ratings closer to the target emotion and

TABLE III
KRUSKAL–WALLIS TEST RESULTS FOR THE **NO-CONTEXT** CONDITION (STUDY I).

Emotion	χ^2	df	p
Happiness	0.133	2	.936
Sadness	4.546	2	.103
Anger	4.593	2	.101
Fear	4.089	2	.129
Disgust	10.127	2	.006**
Surprise	0.193	2	.908

** $p < .01$ (two-sided),$H_A : \mu_{\text{Blink}} \neq \mu_{\text{Full}}$.

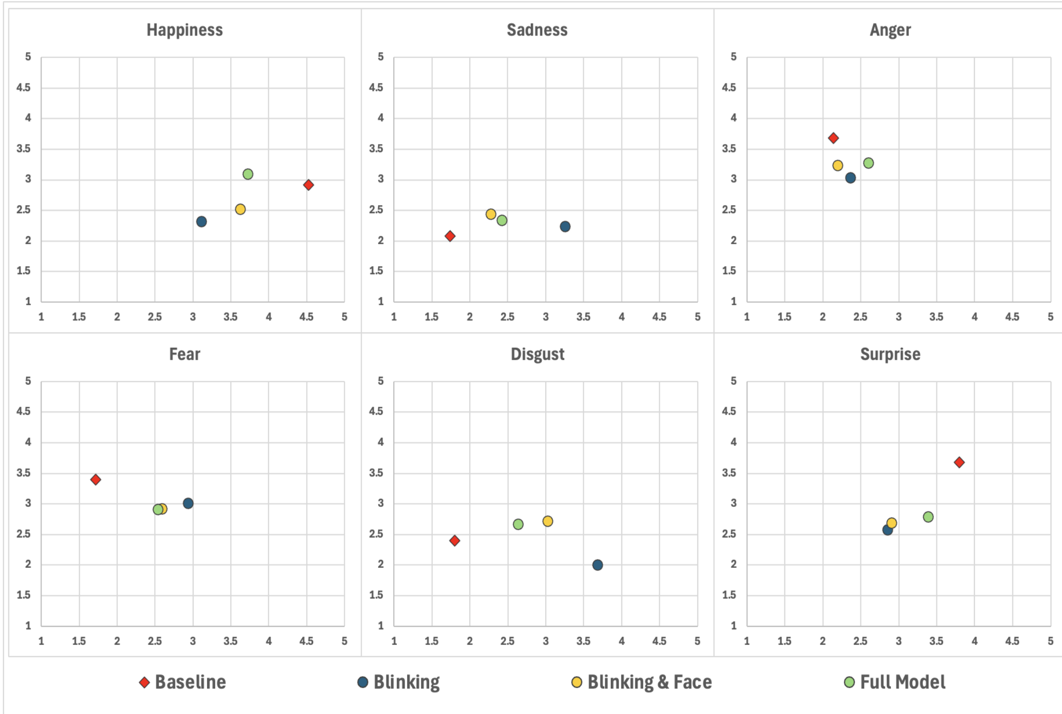

Fig. 6. Average distribution of the emotions in the valence-arousal model. The x-axis represents valence and the y-axis represents arousal.

progressively penalizes distant ones, allowing us to assess both accuracy and consistency at the participant level.

$$y_i = e^{\frac{-\beta}{2\sqrt{n}}} \tag{1}$$

The main idea is to obtain a measure that takes into account the topological relationship between the coordinates of the valence-arousal ratings and the expected baseline values: the closer the coordinates, the higher the activation estimate (from 0 to 1; overlapping coordinates correspond to an activation of 1). This metric accounts for individual participant ratings, complementing analyses based on average valence–arousal coordinates (as in Figure 6; see also Table II). We use this function for each of the rating of the participants, and then compute the average to have an aggregated measure: $A_e = \sum y_i$, where e corresponds to the individual emotion. The results are shown in Figure 7. We can see how, as it can be expected given the inferior number of channels used, the Blinking model has in general the worse performance compared to the other models. However, it quite consistently rates above 0.80 (except for *Disgust*, with 0.77), and exhibits the highest rate among the three models for *Anger* and for *Surprise* (equally to the Full model). The Blinking & Face model outperforms the others for *Happiness*, *Sadness* and *Fear*, while the Full model has its highest rates with *Disgust*.

The aggregated estimates for each models are respectively 0.81 for Blinking, 0.83 for Blinking & Face, and 0.82 for the Full model. By integrating these results with the visual distribution in Figure 6, we can better appreciate how especially the Blinking & Face model and the Full model are perceived often very similarly by the participants.

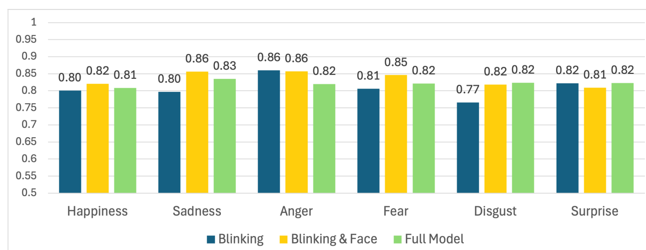

Fig. 7. Evaluation of the spacial relationship between coordinates for each model, measuring how close the participants' ratings were compared to the baseline values. The aggregated estimates for each models are: Blinking=0.81, Blinking & Face=0.83, and Full model=0.82.

B. User Study II

In the second study of this work, we addressed the limitations of the previous one by employing a new set of more pronounced blinking frequencies (see Table IV), chosen empirically based also on the observations from [45]. The primary focus of this second study was to assess whether providing verbal context before the robot displays the emotions (i.e., a sentence describing a situation) improves the participants' perception of the robot's arousal (see Table V). Importantly, the sentences were all delivered in a neutral way to not express any emotional tone that might override the blinking signal. For this second study, we recruited 64 participants (25 females, 38 males, 1 non-binary), aged between 21 and 49 (mean: 29, std: 6).

1) Distribution of Emotions in the Cartesian Space: Observing the arousal values rated by the participants for each emotion (see Figure 8), we can see that the Blinking-

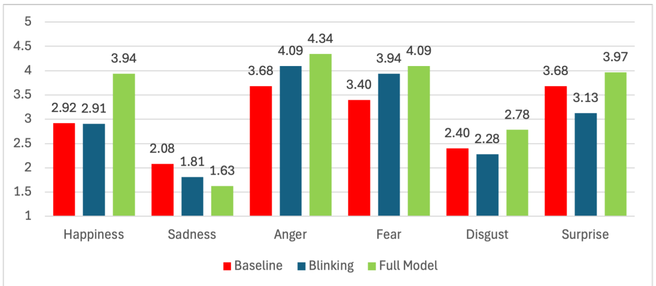

Fig. 8. Average arousal values rated by the participants for each emotion expressed with the models relying only on *Blinking*, and using the *Full model*. Baseline arousal values from [41] are shown for reference.

only model consistently achieved arousal ratings closer to the baseline than the Full model. This is a somewhat surprising result, as we would expect the more complex full model to more precisely capture the target emotional profiles. This suggests that while multimodal cues may facilitate categorical emotion recognition, they may interfere with the perception of arousal, where a single well-calibrated signal could be more effective. Despite its simplicity, blinking alone approximated expected arousal levels for several high-arousal emotions (e.g., anger and fear), while lower-arousal emotions such as sadness and disgust remained underestimated. This suggests that participants could indeed perceive emotional contrasts based solely on how the LEDs behaved (e.g., anger and fear were rated nearly equally to their expected values), and that emotions which are naturally associated with more intense physiological responses, are more readily conveyed through a combination of visual and behavioral cues.

In Table VI we show the results of the Mann–Whitney U tests, comparing distances between participants' ratings and the baseline values for Blinking and Full model. Significant differences emerged for happiness ($U = 289, p = 0.002$) and sadness ($U = 343, p = 0.013$), while no reliable differences were found for anger ($U = 470, p = 0.543$), fear ($U = 493, p = 0.789$) and surprise ($U = 398, p = 0.110$), and a borderline trend for disgust ($U = 377, p = 0.059$). Thus, with contextual framing, we find no reliable advantage of the multimodal model over blinking for most emotions; for happiness and sadness, blinking is closer to the target (significant), while other differences are minor or non-significant.

TABLE IV

MORE PRONOUNCED BLINKING FREQUENCY ASSOCIATED WITH EACH EMOTION TO ASSESS HOW TO BETTER EXPRESS THE EMOTION.

	Happiness	Sadness	Anger	Disgust	Fear	Surprise
Frequencies	3.33Hz	0.9Hz	4Hz	0.7Hz	6.7Hz	3.33Hz

TABLE V

CONTEXTUAL SENTENCES UTTERED BEFORE SHOWING THE ROBOT'S EMOTIONAL BEHAVIOR.

	Contextual Sentence
Happiness	"Hey Temi, you were really helpful to me earlier!"
Sadness	"Hi Temi, it looks like we won't be able to play together today..."
Fear	"Hey Temi, there's a rumor that some thieves are trying to break into the lab."
Disgust	"Temi, do you also smell that bad odor?"
Surprise	"Hi Temi, I heard that some of your friends are going on a mission to the moon!"
Anger	"Temi, they were teasing your friend Pepper earlier."

TABLE VI

MANN–WHITNEY U TEST RESULTS COMPARING THE **BLINKING** AND **FULL MODEL** CONDITIONS WITH CONTEXTUAL FRAMING (STUDY II).

Emotion	U	p
Happiness	289	.002**
Sadness	343	.013*
Anger	470	.543
Fear	493	.789
Disgust	377	.059†
Surprise	398	.110

* $p < .05$, ** $p < .01$, $H_A : \mu_{\text{Blink}} \neq \mu_{\text{Full}}$.

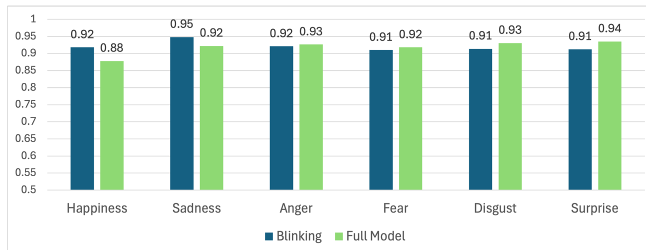

Fig. 9. Evaluation of how close the participants' ratings were compared to the baseline values. The aggregated estimates are 0.921 and 0.918 for the Blinking and the Full model respectively, using the new blinking frequencies and the contextual information.

Due to the increased differentiation in the blinking signals, it is not possible to directly draw comparisons with the results of the first study. However, it is also clear how the results in this study are overall in stark contrast with the observations of the previous one. In fact, the arousal rates recorded demonstrate that the contextual information provided in the form of sentences manages to communicate enough background to the participants for them to more correctly interpret the signals. Moreover, we believe the new frequencies adopted are more clearly distinguishable and allow for a different perception of the participants.

2) Evaluation of the spatial relationship between coordinates: Similarly to the first study, we considered an evaluation of how close the participants' ratings were compared to the baseline values. In Figure 9, we show the results obtained for the Blinking model and the Full model, both with the new frequencies and the contextual information. We can clearly see how the new estimates are significantly improved, with all emotions (except Happiness with the Full model) being rated above 0.9. Interestingly, we can see how Happiness and Sadness are rated higher in the Blinking model compared to the Full model, which is also reflected in the results in Figure 8, where the expected baselines for Happiness and Sadness are closer to the Blinking model's average. This is instead not observed for the rest of the emotions, where the Full model outperforms the Blinking model also in the case of emotions where the latter clearly has closer arousal values to the baseline (i.e., Anger and Fear). The reason for this can be twofold. On the quantitative side, as in the first study, it could be due to higher variance in the Blinking ratings, which is not reflected in the average arousals in Figure 8, but has an impact on the individual estimates of spatial closeness calculated. On the qualitative side, the other channels employed in the Full model might be interfering with the perception of the blinking signals, which are therefore sensed more consistently when presented alone. The aggregated estimates are 0.921 and 0.918 for the Blinking and the Full model, respectively.

C. Discussion

Because multiple expressive parameters vary together within each emotional behavior, conclusions regarding blinking frequency should be interpreted as referring to its role within an integrated expressive bundle rather than as an isolated arousal cue. Disentangling the contribution of individual channels is left to future work.

Overall, the first study shows that blinking LEDs can support basic emotional transparency in a non-humanoid robot (**RQ1**), albeit with limitations. In the absence of contextual information, richer multimodal behaviors were generally rated closer to the expected emotional profiles. The apparent discrepancy between distance-based and kernel-based metrics for the Blinking model can be attributed to higher variance in participants' valence–arousal ratings, likely caused by insufficient differentiation between blinking frequencies. In richer models, other channels appear to override this limitation. In this work, our focus on dimensional ratings reflects the goal of assessing emotional intensity and transparency rather than discrete labeling, while mapping ratings to semantic emotion categories is left to future work.

In Study II, contextual framing and revised blinking frequencies were introduced simultaneously, therefore, their individual effects cannot be disentangled, and improvements should be interpreted as arising from their combined influence. Together, clearer contextual cues and wider frequency spacing compensate for the expressivity limits observed in Study I, supporting the role of context as a key factor for transparent emotion signaling even with a minimal LED modality (**RQ2**). Notably, Blinking and Full models were often rated similarly, and in some cases blinking alone outperformed the Full model (e.g., happiness and sadness), suggesting that once context and frequency separation are sufficient, additional channels do not necessarily improve arousal accuracy and may sometimes interfere.

The LEDs used in this work were simulated on the robot's display, thus, their perceptual impact may differ from that of physical LEDs due to display characteristics and viewing conditions. Future work will validate these findings using external physical LEDs. Importantly, no participants reported discomfort or adverse effects related to blinking frequencies. While arousal was measured via self-reports in online video-based studies, an established and controlled approach for evaluating emotional expressiveness, future in-person experiments with physiological measurements will allow deeper investigation of emotional engagement and ecological validity [46]–[48]. For example, a possible in-person setting could be a storytelling scenario, in which the robot has to appropriately convey its emotion in accordance to the story and the user's reactions. Overall, these results support blinking colored LEDs as a viable and intelligible modality for conveying intended emotional arousal.

VI. Conclusions

This work introduces an affective communication strategy for a non-humanoid robot and implements it on the Temi platform, rendering blinking colored LEDs and facial animations on the tablet. The main aim was to assess the effectiveness of the blinking channel in conveying transparently the emotions. We evaluated the approach in two studies, where we investigated how different non-verbal channels (LEDs only; LEDs + face; Full model with LEDs, face, sounds and motion) convey emotions without contextual framing, then we assessed the impact of adding a brief contextual information to better disambiguate the robot's reaction. In the second study, moreover, we proposed a different blinking frequency to facilitate recognition and arousal mapping.

The results obtained showed that, in lack of contextual disambiguation, blinking can be a useful signal to communicate emotions, but falls clearly behind more complex models (face & blinking, and face; blinking movements, and paraverbal sounds). This was found to be also depending on the blinking frequencies which were not distanced enough among different emotions. When combining a new set of frequencies with the introduction of contextual information, we observed much increased recognition rates as well as arousal mappings that were closer to the intended emotional profiles. Overall, the two studies demonstrated that blinking colored LEDs constitute a promising and intelligible communication modality for conveying robot affect, and that, in the presence of a concise contextual framing, richer multimodal cues can remain valuable for expressiveness, but are not necessarily essential for accurate arousal communication.

Acknowledgments

This work has been supported by the European Union and by the Italian MIMIT, under the action 1.1.4 "*Ricerca collaborativa* del Programma Nazionale Ricerca, Innovazione e Competitività per la transizione verde e digitale 2021-2027", with the SAPORE research project "System for Autonomous Preparation and Optimization with Robotic Execution", no. F/380023/02/X77 - CUP: B69J25000580005.

References

[1] G. Angelopoulos, C. Di Martino, A. Rossi, and S. Rossi, "Unveiling the learning curve: enhancing transparency in robot's learning with inner speech and emotions," in *32nd IEEE RO-MAN*, pp. 1922–1927, 2023.

[2] Y. Takahashi, Y. Kayukawa, K. Terada, and H. Inoue, "Emotional expressions of real humanoid robots and their influence on human decision-making in a finite iterated prisoner's dilemma game," *International Journal of Social Robotics*, vol. 13, no. 7, 2021.

[3] B. J. Moon, J. Choi, and S. S. Kwak, "'pretending to be okay in a sad voice': Social robot's usage of verbal and nonverbal cue combination and its effect on human empathy and behavior inducement," in *2021 IEEE/RSJ IROS*, pp. 854–861, IEEE, 2021.

[4] G. T. Kraft-Todd, D. A. Reinero, J. M. Kelley, A. S. Heberlein, L. Baer, and H. Riess, "Empathic nonverbal behavior increases ratings of both warmth and competence in a medical context," *PLOS ONE*, vol. 12, pp. 1–16, 05 2017.

[5] D. Conti, C. Cirasa, S. Di Nuovo, and A. Di Nuovo, ""Robot, tell me a tale!" a social robot as tool for teachers in kindergarten," *Interaction Studies*, vol. 21, no. 2, pp. 220–242, 2020.

[6] A. L. Cutting and J. Dunn, "Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations," *Child development*, vol. 70, no. 4, pp. 853–865, 1999.

[7] L. Raggioli, R. Esposito, A. Rossi, and S. Rossi, "Exploring the role of robot's movements for a transparent affective communication," *IEEE Robotics and Automation Letters*, vol. 10, no. 5, pp. 4364–4371, 2025.

[8] D. K. Kim, S. Ahn, S. Park, and M. Whang, "Interactive emotional lighting system using physiological signals," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 4, pp. 765–771, 2013.

[9] N. T. Fitter, M. Mohan, K. J. Kuchenbecker, and M. J. Johnson, "Exercising with baxter: preliminary support for assistive social-physical human-robot interaction," *Journal of neuroengineering and rehabilitation*, vol. 17, no. 1, p. 19, 2020.

[10] A. Banh, D. J. Rea, J. E. Young, and E. Sharlin, "Inspector baxter: The social aspects of integrating a robot as a quality inspector in an assembly line," in *Proceedings of the 3rd HAI*, p. 19–26, ACM, 2015.

[11] N. Abdulazeem and Y. Hu, "The impact of robots' facial emotional expressions on light physical exercises," in *15th International Conference, ICSR*, p. 216–229, 2023.

[12] T. Kim, G. Lee, M. Park, H. M. Lee, J.-W. Park, and H.-J. Suk, "User responses to dynamic light in automobiles with eeg and self-assessments," *IEEE Access*, vol. 10, pp. 123847–123857, 2022.

[13] E. Onchi and S. Lee, "Effects of led on emotion-like feedback of a single-eyed spherical robot," *Science of Emotion and Sensibility*, vol. 24, no. 3, pp. 115–124, 2021.

[14] A. Beck, B. Stevens, K. A. Bard, and L. Cañamero, "Emotional body language displayed by artificial agents," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 2, no. 1, pp. 1–29, 2012.

[15] N. Anzabi and H. Umemuro, "Effect of different listening behaviors of social robots on perceived trust in human-robot interactions," *Int. J. of Social Robotics*, vol. 15, no. 6, pp. 931–951, 2023.

[16] I. Leite, G. Castellano, A. Pereira, C. Martinho, and A. Paiva, "Empathic robots for long-term interaction: evaluating social presence, engagement and perceived support in children," *International Journal of Social Robotics*, vol. 6, pp. 329–341, 2014.

[17] C. Tsiourti, A. Weiss, K. Wac, and M. Vincze, "Designing emotionally expressive robots: A comparative study on the perception of communication modalities," in *Proceedings of the 5th international conference on human agent interaction*, pp. 213–222, 2017.

[18] E. Cha, Y. Kim, T. Fong, M. J. Mataric, *et al.*, "A survey of nonverbal signaling methods for non-humanoid robots," *Foundations and Trends in Robotics*, vol. 6, no. 4, pp. 211–323, 2018.

[19] S. Rossi and M. Ruocco, "Better alone than in bad company: effects of incoherent non-verbal emotional cues for a humanoid robot," *Interaction Studies*, vol. 20, no. 3, pp. 487–508, 2019.

[20] J. Urakami and K. Seaborn, "Nonverbal cues in human–robot interaction: A communication studies perspective," *J. Hum.-Robot Interact.*, vol. 12, Mar. 2023.

[21] S. Li and X. Zhang, "Implicit intention communication in human–robot interaction through visual behavior studies," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 4, pp. 437–448, 2017.

[22] A. G. Brooks and R. C. Arkin, "Behavioral overlays for non-verbal communication expression on a humanoid robot," *Autonomous robots*, vol. 22, pp. 55–74, 2007.

[23] A. Beck, L. Cañamero, and K. A. Bard, "Towards an affect space for robots to display emotional body language," in *19th RO-MAN*, pp. 464–469, IEEE, 2010.

[24] B. J. Zhang and N. T. Fitter, "Nonverbal sound in human-robot interaction: A systematic review," *ACM Transactions on Human-Robot Interaction*, vol. 12, Dec. 2023.

[25] S. Rossi and A. Rossi, "Robotic ears as social cues: A first analysis on emotional expressivity," in *Social Robotics + AI* (M. Staffa, J.-J. Cabibihan, B. Siciliano, S. S. Ge, L. Bodenhagen, A. Tapus, S. Rossi, F. Cavallo, L. Fiorini, M. Matarese, and H. He, eds.), (Singapore), pp. 50–57, Springer Nature Singapore, 2026.

[26] S. Song and S. Yamada, "Expressing emotions through color, sound, and vibration with an appearance-constrained social robot," in *Proceedings of the 2017 ACM/IEEE HRI*, pp. 2–11, 2017.

[27] K. Terada, A. Yamauchi, and A. Ito, "Artificial emotion expression for a robot by dynamic color change," in *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, pp. 314–321, IEEE, 2012.

[28] C. Tsiourti, A. Weiss, K. Wac, and M. Vincze, "Multimodal integration of emotional signals from voice, body, and context: Effects of (in) congruence on emotion recognition and attitudes towards robots," *International Journal of Social Robotics*, vol. 11, pp. 555–573, 2019.

[29] C. M. de Melo and K. Terada, "Cooperation with autonomous machines through culture and emotion," *PLOS ONE*, vol. 14, pp. 1–12, 11 2019.

[30] P. Ekman, "Are there basic emotions?," *Psychological Review*, vol. 99, no. 3, p. 550–553, 1992.

[31] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," *Current Psychology*, vol. 14, pp. 261–292, 1996.

[32] J. Pinney, F. Carroll, and P. Newbury, "Human-robot interaction: the impact of robotic aesthetics on anticipated human trust," *PeerJ Computer Science*, vol. 8, p. e837, 2022.

[33] M. e. a. Wairagkar, "Emotive response to a hybrid-face robot and translation to consumer social robots," *IEEE Internet of Things Journal*, vol. 9, no. 5, pp. 3174–3188, 2022.

[34] R. Plutchik, "A general psychoevolutionary theory of emotion," *Emotion: Theory, research, and experience*, vol. 1, 1980.

[35] S. Yoshimoto, J. Garcia, F. Jiang, A. J. Wilkins, T. Takeuchi, and M. A. Webster, "Visual discomfort and flicker," *Vision Research*, vol. 138, pp. 18–28, 2017.

[36] I. Cucciniello, G. L'Arco, A. Rossi, C. Autorino, G. Santoro, and S. Rossi, "Classmate robot: A robot to support teaching and learning activities in schools," in *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pp. 906–911, IEEE, 2022.

[37] A. Kalegina, G. Schroeder, A. Allchin, K. Berlin, and M. Cakmak, "Characterizing the design space of rendered robot faces," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 96–104, 2018.

[38] T. Ozawa, *How to Draw Anime & Game Characters, Vol. 2: Expressing Emotions.* Tokyo: Graphic-Sha, 2001.

[39] N. Dael, M. Goudbeek, and K. R. Scherer, "Perceived gesture dynamics in nonverbal expression of emotion," *Perception*, vol. 42, no. 6, pp. 642–657, 2013. PMID: 24422246.

[40] M. Gácsi, A. Kis, T. Faragó, M. Janiak, R. Muszyński, and Á. Miklósi, "Humans attribute emotions to a robot that shows simple behavioural patterns borrowed from dog behaviour," *Computers in Human Behavior*, vol. 59, pp. 411–419, 2016.

[41] J. Russell and A. Mehrabian, "Evidence for a three-factor theory of emotions," *Journal of Research in Personality*, vol. 11, pp. 273–294, 09 1977.

[42] O. Mitruț, G. Moise, L. Petrescu, A. Moldoveanu, M. Leordeanu, and F. Moldoveanu, "Emotion classification based on biophysical signals and machine learning techniques," *Symmetry*, vol. 12, p. 21, 12 2019.

[43] Y.-S. Seo and J.-H. Huh, "Automatic emotion-based music classification for supporting intelligent iot applications," *Electronics*, vol. 8, no. 2, p. 164, 2019.

[44] A. F. Morse, J. de Greeff, T. Belpeame, and A. Cangelosi, "Epigenetic robotics architecture (era)," *IEEE Transactions on Autonomous Mental Development*, vol. 2, no. 4, pp. 325–339, 2010.

[45] T. Melder and R. J. Savery, "Lightspeak: Conveying human emotions through minimalist led expression," in *Proceedings of the 13th HAI*, HAI '25, p. 461–463, 2026.

[46] V. Herdel and J. R. Cauchard, "Crafting for emotion appropriateness in affective robotics: Examining the practicality of the occ model," *Proc. ACM Hum.-Comput. Interact.*, vol. 8, Sept. 2024.

[47] G. Laban and E. S. Cross, "Sharing our emotions with robots: Why do we do it and how does it make us feel?," *IEEE Transactions on Affective Computing*, pp. 1–18, 2024.

[48] E. Hauser, Y.-C. Chan, S. Modak, J. Biswas, and J. Hart, "Vid2real hri: Align video-based hri study designs with real-world settings," in *2024 33rd IEEE RO-MAN*, pp. 542–548, 2024.