



Integrating colored lights into multimodal robotic storytelling

Sophia C. Steinhäusser, Sophia Maier, Birgit Lugin*^{*}

Socially Interactive Agents, University of Würzburg, Würzburg 97074, Germany.

***Correspondence to:** Sophia C. Steinhäusser, Socially Interactive Agents, University of Würzburg, Am Hubland, 97074 Würzburg, Germany.
E-mail: sophia.steinhaeusser@uni-wuerzburg.de

Received: December 11, 2024 **Accepted:** April 25, 2025 **Published:** May 10, 2025

Cite this article: Steinhäusser SC, Maier S, Lugin B. Integrating colored lights into multimodal robotic storytelling. *Empath Comput.* 2025;1:202404. <https://doi.org/10.70401/ec.2025.0008>

Abstract

Aims: Storytelling has evolved alongside human culture, giving rise to new media such as social robots. While these robots employ modalities similar to those used by humans, they can also utilize non-biomimetic modalities, such as color, which are commonly associated with emotions. As research on the use of colored light in robotic storytelling remains limited, this study investigates its integration through three empirical studies.

Methods: We conducted three studies to explore the impact of colored light in robotic storytelling. The first study examined the effect of emotion-inducing colored lighting in romantic storytelling. The second study employed an online survey to determine appropriate light colors for specific emotions, based on images of the robot's emotional expressions. The third study compared four lighting conditions in storytelling: emotion-driven colored lights, context-based colored lights, constant white light, and no additional lighting.

Results: The first study found that while colored lighting did not significantly influence storytelling experience or perception of the robot, it made recipients feel more serene. The second study showed improved recognition of amazement, rage, and neutral emotional states when colored light accompanied body language. The third study revealed no significant differences across lighting conditions in terms of storytelling experience, emotions, or robot perception; however, participants generally appreciated the use of colored lights. Emotion-driven lighting received slightly more favorable subjective evaluations.

Conclusion: Colored lighting can enhance the emotional expressiveness of robots. Both emotion-driven and context-based lighting strategies are appropriate for robotic storytelling. Through this series of studies, we contribute to the understanding of how colored lights are perceived in robotic communication, particularly within storytelling contexts.

Keywords: Social robots, storytelling, emotions, colors, light, multimodal

1. Introduction

While book sales are declining^[1], interest in audiobooks and podcasts continues to grow^[2]. One reason for the success of this medium might be the rise of digital download collections^[3]. Thanks to providers such as *Spotify* or *Audible* and their respective mobile applications, recipients can listen to storytelling wherever and whenever they want, making it an important part of their everyday lives^[4]. A crucial aspect of storytelling is to tell stories in a way that others can relate to and emphasize^[5]. This is not only achieved by presenting information about characters, settings, and plot^[6], but also requires storytellers to bring the characters' and stories' emotions to life^[7]. To do so, human storytellers use multiple modalities, such as voice, facial expressions, and gestures^[8]. New technologies offer new possibilities for receiving stories. For example, video games allow players to actively experience a story instead of merely witnessing it^[9]. Similarly, the relatively new medium of social robots, which can communicate intuitively and naturally^[10], is well suited for storytelling, particularly due to its multimodal capacities^[11]. While they are capable of using human modalities, so-called biomimetic modalities^[12], such as speech, gaze, or body language, they can also utilize non-biomimetic modalities, such as sound effects and colors, to communicate^[13]. Although colors are not a modality naturally employed by humans, they are associated with emotions^[14]. Robots can easily integrate colors into their emotional expressions using light emitting diodes (LEDs)^[15], thereby operationalizing the color modality in the form of colored light. If this approach proves effective for emotional expression, it would be especially appealing for low-budget robots, since colored lights are more affordable to install than biomimetic options such as motors for facial expressions.

While foundational research in Human-Robot Interaction (HRI) has explored integrating colored light into robotic emotion expressions^[15-17], only a limited number of use cases, such as navigation tasks^[18], have been tested. Many potential application areas,



such as storytelling, remain largely unexplored. Steinhäusser and Lugin^[19] examined the use of emotion-conveying LED eye colors of a storytelling robot, finding that participants in the control group, who interacted with a robot with constant white LED eye colors, had a more engaging storytelling experience. Extending the use of color from the robot's LEDs to the room's illumination, Steinhäusser *et al.* demonstrated a positive effect of colored light on co-presence and the perception of a static robotic storyteller, even in the absence of body language^[20].

Building on HRI research that recommends integrating motion and color for enhanced emotional expression^[17], we conducted a series of studies combining colored lighting with emotion-expressive body language to further investigate the role of emotion-driven light in robotic storytelling. In the first study, we compared robotic storytelling enhanced with colored lighting, designed according to emotion induction guidelines from virtual environments, with storytelling using neutral or no additional lighting. Given the mixed results, we refined our approach and conducted an online study to further explore how colored room illumination influences human perception of robotic emotional expression. This study enabled us to empirically identify combinations of ambient lighting that support accurate recognition of robot-expressed emotions. In a third study, we applied these emotion-based lighting schemes and compared them with a contextual lighting approach, which aligned lighting colors with the environments described in the story. The findings indicate that both emotional and contextual lighting styles are suitable for robotic storytelling; however, participants showed a clear preference for the emotional lighting style, which we therefore recommended for future implementations.

2. Related Work

Storytelling can be defined in three dimensions: by its content, the development of coherent, temporally structured events; by its dialogic nature involving interaction between speaker and audience; and by the participation and responses of recipients^[21,22]. Oral storytelling is inherently multimodal^[23], as storytellers employ their voice, facial expressions, gestures, eye contact, and interactive engagement to connect the narrative with listeners^[8]. These definitions highlight that receiving a story involves more than merely processing information; it is also about experiencing the emotions conveyed through the narrative and empathizing with the characters' feelings and thoughts. The storyteller, in turn, bears the responsibility of crafting and delivering this emotional experience^[7].

2.1 Emotions and storytelling experience

Broadly speaking, emotions are reactions to events^[24,25]. Unlike moods which are typically low in intensity, diffuse in focus, and sustained over time, emotions are intense, short-term responses elicited by specific events^[26]. Notably, even when individuals are aware of the artificial nature of a stimulus, they often react as if it were real^[27,28]. Emotional theories can be categorized into dimensional models such as the *Circumplex Model of Affect*^[29], which maps emotions along two axes: arousal (high or low) and valence (positive or negative); and discrete emotion models^[24]. For instance, Plutchik's *Wheel of Emotions*^[30] proposes eight primary emotions, each with three levels of intensity. Similar emotions, such as ecstasy and admiration, are positioned close together, while opposing emotions, such as acceptance and disgust, are placed opposite one another.

Emotions play a crucial role in understanding a given text, particularly in the context of storytelling^[31]. Moreover, story recipients rely on emotions content not only for comprehension but also for entertainment. For example, when encountering a humorous story, they expect to feel happiness, whereas a thriller or horror story is chosen to evoke feelings of fear or suspense. In essence, humans engage with fictional narratives to be emotionally moved. This emotional engagement is facilitated by the emotions expressed by the story or its characters, which in turn elicit emotional responses in the audience. "Although the emotions of fiction seem to happen to characters in a story, really, all the important emotions happen to [them]"^[32]. These emotional responses are closely linked to the concept of transportation into a narrative^[33]. Transportation is defined as a mental process in which recipients direct their cognitive resources (attention, imagery, and feelings) toward the story, thereby disconnecting from the real world and becoming absorbed in the narrative universe^[34,35]. It has been shown to reduce negative cognitive reactions such as disbelief, enhance the perceived realism of the story, and intensify emotional responses toward characters^[34]. Transportation also contributes to the enjoyment and perceived meaningfulness of a media artifact^[36], making it a fundamental component of the storytelling experience. Importantly, this effect is not limited to traditional media but can emerge across various narrative formats^[34], including novel platforms such as social robots^[11].

2.2 Social robotics

Social robots are designed for natural interaction, enabling communication through both verbal and non-verbal modalities^[10]. In doing so, they elicit social responses from human users not only on a cognitive level but also on an emotional one^[10,37]. With their focus on social interactivity^[38], social robots can be employed in various domains such as education^[39,40] and entertainment^[41,42]. To effectively engage in social interactions, social robots should exhibit anthropomorphic qualities^[43]; that is, their form and behavior should be designed in a human-like manner to evoke anthropomorphism—the tendency to attribute human characteristics, such as gender or personality, to non-human entities like robots^[44]. Therefore, it is recommended that social robots adopt human-like morphology^[45,46], for example, by including features such as arms and legs, and by displaying social communication behaviors such as body language and social gaze^[47,48]. These elements underscore the importance of mimicking human attributes, a design principle

known as biomimetics^[12].

These biomimetic modalities are also essential for storytelling robots. Several studies have shown that robotic storytellers employing emotional facial expressions congruent with the story content can enhance recipients' transportation into the narrative^[11,49], as well as increase the likeability of the robotic storyteller itself^[49]. Moreover, the integration of multiple modalities appears to be as important in HRI as it is in human-human communication. As noted, "When communicating using their full multi-modal expressive potential, speakers can increase communicative efficiency by simultaneously transporting complementary information, and foster robustness by providing redundant information in various modalities"^[50], a concept known as multimodality. In line, Ham *et al.*^[51] found that a robot's persuasiveness in a moral context can be improved by combining gaze with concurrent gestures during speech, thereby demonstrating the applicability of multimodal benefits to HRI.

2.3 Colored light as a modality

In contrast to the biomimetic modalities discussed above—those modeled after human behavior—robots are also capable of expressing emotions through non-biomimetic modalities that lack direct biological analogues, such as the use of colored light^[12]. In a preliminary use case, Rea *et al.*^[52] combined both light and color to enable guests in a café to program a non-humanoid *Roomba* utility robot, equipped with an LED strip, to represent their mood using colored lights. Research further suggests that emotional expression through multimodal robot behaviors can be enhanced by the addition of color^[17], indicating that findings from color psychology may be applicable to HRI^[15]. For example, using the ball-shaped robot *Maru*, Song and Yamada^[13] confirmed associations such as blue with sadness or red and anger in robotic emotional expressions. In a subsequent study, they combined light color and dynamic patterns to create and validate emotion expressions using a *Roomba* robot with an LED strip^[15]. Again, principles from color psychology were validated, for instance, an intensely blinking red light was perceived as conveying hostility. They also found that expressive lighting enhanced the emotional interpretability of in-situ motion cues^[53]. Additional studies employing robots specifically designed for multimodal emotional expression, such as those incorporating colored, blinking lights into robot's ears^[54] or chest^[55], have likewise reported successful emotion communication using these visual cues.

Several robots, such as Sony's *AIBO* robot dog and Aldebaran's humanoid *NAO* robot, are equipped with colored lights, typically installed in their eyes, to support emotional expression. However, studies examining the effectiveness of the *NAO* robot's eye LEDs for conveying emotion have reported rather negative outcomes^[56,57]. In the specific context of storytelling, a preliminary study by Steinhäusser *et al.*^[19] found that the use of emotionally colored eye LEDs on the *NAO* robot negatively affected the storytelling experience. This was reflected in a reduction in cognitive absorption—a state of cognitive and emotional engagement with technology, comparable to flow^[58,59], and characterized by factors such as attention and curiosity^[60]—as well as decrease in the perceived animacy of the robotic storyteller. The authors attribute these effects to the distracting nature of applying a non-biomimetic modality (colored light) to an anthropomorphic body part. They propose expanding the illuminated area to surface such as walls or floors, as demonstrated by Betella *et al.*^[61]. This approach, projecting emotion-associated colors into the environment, has already been explored in audiobook storytelling, where lighting was dynamically adjusted to match the narrator's emotional tone^[62].

However, colors and colored light can not only reinforce emotional expression but can also induce emotions^[63,64]. In films, for example, color is deliberately used to evoke specific emotional responses in viewers^[65,66], a technique that has also been adopted in other media such as video games^[67]. In a study manipulating the background color of a video game, Wolfson and Case^[68] found that players exhibited different cognitive and physiological responses to the colors blue and red. Specifically, while performance steadily improved with a blue background, it stagnated under red. The authors attributed these differences to variations in arousal levels induced by the colors, as evidenced by heart rate measurements. Similar results were reported by Joosten *et al.*^[69], who manipulated ambient light color in a fantasy role-playing game's virtual environment: reddish light elicited negative valenced arousal, whereas yellow light was experienced as positively valenced. Steinhäusser *et al.*^[67] synthesized such findings on the emotional impact of color and light into a set of design guidelines for emotion-inducing virtual environments, which were validated in both desktop and immersive VR settings^[67,70]. Applying these guidelines to storytelling with both a physically embodied and a virtual robot, Steinhäusser *et al.*^[20] demonstrated that the addition of emotion-inducing light enhanced perceptions of the robot's social presence and overall robot perception. These results highlight the potential benefits of emotion-inducing colored light in robotic storytelling as well.

3. Contribution

Given the significant potential of the color modality, realized through colored light, in emotional robotic storytelling, this work seeks to further explore this relatively under-investigated channel. In light of the importance of multimodality and existing recommendations to combine motion with color, we conducted a series of studies examining the integration of colored light and expressive body language in a robotic storytelling context. Our investigation focuses on three key aspects: the storytelling experience, emotion induction, and robot perception.

First (Study I; see chapter 4), we extended the approach of Steinhäusser *et al.*^[20] by implementing emotion-driven colored lighting in

robotic storytelling, guided by design principles from virtual environments and combined with emotional bodily expressions. To evaluate the effects, we conducted a laboratory user study comparing three conditions: robotic storytelling with emotion-driven colored lighting, with constant white lighting, and with no additional lighting. As we were not able to replicate the findings of Steinhäusser *et al.*^[20], we adapted our lighting approach to better suit the context of our expressive social robot, shifting our focus from emotion induction to emotion recognition. Consequently, we conducted an online study (Study II; see chapter 5) aimed at identifying suitable light colors that support the robot's emotional bodily expressions.

The empirically derived emotion-matching light colors were subsequently applied in an evaluative user study (Study III; see chapter 6). In this study, we not only assessed the effectiveness of our emotion-driven lighting approach but also compared it to a context-based lighting strategy. Specifically, we compared four storytelling conditions: one using emotion-driven colored lighting, one using context-based lighting aligned with the story's environmental setting, one with constant white lighting, and one without any additional lighting. Our findings indicate that while overall storytelling experience, emotion induction, and robot perception showed minimal differences across conditions, participants expressed a clear preference for robotic storytelling integrating that incorporated colored lighting over versions with constant or no lighting enhancements. Both lighting strategies, namely emotion-driven and context-based approaches, proved suitable for robotic storytelling; however, qualitative feedback slightly favored the emotion-driven approach. Therefore, we recommend integrating emotion-driven colored lighting into robotic storytelling, while acknowledging the general effectiveness of both approaches.

4. Study I: Embedding Light in Multimodal Robotic Storytelling

To gain initial insights, we conducted a laboratory study examining the combination of expressive body language and emotion-inducing colored light in robotic storytelling. A romantic story was presented under three conditions: colored light, constant white light, and no additional lighting. Colored lighting was hypothesized to enhance the robot's expressiveness and directly influence recipients' emotional states. As prior studies have shown that storytelling experience is improved by the integration of expressive biomimetic modalities in robotic storyteller^[11,49] and that the emotional impact of stories contributes significantly to narrative engagement^[71] which in turn is positively associated with transportation-related effects^[33], we anticipated that adding emotion-inducing colored light would enhance the storytelling experience.

- **H1a:** Transportation will be higher in the adaptive lighting condition compared to the constant white or no lighting conditions.
- **H1b:** Cognitive absorption will be higher in the adaptive lighting condition compared to the constant white or no lighting conditions.

When engaging with a story, recipients often adopt the emotions of the characters, as conveyed by the storyteller, due to empathetic processes^[7,32,72]. Given that we selected a romantic story designed to elicit emotions of positive valence^[32], we expect participants to report increased positive emotions and decreased negative emotions following the storytelling session.

- **H2a:** Positive affect will be higher after receiving the story.
- **H2b:** Negative affect will be lower after receiving the story.
- **H2c:** Joviality will be higher after receiving the story.
- **H2d:** Serenity will be higher after receiving the story.

Previous research has shown that colored lights can enhance a robot's ability to express emotions^[53]. Therefore, we propose that when a robotic storyteller employs multimodal emotional expressions, including colored lighting, to bring the story and its characters to life, recipients' emotional responses and affective states may likewise be influenced.

- **RQ1:** Do changes in emotional differ across the lighting conditions?
- **RQ2:** Does attentiveness vary between the lighting conditions?

As Steinhäusser *et al.*^[20] demonstrated a positive effect of the light modality on the perception of robotic storyteller, we further examine general robot perception across several dimensions, including anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety.

- **RQ3:** Does the perception of the robot differ across the lighting conditions?

4.1 Materials

To examine the impact of lighting on recipients' storytelling experience, emotional responses, and robot perception, we implemented three versions of a romantic story narrated by a social robot: (1) storytelling with adaptive colored lighting based on established guidelines for emotion-inducing lighting in virtual environments^[67], (2) storytelling with constant white lighting, and (3) storytelling without any additional lighting. Representative snapshots of each condition are shown in [Figure 1](#).

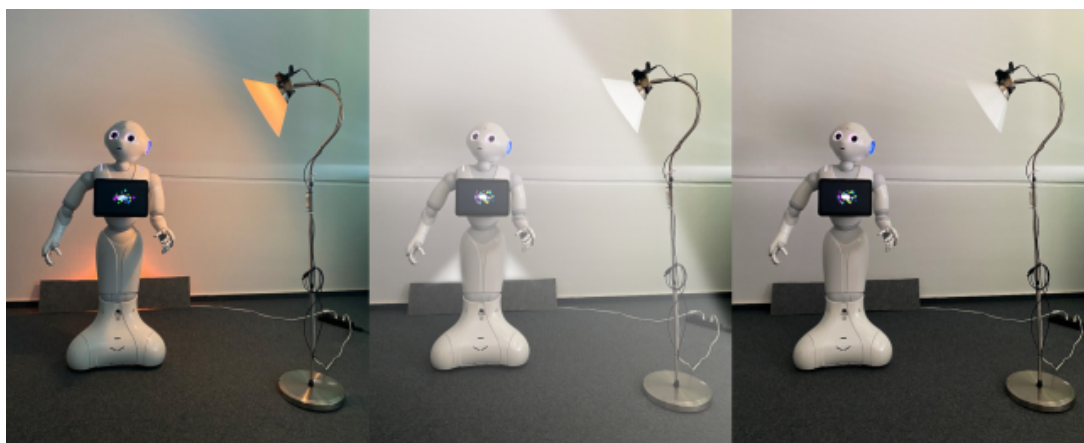


Figure 1. Snapshot of Pepper robot during storytelling expressing admiration, conditions f.l.t.r.: adaptive lighting, constant lighting, no additional light.

4.1.1 Concept

We selected a romantic story due to the high popularity of this genre, particularly among female participants^[73], who were expected to be well-represented in our study sample. The short narrative centers on a girl meeting a boy at the beach, with both gradually falling in love. An English translation of the story is provided in [Supplementary materials](#). The story was delivered by the *Pepper* robot had a duration of approximately six minutes.

First, the story was annotated with respect to emotional content. This annotation was conducted in a prior project, from which we reused the resulting data. The story was tokenized at the clause level, using punctuation marks such as commas or full stops as segmentation points. A total of 89 individual annotators (61 female, 28 male, 0 diverse; age: $M = 21.98$, $SD = 2.09$) labeled each token with the emotion they believed a storyteller should express when narrating that part of the story. We employed the eight primary emotions from Plutchik's *Wheel of Emotions*^[74], as along with an additional “neutral” label and an “I don't know” option. Given the relatively large number of annotators compared to related studies^[20,75,76], we applied a more conservative approach to determine consensus labels. Specifically, we calculated 95% confidence intervals (CIs) for the frequency of each label and retained only the most frequently selected emotion labels whose CIs did not overlap. If no label met this criterion, the token was labeled as “neutral”. Using this method, 55 tokens were assigned specific emotion labels, while 45 tokens were labeled as “neutral”.

Next, we conceptualized the robot's behavior. For body language, we employed emotional bodily expressions for the *Pepper* robot that had been pre-tested by Steinhäusser *et al.*^[75] (recognition rates are presented in [Table 1](#); corresponding images of the body language are provided in [Supplementary materials](#)). As the authors' set of expressions corresponds to the eight primary emotions from Plutchik's *Wheel of Emotions*, we were able to directly align the expressions with the story tokens labeled with the respective emotions.

Table 1. Recognition rates from study II for multimodal and unimodal expressions.

Emotion Label	Color Combination			RR UM ^[89]
	Version	RR MM all	RR MM clean	
Vigilance	1	5.94%	6.82%	26.67%
	2	10.89%	11.96%	
	3	10.89%	12.09%	
Admiration	1	17.82%	18.95%	31.11%
	2	25.74%	27.66%	
	3	23.76%	24.74%	
Amazement	1	30.69%	32.39%	44.44%
	2	33.66%	36.96%	
	3	42.57%	47.25%	
Ecstasy	1	43.56%	44.90%	64.44%
	2	32.67%	33.33%	

	3	49.50%	50.51%	
Loathing	1	11.88%	12.50%	68.89%
	2	12.87%	13.27%	
	3	8.91%	9.09%	
Terror	1	41.58%	43.75%	82.22%
	2	46.53%	51.09%	
	3	49.50%	53.19%	
Rage	1	55.45%	56.57%	60.00%
	2	72.29%	72.28%	
	3	68.32%	70.41%	
Grief	1	64.36%	65.66%	97.78%
	2	79.21%	80.00%	
	3	71.29%	72.00%	
Neutral	1	68.32%	68.32%	31.11%

RR: recognition rates, MM: multimodal, UM: unimodal, all: including all labels, clean: without “i don’t know” option.

Finally, we defined the color settings for the adaptive lighting condition. As prior research^[20] suggests that design guidelines for emotion-inducing lighting in virtual environments can be effectively applied to robotic storytelling, we adopted this approach and conceptualized illumination settings for each emotion label based on the guidelines proposed by Steinhäusser *et al.*^[67]. However, unlike previous studies that used only a spotlight directed at the robot^[20], we additionally incorporated ambient lighting that illuminated the wall behind the robot, as recommended by Steinhäusser and Lugin^[49]. The specific color settings for each emotion are detailed in Table 2. For the neutral label, we used white ambient light (hue 0, saturation 0, brightness 100) and turned off the spotlight. Since amazement can carry both positive and negative valence^[77], we applied the same lighting configuration as for the neutral condition. As vigilance and loathing were not selected during the annotation process, presumably because these emotions were not expressed in the story, no lighting was configured for them.

Table 2. Values for emotion-inducing lighting following guidelines from Steinhäusser *et al.*^[67].

Emotion Label	Spotlight		Ambient Light		Guideline number from ^[67]
	Hue	Saturation	Hue	Saturation	
Ecstasy	35	100	35	100	<i>GLpos12</i> : warm and balanced light, sunset colors
Admiration	60	40	50	100	<i>GLpos8</i> , <i>GLpos9</i> , <i>GLpos11</i> : bright or pastel colors, yellowish sunlight
Terror	180	100	10	100	<i>GLneg5</i> , <i>GLneg9</i> , <i>GLneg10</i> : warm reddish and cool blueish imbalanced lights
Grief	205	40	205	60	<i>GLneg8</i> , <i>GLneg9</i> : blueish and greyish dim light
Rage	0	100	0	100	<i>GLneg5</i> , <i>GLneg9</i> : warm light and red highlights

4.1.2 Implementation

The storytelling sequence was implemented using *Unity* version 2019.1.14f1. We employed the *Pepper Python Unity Toolkit*^[78] to send commands to the robot. The toolkit relies on the robot’s internal speech synthesis engine; thus, no modifications were made to voice modulation. Each story token was transmitted as a text-to-speech command along with encapsulated functions containing the predefined parameters for the corresponding emotional posture, based on the emotion label assigned to the token (Section 4.1.1). This setup constituted the control condition, in which no additional illumination was provided.

To implement the additional lighting, we used two *Tapo* E27 smart bulbs: a spotlight directed at the robot and an ambient light positioned behind it, as illustrated in Figure 1. The smart lights were integrated with *Unity* utilizing a Python 3.0 server, which communicated with *Unity* using the *PyP100* module. In the constant-light condition, a command was sent at the beginning of the storytelling session to set both the spotlight and ambient light to white. In the adaptive, emotion-inducing lighting condition, we defined individual functions for each emotion label, each encapsulating the corresponding spotlight and ambient light color settings. The function matching the emotion label of the current token was called concurrently with sending the token to the robot. For most emotion labels, except ecstasy and rage, which were represented by two yellow and two red lights, respectively, with the spotlight

and ambient light used different colors (Table 2). For instance, grief was represented by two distinct shades of blue, while terror combined a warm reddish ambient light with a cool blueish spotlight. The final stimuli had a total duration of approximately five minutes in length.

4.2 Methods

To investigate the effects of different conditions in robotic storytelling, we conducted a laboratory study. Using a between-subjects design, we compared the influence of (1) adaptive emotion-driven colored lighting, (2) constant white light, and (3) the absence of additional lighting on recipients' storytelling experience, emotional responses, and perception of the robotic storyteller. These conditions were based on the three storytelling sequences described in Section 4.1. The study was reviewed and approved as ethically sound by the local ethics committee.

4.2.1 Measures

To assess **storytelling experience**, transportation was measured using the *Transportation Scale Short Form* (TS-SF)^[79], which consists of six items (e.g., "I wanted to learn how the narrative ended"). Participants responded on a 7-point Likert scale ranging from 1 ("not at all") to 7 ("very much"). Appel *et al.*^[79] reported Cronbach's alpha values between .80 to .87 for the TS-SF; for the current sample, internal consistency was similarly high with a Cronbach's alpha of .86.

Further, the *Cognitive Absorption* (CA) questionnaire was used to measure the recipients' level of involvement in the robotic storytelling experience. Originally developed by Agarwal *et al.*^[58] to investigate engagement in web and software usage, the questionnaire was adapted to the robotic storytelling context following the approach of Steinhäusser and Lugin^[19]. It comprises five scales, the (1) *Temporal Dissociation* scale which originally includes five items but was cut down to three items that refer to the recipients' current feeling of time during the interaction with the robot, e.g., "Time flew while the robot told the story", the (2) *Focused Immersion* scale, comprising five items, e.g., "While listening to the robot, I got distracted by other things very easily", the (3) *Heightened Enjoyment* scale with four items, e.g., "I enjoyed using the robot", the (4) *Control* scale including three items, e.g., "I felt that I have no control while listening to the robot", and the (5) *Curiosity* scale which includes three items, e.g., "Listening to the robot made me curious". All items were rated on a seven-point Likert scale (1: "Strongly disagree", 4: "Neutral", and 7: "Strongly agree"). Agarwal *et al.*^[58] reported reliability values of .93 for the *Temporal Dissociation* as well as the *Heightened Enjoyment* and *Curiosity* scale; .88 for the *Focused Immersion* scale, and .83 for the *Control* scale. In the current sample, Cronbach's alpha was .88 for the *Temporal Dissociation* and *Focused Immersion* scales, .92 for the *Heightened Enjoyment* scale, .62 for the *Control* scale, and .90 for the *Curiosity* scale.

Recipients' **emotions** were measured using the *Positive and Negative Affect Schedule-Expanded Form* (PANAS-X)^[80,81]. In more detail, we utilized the *Positive Affect* scale (10 items, e.g. "enthusiastic") and *Negative Affect* scale (10 items, e.g. "nervous") to obtain a general overview as well as the scales concerning *Joviality* (8 items, e.g. "cheerful"), and *Serenity* scale (3 items, e.g. "relaxed"). Furthermore, we measured *Attentiveness* (4 items, e.g. "attentive") to examine the influence of the robotic storytelling on the participants' attention. Each item was presented with a five-point Likert scale (1: "very slightly or not at all", 2: "a little", 3: "moderately", 4: "quite a bit", 5: "extremely"). Watson and Clark^[81] reported internal consistency of .83 to .88 for *Positive Affect*, .85 to .91 for *Negative Affect*, .93 for *Joviality*, .74 for *Serenity*, and .72 for *Attentiveness* measuring the emotional state *at the moment*. Values for the current study were .90 for *Positive Affect*, .82 to .93 for *Negative Affect*, .92 to .94 for *Joviality*, .75 to .77 for *Serenity*, and .79 to .80 for *Attentiveness*. **Robot perception** was measured using the *Godspeed* questionnaire^[82]. It includes five scales measured on five-point semantic differentials: (1) *Anthropomorphism* (5 items, e.g., "machinelike" versus "humanlike"), (2) *Animacy* (6 items, e.g. "mechanical" versus "organic"), (3) *Likeability* (5 items, e.g., "unfriendly" versus "friendly"), (4) *Perceived Intelligence* (5 items, e.g., "foolish" versus "sensible"), and (5) *Perceived Safety* (3 items, e.g., "anxious" versus "relaxed"). Bartneck *et al.* reported Cronbach's alpha of .88 to .93 for *Anthropomorphism*, .70 for *Animacy*, .87 to .92 for *Likeability*, and .75 to .77 for *Perceived Intelligence*. No internal consistency value was reported for *Perceived Safety*. In the current sample, Cronbach's alpha values were .80 for *Anthropomorphism*, .79 for *Animacy*, .83 for *Likeability*, .80 for *Perceived Intelligence*, and .63 for *Perceived Safety*. Lastly, participants provided demographic data, i.e. age and gender, and were invited to leave a comment.

4.2.2 Procedure

When arriving at the lab, participants first provided informed consent to take part in the study. Next, they completed the first part of the emotion questionnaire. Participants were then randomly assigned to one of the three conditions and received the robotic storytelling either with adaptive, emotion-driven lighting, constant illumination, or no additional lighting. Afterward, they completed the rest of the questionnaire. Lastly, they were thanked and fully debriefed about the aim of the study.

4.2.3 Participants

Overall, 98 participants were recruited. However, six of them had to be excluded due to technical issues with the Pepper robot or the smart lighting system, such as lost connections. Thus, data from 92 participants were analyzed (mean age = 21.99 years, *SD* = 2.44). The majority of 73 participants self-identified as female (age: *M* = 21.82, *SD* = 2.33), whereas only 19 participants self-identified as male

(age: $M = 22.63$, $SD = 2.81$). No participants identified as non-binary or another gender. Being randomly assigned to the conditions, 30 participants received the storytelling with constant lighting (25 female, 5 male; age: $M = 21.53$, $SD = 2.71$), while 31 participants each received the storytelling with adaptive lighting (22 female, 9 male; age: $M = 22.55$, $SD = 2.87$) respectively without any additional lighting (26 female, 5 male; age: $M = 21.87$, $SD = 1.48$).

4.3 Results

All analyses were conducted using JASP^[83] version 0.16, with a significance threshold set at .05. Descriptive statistics are presented in Table 3. First, assumptions checks were performed. Shapiro-Wilk tests indicated violations of the normality assumption for *Transportation*, *Curiosity*, and *Anthropomorphism*. Levene's tests showed violations of the homogeneity assumption for *Negative Affect*, *Perceived Intelligence*, and *Perceived Safety*.

Table 3. Descriptive data for study I.

		Adaptive Lighting		Constant Lighting		No Additional Lighting	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Transportation		4.44	1.33	4.69	1.20	4.56	1.16
CA: Temporal Dissociation ^a		4.31	1.41	4.19	1.42	4.00	1.58
CA: Focused Immersion ^a		4.50	1.14	4.31	1.39	4.43	1.17
CA: Heightened Enjoyment ^a		5.01	1.23	5.12	1.05	4.86	1.21
CA: Control ^a		3.10	1.29	2.80	1.04	2.56	1.03
CA: Curiosity ^a		4.71	1.42	4.96	1.47	4.49	1.24
Positive Affect ^b	pre	2.58	0.76	2.71	0.72	2.67	0.73
	post	2.57	0.66	2.72	0.75	2.60	0.81
Negative Affect ^b	pre	1.37	0.46	1.41	0.38	1.29	0.31
	post	1.27	0.42	1.17	0.20	1.19	0.25
Joviality ^b	pre	2.70	0.83	2.79	0.65	2.82	0.85
	post	2.79	0.80	2.99	0.69	2.90	0.94
Serenity ^b	pre	3.52	0.81	3.31	0.88	3.31	0.77
	post	3.55	0.82	3.59	0.83	3.12	0.70
Attentiveness ^b	pre	2.97	0.73	3.15	0.74	3.11	0.79
	post	2.86	0.70	3.07	0.72	2.86	0.85
Anthropomorphism ^b		2.08	0.77	2.15	0.75	2.08	0.62
Animacy ^b		2.67	0.59	2.71	0.72	2.74	0.65
Likeability ^b		3.92	0.63	4.08	0.49	4.17	0.48
Perceived Intelligence ^b		3.77	0.38	3.79	0.64	3.76	0.79
Perceived Safety ^b		3.47	0.80	3.57	0.76	3.19	0.73

CA: Cognitive Absorption; ^a: Calculated values from 1 to 7. ^b: Calculated values from 1 to 5.

For **storytelling experience**, no significant differences were found when comparing *Transportation* between all three conditions ($\chi^2(2) = 0.31$, $p = .855$). In line, a planned contrast indicated no significant difference between the condition with adaptive lighting condition and the conditions with constant or no lighting ($t(89) = -0.69$, $p = .494$).

Concerning *Cognitive Absorption*, no significant group differences were revealed for *Temporal Dissociation* ($F(2, 89) = 0.36$, $p = .701$), *Focused Immersion* ($F(2, 89) = 0.17$, $p = .842$), *Heightened Enjoyment* ($F(2, 89) = 0.37$, $p = .696$), *Control* ($F(2, 89) = 1.77$, $p = .177$), or *Curiosity* ($\chi^2(2) = 4.51$, $p = .105$). Planned contrasts between the condition with adaptive lighting and the conditions with constant or no lighting also did not reveal significant differences for *Temporal Dissociation* ($t(89) = 0.67$, $p = .504$), *Focused Immersion*

($t(89) = 0.46, p = .651$), Heightened Enjoyment ($t(89) = 0.07, p = .943$), Control ($t(89) = 1.68, p = .097$), and Curiosity ($t(89) = 0.02, p = .981$).

Regarding **emotions**, a mixed ANOVA indicated no significant main effect of lighting condition ($F(2, 89) = 0.32, p = .726$) or time of measurement ($F(1, 89) = 0.37, p = .605$) on *Positive Affect*. Also, no significant interaction effect was observed ($F(2, 89) = 0.22, p = .806$), and a planned contrast indicated no difference between the condition with adaptive lighting and the conditions with constant or no lighting ($t(89) = -0.66, p = .509$). For *Negative Affect*, no significant main effect of lighting condition was found ($F(2, 89) = 0.42, p = .656$), whereas a significant main effect of time was observed ($F(1, 89) = 27.23, p < .001, \omega^2 = .04$), indicating a general decrease of *Negative Affect* after the storytelling. No significant interaction effect emerged ($F(2, 89) = 2.64, p = .077$), and a planned contrast indicated no difference between the condition with adaptive lighting and the conditions with constant or no lighting ($t(89) = 0.69, p = .491$). For *Joviality*, again, no significant main effect of lighting condition was found ($F(2, 89) = 0.33, p = .717$), but a significant main effect of time emerged ($F(1, 89) = 4.38, p = .039, \omega^2 = .01$), with a general increase in *Joviality* after the storytelling. No significant interaction effect was indicated ($F(2, 89) = 0.40, p = .673$), and the planned contrast indicated no difference between the condition with adaptive lighting and the conditions with constant or no lighting ($t(89) = -0.81, p = .421$). For *Serenity*, no significant main effects of lighting condition ($F(2, 89) = 1.57, p = .215$) or time ($F(1, 89) = 0.33, p = .569$) were identified, but a significant interaction effect was identified ($F(2, 89) = 4.12, p = .019$). As displayed in [Figure 2d](#), *Serenity* increased descriptively in the constant light condition, decreased in the no-light condition, and slightly increased in the adaptive light condition. However, Bonferroni-corrected post-hoc tests did not reveal any significant pairwise differences ($ps > .05$), and the planned contrast also showed no significant difference between adaptive and the other conditions ($t(89) = 1.24, p = .220$). For *Attentiveness*, no significant main effect of lighting condition was found ($F(2, 89) = 0.58, p = .560$), but again a significant main effect of time was identified ($F(1, 89) = 6.56, p = .012$). As displayed in [Figure 2e](#), attentiveness decreased descriptively in all lighting conditions, with the largest decrease occurring in the no-light condition. However, post-hoc pairwise comparisons revealed no significant differences ($ps > .05$). A planned contrast did not indicate a significant difference between the adaptive lighting condition and the conditions with constant or no lighting ($t(89) = -0.86, p = .395$).

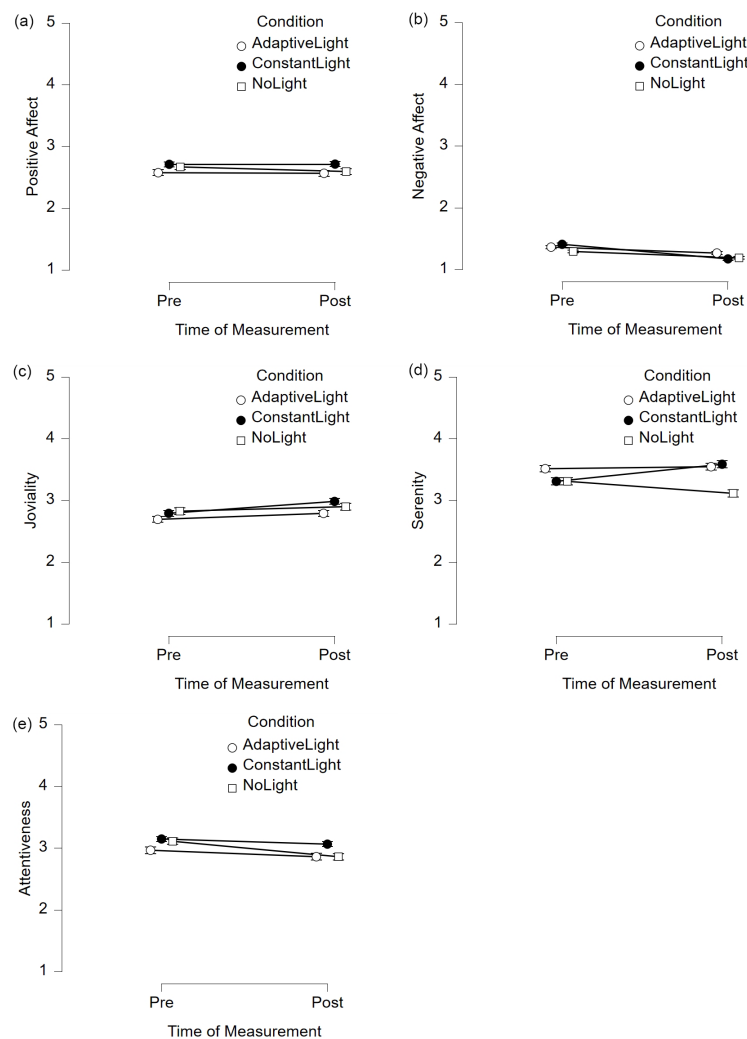


Figure 2. Plots for pre- and post-measured emotions of Study I. (a) Descriptive plot for *Positive Affect*; (b) Descriptive plot for *Negative Affect*; (c) Descriptive plot for *Joviality*; (d)

Descriptive plot for *Serenity*; (e) Descriptive plot for *Attentiveness*. AdaptiveLight: condition with adaptive lighting; ConstantLight: condition with constant lighting; NoLight: condition without additional lighting; Error bars represent the standard error.

Comparing **robot perception**, no significant group differences were found for *Anthropomorphism* ($\chi^2(2) = 0.21, p = .899$), *Animacy* ($F(2, 89) = 0.08, p = .927$), *Likeability* ($F(2, 89) = 0.20, p = .195$), *Perceived Intelligence* ($\chi^2(2) = 0.16, p = .925$), or for *Perceived Safety* ($\chi^2(2) = 3.63, p = .163$). Planned contrasts revealed no significant differences between the adaptive lighting condition and the conditions with constant or no lighting for *Anthropomorphism* ($t(89) = -0.20, p = .843$), *Animacy* ($t(89) = -0.34, p = .736$), *Likeability* ($t(89) = -1.70, p = .092$), *Perceived Intelligence* ($t(89) = -0.02, p = .983$), or for *Perceived Safety* ($t(89) = 0.55, p = .584$).

4.4 Discussion

We conducted a laboratory study to examine the combination of the well-researched biomimetic modality of body language together with the rather unexplored non-biomimetic modality of colored light in a robotic storytelling scenario. We used the *Pepper* robot, which employed emotion-conveying body language to tell a romantic story in three versions, (1) with emotion-driven colored lights, (2) with constant white lighting, (3) without additional illumination.

Concerning storytelling experience, we found no differences between the lighting conditions in terms of transportation or cognitive absorption, therefore **H1a** and **H1b** must be rejected. While our findings for transportation are in line with results reported by Steinhäusser and Lugin^[19], who used the NAO robot's eye LEDs as an additional colored modality, we were able to mitigate the negative effect they yielded for cognitive absorption. In addition, the values of transportation and cognitive absorption of the adaptive lighting condition were apparently higher than for the colored LED group. Thus, we conclude that even though our implementation of colored light for robotic storytelling did not yet positively influence recipients' perception of the storytelling experience, it did not confuse them in contrast to colored eye LEDs.

Regarding emotions, we found no main effect of time on *Positive Affect* or *Serenity* over the storytelling, leading to the rejection of **H2a** and **H2d**. However, we did observe a significant decrease in *Negative Affect* as well as a significant increase in *Joviality* across all tested conditions. Thus, we can accept **H2b** and **H2c**, revealing a decrease of negative emotions and an increase of particular positive emotions due to robotic storytelling of romantic stories. Nevertheless, we cannot determine whether the positive emotions inherent in the romantic story were adopted by the recipients as is suggested for human oral storytelling^[32], or whether robotic storytelling in general has a positive effect on recipients' emotions regardless of the story genre. Future research utilizing different genres should be carried out to investigate both possibilities: the role of story genre on changes in specific emotions, and the overall positive effects of robotic storytelling on emotions compared to traditional human storytelling.

Concerning the effect of lighting conditions on altered emotions, we identified a relationship between changes in emotion from pre- to post-measurement and lighting condition for *Serenity*, but not for the other examined emotions (**RQ1**). *Serenity* tended to increase in the conditions with colored and constant light, but tended to decrease in the condition without any additional lighting. *Serenity* is often referred to as a state of inner peace^[84] that is closely connected to harmony^[85]. As harmony is the final state that is achieved at the end of our romantic story the additional lighting might have reinforced this feeling in the recipients. However, since we did not find a difference between the constant and colored lighting, this effect might be explained by an increase in attention toward the robotic storyteller triggered by the presence of light, independent of its color. Our results reveal that although attentiveness decreased in all three conditions (**RQ2**), the decrease was descriptively strongest in the condition without any additional light which is in line with this explanation. Thus, it is possible that even minimal additional lighting placed near the robot can enhance the storytelling experience.

Finally, we found no influence of lighting conditions on robot perception (**RQ3**). Comparing this result with the findings of Steinhäusser and Lugin^[19], they reported a negative effect of applying colored light in the robot's eye LEDs on Animacy. This negative effect disappeared when expanding the illuminated surface, reinforcing the suggestion that room illumination is better suited as a color and light modality for robotic storytelling than usage of eye LEDs. However, while Steinhäusser *et al.*^[20] revealed a positive effect of colored light on robot perception in terms of perceived competence, we were not able to indicate a positive influence on robot perception in terms of *Anthropomorphism*, *Animacy*, *Likeability*, *Perceived Intelligence*, and *Perceived Safety*. While this finding is surprising given the apparent similarity of competence and perceived intelligence at first glance, Scheunemann *et al.*^[86] also reported a lack of correspondence between the two variables. It seems that, as storytelling is a social activity, the exploration of more socially oriented variables is more suitable. Future work should thus take further facets of robot perception into account to focus on the social aspects of the robotic storyteller. Similarly, the Anthropomorphism scale of the *Godspeed* questionnaire was already criticized for not covering subtle changes^[87] and being correlated to its other scales^[88], therefore other scales assessing anthropomorphism should be taken into account. Furthermore, our restrictive annotation process might have influenced our results. As fewer tokens were annotated with emotion labels and the majority of them were labeled as neutral and thus matched to white light only, less colored light was integrated into the storytelling compared to previous work by Steinhäusser *et al.*^[20]. It can therefore be suggested that our conservative consensus-based labeling approach may not have been suitable for this use case. Future iterations might rely on a lower number of annotators but consider all of their decisions.

In summary, while we found no significant differences between the lighting conditions on storytelling experience and robot perception, we observed a trend towards increased serenity in the colored and constant light conditions but a decrease in the condition without additional light. This suggests that emotional content from the story may have been transferred to recipients more effectively in the illuminated conditions, potentially due to enhanced focus induced by the lighting. Moreover, we indicated a general decrease of negative emotions and an increase of specific positive emotions over the storytelling in all conditions, suggesting a general positive effect of robotic storytelling on emotions, which is in line with related studies^[76,89] and highlights the potential of this field. However, regarding the integration of emotion-inducing colored light, applying guidelines from virtual environment research alone appears insufficient for robotic storytelling. Therefore, future work should empirically determine appropriate light colors that enhance and complement the emotional bodily expressions of a robotic storytellers, emphasizing the expressive potential of colored light.

5. Study II: Determining Light Colors for Emotion Expression

As integrating emotion-inducing colored light based on guidelines developed for virtual environments proved insufficient for improving robotic storytelling, we focused on reinforcing the expressiveness of the colored light in multimodal storytelling, facilitating the emotion recognition of the robotic storyteller. Therefore, we conducted an online study empirically derive light colors supporting emotional expressions of our robotic storyteller *Pepper*. While the video-based approach is not uncommon in related studies^[55,90-92], the advantages of robots' physical embodiment do particularly unfold in live and in-person interaction^[93-96]. In line, participants tend to prefer in-person studies over live studies with robots^[97]. However, video trials have been shown to be representative for in-person studies^[97-100].

We based our approach on works by Song and Yamada^[53] and Steinhäusser *et al.*^[89,91]. Working with a *Roomba* vacuum cleaner robot, Song and Yamada first derived emotional expressions separately via motion and light color. In a subsequent study both modalities were combined to compare the initial recognition rates of the unimodal expressions with the ones from the combined multimodal expressions, revealing that while some emotions were better recognized from the multimodal expression, some were better expressed using only motion. We adapted this approach by combining previously established emotional bodily expressions for the *Pepper* robot^[89] with light colors derived from the literature and comparing their recognition rates with those from the original study.

H3: Adding colored light facilitates emotion recognition from robotic expressions.

5.1 Materials

To acquire comparable recognition rates, we followed the empirical online approach of Steinhäusser *et al.* using pictures of the robot's emotional expressions illuminated by colored lights^[89,91].

5.1.1 Concept for emotion-driven light colors

In general, warm, i.e. more reddish, colors are suggested to be more stimulative^[101,102] and are therefore also associated with more active emotions of higher arousal^[103], whereas cool colors seem to be more calming and associated with passive moods^[102], but at the same time they attract more attention than warm colors^[101]. These effects were already transferred to new media such as mixed reality, as Betella *et al.*^[61] reported warm colors inducing more arousal than cold colors when used in an interactive colored floor. The associations between colors and emotions seem to be consistent across different media^[104], thus also findings on emotion-color-associations from other presentation modes than light are taken into account for conceptualizing our stimulus materials as they might transfer to light in supporting emotion recognition within robotic storytelling. Therefore, we combined findings from color theory, related study results, and findings from other media to conceptualize our color combinations for the online study. All color combinations derived from the literature are displayed in [Table 4](#).

Table 4. Color combinations derived from literature.

Emotion Label	Color Combination			Associations
	Version	Spotlight	Ambient Light	
Vigilance	1	white	pale blue	White: cleanliness and honesty ^[110] ; Blue: comfortable for vision ^[109]
	2	white	pale orange	Orange: upbeat ^[113] , energetic ^[102,103] , vitality ^[101] , vigilance ^[74]
	3	pale blue	pale yellow	Yellow: attention grabbing caution sign ^[102,111] , awareness ^[112] , focusing, feeling awake ^[109]
Admiration	1	pale green	warm yellow	Yellow: cheerful, celestial ^[102] ; Green: health ^[115] , trust ^[110] , admiration ^[74]

Amazement	2	warm orange	warm yellow	Orange: fanciness and beauty ^[101]
	3	pink	warm yellow	Pink: love, temptation ^[101] , and romance ^[114]
	1	pale blue	white	Blue: amazement ^[74] ; White: excitement ^[103]
	2	pale purple-blue	pale blue	Purple: magical and spiritual ^[116]
Ecstasy	3	warm yellow	white	Yellow: inspiring ^[102] , surprise ^[14]
	1	warm orange	bright yellow	Yellow: sun ^[111] , joy ^[17,67] , jolliness ^[101] , and happiness ^[14,54,108,109]
	2	pale yellow	bright yellow	Orange: high arousal and happiness ^[103] , hilarity and exuberance ^[102]
Loathing	3	green	bright yellow	Green: health ^[115] , refreshing ^[102]
	1	green-yellow	violet	Green & Yellow: jealousy ^[101,118,119] , envy ^[117,120] , toxicity ^[115] , disease ^[102] , ghastliness ^[102]
	2	orange	green-yellow	Orange: distress, being upset ^[123]
Terror	3	pink	orange	Violet: hatred ^[122] ; Pink: controversial ^[124]
	1	purple	blue	Red: danger ^[101,102] , signal to stop ^[126]
	2	blue	red	Blue: powerlessness ^[106]
Rage	3	red	purple	Purple: death ^[116] , insecurity ^[101] , loneliness and desperation ^[102]
	1	red	dark blue	Red: rage ^[102] , anger ^[14,107] , flushing with aggression ^[54] , hostility ^[15]
	2	red	red	Blue: powerlessness ^[106] , coldness ^[106]
Grief	3	red	white	White: emptiness and loneliness ^[108]
	1	white	gray	Blue: coldness ^[106] , powerlessness ^[106] , gloom ^[102] , sadness ^[14,103,107] , and sorrow ^[101]
	2	pale blue	gray	Gray: sadness ^[14]
	3	dark blue	dark blue	White: emptiness and loneliness ^[108]

One of the most popular color-emotion associations is the one between blue and sadness-related subemotions, which is also utilized in Disney's Pixar movie *Inside Out*^[105] for the figure Sadness. Blue is the coldest color of the spectrum^[106] and is referred to as the “quintessential color for powerlessness”^[106]. It gives the impression of gloom^[102] and is associated with low arousal^[103] as found in the associated subemotions of sadness^[103] and sorrow^[101]. Regarding its effects, blue tends to induce inertia^[106], and in the form of light, it can reduce pleasantness^[67]. In studies with human facial expressions, sad faces were associated with blueish^[14,107] but also grayish colors^[14]. In the context of robots, blue also was reported to be appropriate for depicting sadness^[13,17] and blue-purple for grief^[16]. Therefore, we created three versions of light combinations for grief integrating blue, gray, and white light—due to their association with emptiness and loneliness^[108] which are connected to grief.

However, blue is a manifold color, and “the slightest change in that color, therefore, can completely alter how you respond to it”^[106]. While dark blue often gives negative impressions, blue is also associated with the ocean and the sky, which can have comforting effects^[108]. Pale blue was also shown to be comfortable for vision^[109], supporting attention. We combined two of the versions with white, as it is associated with cleanliness and honesty^[110]. In contrast to pale blue, yellow is an attention-grabbing color, used as a caution sign not only by humans in industrial societies but also in nature^[102,111]. Moreover, yellow is associated with awareness^[112]. For vigilance, we tested both blue and yellow as well as their combination due to these attention-related effects. However, to make the yellow aesthetically more pleasing, we turned it into a pale yellow^[111]—a color that was already reported to benefit learning performance, aid task focus, and promote alertness^[109]. Lastly, as orange was referred to as upbeat^[113] and energetic^[102,103], is associated with vitality^[101], and also depicts vigilance in Plutchik's *Wheel of Emotions*^[74], so we also implemented a version with orange light.

Orange is also associated with fanciness and beauty^[101], thus we also used it for depicting admiration. We combined it with yellow, which has been reported to give cheerful and celestial impressions^[102] and was previously successfully used for a robot to display admiration. For another version, yellow was also combined with pink which is psychologically associated with love, temptation^[101],

and romance^[114]. Lastly, we combined yellow with green, as it signals health^[115], can be associated with trust^[110], and is used for admiration in Plutchik's *Wheel of Emotions*^[74].

For amazement, we again used the respective color from Plutchik's *Wheel of Emotions*—blue. We combined it with white, which is associated with excitement^[103], as well as purple which is referred to as magical and spiritual^[116]. Lastly, we combined white with inspiring yellow^[102], which has been reported to be associated with surprised faces^[14] and to increase the intensity of surprise in robots' faces^[54].

Yellow is furthermore identified with the Sun^[111] and therefore evolutionarily associated with joy^[17,67], jolliness^[101], and happiness^[14,54,108]. Yellow-colored environments even increase our heartbeat^[109]. Thus, it is not surprising that Disney used yellow for the character Joy in the Pixar movie *Inside Out*^[105]. We used yellow for our implementation of the most intense subemotion of joy, namely *ecstasy*, in our social robot, too, as suggested by Terada *et al.*^[16]. Again, the color's alarming effect can be moderated by desaturation^[111], as pale yellow is still associated with happiness^[109], leading to our first combination. Next, orange is also associated with high arousal emotions such as happiness^[103], but also with hilarity and exuberance^[102], providing the base for our second color combination for *ecstasy*. Lastly, we combined yellow with green, which signals health^[115] and gives a refreshing impression^[102]. All color combinations were balanced in their tone^[67].

However, yellow is also associated with negative traits such as jealousy^[101] and envy in some cultures^[117]. More popular is the connection between green and jealousy or related emotions^[118,119], due to the proverbial “green-eyed monster” or “green with envy”^[120]. Green also signals poison or toxicity, for instance in spoiled meat^[115], and disease^[102], a connection again utilized in Disney's Pixar movie *Inside Out*^[105] for the character design of Disgust as well as for the people-hating Grinch^[121]. Also, human faces depicting disgust were shown to be associated with both yellow and green^[14]. Moreover, green is connected to ghastliness^[102]. Given these associations, we used a green–yellow tone for the most intense sub-emotion of disgust, namely loathing. We combined it with violet, as it is reported to be associated with hatred^[122], and is used within Plutchik's *Wheel of Emotions* for illustrating loathing. In a second version, we combined it with orange, which is associated with distress and being upset^[123]. In our last version, we combined orange with pink, as it is one of the most controversial colors^[124], evoking strong feelings^[125], which might reinforce the depicted emotion.

For terror, we created two light color combinations using red, as it is directly connected to danger^[101,102], signaling us to stop because of dangerous circumstances^[126]. This connection can even be visualized as the human heartbeat increases in red environments^[109,126]. As the combination with unbalanced light temperatures creates imbalance and tension^[67], we combined red with blue, which forms the other end of the light spectrum^[102] and is associated with powerlessness^[106]. Moreover, we combined red with purple, which is associated with death, especially in movies^[116], and gives the impression of insecurity^[101], loneliness and desperation^[102]. Furthermore, purple was already used to display terror in a robot^[16].

For rage, again red is a prominent color^[102], that is associated with anger across diverse cultures^[14,107]. This can easily be explained by the human face getting red when in anger and aggression^[54]. But red is not only a sign of anger and related feelings, it also induces it^[126]. Also in HRI, several studies reported that adding the color red to a robot's expression helps to recognize the intended emotion of anger^[13,17] or rage^[16], increases the perceived intensity of the depicted emotion^[54], and shows hostility^[15]. Therefore, we used red light for all tested variations of colored light combinations. For diversity, we also combined it with white and blue light given the above-described associations with negative states. For the neutral expression, only one version of colored light was created. We used white light, as the color was already used for expressing neutrality within a robot^[107]. Furthermore, the lack of light was viewed unfavorably in our initial study, so we proceeded with white lighting to maintain consistent brightness.

5.1.2 Implementation

We created three versions of each of the eight original emotional expressions^[89] for the eight inner subemotions in Plutchik's model, using the different light colors determined in Section 5.1.1. We used the same lighting setup as in Study I—one spotlight and one ambient light—, the original pose library by Steinhäusser *et al.*^[89] within *Choregraphe* 2.5.10.7^[127], and two *Tapo* E27 bulbs controlled with the *Tapo-Link* smartphone application. Following this approach, 25 pictures of the robot—three for each emotion and one for “neutral”—executing the original bodily expressions illuminated by the colored lights were taken with constant camera setup and background. All 25 pictures can be found in our online repository (<https://dx.doi.org/10.58160/sr1wkzg92vewvry4>) as well as in appendix B. Representative pictures are displayed in Figure 3.

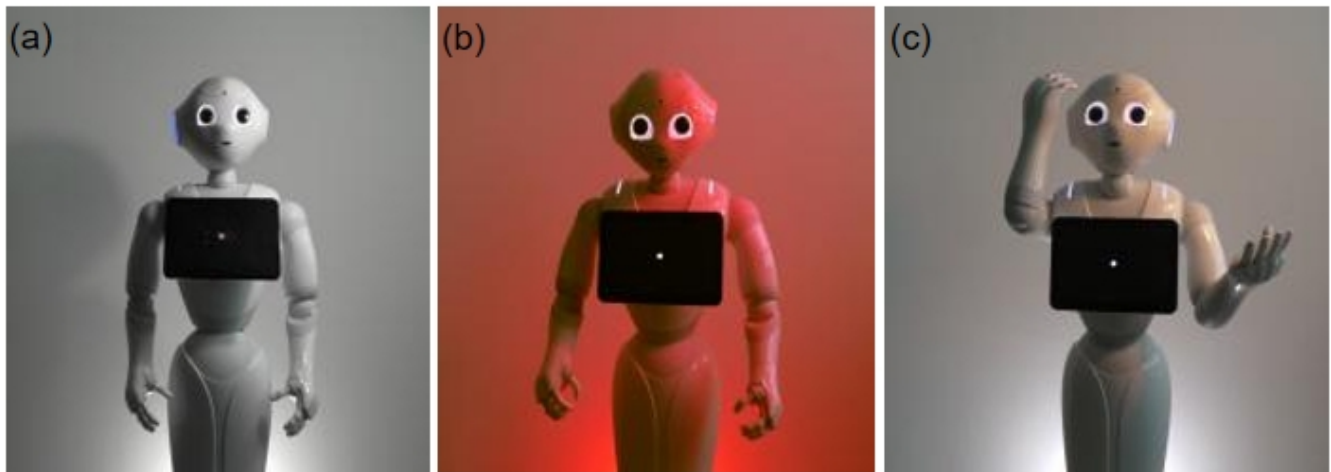


Figure 3. Exemplary pictures of stimulus material from Study II. (a) Picture with neutral expression and white ambient and spotlight; (b) Picture with rage expression and red ambient and spotlight; (c) Picture with amazement expression and white ambient and yellow spotlight.

5.2 Methods

In an online survey, participants were asked to assign one emotion label to each of the pictures showing bodily expressions combined with colored light for validation.

5.2.1 Measures and procedure

Upon entering the webpage hosted using *LimeSurvey*^[128], participants first gave informed consent. Within the survey, one picture of a combined expression was presented without any contextual information, along with a list of the eight inner Plutchik emotions and two other options; a “neutral” label and the answer “I don’t know”. Participants were asked to choose what the depicted expression might represent. After they chose the emotion label they thought fit best, they proceeded to the next one. At the end of the survey, participants were asked to provide demographic data.

5.2.2 Participants

In total, 101 persons participated in the survey. Twenty participants self-reported as male (age: $M = 23.45$, $SD = 2.19$), eighty as female (age: $M = 20.85$, $SD = 1.62$) and one participant identified as diverse (age: 21). The overall mean age was 21.37 years ($SD = 2.01$).

5.3 Results

All analyses were carried out using *Microsoft Excel 2016* and *JASP*^[83] (version 0.16), and a significance level of .05. Recognition rates achieved in Study II are presented in [Table 1](#), including rates with all labels (column 3) and rates excluding the “I don’t know” option (column 4), in comparison to the original study, which used the bodily expressions without light (column 5). In general, multinomial tests showed that label assignment frequencies differed significantly across all 25 pictures, $ps < .001$. In addition, we statistically compared the assignment frequency of the intended emotion versus other emotion labels across the three color combinations for each target emotion using χ^2 -tests. In addition, we statistically compared the assignment frequency of the intended emotion versus other emotion labels across the three color combinations for each target emotion using χ^2 -tests.

For **vigilance**, none of the color combinations was assigned to the respective label as the most frequently. For version 1, most participants chose the label “neutral” ($n = 22$), while for version 2 “amazement” ($n = 20$) and for version 3 “terror” ($n = 22$) were chosen most frequently. We found no significant relationship between recognition rates and color combination, $\chi^2(2) = 1.97$, $p = .374$. Similarly, all versions of the **admiration** expression scored higher on other emotion labels than on “admiration”. Version 1 was most frequently associated with “vigilance” ($n = 30$), version 2 with “amazement” ($n = 28$), and version 3 again with “vigilance” ($n = 29$). Again, no significant relationship between recognition rates and color combination was obtained, $\chi^2(2) = 1.97$, $p = .373$. Lastly, also all color combinations created for **loathing** were not assigned to the respective label most frequently, all versions scored highest on “terror” ($n_{v1} = 56$, $n_{v2} = 57$, $n_{v3} = 57$). No significant association between recognition rates and color combination was found, $\chi^2(2) = 0.86$, $p = .650$.

For **amazement**, all color combinations were most frequently assigned to the respective label, with version 3 exceeding the original recognition rate of the unimodal bodily expression when excluding the “I don’t know” label from the calculation. However, our test results indicated no significant relationship between recognition rates and color combination, $\chi^2(2) = 3.67$, $p = .186$. For **rage**, also all combinations were mostly recognized as “rage”, with version 2 exceeding the original recognition rate from the unimodal

expression, and a significant relationship between recognition rate and color combination, $\chi^2(2) = 6.91, p = .032$. In contrast, for **ecstasy**, again all color combinations were assigned to the respective emotion label most frequently with a significant relationship of recognition rates and color combination ($\chi^2(2) = 6.05, p = .049$), but the recognition rates were lower than in the original study. The same pattern was found for grief, however, the relationship between the recognition rates and the color combinations did not reach significance, $\chi^2(2) = 5.49, p = .064$. Lastly, the **neutral** multimodal expression was most often assigned to the “neutral” label, exceeding the original unimodal recognition rate.

5.4 Discussion

We conducted an online survey to determine which light colors supporting the recognition of emotions expressed by our robotic storyteller. Our results show that only three examined expressions achieved improved recognition rates compared to the initial unimodal expressions^[89]. This was the case for the two emotional expressions of amazement and rage as well as the neutral expression not conveying any emotion. For vigilance, admiration, ecstasy, terror, and grief the recognition rates declined when adding the light modality with the highest drop for terror. For loathing, the added lighting even led participants to assign a completely different emotion. Therefore, H3 can be partly accepted.

The significant difference in recognition rates among the three versions of rage reflects the increased recognition rate compared to the unimodal expression being only evident in versions two and three, in which red light was added to the expression. This finding is in line with related works. Song and Yamada^[53] as well as Lffler *et al.*^[17] already revealed the high impact of the colored lighting for this emotion. This might be explained by the human bodily reaction of facial flushing when enraged^[17,54]. Our extension of the color to the room’s illumination might have further resembled the visual alarm signal associated with red^[126], which is even more visually intense^[126] than when it was only attached to the robot. In contrast, no significant differences were found between recognition rates for the three versions of the sub-emotion amazement, the most intense level of the basic emotion surprise^[24,74]. However, slightly improved recognition rates were observed when being depicted with yellowish spotlight and white ambient light, an association in line with findings for colored framing of human facial expressions^[14], although such a color cue has no direct analogue in human communication. This finding is even more interesting given that while rage or the related basic emotion of anger is easily recognized from a robot’s bodily expressions^[57,91,129], the recognition of amazement seems to be rather difficult^[91]. Also, Terada *et al.* reported that when using only the light modality amazement often “might be recognized as surprise, joy, or expectation”^[16]. Although our recognition rates showed only minor improvements after excluding uncertain answers from the dataset, this type of confounding did not occur in our study, strengthening the idea of combining bodily expressions with colored light.

In a related study, Löffler *et al.*^[17] also reported that joy, fear, and sadness to benefit from the color modality which was not the case with the related emotions investigated in our survey. In addition, although there was a significant relationship between recognition rate and color combination for ecstasy, suggesting a better suitability of the green/yellow combination over the others, the recognition rate was lower than for the unimodal expression. One possible explanation might be the intensities of the emotions in question, which were higher for our sub-emotions utilized in our study. Most surprisingly given the literature is the decrease of the recognition rate for grief when adding colored light. Related works all agree that blue is closely connected to sadness and sorrow^[14,16,17,101] and has already been shown to be sufficient for emotion expression with robots^[13,53]. In our case, the dark blue light did not improve emotion recognition. However, next to the higher intensity inherent to our emotion label, the recognition rate of the original unimodal expression for grief was already so high, that it was hard to surpass.

Lastly, the notable increase in recognition for the neutral expression combined with white light aligns well with our expectations. Participants seemed to interpret white as the default illumination, probably because the modality of light cannot be omitted when the robot must remain visible. Similar to other modalities such as body language, light may be considered a modality that “cannot not communicate”^[130]. It seems to convey a message, even if it is only improving emotion recognition for single states.

Given these results, we recommend the use of colored lighting to support emotion recognition from robotic bodily expressions for the categories of amazement, rage, and neutral. The exact color combinations and HSB values can be retrieved from Table 5.

Table 5. HSB values for finally chosen color combinations. Brightness was always set to 100.

	Spotlight		Ambient Light	
	Hue	Saturation	Hue	Saturation
Amazement	47	90	0	0
Rage	2	100	2	100
Neutral	0	0	0	0

6. Study III: Evaluating Colored Lights in Robotic Storytelling

While related studies show positive effects of integrating colored light in robotic storytelling^[20], the results from our first user study in this paper were relatively inconclusive, showing only a positive effect on individual emotional states. Within both the former work by Steinhäusser *et al.*^[20] and our own first user study, decisions on colors associated with story emotions were based on guidelines for designing emotion-inducing virtual environments. Nevertheless, given our results, this approach does not appear to be effective with the Pepper robot when using emotional body language. Therefore, we revised our color-emotion associations within an empirical online study. As the light-supported multimodal expressions had so far only been examined in isolation in the online setting, we conducted a follow-up laboratory user study to examine them in the context of robotic storytelling.

It might be the case that not the color-emotion-associations themselves, but the entire concept of emotion-driven lighting is not appropriate for improving robotic storytelling. While studies show that robotic storytellers should display emotions^[11,49], and that colored lights can improve a robot's expressiveness^[13,17], context-based room illumination might improve the storytelling experience more effectively than emotion-driven lighting. For audiobook reception, studies show that visual anchors depicting story context such as persons, actions, or environments added by augmented reality can improve not only recall but also narrative engagement without influencing users' cognitive load^[134]. While colored light cannot depict persons or actions, it can represent environments, for example illustrating the greenery of a forest or the blueness of the ocean^[103]. Therefore, we compared our emotion-based approach for integrating colored light into robotic storytelling to a context-based approach using colored light to illustrate story environments. As control conditions, we again used a constant white light as well as a version without any additional lighting. Therefore, we implemented four conditions: (1) emotion-based lighting, (2) context-based lighting, (3) constant lighting, and (4) no additional lighting. To allow recipients to express their liking for the different approaches in comparison, we used a within-subjects design in a laboratory in-person study. All sub-studies were deemed ethically sound by the local ethics committee.

Based on our previous studies and related work we postulate the following hypotheses:

- **H4:** Adding emotion-based or context-based colored lighting will enhance storytelling experience compared to constant or no additional lighting.
- **H5:** Adding emotion-based or context-based colored lighting will increase emotion induction compared to constant or no additional lighting.
- **H6:** Adding emotion-based or context-based colored lighting will improve robot perception compared to constant or no additional lighting.

Moreover, we were interested in the differences between the emotion-based and context-based approach:

- **RQ4:** Does context-based colored lighting outperform emotion-driven colored lighting?
- **RQ5:** Which presentation style is preferred by recipients?

6.1 Materials

This time, we utilized stories from the fantasy genre, as the importance of the environmental setting is higher in this genre^[132], and thus fantasy stories typically include more details on the places and environments in which the story unfolds.

6.1.1 Concept

Due to the widespread popularity^[133] and the positive remarks reported in related studies^[89], we selected the Harry Potter universe as the narrative framework for our study. We used short stories from the *Wizarding World's* website (former Pottermore) written by J.K. Rowling. These stories provide descriptions of the lesser-known wizarding schools.

Story Selection. Since we used a within-participants design, four different stories were assigned to the four conditions. We conducted an online survey to ensure that the four stories did not differ in story liking, storytelling experience, or emotion induction. To ensure we had four comparable stories, we tested five stories, keeping one as a backup. All five stories, *Uagadou*^[134], *Beauxbatons*^[135], *Castelobruxo*^[136], *Durmstrang*^[137], and *Mahoutokoro*^[138], were approximately equal in length.

The survey followed a within-subjects design. When accessing, participants provided informed consent. They then read the first short story presented as plain text. Afterwards, they answered an attention check question about story detail, completed the previously used questionnaires on *Transportation*^[79] and emotions^[80] (i.e. *Positive* and *Negative Affect*), rated their liking for the story ("I liked the story.") on a five-point Likert scale (1: "strongly disagree" and 5: "strongly agree"), and stated whether they knew the story before. This process was repeated for each story with the order of stories randomized. Lastly, participants provided demographic data and were thanked for their participation.

Fifteen participants with a mean age of 21.93 years ($SD = 2.28$) took part in the study. Four participants self-identified as male

(age: $M = 23.00$, $SD = 1.83$), while ten participants self-identified as female (age: $M = 21.60$, $SD = 2.50$). One participant identified as diverse gender (age: 21).

A Bayesian RM-ANOVA was conducted for Transportation, the Bayes factor indicated that data best support the null model ($BF_{10} = 1.00$). The model including the main effect of story was less likely ($BF_{10} = 0.15$), suggesting evidence for H_0 . The same pattern was found for both *Negative Affect* ($BF_{10} = 0.21$ for the model including the effect of story), and *Positive Affect* ($BF_{10} = 0.08$), and story liking ($BF_{10} = 0.09$). Thus, we conclude that the data do not support the existence of differences between the stories. Descriptive data are displayed in Table 6.

Table 6. Descriptive data of the pre-study on story selection.

	Beauxbaton		Castelobruxo		Durmstrang		Mahoutokoro		Uagadou	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Transportation ^a	4.27	1.32	4.11	1.11	4.61	1.35	4.19	1.33	4.21	1.54
Positive Affect ^b	2.42	1.03	2.37	0.76	2.36	0.99	2.49	0.73	2.36	0.77
Negative Affect ^b	1.09	0.17	1.14	0.29	1.18	0.26	1.08	0.11	1.13	0.33
Story Liking ^b	3.47	1.30	3.53	1.06	3.73	1.10	3.67	1.05	3.47	1.13

^a: Calculated values from 1 to 7; ^b: calculated values from 1 to 5.

We discarded the short story about the *Durmstrang Institute* due to its more negative tone compared to the others. Thus, we used the stories about *Uagadou*, *Beauxbatons*, *Castelobruxo*, and *Mahoutokoro* within our study.

Emotional Annotation. For the process of annotating emotions to the four short stories, we followed the manual annotation approach by Steinhäusser *et al.*^[89,91]. We used the same tokenization method and emotion labels as in Study I (Section 4.1.1). Thirteen annotators with a mean age of 22.15 years ($SD = 2.15$) were acquired. Eleven annotators self-identified as female (age: $M = 21.64$, $SD = 1.86$), while two annotators self-identified as male (age: $M = 25.00$, $SD = 1.41$). No annotators identified as diverse gender.

They annotated the stories using an online survey. Again, the stories were presented individually in random order. Each story was displayed on one page with tokens presented line-by-line in the order they appeared, each accompanied by a drop-down menu with the eight emotion labels, a “neutral” label, and an “I don’t know” option. Annotators were asked to choose the emotion label they would want a storyteller to act out while speaking the token. We withdrew from using the restrictive consensus label-building process of Study I, as it led to a huge amount of “neutral” labels and, thus, less emotional storytelling. This resulting lack of emotions might have negatively influenced our results^[14]. Instead, we utilized a majority decision as a simpler process to derive the final annotations. As suggested by Steinhäusser *et al.*^[89] only emotion labels with higher confidence than random assignment value were considered. If two labels were assigned equally frequently, the token was labeled as “neutral”. The frequency of each emotion label’s final annotation can be retrieved from Table 7. The average agreement between the raters for the selected consensus labels, as well as their inter-rater agreement, is also displayed in Table 7. Fleiss’ Kappa indicated little agreement among the raters^[139], which is comparable to similar studies^[89,140] as emotion annotation is a highly individual and complex process^[89,91].

Table 7. Annotated emotion labels per story.

	Beauxbatons	Castelobruxo	Mahoutokoro	Uagadou
Vigilance	1	1	0	1
Admiration	8	8	4	10
Amazement	0	2	3	5
Ecstasy	1	0	2	1
Loathing	1	0	1	0
Terror	1	0	2	4
Rage	0	0	0	0
Grief	0	2	2	0
Neutral	21	27	21	29

Agreement consensus labels	51.28%	49.23%	44.66%	48.15%
Fleiss Kappa	.16	.09	.09	.11
<i>n</i> Tokens	33	40	36	50

Context-based Annotation. The story *Mahoutokoro* was randomly selected for context-sensitive light color annotation. The process for creating the context-based annotations of colors representing the described story environments followed the same survey method as the emotional annotation. Again, individual clauses were used for tokenization. As possible labels we used the colors from Itten's color wheel^[141] and added the options of brown and white. In total, 17 annotators with a mean age of 25.18 years ($SD = 9.46$) were recruited. While 15 of them self-identified as female (age: $M = 24.60$, $SD = 9.87$), only two annotators self-identified as male (age: $M = 29.50$, $SD = 4.95$). No one indicated being diverse gender.

The story was displayed on one page with a line-by-line presentation of the tokens in the order they appeared in the story. Each token was presented next to a drop-down menu with the color labels, the label "no color", and an "I don't know" option. Annotators were asked to choose the color from the list that might illustrate the locations and settings described in the story's text. They were given the example association, such as green with forest and blue with sky. Further, they were asked not to use emotional associations such as red with love. To create consensus labels, majority decision was used. If two or more colors were assigned equally frequently, the token was labeled as "no color". Doing so, 9 out of the 36 tokens had a color assigned, while 27 tokens were labeled as "no color". The colors assigned were white ($n = 3$), red-orange ($n = 1$), magenta ($n = 1$), dark blue ($n = 2$), brown ($n = 1$), and cyan ($n = 1$). The average agreement for the finally chosen consensus labels was 56.21%, while the inter-rater agreement was Fleiss' Kappa = 0.15.

6.1.2 Implementation

The storytelling sequences were again implemented for the *Pepper* robot. As technical issues arose using the implementation approach via *Unity*, we programmed both the robot and the smart lights using *Python* this time. We used the *NAOqi* *Python* module together with *Python 2.7* to control the *Pepper* robot and the *P100* library to control the *Tapo* smart lights. The module utilizes the robot's internal speech synthesis; no adjustments in voice modulation were made.

Version without Additional Light. We used the emotion-based annotation to match the robot's bodily expressions to the story tokens. For each token, a text-to-speech command was sent to the robot together with the respective bodily expression's values. We used the same expressions as in the former sub-studies. Following this approach, we created the base of all four storytelling versions. The story *Castelobrujo* was randomly chosen and not further modified to function as the first control condition without additional lighting. The resulting storytelling was about 1 minute and 50 seconds in length.

Version with Constant White Light. The short story about *Beauxbatons* was randomly chosen, and we implemented constant white light for both the spot and the ambient light. Therefore, we used 0 as the value for both hue and saturation and the maximum value of 100 for brightness. The resulting storytelling was of about 1 minute and 40 seconds in length.

Emotion-driven Version. For the storytelling with emotion-driven colored lighting, we randomly chose the short story about *Uagadou* from remaining stories. We added the color combinations empirically validated in Study II to the storytelling. Based on the annotations displayed in Table 7, the combination for amazement was used three times and the combination for neutral was applied for 21 tokens. As the emotion label of rage was not annotated to the story, the respective light color combination was also excluded from the implementation. For the rest of the tokens, the smart lights were turned off. The resulting storytelling was of about 2 minutes and 20 seconds in length.

Context-based Version. For the context-based condition in which the light colors were based on environments' descriptions within the story the remaining short story about *Mahoutokoro* was chosen. As described in Section 6.1.3, six colors were assigned to the story in total. The HSB values utilized per color can be retrieved from Table 8. They were used for both the ambient and spotlight. The resulting storytelling was of about 2 minutes and 20 seconds in length.

Table 8. Utilized HSB values for light colors in the context-based version.

	Hue	Saturation	Brightness	<i>n</i>
White	0	0	100	3
Red-Orange	15	100	100	1
Magenta	290	100	100	1
Dark Blue	245	91	50	2

Brown	30	100	30	1
Cyan	180	100	100	1
No Color	-	-	-	27

6.2 Methods

A within-subjects design with the four conditions: emotion-driven light (*EmoLight*), context-based light (*CxtLight*), constant white light (*ConLight*), and no additional light (*NoLight*) was applied to investigate effects of the different lighting approaches on storytelling experience, participants' emotions, and robot perception.

6.2.1 Measures

Storytelling experience was again assessed using the TS-SF^[79] measuring Transportation into the story and the CA questionnaire^[58] measuring *Cognitive Absorption*, as described in Study I. Reliabilities for the current study were .77 to .85 for *Transportation*. For the *Cognitive Absorption* subscales, Cronbach's alpha of .83 to .95 was achieved for *Temporal Dissociation*, .82 to .88 for *Focused Immersion*, .83 to .96 for *Heightened Enjoyment*, .36 to .83 for *Control*, and .83 to .90 for *Curiosity*, respectively. In addition, we used the *Narrative Engagement Scale*^[142] in its German version. The questionnaire comprises four subscales, each including three items. The subscale (1) *Narrative Understanding* assesses difficulties in grasping the storyline and characters (e.g., "I had a hard time recognizing the thread of the story."), whereas the subscale (2) *Attentional Focus* captures recipients' attention on the storytelling (e.g., "I found my mind wandering while the storytelling was on."). Further, the subscale (3) *Narrative Presence* measures recipients' perceived presence in the world created by the story told (e.g., "At times during the storytelling, the story world was closer to me than the real world."). Lastly, the (4) *Emotional Engagement* subscale assesses the storytelling's effects on recipients' emotions (e.g., "The story affected me emotionally."). All items were presented with a seven-point Likert scale anchored by 1: "strongly disagree" and 7: "strongly agree". Regarding reliabilities, the authors reported Cronbach's alpha of .58 to .78 for *Narrative Understanding*, .79 to .85 for *Attentional Focus*, .70 to .84 for *Emotional Engagement*, .69 to .86 for *Narrative Presence*, and .80 to .86 for the overall value^[142]. For the current sample values of .74 to .90 for *Narrative Understanding*, .89 to .93 for *Attentional Focus*, .62 to .83 for *Narrative Presence*, .75 to .84 for *Emotional Engagement*, and .89 to .93 for the overall value were computed.

Participants' **emotions** were again queried using the German PANAS-X^[143] from Study I. However, in addition to the subscales utilized in Study I (*Positive Affect*, *Negative Affect*, *Joviality*, *Serenity*, and *Attentiveness*), we further used the subscales measuring *Fear* (6 items, e.g. "jittery", $\alpha = .87$ ^[81]), *Sadness* (5 items, e.g. "blue", $\alpha = .86$), and *Surprise* (3 items, e.g. "amazed", $\alpha = .80$) to cover the potentially more diverse emotions evoked by the fantasy stories. Cronbach's alpha obtained for the present sample were .83 to .90 for *Positive* respectively .69 to .86 for *Negative Affect*, .94 to .96 for *Joviality*, .40 to .90 for *Serenity*, .31 to .76 for *Fear*, .34 to .63 for *Sadness*, .82 to .89 for *Surprise*, and .66 to .81 for *Attentiveness*.

Robot perception was again measured using selective scales from the *Godspeed* questionnaire^[82] from Study I. Due to criticism on its Anthropomorphism scale, we additionally applied the questionnaire on perceived Robot Morphology regarding Anthropomorphism (RoMo-A)^[87]. It includes four scales each comprising four items, referring to the robot's (1) *Appearance*, e.g., "How human-like is the robot's external appearance?", (2) *Motion*, e.g., "How human-like is the speed of the robot?", (3) *Communication*, e.g., "How human-like is the speech rhythm of the robot?", and (4) *Context*, e.g., "How human-like is the robot's task?". All items were answered using a slider anchored by 0%: "not at all" and 100%: "fully". Cronbach's alpha reported by the authors are .90 for *Appearance*, .84 for *Motion*, .88 for *Communication*, and .81 for *Context*. For the current sample, computed values were .91 to .94 for *Appearance*, .51 to .64 for *Motion*, .59 to .81 for *Communication*, and .83 to .90 for *Context*. From the *Godspeed* questionnaire series^[82], only the scales *Animacy* and *Likeability* were used. Calculated reliabilities were .65 to .72 and .85 to .91, respectively. To further cover aspects of the social perception of the storyteller, we additionally used the *Robotic Social Attributes Scale* (RoSAS) by Carpinella *et al.*^[144]. It measured the social perceptions of robots on the three scales (1) *Warmth* (e.g., "feeling"), (2) *Competence* (e.g., "reliable"), and (3) *Discomfort* (e.g., "awkward") with six items each. All items are presented with a nine-point Likert scale anchored by 1: "definitely not associated" to 9: "definitely associated". We used a German translation that was previously used in other works^[20,145]. Reliabilities reported by Carpinella *et al.* are .91 to .92 for *Warmth*, .84 to .95 for *Competence*, and .82 to .90 for *Discomfort*. Cronbach's alpha for the current study were .80 to .83, .79 to .88, and .82 to .84, respectively.

Additionally, we asked questions on participants' former knowledge of each story ("I already knew the story", yes/no), liking of the story ("I liked the story", 1: "Do not agree" to 5: "Strongly agree") and its presentation ("I liked the presentation style", 1: "Do not agree" to 5: "Strongly agree") and they were invited to comment on their answers.

At the end of the questionnaire, we applied a manipulation check on the different lighting strategies ("Did you recognize the different lighting strategies", yes/no), asked which strategy the participants liked best, and whether they found them appropriate. Moreover, we asked which story they liked best, and asked them to provide demographic data (e.g., age, gender, liking of fantasy stories).

6.2.2 Procedure

After arriving at the laboratory, participants first provided informed consent to participate in the study. Next, they filled in the pre-questionnaire on their current emotional state. Following a counterbalanced design, with each possible order being conducted once, they received the first storytelling condition and afterward answered the questionnaires on storytelling experience, evoked emotions, robot perception, and questions on the story and its presentation itself. This process was repeated for all four conditions. After receiving the last condition, participants also answered the manipulation check and the single questions described above and provided demographic data and optional comments. Lastly, they were thanked for their participation and farewelled.

6.2.3 Participants

In total, 24 persons with a mean age of 21.67 ($SD = 2.04$) took part in the study. 9 of self-identified as male (age: $M = 22.33$, $SD = 1.41$), whereas 15 self-identified as female (age: $M = 21.27$, $SD = 2.28$). No one self-identified as diverse gender.

Regarding their media reception behavior, 21 of the participants stated they like to receive fantasy stories. Participants expressed liking for the Harry Potter universe with 4.21 ($SD = 1.18$) on a five-point scale.

6.3 Results

All analyses were carried out using JASP version 0.19.1 and a significance level of .05. Descriptive data are displayed in Table 9. Test results, except for post-hoc analyses are displayed in Table 10. Calculated Shapiro-Wilk tests indicated violation of the normality assumption for all subscales except for *Transportation*, the overall *Narrative Engagement* score, *Focused Immersion*, *Joviality*, *Attentiveness*, anthropomorphic *Appearance* and *Context*, *Animacy*, *Likeability*, *Warmth*, and *Competence*. For all other variables, non-parametric tests were calculated. Further, Mauchly's tests indicated sphericity only for *Transportation*, the overall *Narrative Engagement* score, *Focused Immersion*, anthropomorphic *Appearance* and *Context*, *Likeability*, and *Competence*. For the other variables Greenhouse-Geisser correction was applied. Concerning former story knowledge, only two people indicated to have previously received the story about *Beauxbatons*.

Table 9. Descriptive data for study III.

	Emotion-based Light		Context-based Light		Constant Light		No Additional Light	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Transportationa	3.13	1.09	3.03	1.29	3.05	1.35	2.93	1.17
NE: Narrative Understandinga	3.49	1.30	3.25	1.89	3.43	1.80	3.24	1.60
NE: Attentional Focusa	3.75	1.63	3.86	1.76	4.29	1.95	3.93	1.75
NE: Narrative Presencea	2.46	1.00	2.42	1.24	2.43	1.28	2.29	1.33
NE: Emotional Engagementa	2.04	1.10	2.17	1.11	2.04	1.15	2.11	1.11
Overall NEa	2.93	1.02	2.92	1.24	3.05	1.31	2.89	1.14
CA: Temporal Dissociationa	3.56	1.76	3.81	1.69	3.81	1.50	3.50	1.53
CA: Focused Immersiona	3.58	1.30	3.63	1.34	3.73	0.97	3.87	1.31
CA: Heightened Enjoymenta	4.69	1.72	4.58	1.59	4.30	1.82	4.27	1.79
CA: Controla	2.51	1.16	2.37	0.91	2.64	1.00	2.22	0.93
CA: Curiositya	4.60	1.64	4.67	1.48	4.08	1.59	4.50	1.45
Positive Affect b	2.52	0.67	2.58	0.76	2.52	0.83	2.46	0.78
Negative Affect b	1.20	0.29	1.25	0.35	1.25	0.44	1.25	0.36
Joviality ^b	2.73	0.88	2.82	1.02	2.66	1.14	2.54	0.99
Serenity ^b	3.78	0.81	3.65	0.73	3.67	0.94	3.56	0.88
Fear b	1.27	0.30	1.40	0.48	1.34	0.46	1.31	0.41
Sadness ^b	1.19	0.25	1.23	0.32	1.36	0.52	1.26	0.34
Surprise ^b	2.25	0.98	2.43	1.08	2.11	0.84	2.21	1.08

Attentiveness ^b	2.92	0.75	3.02	0.70	2.92	0.87	2.98	0.80
Anthro. Appearance ^c	30.14	21.87	29.49	20.75	29.46	21.60	28.99	19.95
Anthro. Motion ^c	21.06	12.39	19.07	12.63	20.07	13.12	20.61	12.93
Anthro. Communication ^c	24.97	19.91	28.10	20.22	22.85	16.13	20.70	14.26
Anthro. Context ^c	52.46	23.51	51.54	24.76	52.10	24.72	51.22	25.34
Animacy ^b	2.50	0.57	2.48	0.50	2.26	0.63	2.42	0.57
Likeability ^b	3.74	0.92	3.94	0.72	3.73	0.85	3.65	0.88
Warmth ^d	3.83	1.52	3.81	1.32	3.41	1.45	3.53	1.53
Competence ^d	5.88	1.49	6.07	1.64	6.00	1.75	5.76	1.79
Discomfort ^d	2.40	1.38	2.40	1.33	2.50	1.43	2.24	1.27
Story Liking ^b	3.50	0.93	3.29	0.86	3.29	1.27	3.13	1.12
Presentation Liking ^b	2.88	1.12	3.04	1.16	2.75	0.99	2.29	1.00

NE: Narrative Engagement; CA: Cognitive Absorption; Anthro.: Anthropomorphi; a: Calculated values from 1 to 7; b: Calculated values from 1 to 5; c: Calculated values from 0 to 100; d: Calculated values from 1 to 9.

Table 10. Test results for study III with planned contrasts between experimental (*EmoLight*, *CxtLight*) and control (*Con-Light*, *NoLight*) conditions.

	Repeated measures test			Planned contrast		
	Statistic1	<i>p</i>	effect size2	<i>t</i>	<i>p</i>	<i>d</i>
Transportation ^a	0.32	.813	.00	-0.69	.499	.16
NE: Narrative Understanding ^c	1.08	.783	.02	-0.21	.839	.04
NE: Attentional Focus ^c	7.27	.064	.10	1.32	.200	.35
NE: Narrative Presence ^c	0.93	.818	.01	-0.53	.603	.13
NE: Emotional Engagement ^c	0.10	.992	.00	-0.20	.847	.05
Overall NE ^a	0.20	.895	.00	0.31	.761	.07
CA: Temporal Dissociation ^c	2.47	.481	.03	-0.14	.890	.03
CA: Focused Immersion ^a	0.56	.643	.00	1.25	.226	.31
CA: Heightened Enjoyment ^c	2.48	.479	.03	-1.69	.105	.40
CA: Control ^c	6.27	.099	.09	-0.10	.920	.03
CA: Curiosity ^c	6.63	.085	.09	-1.96	.062	.44
Positive Affect ^c	6.27	< .001***	.20	-0.72	.478	.16
Negative Affect ^c	3.23	.521	.03	0.72	.480	.12
Joviality ^b	6.77	.001**	.04	-1.67	.110	.36
Serenity ^c	4.70	.320	.05	-1.02	.320	.26
Fear ^c	1.64	.802	.02	-0.14	.889	.03
Sadness ^c	16.05	.003**	.17	1.47	.155	.52
Surprise ^c	21.56	< .001***	.23	-1.36	.186	.38
Attentiveness ^b	6.77	.005**	.04	-0.24	.814	.06
Anthro. Appearance ^a	0.18	.912	.00	-0.44	.662	.06

Anthro. Motion ^c	1.91	.592	.03	0.20	.846	.04
Anthro. Communication ^c	3.77	.288	.04	-2.22	.037	.54
Anthro. Context ^a	0.16	.925	.00	0.25	.805	.03
Animacy ^b	2.17	.131	.02	-1.41	.172	.53
Likeability ^a	2.20	.096	.01	-1.41	.171	.36
Warmth ^b	1.73	.189	.01	-1.70	.102	.49
Competence ^a	1.28	.287	.00	-0.82	.420	.11
Discomfort ^c	0.82	.844	.01	-0.33	.747	.05

NE: Narrative Engagement; CA: Cognitive Absorption, Anthro.: Anthropomorphic; a: Calculated RM-ANOVA; b: Calculated RM-ANOVA with Greenhouse-Geisser correction; c: Calculated Friedman test; 1: F provided for RM-ANOVAs, χ^2 provided for Friedman tests. 2: ω^2 provided for RM-ANOVAs, W provided for Friedman tests; * $p < .05$; ** $p < .01$; *** $p < .001$.

Regarding **storytelling experience**, a repeated-measures ANOVA indicated no significant differences between the four conditions in terms of *Transportation*. A planned contrast comparing both experimental conditions (*EmoLight* and *CxtLight*) to the control conditions (*ConLight* and *NoLight*) also revealed no significant difference. Concerning *Narrative Engagement*, calculated tests also showed no significant differences on any of the subscales nor for the overall *Narrative Engagement* score. Planned contrasts also showed no significant differences comparing experimental and control conditions. The same pattern was found for all subscales of *Cognitive Absorption*.

Concerning participant's **emotions** depicted in Figure 4, for *Positive Affect*, a significant effect of time of measurement was identified. Bonferroni-holm-corrected Conover's post hoc tests indicate a significant decrease of *Positive Affect* from the pre-measurement ($M = 2.97$, $SD = 0.69$) to all other conditions, namely *EmoLight* ($p = .011$), *CxtLight* ($p = .011$), *ConLight* ($p = .003$), and *NoLight* ($p < .001$). No other paired comparisons yielded significant differences. Alike, a planned contrast between the experimental and control conditions revealed no significant difference. Regarding *Negative Affect*, no significant differences between the conditions or to the pre-measurement ($M = 1.23$, $SD = 0.22$) were found. Also the planned contrast between experimental and control conditions yielded no significant difference. For *Joviality*, again a significant effect of time of measurement was revealed. Again, Bonferroni-holm-corrected Conover's post hoc tests showed a significant decrease from pre-measurement ($M = 3.18$, $SD = 0.77$) to all of the four conditions, *EmoLight* ($p = .011$), *CxtLight* ($p = .011$), *ConLight* ($p = .043$), and *NoLight* ($p = .004$). No further significant differences between conditions were found. Contrasting the experimental and control conditions, again no significant difference was obtained. In contrast, no significant main effect between conditions or pre-measurement ($M = 3.93$, $SD = 0.60$) was indicated for *Serenity*. In addition, a planned contrast between experimental and control conditions yielded no significant difference. Similarly, no significant differences were obtained for *Fear* between the conditions or pre-measurement ($M = 1.33$, $SD = 0.37$). Further, a planned contrast between experimental and control conditions indicated no significant difference. In contrast, a significant main effect was found for *Sadness*. Bonferroni-holm-corrected Conover's post hoc tests revealed significant higher values of pre-measured *Sadness* ($M = 1.40$, $SD = 0.32$) compared to after the conditions *EmoLight* ($p = .002$), *CxtLight* ($p = .018$) and *NoLight* ($p = .039$). No other paired comparisons yielded significant differences. In addition, a planned contrast between experimental and control conditions yielded no significant difference. Alike, a significant main effect was obtained for *Surprise*. Bonferroni-holm-corrected Conover's post hoc tests showed significant increases from pre-measurement ($M = 1.43$, $SD = 0.72$) to all conditions, *EmoLight* ($p = .001$), *CxtLight* ($p < .001$), *ConLight* ($p = .003$), and *NoLight* ($p = .002$). No further pairwise comparisons yielded significant differences. A planned contrast comparing experimental and control conditions indicated no significant difference. Lastly, for *Attentiveness* again a significant main effect was obtained. Bonferroni-holm-corrected Conover's post hoc tests revealed significantly lower values of *Attentiveness* for *EmoLight* compared to the pre-measurement ($M = 3.41$, $SD = 0.66$), $p = .012$. No further significant differences in pairwise comparisons were found. Further, a planned contrast between experimental and control conditions yielded no significant difference.

Regarding **robot perception**, neither anthropomorphic *Appearance*, *Motion*, *Communication*, nor *Context* differed significantly between the conditions. While planned contrasts comparing the experimental to the control conditions yielded no significant differences for anthropomorphic *Appearance*, *Motion*, and *Context*, a significant difference was found contrasting experimental and control conditions in terms of anthropomorphism in *Communication* with higher values in the experimental conditions with colored light (*EmoLight* and *CxtLight*) compared to the control conditions (*ConLight* and *NoLight*). Alike, for *Animacy* as well as for *Likeability* no significant group differences were found. Similarly, the planned contrasts between experimental and control groups yielded no significance for both *Animacy* and *Likeability*. Lastly, regarding the robot's social attributes, no significant differences between the conditions were obtained for *Warmth*, *Competence*, or *Discomfort*. Alike, planned contrasts comparing the experimental and control groups did not reveal significant differences, neither for *Warmth*, nor for *Competence*, or *Discomfort*.

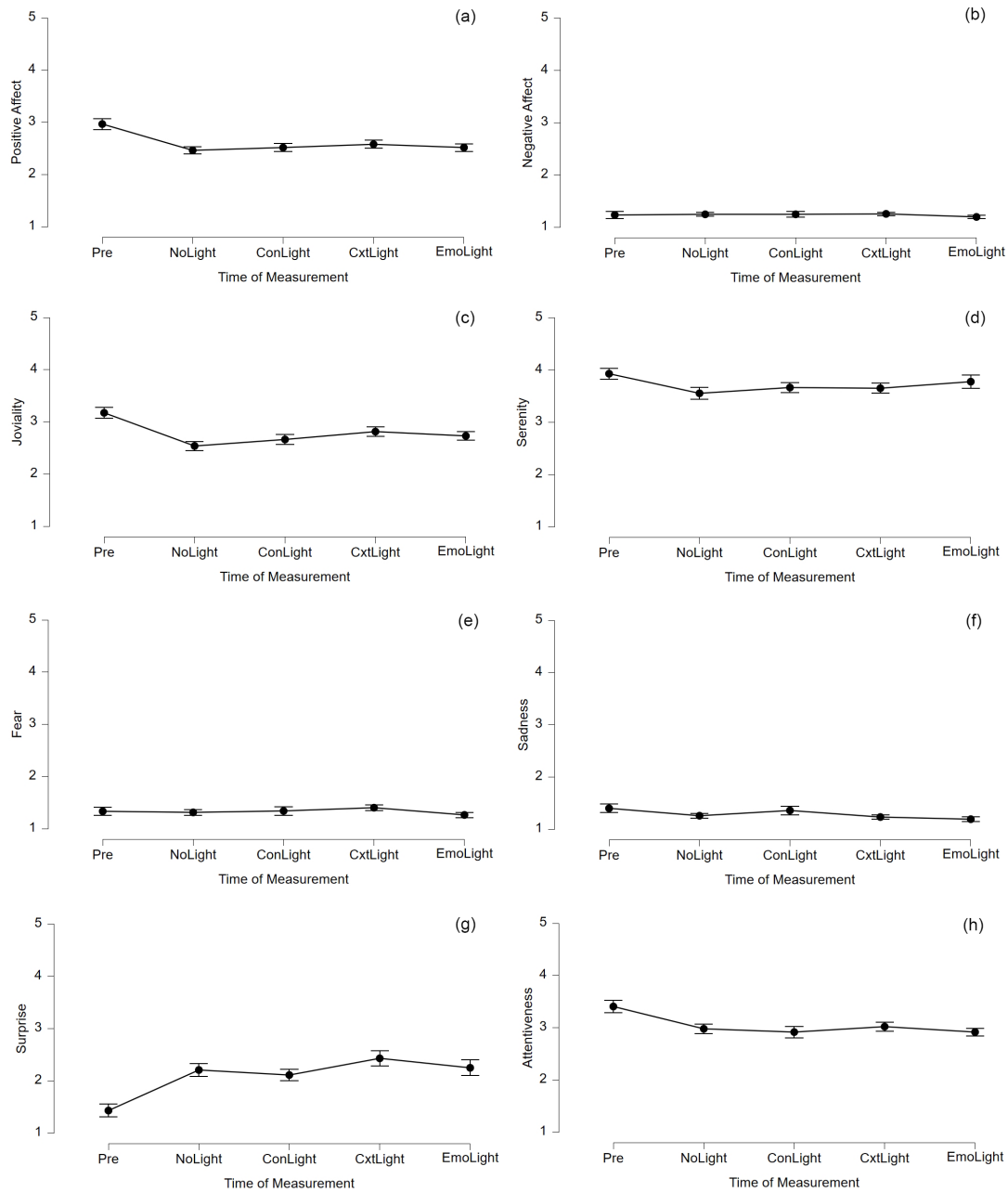


Figure 4. Plots for emotion induction measured throughout Study III. Order of the conditions was counterbalanced. (a) Descriptive plot for *Positive Affect*; (b) Descriptive plot for *Negative Affect*; (c) Descriptive plot for *Joviality*; (d) Descriptive plot for *Serenity*; (e) Descriptive plot for *Fear*; (f) Descriptive plot for *Sadness*; (g) Descriptive plot for *Surprise*; (h) Descriptive plot for *Attentiveness*. Pre: pre-measured baseline; error bars represent the standard error.

Being asked to choose their favorite story independently from its presentation after receiving all four stories (Figure 5), the majority of the participants liked the story about Mahoutokoro, which was presented with the *CxtLight* condition, as their favorite ($n = 10$). This was followed by *Beauxbatons* (*ConLight*, $n = 5$) and *Castelobrujo* (*NoLight*, $n = 5$), and *Uagadou* (*EmoLight*, $n = 4$). However, there was no significant difference in frequency of choice from chance, $\chi^2(3) = 3.67$, $p = .300$. Regarding the participants' expressed story liking queried directly after each story, the comparison results calculated by a Friedman test were in line with this finding, $\chi^2(3) = 4.08$, $p = .253$, $W = .06$. Looking at the comments provided, both positive and negative comments were provided. While some participants found the stories interesting ($n_{EmoLight} = 10$, $n_{CxtLight} = 6$, $n_{ConLight} = 3$, $n_{NoLight} = 5$), easy to follow ($n_{EmoLight} = 2$) and imagine ($n_{EmoLight} = 1$, $n_{CxtLight} = 1$), humorous ($n_{EmoLight} = 1$) and especially liked the *Harry Potter* theme ($n_{EmoLight} = 7$, $n_{CxtLight} = 6$, $n_{ConLight} = 10$, $n_{NoLight} = 4$), others referred to the stories being boring ($n_{EmoLight} = 2$, $n_{CxtLight} = 2$, $n_{ConLight} = 4$, $n_{NoLight} = 4$), lacking action ($n_{EmoLight} = 3$, $n_{CxtLight} = 1$, $n_{NoLight} = 1$) or emotions ($n_{ConLight} = 1$), and being hard to follow ($n_{CxtLight} = 3$, $n_{NoLight} = 1$).

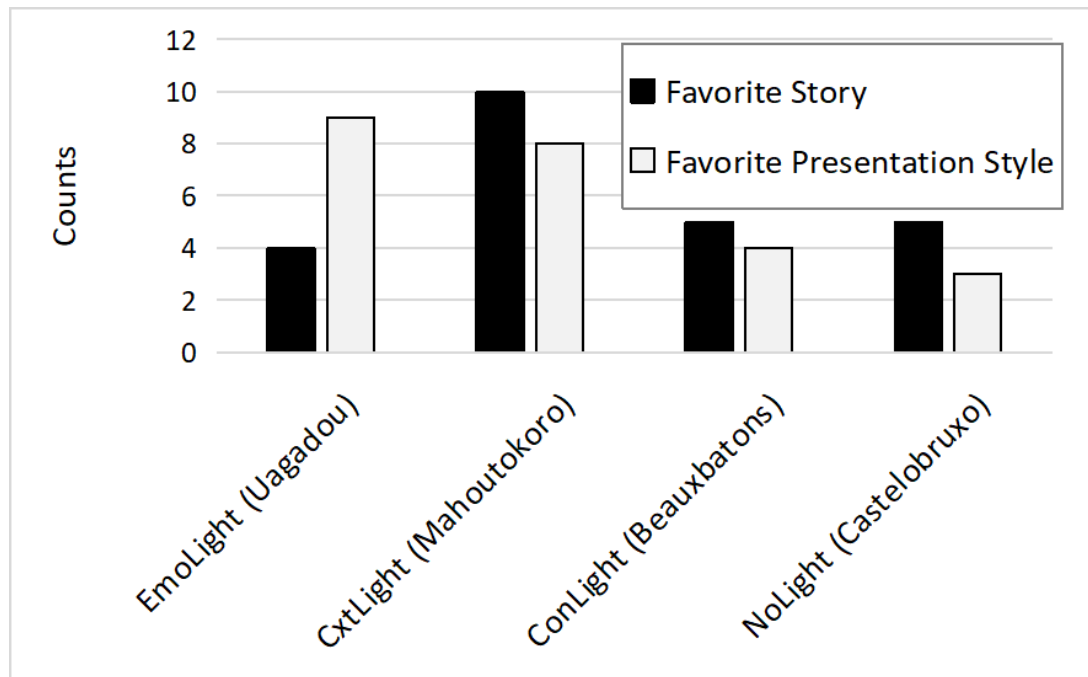


Figure 5. Best liked story and presentation style.

Choosing their favorite presentation approach (Figure 5), participants descriptively favored the *EmoLight* condition ($n = 9$), followed by the *CxtLight* condition ($n = 8$), *ConLight* ($n = 4$), and *NoLight* ($n = 3$). However, no significant difference was indicated for frequency of choice from chance, $\chi^2(3) = 4.33$, $p = .228$. Contrastingly, comparing the presentation liking expressed after each story, a significant difference was revealed by a calculated Friedman tests ($\chi^2(3) = 10.40$, $p = .015$, $W = .14$), with descriptively highest liking expressed for the *CxtLight* condition, followed by the *EmoLight* condition, *ConLight*, and last *NoLight*. According to Bonferroni-holm-corrected Conover's post hoc tests, however, the only significant difference was between *CxtLight* and *NoLight*, $p = .009$.

These results are also supported by participants' comments. For the *NoLight* condition, most of the comments on presentation style were negative (84.85%). For the *ConLight* and *CxtLight* conditions, 67.86% and 56.25% of the comments were negatively valenced, respectively. Only for the *EmoLight* condition less negative than positive comments were provided (47.83%). Arxc test revealed a significant relationship between condition and valence of the comments, $\chi^2(3) = 9.93$, $p = .019$.

Most negative comments focused on the robot's voice, describing it as too quiet ($n_{CxtLight} = 1$, $n_{ConLight} = 1$, $n_{NoLight} = 5$), hard to understand ($n_{EmoLight} = 3$, $n_{CxtLight} = 5$, $n_{ConLight} = 7$, $n_{NoLight} = 6$), unnatural, discomforting or strange ($n_{EmoLight} = 1$, $n_{CxtLight} = 2$, $n_{ConLight} = 2$, $n_{NoLight} = 6$). Some participants also criticized the robot for making unnatural speech pauses ($n_{EmoLight} = 2$, $n_{CxtLight} = 3$, $n_{ConLight} = 1$, $n_{NoLight} = 2$). Regarding the robot's motion, several participants mentioned the body language being unnatural ($n_{EmoLight} = 1$, $n_{CxtLight} = 3$, $n_{ConLight} = 4$, $n_{NoLight} = 3$), inappropriate ($n_{CxtLight} = 1$, $n_{ConLight} = 3$, $n_{NoLight} = 3$), or distracting ($n_{EmoLight} = 1$). One participant in the *NoLight* condition also mentioned that the robot's motors sounds were too loud during movement. Concerning the robot itself, participants felt the robot wasn't energetic enough ($n_{NoLight} = 1$), boring ($n_{ConLight} = 1$), or distracting ($n_{CxtLight} = 1$). Regarding positive comments, several participants referred to the robot's body language as appropriate ($n_{CxtLight} = 3$, $n_{ConLight} = 1$, $n_{NoLight} = 3$) and natural ($n_{ConLight} = 1$), or liking them ($n_{NoLight} = 1$). Several participants positively evaluated the robot's pronunciation ($n_{CxtLight} = 1$, $n_{ConLight} = 3$, $n_{NoLight} = 1$) and stated that they liked listening to it ($n_{CxtLight} = 1$).

Regarding the lighting approaches, few participants expressed liking for both approaches ($n_{EmoLight} = 3$, $n_{CxtLight} = 3$). Positive remarks highlighted that the added lights were pleasing ($n_{EmoLight} = 1$, $n_{ConLight} = 2$) or interesting ($n_{CxtLight} = 1$), helped attract attention ($n_{CxtLight} = 1$), supported understanding ($n_{CxtLight} = 1$), and aided in a more vivid mental image ($n_{CxtLight} = 1$). In both the *EmoLight* and the *CxtLight* condition, two participants each found the light colors were appropriate. However, several critical remarks were made as well. Participants claimed the light switching being too harsh and abrupt ($n_{EmoLight} = 2$, $n_{CxtLight} = 1$) or the colored lighting overwhelming ($n_{CxtLight} = 1$), and not beneficial for the storytelling ($n_{EmoLight} = 1$). Interestingly, one participant in the *ConLight* condition especially mentioned the lighting to be not distracting, while another participant in the *NoLight* condition wished the lighting had been used for this story as well.

After experiencing all four storytelling conditions,, participants assessed their general liking for integrating colored light across all approaches at 3.33 ($SD = 1.31$). Arguments against light integration were distraction ($n = 5$)—with one participant specifically referencing the *CxtLight* condition—and concerns about excessive switching between colors ($n = 3$). However, participants generally

appreciated the idea of integrating colored lighting into robotic storytelling ($n = 4$), including for setting an emotional tone ($n = 3$) illustrating the stories' context ($n = 2$), and making the storytelling less monotonous ($n = 1$). In more detail, their perceived appropriateness ratings were 3.29 ($SD = 1.33$) for the emotion-based lighting and 3.83 ($SD = 0.96$) for the context-based lighting. Participants provided seven positive comments each for *EmoLight* and *CxtLight* conditions. *EmoLight* received two negative comments, while *CxtLight* received only one. Regarding negative feedback, complaints about abrupt light switching were mentioned again ($n_{EmoLight} = 1$, $n_{CxtLight} = 1$), and the lighting was perceived as distracting ($n_{EmoLight} = 1$). In contrast, positive comments supporting both the *EmoLight* and *CxtLight* highlighted the appropriateness of the light colors ($n_{EmoLight} = 2$, $n_{CxtLight} = 2$) and their contribution to creating a more vivid mental image ($n_{EmoLight} = 1$, $n_{CxtLight} = 3$). One participant specifically noted the enhanced ability to empathize in the *EmoLight* condition. Further, both *EmoLight* and *CxtLight* were mentioned to be rare positively unobtrusive ($n_{EmoLight} = 1$, $n_{CxtLight} = 1$), although one participant stated that lighting in the *EmoLight* condition was more noticeable and thus more appealing. Comments favoring the other two conditions (*ConLight* and *NoLight*) were all based on the reduced distraction resulting from the lack of color changes ($n_{ConLight} = 2$, $n_{NoLight} = 2$).

6.4 Discussion

We carried out a lab-based within-subjects design study to analyze the effects of integrating colored light—either based on story emotions (*EmoLight*) or story context (*CxtLight*)—on participants' storytelling experience, emotions, and robot perception. For our control conditions, we utilized constant white light (*ConLight*) respectively no additional light (*NoLight*). Regarding recipients' **storytelling experience**, no significant effect of the lighting approach was found, neither for *Transportation*, nor for *Narrative Engagement*, or *Cognitive Absorption*. Therefore, we reject H4. This finding is in line with both the results from our first study as well as the prior work^[20], despite our refinement of the lighting strategy. The light modality does not seem to influence the storytelling experience at all. While emotional body language and facial expressions improve robotic storytelling experiences compared to uni-modal speech-based robotic storytelling^[11,49], the adding of only light to an unimodal speech-based storytelling^[20] as well as the addition of light strategies to multimodal storytelling with speech and emotional body language does not improve the storytelling experience. One noticeable difference between the modalities of body language/facial expressions and (colored) light is the nature of the modality—whether biomimetic or not. Furthermore, the source of the modality differs. Since the light was not coming directly from the robot, it might divert attention away from the storytelling. This is reflected in some of the negative comments from participants, who claimed that the colored lights were distracting. This finding is surprising, given the substantial emphasis placed on lighting in contexts such as musical performances, where audience evaluations greatly depend on the quality of the light presentation^[146]. Distraction has been shown to reduce the storytelling experience in terms of transportation^[34,147]. In our study, the conditions with colored light did not score significantly lower than those without changing lights. Potentially, the distracting and enhancing effects of the colored light modality might have balanced each other out. Future work is needed to provide more detailed insights into the connections between underlying processes.

Concerning participants' *emotions*, no significant differences were found between the conditions using colored light and the control conditions with either constant white or no additional light, leading to the rejection of H5. This finding is in line with prior works utilizing horror stories^[20]. While colored lights seem to help emotion recognition from the robot's expression—as demonstrated in related studies^[16,17,53], they do not seem to influence emotions evoked in participants by the stories narrated. As our results show that storytelling in general did affect participants' individual emotions, for instance, participants expressed increased *Surprise* or decreased *Sadness* after the storytelling compared to the pre-measurement, this finding cannot be attributed to a general failure of the robot to evoke emotion. Rather, it suggests that the addition of colored light does not further enhance the emotional impact already achieved by multimodal storytelling using speech and emotional body language. Future work may include the testing of different story genres. At the moment, three studies investigating the integration of the colored light modality into robotic storytelling have explored three story genres: horror^[20], romance (see Study I), and fantasy (see Study III). Differences between the studies' results—such as the variation of *Serenity* independence of the lighting approach in Study I which was not found in the current study—might be due to the different genres utilized, as they tackle different emotions, which seem to respond differently to the colored lights.

Furthermore, we found no significant differences in robot perception in terms of *Animacy* or *Likeability* of the robot as well as its social attributes. This result conflicts with the increased *Competence* of a robotic storyteller using emotionally adaptive colored light by Steinhäusser *et al.*^[20]. One potential explanation might be the additional emotional body language utilized in our study compared to the static storyteller in the prior work. While for a static storyteller the addition of a second modality, the light modality, to the unimodal speech-based storytelling may enhance the perceived competence of the story-teller^[20], the added colored light to an already multimodal system (combining speech with emotional body language) appears to offer no benefit in terms of perceived expressiveness or competence. This outcome aligns with our earlier theoretical reasoning. However, we found a significant positive influence of the colored light modality on the perceived anthropomorphic morphology, but only on the subscale of anthropomorphic *Communication*, leading to partly accepting H6. It is kind of surprising that of all the perceived anthropomorphism of the robot is increased when adding a non-biomimetic modality. While red light could be associated with human blushing^[54], there are no human-like associations with the rest of the utilized colors. Potentially, the non-biomimetic modality might be expressive enough to

substitute lacking biomimetic ones in the robot, leading to a generally higher perceived anthropomorphic communication. Given the positive effects of a high perceived anthropomorphism^[148,149], this finding is especially interesting for low-budget robots as colored light is rather affordable to install compared to biomimetic modalities such as motors for facial expressions.

Overall, in terms of storytelling experience, evoked emotions, and robot perception, no significant differences were found between our two approaches for integrating colored light, either focusing on story-relevant emotions or context in the form of described environments (RQ4). This finding is mainly reflected by participants' subjective choice of their favorite presentation style. While their rating at the end of the study tended to favor the *EmoLight*, when asked after each individual presentation, they favored the *CxtLight*—though this did not differ significantly from the preference for the *EmoLight* condition. As shown in Figure 5, descriptively, both *EmoLight* and *CxtLight* outperform control conditions, *ConLight* and *NoLight*. The lack of statistical significance might be due to the small sample size of the study. The preference is also backed by the comments provided by the participants which were significantly associated with the conditions, being most negative for the *NoLight* condition, followed by the *ConLight* condition. Further, the comments shed more light on the comparison between *EmoLight* and *CxtLight*, with the most positive for the *EmoLight* condition and *CxtLight* being just on the fence. Thus, we suggest that—given the limitation of our small sample size—participants seem to prefer storytelling integrating colored light over storytelling with constant or no additional light (RQ5), although no noticeable differences were found in storytelling experience and induced emotions. Further, the subjective evaluation of the *EmoLight* approach seems to be slightly more favorable than the *CxtLight* approach. However, future work is needed to support this qualitative finding.

7. General Discussion and Limitations

Within this work, we presented three studies on integrating colored lights into multimodal robotic storytelling. As a non-biomimetic modality, colored light remains less extensively examined in HRI research compared to human-like modalities such as facial expressions or gestures^[15], while a great amount of literature deals with the psychological associations between colors and emotions (Section 5.1.1). Former studies integrating colored light into robotic storytelling indicated its potential to improve the storyteller's social presence and increase its perceived competence^[20].

Within Study I, we combined lighting strategies for emotion induction with expressive body language shown by the robotic storyteller. Results showed that the positive effects of colored lighting reported by Steinhäusser *et al.*^[20] were diminished when emotional bodily expressions were added. This suggests that applying emotion-inducing light was not sufficient in improving the robotic storytelling. Given the vague results of study I, we conducted study II to deduce light colors supporting emotion expression from the robot. In doing so, we found that certain light colors improved the recognition of the emotions, including *amazement*, *rage*, and the *neutral* state (Table 5). These findings informed the design of our final study III, in which we compared two colored light approaches—an emotion-based approach using the color associations from study II, and a context-driven approach based on environment descriptions within the story text—with constant white light and the absence of additional lighting. Although no differences were found in terms of storytelling experience, emotion induction, and robot perception, subjective feedback generally favored the use of colored lights, regardless of the specific approach.

While our findings provide valuable insights into the implementation of colored lights as a modality in robotic storytelling, future work is needed to enhance their generalizability. First of all, we focused our work on the Pepper robot. While it is one of the most widespread robot models^[150], results from our studies might not be transferrable to other robot models as they differ not only in size and style, but also their capabilities. For instance, the Pepper robot with 1.2 meters^[151] is way taller than the most popular NAO robot^[150], being only child-sized^[152], while the NAO robot with 25 degrees of freedom^[152] bears more options for bodily expressions than the Pepper robot with only 20 degrees of freedom^[151]. In line, the two robots also differ in their perception, with NAO being perceived as more anthropomorphic than Pepper^[153]. This might have influenced our results, as Thorstenson and Tian^[154] showed that color-emotion associations interact with a robot's perceived human-likeness. Additionally, robot morphology can affect attributed communal traits and perceived task suitability^[155]. Furthermore, other robot models are capable of facial expressions, such as the Reeti robot^[156]. In HRI, facial expressions are generally considered more important than bodily expressions. While recognition rates of expressed emotions appear to be comparable between both modalities^[157], their interaction with colored light might differ. Thus, future work incorporating other robot models is needed to provide more insights into the links between these capabilities and perceptions.

Similarly, generalizability of our findings is limited due to our small, young, and gender-skewed samples. Future research should not only aim for larger sample sizes but also consider age and gender effects. For example, females have been shown to relate “negatively to Robot Liking and positively to Robotphobia” on average^[158], and they also tend to be more sensitive to potential eeriness^[159]. Regarding users' age, “older people [are] more willing to overlook defects in a robot by giving it a higher rating for humanlike”^[159]. Therefore, a more balanced sample might have yielded different results, as minor shortcomings might have been more easily overlooked, leading participants to concentrate more on the actual manipulation in the robotic storytelling. However, addressing such shortcomings remains essential. Most apparent given participants' comments is the technical limitation of the robot's voice due to using the robot's built-in text-to-speech module without adjustments. Speaker voices and their acting are important when

receiving audio books^[160,161], thus it is not surprising that participants criticized the absence of expressive voice modulation in robotic storytelling. Following the approach of Zabala *et al.*^[162] a robotic storyteller should alter its pitch, speed, and volume based on the annotated emotions, using additional tools to create timestamps for individual words.

Lastly, methodological extensions might generate additional insights into the influence of different approaches of adding colored light to robotic storytelling. In our studies, we focused on subjective measures acquired using questionnaires. Particularly for affective responses such as attention and emotions, objective indicators—including physiological data and behavioral observations—can yield complementary findings. Although prior work^[20] found only marginal effects of colored lights on physiological data (e.g., heart rate), such data can still complement self-report questionnaires, as they are less prone to manipulation, less affected by memory biases, and allow for data collection over time^[163]. In addition, behavioral measures can provide further insights into the experience, for example, personal distance kept to the robot is connected to comfort^[164] and trust^[48], while blink rates correlate with attention^[165]. Next to these supplementary measures, also the experimental time span could be extended. Our studies investigated the short-term effects of integrating colored light into robotic storytelling. However, due to novelty effects, users may focus more on the robot itself than the information it conveys^[166]. Smedegaard^[167] argues that experiencing a robot being socially capable might be novel. Similarly, a social robot capable of emotionally expressive multimodal storytelling might have been perceived as surprising, as suggested by the increase of surprise values in Figure 4g. This novelty may have distracted from the story and the lighting approaches, potentially masking condition differences in study I and III. Long-term studies including several storytelling occasions could mitigate this novelty effects and elicit a mere exposure effect, whereby repeated interactions increase user liking^[168]. However, Fiolka *et al.*^[169] demonstrated that the mere exposure effect can also be achieved through consecutive exposures occurring one after the other within a short time span, as by the within-subjects design of study II. Nevertheless, this effect might have been covered by counterbalancing the conditions. Thus, a between-designed longitudinal study might provide further insights on colored lights integration into multimodal robotic storytelling.

8. Conclusions

Color is a modality yielding many associations, particularly with emotions. In addition, it is a modality that robots are capable of while humans are not. Thus, many works tried to use color or colored light to express emotions in social robots. Nonetheless, the integration of colored light into robotic storytelling has received comparatively little attention, despite the demonstrated benefits of emotion-inducing colored lighting in other narrative media such as film and video games. In this work, we initially focused on the emotion-inducing effects of colored light. Therefore, we integrated colored room illumination into romantic robotic storytelling, complemented by emotional bodily expressions, and informed by design guidelines for emotion-inducing virtual environments. However, our results showed only limited effects compared to previous research, suggesting that focusing on the emotion-inducing effect of colors may not significantly enhance the robotic storytelling experience. In consequence, we focused on the emotion-expressive effect next.

We conducted an online study to determine whether light colors can facilitate emotion recognition from robotic bodily expressions. These expressions were used in a third study to evaluate and compare the emotion-driven colored light to context-based colored light based on described environments in the story. Both of these information-conveying light approaches were compared to against two control conditions: constant white light and no additional lighting. Surprisingly, results show that the addition of colored lights does not further improve multimodal robotic storytelling in terms of storytelling experience or induced emotions, despite a positive impact on the robot's perceived communicative anthropomorphism. However, participants demonstrated a clear preference for robotic storytelling integrating colored lights over control conditions with constant or no additional lighting, independent of the approach. Future work should incorporate more diverse story genres, robot models, and larger, more demographically balanced sample sizes to improve generalizability. Thus, both emotion-based and context-based lighting strategies seem to be appropriate for enhancing robotic storytelling.

Supplementary materials

The supplementary material for this article is available at: [Supplementary materials](#).

Declarations

Acknowledgements

The authors would like to thank Stephanie Gaman for providing the story for study I, Matthias Popp for helping with the implementation for study I, Lisa Kreis for helping in data acquisition for study I, Valeria Freese for producing the stimulus material for study II and helping implementing the Python framework.

Authors contribution

Steinhaeuser SC: Developed the research aims and questions, made substantial contributions to conception and design of the study and

materials, performed data acquisition, analysis and interpretation, wrote the manuscript.

Maier S: Implemented the materials.

Lugrin B: Supervised the work, provided administrative support, contributed to the scientific framing, provided feedback to the draft.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval

All studies were deemed ethically sound by the local ethics committee of the human-computer-media institute of the University of Würzburg (#020523, #250324, #260924).

Consent to participate

All participants provided freely-given informed consent.

Consent for publication

Not applicable.

Availability of data and materials

The data and materials could be obtained from the corresponding author upon request.

Funding

None.

Copyright

©The Author(s) 2025.

References

1. Statista. Konsumenten & Marken—Lesen in Deutschland [Internet]. 2023. Available from: <https://de.statista.com/statistik/studie/id/44344/dokument/lesen-in-deutschland/>
2. Audible. Wie oft haben Sie innerhalb der letzten 12 Monate Hörbücher, Hörspiele oder Podcasts gehört? [Graph] [Internet]. 2023. Available from: <https://de.statista.com/statistik/daten/studie/1274205/umfrage/anteil-der-nutzer-von-digitalen-hoerbuechern-und-podcasts/>
3. Moyer JE. Audiobooks and e-books: A literature review. *Ref User Serv Q*. 2012;51(4):340-354. [DOI]
4. Tattersall Wallin E, Nolin J. Time to read: Exploring the timespaces of subscription-based audiobooks. *New Media Soc*. 2020;22(3):470-488. [DOI]
5. Adamczyk G. Storytelling: mit Geschichten überzeugen. 2nd ed. Germany: Haufe-Lexware; 2018.
6. Fludernik M. Erzähltheorie: Eine Einführung. Germany: Wissenschaftliche Buchgesellschaft; 2013.
7. Choo YB, Abdullah T, Nawi AM. Digital storytelling vs. oral storytelling: An analysis of the art of telling stories now and then. *Univers J Educ Res*. 2020;8(5A):46-50. [DOI]
8. Hsu Y. The influence of English storytelling on the oral language complexity of EFL primary students[dissertation]. Yunlin: National Yunlin University of Science & Technology; 2010.
9. Adams E. Sandbox storytelling [Internet]. 2010. Available from: http://www.designersnotebook.com/Columns/106_Sandbox_Storytelling/106_sandbox_storytelling.htm
10. Breazeal C, Dautenhahn K, Kanda T. Social robotics. In: Siciliano B, Khatib O, editors. *Springer Handbook of Robotics*. Cham: Springer; 2016. p. 1935-1972.
11. Striepe H, Lugrin B. There once was a robot storyteller: Measuring the effects of emotion and non-verbal behaviour. In: Kheddar A, Yoshida E, Ge SS, Suzuki K, Cabibihan JJ, Eyssel F, He H, editors. *Social Robotics*. Cham: Springer; 2017. p. 22-24.10652. 2017. [DOI]
12. Collins EC, Prescott TJ, Mitchinson B. Saying it with light: A pilot study of affective communication using the mirorobot. In: Wilson SP, Verschure PFMJ, Mura A, editors. *Living Machines 2015: Proceedings of the 4th International Conference on Biomimetic and Biohybrid Systems*; 2015 Jul 28-31; Barcelona: Spain. Berlin: Springer; 2015. p. 243-255. [DOI]
13. Song S, Yamada S. Expressing emotions through color, sound, and vibration with an appearance-constrained social robot. In: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*; 2017 Mar 6-9; Vienna, Austria. New York: Association for Computing Machinery; 2015. p. 243-255. [DOI]
14. Da Pos O, Green-Armytage P. Facial expressions, colours and basic emotions. *J Int Colour Assoc*. 2007;1(1):2. Available from: <https://d1wqtxts1xzle7.cloudfront.net/81197228>

15. Song S, Yamada S. Bioluminescence-inspired human-robot interaction. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction; 2018 Mar 5-8; Chicago, USA. New York: Association for Computing Machinery; 2018. p. 224-232. [DOI]
16. Terada K, Yamauchi A, and Ito A. Artificial emotion expression for a robot by dynamic color change. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction; 2018 Mar 5-8; Chicago, USA. New York: Association for Computing Machinery; 2018. p. 224-232. [DOI]
17. Löffler D, Schmidt N, Tscharn R. Multimodal expression of artificial emotion in social robots using color, motion and sound. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction; 2018 Mar 5-8; Chicago: USA. New York: Association for Computing Machinery; 2018. p. 334-343. [DOI]
18. Baraka K, Veloso MM. Mobile service robot state revealing through expressive lights: Formalism, design, and evaluation. *Int J Soc Robot.* 2018;10:6592 [DOI]
19. Steinhäusser SC, Lugrin B. Effects of colored led in robotic storytelling on storytelling experience and robot perception. In: 2022 17th ACM/IEEE International Conference on Human-Robot Interaction; 2022 Mar 7-10; Sapporo, Japan. Piscataway: IEEE; 2022. p. 1053-1058. [DOI]
20. Steinhäusser SC, Ganai E, Yalcin M, Latoschik ME, Lugrin B. Binded to the lights: Storytelling with a physically embodied and a virtual robot using emotionally adapted lights. In: 2024 33rd IEEE International Conference on Robot and Human Interactive Communication; 2024 Aug 26-30; Pasadena, USA. Piscataway: IEEE; 2024. p. 2117-2124. [DOI]
21. Kleine Wieskamp P. Storytelling: digital-multimedial-social. Wiesbaden: Springer VS; 2016. [DOI]
22. ZagaloN, Barker A, Branco V. Story reaction structures to emotion detection. In: Proceedings of the 1st ACM workshop on Story representation, mechanism and context; 2004 Oct 15; New York, USA. New York: Association for Computing Machinery; 2004. p. 33-38. [DOI]
23. Marlar Lwin S. Capturing the dynamics of narrative development in an oral storytelling performance: A multimodal perspective. *Lang Lit.* 2010;19(4):357-377. [DOI]
24. Ekman P. Basic emotions. In: Dalgleish T, Power MJ, editors. *Handbook of Cognition and Emotion*. Chichester: John Wiley & Sons, Ltd; 1999. p. 45-60. [DOI]
25. Frijda NH. *The emotions*. Cambridge: Cambridge University Press; 1986.
26. Parkinson B, Totterdell P, Briner RB. *Changing moods: The psychology of mood and mood regulation*. London: Longman; 1996.
27. Slater M. A note on presence terminology. *Presence Connect.* 2003;3(3):1-5. Available from: https://www.researchgate.net/publication/242608507_A_Note_on_Presence_Terminology
28. Slater M. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Phil Trans R Soc B.* 2009;364(1535):3549-3557. [DOI]
29. Russell JA. A circumplex model of affect. *J Pers Soc Psychol.* 1980;39(6):1161-1178. [DOI]
30. Plutchik R. The nature of emotions. *Am Sci.* 2001;89(4):344-350. [DOI]
31. Park SH, Bae BC, Cheong YG. Emotion recognition from text stories using an emotion embedding model. In: 2020 IEEE International Conference on Big Data and Smart Computing; 2020 Feb 19-22; Busan, Korea. Piscataway: IEEE; 2020 p. 579-583. [DOI]
32. Oatley K. *The passionate muse: Exploring emotion in stories*. Oxford: Oxford University Press; 2012.
33. Green MC, Clark JL. Transportation into narrative worlds: implications for entertainment media influences on tobacco use. *Addiction.* 2013;108(3):477-484. [DOI]
34. Green MC, Brock TC. The role of transportation in the persuasiveness of public narratives. *J Pers Soc Psychol.* 2000;79(5):701-721. [DOI]
35. Green MC, Appel M. Chapter One - Narrative transportation: How stories shape how we see ourselves and the world. *Adv Exp Soc Psychol.* 2024;70:1-82. [DOI]
36. Hall A, Zwarun L. Challenging entertainment: Enjoyment, transportation, and need for cognition in relation to fictional films viewed online. *Mass Commun Soc.* 2012;15(3):384-406. [DOI]
37. de Graaf MMA, Allouch SB, van Dijk JAGM. Long-term acceptance of social robots in domestic environments: Insights from a user's perspective. In: *Enabling Computing Research in Socially Intelligent Human-Robot Interaction: A Community-Driven Modular Research Platform*; 2016 Mar 21-23; California, USA. California: AAAI Press; 2016.
38. Lugrin B. Introduction to socially interactive agents. In: Lugrin B, Pelachaud C, Traum D. editors. *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics*. New York: Association for Computing Machinery; 2021. p. 1-20. [DOI]
39. Lane HC, Schroeder NL. Pedagogical agents. In: Lugrin B, Pelachaud C, Traum D. editors. *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics*. New York: Association for Computing Machinery; 2022. p. 307-330. [DOI]
40. Donnermann M, Lugrin B. Integration of robot-supported tutoring in higher education-an empirically based concept. In: Proceedings of the 2024 the 16th International Conference on Education Technology and Computers; 2024 Sep 18-21; Porto Vlaams-Brabant, Portugal. New York: Association for Computing Machinery; 2024. p. 1-7. [DOI]
41. Prada R, Rato D. Socially interactive agents in games. In: *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics*. New York: Association for Computing Machinery; 2022. p. 493-526. [DOI]
42. Aylett R. Interactive narrative and story-telling. In: Lugrin B, Pelachaud C, Traum D, editors. *The Handbook on Socially Interactive*

- Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics. New York: Association for Computing Machinery; 2022. p. 463-492. [DOI]
43. Duffy BR. Anthropomorphism and the social robot. *Robot Auton Syst*. 2003;42(3-4):177-190. [DOI]
44. Fink J. Anthropomorphism and human likeness in the design of robots and human-robot interaction. In: Ge SS, Khatib O, Cabibihan JJ, Simmons R, Williams MA, editors. *Social Robotics, 4th International Conference*; 2012 Oct 29-31; Chengdu, China. Berlin: Springer; 2012. p. 199-208. [DOI]
45. Breazeal C, Scassellati B. How to build robots that make friends and influence people. In: *Proceedings 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human and Environment Friendly Robots with High Intelligence and Emotional Quotients*; 1999 Oct 17-21; Kyongju, Korea. Piscataway: IEEE; 1999. p. 858-863. [DOI]
46. Gao Y, Chang Y, Yang T, Yu Z. Consumer acceptance of social robots in domestic settings: A human-robot interaction perspective. *J Retail Consum Serv*. 2025;82:104075. [DOI]
47. Salem M, Eyssel F, Rohlfing K, Kopp S, Joublin F. Effects of gesture on the perception of psychological anthropomorphism: A case study with a humanoid robot. In: Mutlu B, Bartneck C, Ham J, Evers V, editors. *Proceedings of the Third international conference on Social Robotics*; 2011 Nov 24-25; Amsterdam, The Netherlands. Berlin: Springer; 2011. p. 31-41. [DOI]
48. Babel F, Kraus J, Miller L, Kraus M, Wagner N, Minker W, et al. Small talk with a robot? the impact of dialog content, talk initiative, and gaze behavior of a social robot on trust, acceptance, and proximity. *Int J of Soc Robotics*. 2021;13(6):1485-1498. [DOI]
49. Appel M, Lugrin B, Kühle M, Heindl C. The emotional robotic storyteller: On the influence of affect congruency on narrative transportation, robot perception, and persuasion. *Comput Hum Behav*. 2021;120:106749. [DOI]
50. Wagner P, Malisz Z, Kopp S. Gesture and speech in interaction: An overview. *Speech Commun*. 2014;57:209-232. [DOI]
51. Ham J, Bokhorst R, Cuijpers R, van der Pol D, Cabibihan JJ. Making robots persuasive: The influence of combining persuasive strategies (gazing and gestures) by a storytelling robot on its persuasive power. In: Mutlu B, Bartneck C, Ham J, Evers V, Kanda T, editors. *Social Robotics: Third International Conference on Social Robotics, ICSR 2011*; 2011 Nov 24-25; Amsterdam, The Netherlands. Berlin: Springer; 2011. p. 71-83. [DOI]
52. Rea DJ, Young JE, Irani P. The roomba mood ring: an ambient-display robot. In: *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*; 2012 Mar 5-8; Boston Massachusetts, USA. New York: Association for Computing Machinery; 2012. p. 217-218. [DOI]
53. Song S, Yamada S. Designing expressive lights and in-situ motions for robots to express emotions. In: *Proceedings of the 6th International Conference on Human-Agent Interaction*; 2018 Dec 15-18; Southampton, United Kingdom. New York: Association for Computing Machinery; 2018. p. 222-228. [DOI]
54. Kim Mg, Lee HS, Park JW, Jo SH, Chung MJ. Determining color and blinking to support facial expression of a robot for conveying emotional intensity. In: *RO-MAN - The 17th IEEE International Symposium on Robot and Human Interactive Communication*; 2008 Aug 1-3; Munich, Germany. Piscataway: IEEE; 2008. p. 219-224. [DOI]
55. Fernández-Rodicio E, Maroto-Gómez M, Castro-González Á, Malfaz M, Salichs MÁ. Emotion and mood blending in embodied artificial agents: Expressing affective states in the mini social robot. *Int J of Soc Robotics*. 2022;14(8):1841-1864. [DOI]
56. Hong A, Lunscher N, Hu T, Tsuboi Y, Zhang X, dos Reis Alves SF. A multimodal emotional human-robot interaction architecture for social robots engaged in bidirectional communication. *IEEE Trans Cybern*. 2021;51(12):5954-5968. [DOI]
57. Häring M, Bee N, André E. Creation and evaluation of emotion expression with body movement, sound and eye color for humanoid robots. In: *2011 RO-MAN*; 2011 Jul 31-Aug 3; Atlanta, USA. Piscataway: IEEE; 2011. p. 204-209. [DOI]
58. Agarwal R, Karahanna E. Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Q*. 2000;24(4):665-694. [DOI]
59. Weniger S, Loebbecke C. Cognitive absorption: Literature review and suitability in the context of hedonic IS usage [dissertation]. Germany: University of Cologne; 2011.
60. Trevino LK, Webster J. Flow in computer-mediated communication: Electronic mail and voice mail evaluation and impacts. *Commun Res*. 1992;19(5):539-573. [DOI]
61. Betella A, Inderbitzin M, Bernardet U, Verschure PMFG. Non-anthropomorphic expression of affective states through parametrized abstract motifs. In: *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*; 2013 Sep 2-5; Geneva, Switzerland. Piscataway: IEEE; 2013. p. 435-441. [DOI]
62. Orloff AM, Güntner L, Windl M, Schmidt T, Kocur M, Wolff C. In: *Proceedings of Mensch und Computer 2019*; 2019 Sep 8-11; Hamburg, Germany. New York: Association for Computing Machinery; 2019. p. 863-866. [DOI]
63. Roohi S, Forouzandeh A. Regarding color psychology principles in adventure games to enhance the sense of immersion. *Entertain Comput*. 2019;30:100298. [DOI]
64. Wilms L, Oberfeld D. Color and emotion: effects of hue, saturation, and brightness. *Psychol Res*. 2018;82(5):896-91 [DOI]
65. Kennedy AJ. The effect of color on emotions in animated films [dissertation]. West Lafayette: Purdue University; 2014.
66. Ionescu B, Coquin D, Lambert P, Buzuloiu V. A fuzzy color-based approach for understanding animated movies content in the indexing task. *EURASIP J Image Video Process*. 2008;2008:1-17. [DOI]
67. Steinhäusser SC, Oberdörfer S, von Mammen SA, Latoschik ME, Lugrin B. Joyful adventures and frightening places—designing emotion-inducing virtual environments. *Front Virtual Real*. 2022;3:919163. [DOI]
68. Wolfson S, Case G. The effects of sound and colour on responses to a computer game. *Interacting Comput*. 2000;13(2):183-192. [DOI]

69. Joosten E, van Lankveld G, Spronck PHM. Colors and emotions in video games. In: Ayeshe A, editor. Proceedings on the 11th International Conference on Intelligent Games and Simulation GAME-ON 2010. United Kingdom: Eurosis, Ostend; 2021. p. 61-65.
70. Oberdörfer S, Steinhäusser SC, Najjar A, Tümmers C, Latoschik ME. Pushing yourself to the limit-influence of emotional virtual environment design on physical training in VR. *Games Res Pract.* 2024;2(4):1-26. [DOI]
71. Durkin S, Wakefield M. Interrupting a narrative transportation experience: program placement effects on responses to antismoking advertising. *J Health Commun.* 2008;13(7):667-680. [DOI]
72. Keen S. A theory of narrative empathy. *Narrative.* 2006;14(3):207-236. [DOI]
73. Statista. Umfrage in Deutschland zu gelesenen Genres von Büchern 2020 [Internet]. Hamburg: Statista. Available from: <https://de.statista.com/statistik/daten/studie/1189038/umfrage/gelesene-genres-von-buechern/>
74. Plutchik R. A psychoevolutionary theory of emotions. *Soc Sci Inf.* 1982;21(4-5):529-553. [DOI]
75. Steinhäusser SC, Piller R, Lugrin B. Combining emotional gestures, sound effects, and background music for robotic storytelling - effects on storytelling experience, emotion induction, and robot perception. In: 2024 19th ACM/IEEE International Conference on Human-Robot Interaction (HRI); 2024 Mar 11-14; Boulder, USA. Piscataway: IEEE; 2024. p. 687-696. [DOI]
76. Steinhäusser SC, Knauer L, Lugrin B. What a laugh!-effects of voice and laughter on a social robot's humorous appeal and recipients' transportation and emotions in humorous robotic storytelling. In: 2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN); 2024 Aug 26-30; Pasadena, USA. Piscataway: IEEE; 2024. p. 2131-2138. [DOI]
77. Landn B. Endlich seid ihr da! Zu Erscheinungsformen von Emotionen in Grammatikbüchern. In: Beiträge zur 12. Arbeitstagung schwedischer Germanistinnen und Germanisten; 2018; Schweden. Stockholm: Stockholm University; 2018. p. 37-59. Available from: <https://su.diva-portal.org/smash/record.jsf?pid=diva2%3A1200781>
78. Ganai E, Siol L, Lugrin B. Peput: A unity toolkit for the social robot pepper. In: 2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN); 2023 Aug 28-31; Busan, Korea. Piscataway: IEEE; 2023. p. 1012-1019. [DOI]
79. Appel M, Gnams T, Richter T, Green MC. The transportation scale-short form (TS-SF). *Media Psychol.* 2015;18(2):243-266. [DOI]
80. Gruhn D, Kotter-Gruhn D, Röcke C. Discrete affects across the adult lifespan: Evidence for multidimensionality and multidirectionality of affective experiences in young, middle-aged and older adults. *J Res Pers.* 2010;44(4):492-500. [DOI]
81. Watson D, Clark LA. The PANAS-X: Manual for the positive and negative affect schedule-expanded form. University of Iowa; 1994. [DOI]
82. Bartneck C, Kulić D, Croft E, Zoghbi S. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *Int J of Soc Robotics.* 200;1(1):71-81. [DOI]
83. JASP Team. JASP [Internet]. 2021. Available from: <https://jasp-stats.org/>
84. Roberts KT, Aspy CB. Development of the serenity scale. *J Nurs Meas.* 1993;1(2):145-164. [PubMed]
85. Floody DR. Serenity and inner peace: Positive perspectives. In: Sims G, Nelson L, Puopolo M, editors. *Personal Peacefulness*. New York: Springer; 2014. p. 107-133. [DOI]
86. Scheunemann MM, Cuijpers RH, Salge C. Warmth and competence to predict human preference of robot behavior in physical human-robot interaction. In: 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN); Naples, Italy. Piscataway: IEEE; 2020. p. 1340-1347. [DOI]
87. Roesler E, zur Kammer K, Onnasch L. Multidimensionale Fragebögen zur Erfassung der wahrgenommenen Robotermorphologie (RoMo) in der Mensch-Roboter-Interaktion. *Z Arb Wiss.* 2023;77(4):609-628. [DOI]
88. Ho CC, MacDorman KF. Revisiting the uncanny valley theory: Developing and validating an alternative to the godspeed indices. *Comput Hum Behav.* 2010;26(6):1508-1518. [DOI]
89. Steinhäusser SC, Zehe A, Maier S. Development of the Fully Automatic Robotic Storyteller—Comparing Manual, Semi-Automatic, and LLM-Based Emotional Story Annotation for Multimodal Robotic Storytelling. 2025. Available from: https://www.uni-wuerzburg.de/fileadmin/10030500/2025/Annotations_2_preprint.pdf
90. Baraka K, Paiva A, Veloso M. Expressive lights for revealing mobile service robot state. In: Reis LP, Moreira AP, Lima PU, Montano L, Muñoz-Martinez V, editors. *Robot 2015: Second Iberian Robotics Conference*. Cham: Springer; 2015. p. 107-119. [DOI]
91. Steinhäusser SC, Zehe A, Schnetter P, Hotho A, Lugrin B. Towards the development of an automated robotic storyteller: comparing approaches for emotional story annotation for non-verbal expression via body language. *J Multimodal User Interfaces.* 2024;18:1-23. [DOI]
92. Westhoven M, van der Grinten T, Mueller S. Perceptions of a help-requesting robot-effects of eye-expressions, colored lights and politeness of speech. In: Alt F, Bulling A, Döring T, editors. *Proceedings of Mensch und Computer 2019*; 2019 Sep 8-11; Hamburg, Germany. New York: Association for Computing Machinery; 2019. p. 43-54. [DOI]
93. Bainbridge WA, Hart J, Kim ES, Scassellati B. The effect of presence on human-robot interaction. In: *RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication*; 2008 Aug 1-3; Munich, Germany. Piscataway: IEEE; 2008. p. 701-706. [DOI]
94. Kiesler S, Powers A, Fussell S, Torrey C. Anthropomorphic interactions with a robot and robot-like agent. *Soc Cogn.* 2008;26(2):169-181. [DOI]
95. Lee KM, Jung Y, Kim J, Kim SR. Are physically embodied social agents better than disembodied social agents? The effects of physical embodiment, tactile interaction, and people's loneliness in human-robot interaction. *Int J Hum Comput Stud.* 2006;64(10):962-973. [DOI]
96. Liang N, Nejat G. A meta-analysis on remote hri and in-person hri: What is a socially assistive robot to do? *Sensors.* 2022;22(19):7155.

97. Woods S, Walters ML, Koay KL, Dautenhahn K. Comparing human robot interaction scenarios using live and video based methods: towards a novel methodological approach. In: 9th IEEE International Workshop on Advanced Motion Control; 2006 Mar 27-29; Istanbul, Turkey. Piscataway: IEEE; 2006. p. 750-755. [DOI]
98. Gittens CL. Remote HRI: a methodology for maintaining covid-19 physical distancing and human interaction requirements in HRI studies. *Inf Syst Front*. 2022;26:91-106. [DOI]
99. Steinhäusser SC, Heckel M, Lugin B. The way you see me-comparing results from online video-taped and in- person robotic storytelling research. In: Companion of the 2024 ACM/IEEE International Conference on Human - Robot Interaction; 2024 Mar 11-15; Boulder, USA. Piscataway: IEEE; 2024. p. 1018-1022. [DOI]
100. Donnermann M, Heinzmann P, Lugin B. Meet or call my robotic tutor? - The effect of a physically vs. virtually present social robot on learning outcomes, engagement and perception. In: Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction; 2021 Mar 11; Boulder, USA. Piscataway: IEEE; 2024. p. 422-426. [DOI]
101. Choi Y, Kim J, Pan P, Jeung J. The considerable elements of the emotion expression using lights in apparel types. In: Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology; 2007 Sep 10-12; Singapore. New York: Association for Computing Machinery; 2007. p. 662-666. [DOI]
102. Birren F. Color psychology and color therapy; a factual study of the influence of color on human life. Maryland: Pickle Partners Publishing; 2016.
103. Clarke T, Costall A. The emotional connotations of color: A qualitative investigation. *Color Res Appl*. 2008;33(5):406-410. [DOI]
104. Suk HJ, Irtel H. Emotional response to color across media. *Color Res Appl*. 2010;35(1):64-77. [DOI]
105. Son E. Visual, Auditory, and Psychological Elements of the Characters and Images in the Scenes of the Animated Film, Inside Out. *Q Rev Film Video*. 2022;39(1):225-240. [DOI]
106. Bellantoni P. Blue the detached color. In: *If It's Purple, Someone's Gonna Die*. Amsterdam: Elsevier; 2005. p. 81-84. [DOI]
107. Takahashi F, Kawabata Y. The association between colors and emotions for emotional words and facial expressions. *Color Res Appl*. 2018;43(2):247-257. [DOI]
108. Kaya N, Epps HH. Relationship between color and emotion: A study of college students. *Coll Stud J*. 2004;38(3):396-405. Available from: <https://psycnet.apa.org/record/2004-19149-009>
109. AL-Ayash A, Kane RT, Smith D, Green-Armytage P. The influence of color on student emotion, heart rate, and performance in learning environments. *Color Res Appl*. 2016;41(2):196-205. [DOI]
110. Demir Ü. Investigation of color-emotion associations of the university students. *Color Res Appl*. 2020;45(5):871-884 [DOI]
111. Bellantoni P. Yellow the contrary color. In: *If It's Purple, Someone's Gonna Die*. Amsterdam: Elsevier; 2005. p. 41-44. [DOI]
112. Nijdam NA. Mapping emotion to color. University of Twente; 2009. p. 2-9.
113. Bellantoni P. Orange the sweet and sour color. In: *If It's Purple, Someone's Gonna Die*. Amsterdam: Elsevier; 2005. p. 111-114. [DOI]
114. Manav B. Color-emotion associations and color preferences: A case study for residences. *Color Res Appl*. 2007;32(2):144-150. [DOI]
115. Bellantoni P. Green the split personality color. In: *If It's Purple, Someone's Gonna Die*. Amsterdam: Elsevier; 2005. p. 159-162. [DOI]
116. Bellantoni P. Purple the beyond-the-body color. In: *If It's Purple, Someone's Gonna Die*. Amsterdam: Elsevier; 2005. p. 189-192. [DOI]
117. Hupka RB, Zaleski Z, Otto J, Reidl L, Tarabrina NV. The colors of anger, envy, fear, and jealousy: A cross-cultural study. *J Cross-Cult Psychol*. 1997;28(2):156-171. [DOI]
118. Güneş E, Olguntürk N. Color-emotion associations in interiors. *Color Res Appl*. 2020;45(1):129-141. [DOI]
119. Oberascher L, Gallmetzer M. Colour and emotion. In: Nieves JL, Hernández-Andrés J, editors. *AIC 2005, Proceedings of the 10th Congress of the International Color Association*; 2005 May 9-13; Granada, Spain. Amsterdam: John Benjamins Publishing; 2003. p. 370-374.
120. Fugate JMB, Franco CL. What color is your anger? Assessing color-emotion pairings in english speakers. *Front Psychol*. [DOI]
121. Hammond B. Greenface-Exploring green skin in contemporary Hollywood cinema. *NECSUS*. 2013;2:213-232. [DOI]
122. Hanada M. Correspondence analysis of color-emotion associations. *Color Res Appl*. 2018;43(2):224-237. [DOI]
123. Wexner LB. The degree to which colors (hues) are associated with mood-tones. *J Appl Psychol*. 1954;38(6):432-435. [DOI]
124. Tarajko-Kowalska J, Kowalski P. "pretty in pink"—The pink color in architecture and the built environment: Symbolism, traditions, and contemporary applications. *Arts*. 2023;12(4):161. [DOI]
125. Blegvad K. The pink book: An illustrated celebration of the color, from bubblegum to battleships. California: Chronicle Books; 2019.
126. Bellantoni P. Red the caffeinated color. In: *If It's Purple, Someone's Gonna Die*. Amsterdam: Elsevier; 2005. p. 1-4. [DOI]
127. Aldebaran Robotics. Choregraphe. 2016. Available from: <https://aldebaran.com/en/support/kb/category/pepper/downloads-softwares/>.
128. LimeSurvey GmbH. LimeSurvey [Internet]. 2021. Available from: <https://www.limesurvey.org/de/>
129. McColl D, Nejat G. Recognizing emotional body language displayed by a human-like social robot. *Int J of Soc Robotics*. 2014;6(2):261-280. [DOI]
130. Trunk T, Watzlawick P. Man kann nicht nicht kommunizieren: Das Lesebuch. 1st ed. Bern: Verlag Hans Huber; 2011.
131. Tan FFY, Xu P, Ram A, Suen WZ, Zhao S, Huang Y, et al. Audiostend: Assisted reality visual accompaniments for audiobook storytelling during everyday routine tasks. In: Mueller FF, Kyburz P, Williamson JR, Sas C, Wilson ML, Dugas PT, Shklovski, editors. *Proceedings of the CHI Conference on Human Factors in Computing Systems*; 2024 May 11-16; Honolulu, USA. New York: Association for Computing

- Machinery; 2024. p. 1-22. [DOI]
132. Ekman S. Here be dragons: Exploring fantasy maps and settings. Middletown: Wesleyan University Press; 2013.
 133. Loesche D. Die lukrativsten Kinoserien [Internet]. 2017. Available from: <https://de.statista.com/infografik/10248/die-lukrativsten-kinoserien-nach-einspielergebnissen-usa/>
 134. Rowling JK. Uagadou [Internet]. 2016. Available from: <https://www.wizardingworld.com/de/writing-by-jk-rowling/uagadou>
 135. Rowling JK. Beauxbatons Academy of Magic [Internet]. 2015. Available from: <https://www.wizardingworld.com/de/writing-by-jk-rowling/beauxbatons-academy-of-magic>
 136. Rowling JK. Castelobruxo [Internet]. 2016. Available from: <https://www.wizardingworld.com/de/writing-by-jk-rowling/castelobruxo>
 137. Rowling JK. Durmstrang Institute [Internet]. 2016. Available from: <https://www.wizardingworld.com/de/writing-by-jk-rowling/durmstrang-institute>
 138. Rowling JK. Mahoutokoro [Internet]. 2016. Available from: <https://www.wizardingworld.com/de/writing-by-jk-rowling/mahoutokoro>
 139. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159. [DOI]
 140. Francisco V, Hervás R, Peinado F, Gervás P. EmoTales: creating a corpus of folktales with emotional annotations. *Lang Resour Eval*. 2012;46(3):341-381. [DOI]
 141. Itten J. The art of color: the subjective experience and objective rationale of color. Hoboken: Wiley; 1961.
 142. Busselle R, Bilandzic H. Measuring narrative engagement. *Media Psychol*. 2009;12(4):321-347. [DOI]
 143. Röcke C, Gruhn D. German translation of the PANAS-X [Interent]. 2003. Available from: <https://acelab.wordpress.ncsu.edu/files/2019/07/PANAS-X-German.pdf>
 144. Carpinella CM, Wyman AB, Perez MA, Stroessner SJ. The robotic social attributes scale (RoSAS). In: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*; 2017 Mar 6-9; Vienna, Austria. New York: Association for Computing Machinery; 2017. p. 254-262. [DOI]
 145. Steinhäusser SC, Lein M, Donnermann M, Lugin B. Designing social robots' speech in the hotel context-a series of online studies. In: *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*; 2022 Aug 29-2022 Sep 1; Napoli, Italy. Piscataway: IEEE; 2022. p. 163-170. [DOI]
 146. Platz F, Kopiez R. When the eye listens: A meta-analysis of how audio-visual presentation enhances the appreciation of music performance. *Music Percept*. 2012;30(1):71-83. [DOI]
 147. Tal-Or N, Cohen J. Unpacking Engagement: Convergence and Divergence in Transportation and Identification. *Commun Yearb*. 2016;40(1):33-66. [DOI]
 148. Tussyadiah IP, Park S. Consumer evaluation of hotel service robots. In: *Stangl B, Pesonen J, editors. Information and Communication Technologies in Tourism 2018*. Cham: Springer; 2018. p. 308-320. [DOI]
 149. de Graaf MMA, Allouch SB. Exploring influencing variables for the acceptance of social robots. *Robot Auton Syst*. 2013;61(12):1476-1486. [DOI]
 150. Youssef K, Said S, Alkork S, Beyrouthy T. A survey on recent advances in social robotics. *Robotics*. 2022;11(4):75. [DOI]
 151. Pandey AK, Gelin R. A mass-produced sociable humanoid robot: Pepper: The first machine of its kind. *IEEE Robot Autom Mag*. 2018;25(3):40-48. [DOI]
 152. Pot E, Monceaux J, Gelin R, Maisonnier B. Choregraphe: a graphical tool for humanoid robot programming. In: *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*; 2009 Sep 27-Oct 2; Toyama, Japan. Piscataway: IEEE; 2009. p. 46-51. [DOI]
 153. Buchem I, Tutul R, and Bäcker N. Same task, different robot. comparing perceptions of humanoid robots nao and pepper as facilitators of empathy mapping. In: *Biele C, Kacprzyk J, Kopeć W, Możaryn J, Owsirski JW, Romanowski A, Sikorski M, editors. Digital Interaction and Machine Intelligence*. Cham: Springer; 2024. p. 133-143. [DOI]
 154. Thorstenson CA, Tian Y. Dimensional approach for using color in social robot emotion communication. *Color Res Appl*. 2025. [DOI]
 155. Bernotat J, Eyssel F, Sachse J. Shape it-The influence of robot body shape on gender perception in robots. In: *Kheddar A, Yoshida E, Ge SS, Suzuki K, Cabibihan JJ, Eyssel F, He H, editors. Social Robotics: 9th International Conference, ICSR 2017*; 2017 Nov 22-24; Tsukuba, Japan. Cham: Springer; 2017. p. 75-84. [DOI]
 156. Steinhäusser SC, SiolL, Ganai E, Maier S, Lugin B. The narrobot plugin - Connecting the social robot reeti to the unity game engine. In: *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*; 2023 Mar 13-16; Stockholm, Sweden. New York: Association for Computing Machinery; 2023. p. 65-70. [DOI]
 157. Stock-Homburg R. Survey of emotions in human-robot interactions: Perspectives from robotic psychology on 20 years of research. *Int J of Soc Robotics*. 2022;14:389-411. [DOI]
 158. Halpern D, Katz JE. Unveiling robotophobia and cyber-dystopianism: the role of gender, technology and religion on attitudes towards robots. In: *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*; 2012 Mar 5-8; Boston, USA. New York: Association for Computing Machinery; 2012. p. 139-140. [DOI]
 159. Ho CC, Macdorman KF, Pramono ZADD. Human emotion and the uncanny valley: a GLM, MDS, and Isomap analysis of robot video ratings. In: *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*; 2008 Mar 12-15; Amsterdam, Netherlands. New York: Association for Computing Machinery; 2008. p. 169-176. [DOI]
 160. Spjeldnæs K, Karlsen F. How digital devices transform literary reading: The impact of e-books, audiobooks and online life on reading habits. *New Media Soc*. 2024;26(8):4808-4824. [DOI]

161. Ji D, Liu B, Xu J, Gong J. Why do we listen to audiobooks? The role of narrator performance, bgm, telepresence, and emotional connectedness. *Sage Open*. 2024;14(2):21582440241257357. [DOI]
162. Zabala U, Diez A, Rodriguez I, Augello A, Lazkano E. Attainable digital embodied storytelling using state of the art tools, and a little touch. In: Ali AA, Cabibihan JJ, Meskin N, Rossi S, Jiang W, He H, Ge SS, editors. *Social Robotics: 15th International Conference; 2023 Dec 3-7; Doha, Qatar*. Singapore: Springer; 2023. p. 68-79. [DOI]
163. Döring N, Bortz J. *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin: Springer; 2016.
164. Walters ML, Syrdal DS, Koay KL, Dautenhahn K, te Boekhorst R. Human approach distances to a mechanical-looking robot with different robot voice styles. In: *RO-MAN 2008—The 17th IEEE International Symposium on Robot and Human Interactive Communication; 2008 Aug 1-3; Munich, Germany*. Piscataway: IEEE; 2008. p. 707-712. [DOI]
165. Oh J, Jeong SY, Jeong J. The timing and temporal patterns of eye blinking are dynamically modulated by attention. *Hum Mov Sci*. 2012;31(6):1353-1365. [DOI]
166. Lein M, Donnermann M, Steinhäusser SC, Lugin B. Reiteratively designing a robotic concierge with different stakeholders- A multi-methods field study. In: *2024 IEEE International Conference on Advanced Robotics and Its Social Impacts (ARSO); 2024 May 20-22; Hong Kong*. Piscataway: IEEE; 2024. p. 55-60. [DOI]
167. Smedegaard CV. Reframing the role of novelty within social HRI: from noise to information. In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI); 2019 Mar 11-14; Daegu, Korea*. Piscataway: IEEE; 2019. p. 411-420. [DOI]
168. Bornstein RF, Craver-Lemley C. Mere exposure effect. In: Pohl RF, editor. *Cognitive Illusions*. London: Routledge; 2022. p. 241-258.
169. Fiolka AK, Donnermann M, Lugin B. Investigating the mere exposure effect in relation to perceived eeriness and humaneness of a social robot. In: *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction; 2024 Mar 11-15; Boulder, USA*. New York: Association for Computing Machinery; 2024. p. 453-457. [DOI]

Publisher's Note

Science Exploration remains a neutral stance on jurisdictional claims in published maps and institutional affiliations. The views expressed in this article are solely those of the author(s) and do not reflect the opinions of the Editors or the publisher.