# Drone in Love: Emotional Perception of Facial Expressions on Flying Robots

Viviane Herdel[1,2]
herdel@post.bgu.ac.il
[1]Magic Lab, Industrial Engineering & Management
Ben Gurion University of the Negev
Be'er Sheva, Israel

Anastasia Kuzminykh[1,3*]
anastasia.kuzminykh@utoronto.ca
[3]Faculty of Information
University of Toronto
Toronto, Canada

Andrea Hildebrandt[2]
andrea.hildebrandt@uol.de
[2]Department of Psychology
Carl von Ossietzky Universität Oldenburg
Oldenburg, Germany

Jessica R. Cauchard[1]
jcauchard@acm.org
[1]Magic Lab, Industrial Engineering & Management
Ben Gurion University of the Negev
Be'er Sheva, Israel

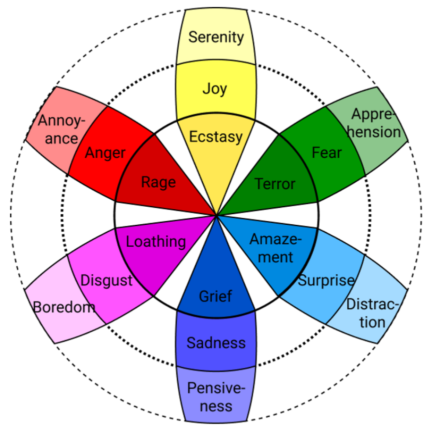



**Figure 1: "The drone looks like it is in love." Example of a participant's reasoning when deciding on the drone's emotional state based on its facial expression (left). Wheel of six emotions derived from Plutchik's theory of emotion [65] (right).**

## ABSTRACT

Drones are rapidly populating human spaces, yet little is known about how these flying robots are perceived and understood by humans. Recent works suggested that their acceptance is predicated upon their sociability. This paper explores the use of facial expressions to represent emotions on social drones. We leveraged design practices from ground robotics and created a set of rendered robotic faces that convey basic emotions. We evaluated individuals' response to these emotional facial expressions on drones in two empirical studies ($N$ = 98, $N$ = 98). Our results demonstrate that individuals accurately recognize five drone emotional expressions, as well as make sense of intensities within emotion categories. We describe how participants were emotionally affected by the drone, showed empathy towards it, and created narratives to interpret its emotions. As a consequence, we formulate design recommendations for social drones and discuss methodological insights on the use of static versus dynamic stimuli in affective robotics studies.

## CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous and mobile devices**; **Empirical studies in collaborative and social computing**; • **Computer systems organization** → *Robotics*.

## KEYWORDS

Human-Drone Interaction, Affective Computing, Emotion Recognition, Facial Expressions, Anthropomorphism, Robot, UAV.

**ACM Reference Format:**
Viviane Herdel[1,2], Anastasia Kuzminykh[1,3], Andrea Hildebrandt[2], and Jessica R. Cauchard[1]. 2021. Drone in Love: Emotional Perception of Facial

[*]Also with Cheriton School of Computer Science, University of Waterloo.


*CHI '21, May 08–13, 2021, Yokohama, Japan*

ACM ISBN 978-1-4503-8096-6/21/05…$15.00
https://doi.org/10.1145/3411764.3445495

Expressions on Flying Robots. In *CHI Conference on Human Factors in Computing Systems (CHI '21), May 8–13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 20 pages. https://doi.org/10.1145/3411764.3445495

# 1 INTRODUCTION

Over the last few years, small-size drones have become increasingly popular, being used in a wide range of applications from photo and videography to deliveries and search and rescue [78]. Early works on collocated Human-Drone Interaction (HDI) observed that people interacted with the drone as with a person or an animal [18, 24]. Recently, researchers have highlighted novel opportunities created by *social drones* that operate in human spaces and can support people in their daily lives, such as when exercising [57], to get home safely [43], as a navigational aid [2, 14, 20], and even as a personal companion [41, 45]. Yet, designing a social drone [7] is not trivial, and we are only at the beginning of understanding which factors influence people's perception of drones [81]. However, the literature on interacting with ground robots is rich and teaches us that social robots can communicate with people using emotions and expressive behaviors designed around features such as: facial expression, posture, gesture, and voice [10]. This results in interactions that are informative, "human-like", and pleasant [32]. Unfortunately, findings from ground robotics cannot be directly translated into drones [81]. For instance, prior work showed that robots with eyes and no mouth are perceived as unfriendly [40], while drones with equivalent facial features are perceived as likable and warm [70].

We address this gap in the literature by designing facial expressions to convey emotional states on a social drone. Our focus on facial expressions is motivated by their significance as a non-verbal communication channel in human-robot [40] and in human-human communication [61, 64], as they trigger the tendency to read emotions and interpret intentions and personality traits [59, 77]. Thus, the use of facial expressions in social drones might have a potential to, not only communicate the drone's state, but also to elicit particular reactions and behaviors from a user.

We designed six drone facial expressions and evaluated them in two online user studies ($N = 98$, $N = 98$) where we investigated how people recognize and interpret a drone's emotional state using both static and dynamic stimuli. Our results show that 5 different emotions (*Joy*, *Sadness*, *Fear*, *Anger*, *Surprise*) can be recognized with high accuracy in static stimuli and 4 emotions (*Joy*, *Surprise*, *Sadness*, *Anger*) in dynamic videos. Surprisingly, participants created narratives around the drone's emotional states, and either imagined that the drone's state is caused by external factors (environment or people) or that the drone affects its environment or people within it. Participants further envisioned themselves as involved in the scene, they described empathy towards the drone, which then triggered mentions of prosocial behaviors.

Our work contributes in the following:

- A set of five rendered robotic faces representing *Joy, Sadness, Fear, Anger,* and *Surprise.*
- Two user studies ($N = 196$) showing how people recognize, interpret, and are affected by emotions on drones.
- Design recommendations for social drones using emotions and facial features.
- Methodological insights on the use of static vs. dynamic stimuli in affective robotics.

# 2 RELATED WORK

We present the state of the art on emotional robotics for both ground and aerial robots, and discuss the use of facial expressions in conveying emotions.

## 2.1 Affective Robotics

The enriching effects of integrating convincing emotions into non-human agents has been extensively researched in the robotics domain [5, 12, 75]. Eyssel et al. [29] showed that users tend to perceive the interaction with an emotional robot as more pleasant, feel closer to it, and ascribe human attributes to it, such as intentionality. Attribution of intentions, in turn, can foster feelings of social connection, empathy, and prosociality [44, 80]. Furthermore, there is evidence that robots' abilities to display emotions contribute, not only to their overall perceived relatability and sociability, but also to the effectiveness of their communication with a user. Leite et al. [50] found that when a robot acts as a game companion, their emotional behavior can help users better understand the game. In their recent work, Fischer et al. [31] showed that people expect robots to express emotions reactively, similarly to the conventionally required display of emotions between humans [73]. Indeed, we know from the research on human-human communication that displaying emotions has crucial communicative and social functions [33, 46], such as forming guidelines for future behavior to avoid or elicit emotional outcomes [6]. Research showed that affective robots similarly induce people to make sense of their intentions to guide human behaviors [29, 66]. The communicative and social functions of displaying emotions motivates the interest in the exploration of its application to the design of social drones [7, 19]. In particular, it was recently argued that, in human spaces, drones need to present social features [7]. The researchers name the intuitive comprehension of drones' intentions – to what degree people are able to interpret intentions that the drone is trying to convey via the interaction – as one of the major human-centered concerns in the design of social drones. One major challenge is then to investigate how to appropriately embed the display of emotions into the interaction design of social drones, including *how to display* emotions and *what emotions* are reasonable to display.

## 2.2 Conveying Emotion in Aerial Robotics

In both human-human and human-robot communication, the display of emotions occurs though external, i.e., visible and audible, behavioral manifestations. Such manifestations can take diverse forms [46], such as verbal and non-verbal elements of language, gesture, posture, and gaze. Correspondingly, in robotics, researchers have explored diverse ways to convey emotions. For instance, there is a large body of work on affective perceptions of robots' body movements, sound and color, and diverse combinations of these elements [37, 38, 52]. Furthermore, researchers found that people's ability to anthropomorphize objects [23, 30, 68] provides a supporting mechanism in the design of social robots, allowing to translate some of the social conventions, expectations, and perception mechanisms from human-human to human-robot contexts [22, 47, 48]. Drones, however, because of their flight capability tend to be non-anthropomorphic in nature, and as such, their design differs from their ground counterparts [81].

In drone design, the communication of emotions has been predominantly explored through flight path [19, 74]. Sharma et al. [74] proposed expressive flights following the Laban Effort System [60] and showed that people can differentiate between drone states along the valence and arousal dimensions. Cauchard et al. [19] later defined an emotional model space for drones and showed that humans can accurately associate a drone's movements and behavior to an emotional state corresponding to a personality model. Arroyo et al. [1] investigated other means to convey a drone's emotional state using head movement, eye color, and the propellers' noise. They demonstrated that different states of these elements are being selectively associated with positive, negative, and neutral emotions. Additional efforts have been conducted towards establishing design recommendations for social drones, suggesting the suitability of faces [41] that could convey friendly features [84], such as Kim et al. [45] who further proposed that an ideal companion drone should present "adorability" features. Wojciechowska et al. [81] quantified how physical properties, such as facial features, influence people's perception of drones, such as eyes, which increase perceived friendliness, likeability, and a person's willingness to interact with the drone. Although these works did not investigate emotional expression per se, they open the space to the use of facial features on drones.

### 2.3 Facial Expressions and Emotions

In human-human communication, information conveyed through one's face plays a fundamental role in interactions [25], leading to the human's strong ability to use the facial information to infer emotional states, personality traits, and intentions [35, 61]. In affective robotics, while facial expressions received much attention [4, 11, 15], the consensus on their recognition and interpretation is still lacking [76]. Prior works further highlighted that the accuracy of emotion recognition for artificial emotions differ across robot faces' designs [11, 17] and that it might depend on the displayed intensity of the emotion [4]. Moreover, it is unclear how findings from affective ground robotics apply to drones given that there is evidence that robots with different shapes trigger different emotions in humans [39], and given that drones tend to be non-anthropomorphic in nature [81].

Facial features are a key factor in creating affective robots, and it was shown that robots without a face are perceived as less sociable and amiable compared to robots with a face [13]. This is in line with findings from drone literature showing that the presence of facial features influence the perception of drones as more likable, trustworthy, and intelligent [70, 81] compared to drones without facial features. While a first attempt has been made at designing drone facial features (eyes) to enhance non-verbal communication with humans [79], the question of the recognition and interpretation of emotions displayed by drones remains open. This further highlights the challenge of identifying what emotions are relevant and appropriate to display in HDI. For instance, research in psychology shows that, inter-personally, some emotions are better recognized than others [27], the recognition speed differs for positive and negative emotions [51], and similar expressions are often confused [69, 71].

For the purpose of this work, we focused on the six basic emotions [25]: *Joy, Sadness, Fear, Anger, Surprise,* and *Disgust*, along three levels of intensity: Low, Medium, and High. While some robotics researchers have used extended sets of emotions (e.g., [3, 11, 31]), our rationale for using basic emotions is two-fold. First, there is convincing evidence that the basic emotions are universally recognized by humans, independently of their culture [25]. Second, this set of six primary emotions is commonly applied in research on emotional perception of robots [4, 11, 17]. Building upon prior work, this paper explores the possibility to effectively convey drone's emotions through facial expressions. In the following section, we discuss our approach to the development of facial expressions and describe our design choices to display basic emotions on drones.

## 3 DESIGNING EMOTIONS FOR DRONES THROUGH FACIAL FEATURES

While affective robotics offers several developed sets of faces with emotion expressions (e.g., [4, 11, 17]), the emotion recognition rates between these existing faces are inconsistent [76]. Additionally, there is an open question of whether these sets of faces are appropriate for drones. These considerations motivated our decision to develop a novel set of drone faces that would allow us to explore the perception of emotional facial expressions on drones for six basic emotions each with three intensity levels. In this section, we describe the corresponding design process, including the chosen representation style, facial features, face canvas, and approach to the construction of emotional expressions.

### 3.1 Representation Style

The first design choice we faced was concerned with the style of the faces, which reflects on the degree of realism. Several approaches exist to the design of facial expressions to communicate emotions [40, 66]. Some artificial faces attempt to mimic the look of the human face, e.g., by using the Delsarte's code of facial expressions [15] or by establishing pseudo-facial expressions by "cloning" real human expressions onto an avatar [85]. Others choose to develop animal-like designs [11, 76] or cartoon-like facial expressions [4] inspired by principles of animations [67], often minimizing the amount of facial elements [66]. Our design used a cartoon-like 2D format since such faces led to higher emotion recognition compared to photo-realistic faces in prior work [42, 86], and to minimize the risk of falling into the uncanny valley [54], which can trigger undesired emotional reaction [53, 56].

### 3.2 Constructing the Face

The chosen cartoon-like format allowed us to minimize the number of included facial features which contribute to reducing the cognitive efforts needed for a person to process the resulting facial expressions [42]. Our next step was to identify the minimal set of features required to convey the chosen set of emotions. We used the well-established Facial Action Coding System (FACS) [26], which documents single muscle units required to create universally recognizable basic emotions (Table 1). This provided us with the necessary systematical approach needed for creating emotional facial expressions for drones. Furthermore, we chose to include pupils, as rendered robot faces without pupils are perceived negatively [40]. Our final resulting set included: eyes, eyebrows, pupils, and mouth. Finally, to assemble the set of chosen facial features into a face, we needed to decide on the face canvas. We turned to the existing body of work on rendered robot faces to find a suitable base for our work [40]. We opted for the face of Omate Yumi[1], chosen as best-suited for a domestic robot, as it presented high likeability and fit our chosen set of facial features.

**Table 1: Colored cells represent Action Units (AU) from the Facial Action Coding System (FACS) [26] required to create specific basic emotions.**

| AU | FACS Name | Joy | Sadness | Fear | Anger | Surprise | Disgust |
|---|---|---|---|---|---|---|---|
| 1 | Inner brow raiser | | ■ | ■ | | ■ | |
| 2 | Outer brow raiser | | | ■ | | ■ | |
| 4 | Brow lowerer | | ■ | □ | ■ | | |
| 5 | Upper lid raiser | | | □ | □ | □ | |
| 6 | Cheek raiser | □ | | | | | |
| 7 | Lid tightener | | | ■ | ■ | | |
| 9 | Nose wrinkler | | | | | | □ |
| 12 | Lip corner puller | ■ | | | | | |
| 15 | Lip corner depressor | | ■ | | | | ■ |
| 16 | Lower lip depressor | | | | | | ■ |
| 20 | Lip stretcher | | | ■ | | | |
| 23 | Lip tightener | | | | ■ | | |
| 26 | Jaw drop | | | ■ | | ■ | |

*Notes.* AU corresponding to dark colored cells were used to design the drone faces, while AU corresponding to light colored cells were not manipulated.

## 3.3 Designing Emotions on the Face

For each of the basic emotions, we designed representations with three levels of intensity by intensifying the corresponding Action Units. We put special attention in the design of each feature to represent the emotions by extensively surveying robot faces in the literature and on the market. The resulting core set of rendered faces (Figure 2) includes 18 images of cartoon-like facial expressions (6 basic emotions $\times$ 3 intensity levels). We additionally designed a neutral face.

# 4 METHODOLOGY

To explore the recognition, interpretation, and reactions to the emotional facial expressions on drones, we conducted two empirical studies, both employing a mixed-methods approach. Study I explored the perception of emotional facial expressions of different intensities presented statically (image-based). Study II was conducted four months after Study I and addressed the perception of animated emotional facial expressions on drones presented dynamically (video-based). We investigated both static and dynamic stimuli, as prior work had discussed [40] and proved [11, 70] that these stimuli can elicit different responses in humans. In this section, we describe the participants, stimuli, tasks, study procedures, and the process of data preparation, and analysis.

## 4.1 Participants

Participants were recruited using the Amazon mechanical Turk platform. The recruitment selection was based on HIT rate ($\geq$ 97) and approved number of HITs ($\geq$ 100) with all participants located in the US. The resulting samples included $N_1$ = 98 (Study I) and $N_2$ = 98 (Study II). We found that 31 individuals participated in both studies; yet, their recognition rate did not significantly differ from other participants. Additional demographic information are provided in Table 2. Participants were remunerated US$4.90 (Study I) and US$3.70 (Study II) with a bonus of respectively $0.5-$1 and $0.5-$2 for elaborate answers on the open-ended questions. The surveys took in average respectively 30 and 25 minutes to complete. Note that we initially recruited 100 participants and that 2 were discarded in each study due to technical issues.

**Table 2: Demographic information from participants in Study I and II.**

| | | Study I | Study II |
|---|---|---|---|
| **Number of Participants** | | 98 | 98 |
| **Gender** | Male | 49 | 49 |
| | Female | 48 | 48 |
| | Prefer not to say | 1 | 1 |
| **Age** | *M* | 40 | 42 |
| | *SD* | 8.7 | 9.9 |
| | Range | 23-69 | 25-68 |
| **Education (%)** | College degree or higher | 69 | 76 |

## 4.2 Stimuli

Image and video stimuli are commonly used in perception studies in the robotics literature [31, 40, 53, 81]. We presented the developed set of faces (Figure 2) on a screen display embedded on the DJI Phantom 3 body[2]. Since design elements have been demonstrated to affect drone perception, and to avoid potential biases in emotion perceptions [62], the body was chosen for its average scores (e.g., on friendliness) in a prior perception study [81]. In Study I, we presented 18 stimuli images (6 emotions $\times$ 3 intensities) each displaying one of the developed facial expressions on the drone's body (e.g., Figure 1 left). In Study II, the drone was presented in 16.6 seconds video clips (e.g., Figure 3). The drone was shown approaching in a straight line for 10 seconds. While the drone moved, its face displayed a neutral expression. Once stopped, its face changed from neutral to low, medium, and high intensity of emotion (in 600 ms). We used the developed set of faces (Figure 2) as key-frames and blended them in the Unity[3] game-engine to create the animated facial expressions. After reaching high intensity, the drone remained on display for an additional 6 seconds. The chosen speed, distance, and movement of the drone are based on the literature on comfortable approach [82]. In total, Study II included five video stimuli, one per emotion of: *Joy*, *Sadness*, *Fear*, *Anger*, and *Surprise*. *Disgust* was omitted based on low accuracy results from Study I.

## 4.3 Tasks Description

This section describes the main tasks of Study I and II.

*4.3.1 Study I △.* It consisted in two main tasks for each image stimulus which investigated how well static facial expressions of emotions of different intensities on a drone can be recognized and differentiated.

[1]https://www.omate.com/

[2]https://www.dji.com/phantom3-4k

[3]https://unity.com/

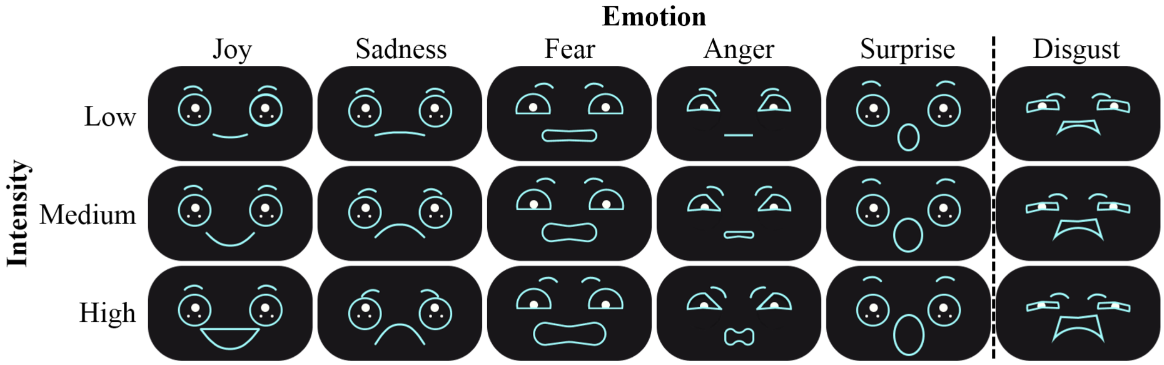



**Figure 2: Set of rendered faces representing six basic emotions in three different intensity levels ©Viviane Herdel. The faces use four core facial features: eyes, eyebrows, pupils, and mouth based upon FACS (Table 1). All emotion categories performed well, only *Disgust* did not perform as well as the other emotion categories.**

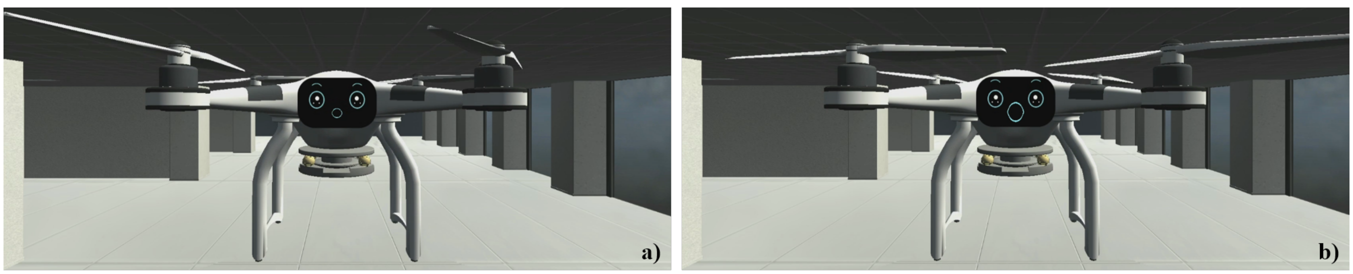



**Figure 3: Study II. Screenshots from video stimuli of drone displaying *Surprise*. The emotion intensity increases from a) to b).**

In Task 1, participants first chose an emotion label to best describe the expression on the drone's face (see Figure A.1). We used a forced choice paradigm [11, 17, 19] to facilitate homogeneity and comparability of answers, given the wide variations in language that people use to define emotions. The set of emotion labels was provided using a modified version of Plutchik's wheel of emotions [65], presented side-by-side with each stimuli to avoid memory biases (Figure 1). Our wheel shows the six basic emotion categories [25] with labels describing high (inner circle), medium (middle circle), and low intensities (outer circle) of emotions corresponding to the 18 stimuli images. With respect to their best-choice answer, participants were then asked to justify their choice (free-form answer), specify how confident they were with their answer (7-point Likert scale), and rate the intensity of the facial expression (7-point Likert scale). Finally, we asked participants to select any additional emotion labels that could also apply the drone's facial expression.

In Task 2, the 18 image stimuli were presented all at once, alongside with a table of 18 cells labeled with six emotions and three intensities each (e.g., "Sadness + Low Intensity"). Participants were asked to drag and drop each image into the cell that best fit the drone's facial expression. Note that participants did not receive any feedback regarding the validity of their answers.

*4.3.2 Study II* □. There were two main tasks for each video stimulus. Task 1 aimed to assess the participants' emotional response towards the drone and Task 2 measured how well dynamic facial expressions of emotions could be recognized, and how they would be interpreted by participants.

In Task 1, participants were first presented with a video stimulus, which they could watch multiple times (see Figure A.2). They were then asked to rate how they felt towards the drone using the Self-Assessment Manikin (SAM) [9], a 9-point Likert scale using sketches of a manikin to measure emotions along three dimensions: Valence (from negative to positive emotions), Arousal (from low to high intensity), and Dominance (from submissive to in control).

In Task 2, participants were asked to watch the same video again and to select the emotion category that best matches the drone's facial expression on the original Plutchik's wheel of eight emotions [65]. As per Study I, participants then had to justify their best-choice answer (free-form answer) and specify their level of confidence (7-point scale). Finally, participants had the opportunity to select any other additional emotion categories that could also apply.

## 4.4 Procedure

Both studies followed a within-subject design and were conducted as online surveys. In both studies, participants were first presented with the study description and asked to sign a consent form.

*4.4.1 Study I* △. After reading the instructions, participants performed a training session using the neutral face to get familiarized with Task 1. Participants then filled in a demographic questionnaire. The study proceeded to the first block of tasks where they performed Task 1 for each of the 18 (static) stimuli images presented to them one-by-one in a random order. This first block of tasks was followed by a control, attention-check question, and by the second task thereafter. The study concluded with additional questions including participants' prior experience with drones.

*4.4.2 Study II* □. As per Study I, the instructions were followed by a training session using the drone's neutral facial expression which provided a baseline for Tasks 1 and 2. The study then proceeded to the core phase where participants were presented with the five video stimuli one-by-one in a random order. For each video, participants completed first Task 1 and then Task 2. Half way through participants answered a control question. The study concluded with a demographic questionnaire and a series of additional questions such as what participants thought about the drone in the videos and their prior experience with drones.

## 4.5 Dependent Variables & Analysis

The two studies generated a consequent amount of data. The dependent variables collected from both studies that are further described in the result section are summarized in Table 3. The measures of recognition rate and intensities (I a, I b, II a) are standard in the literature [4, 11, 17, 76]. In addition, we measured how participants felt towards the drone (II b). Furthermore, the qualitative data gathers participants' reasoning on their choice of emotion label (I c, II c) and opinion of the drone (II d).

*4.5.1 Qualitative Analysis.* The qualitative data was analyzed using a thematic analysis for all free-form answers. This exploratory method strives to identify patterns of themes that depend on the related data (as in [45, 83]). First, incremental open coding [21] was performed by the primary author, and codes were discussed and refactored in consultation with additional members of the research team. Next, we performed axial coding, seeking relationships between the lower-level coding. This resulted in two coding schemes: one used for both studies, justification for the best-choice of emotion label (see Table 4), and one for the participants' opinion of the drone provided in Study II (see Table 5). After the primary researcher coded the entire dataset, a second member of the team independently reviewed the coded dataset. Disagreements were resolved in discussions amongst team members. Quotes were separated into elements respective to the categories. For each element that applied to a specific category, we incremented the occurrences by 1 in the respective category.

The within-subject study design of both studies led to paired samples, as all participants evaluated all stimuli. Thus, we used a linear mixed-effects model which is appropriate for paired data with the advantage that a Poisson link function can be used. This was here necessary for count dependent variables (non-negative integers, heteroscedastic, skewed). With the described model, we analyzed whether categories differed significantly within emotions (e.g., whether external factors were significantly named more often than internal factors within each emotion category) or across emotions (e.g., whether external factors were significantly named more often in *Surprise* compared to *Sadness*). We did not perform a statistical analysis across studies as they collected uneven data and contained a different number of stimuli (18 images with different emotion intensities vs. 5 video clips created upon the images), thus generating a different amount of count-data.

*4.5.2 Quantitative Analysis.* For Study I, we calculated the proportion of how often images belonging to an emotion category were associated with that category (e.g., images: serenity, joy, and ecstasy; participant's choice: serenity or joy or ecstasy) (see Section 5.2.1). In addition, we calculated the percentages of how often participants selected both the correct emotion category and intensity (e.g., image: serenity; participant's choice: serenity) and how often they selected the wrong emotion category or the correct emotion category but wrong intensity (e.g., image: serenity; participant's choice: apprehension) for each of the 18 images of drones (see Section 5.3.1). Two resulting confusion matrices were computed, one for emotion category recognition, and one for emotion intensity label. For each confusion matrix, we analyzed whether the emotion/intensity were selected above random choice. To test whether the participants' best-choice answers were significantly above random choice, we used a binomial test for choosing both the correct emotion intensity label (random choice = 1/18) and the correct emotion category (random choice = 1/6). We used the Benjamini-Hochberg procedure to control false discovery rates (FDR) and thus the proportion of Type I errors. When more than one emotion intensity, or emotion category, was significantly above random choice for a stimulus, we additionally tested whether the number of best-choice answers differed significantly from each other ($\chi^2$-goodness of fit test). We used the same statistical procedure to analyze emotion classification accuracy in Study II.

To analyze the SAM data in Study II, we applied a difference score path model (DSM) [55] for each of the three scales: Valence, Arousal, and Dominance. The aim was to quantify differences in participants' emotional responses after seeing a video of a drone with one of the five emotions (emotional condition) as compared with a drone displaying a neutral expression (baseline). This study design led to paired samples, as all participants were exposed to the baseline, as well as to all emotional stimuli. In DSMs, the variance of the comparison condition (i.e., emotional condition) is decomposed onto baseline (i.e., neutral condition) variance and difference to the baseline. Variance decomposition is achieved by model constraints, specifically on the regression weights of the baseline and the difference score predicting the comparison condition (fixed to 1). Additionally, the residual variance of the comparison condition is being constrained to 0. All five emotions were included in DSMs fitted separately for Valence, Arousal, and Dominance. Thus, each DSM contained five difference scores (emotional vs. neutral condition), which allowed estimating sample average differences in ratings of emotional as compared to neutral drones, as well as individual differences therein.

# 5 RESULTS

We report the results of Study I △ and Study II □. All statistical tests discussed as *demonstrating statistically significant results* have a $p$-value $< .05$. Example quotes provide additional information about the participant number and whether the participant belong to Study I: S (Static △) or to Study II: D (Dynamic □) stimuli (e.g., S44 corresponds to participant 44 from Study I). We describe the results indicating how individuals perceive facial expressions in drones,

**Table 3: Dependent variables collected in both studies that are further described in the results section.**

| | |
|---|---|
| **Study I** △ | Ia) Recognition rate of six emotion categories: *Joy, Sadness, Fear, Anger, Surprise, Disgust.*<br>Ib) Recognition rate of three intensities: low, medium, and high, for each emotion category.<br>Ic) Free-form answers justifying the best-choice of emotion. |
| **Study II** ◻ | IIa) Recognition rate of five emotion categories: *Joy, Sadness, Fear, Anger, Surprise.*<br>IIb) Self-Assessment Manikin (SAM) answers in Valence, Arousal, and Dominance.<br>IIc) Free-form answers justifying the best-choice of emotion.<br>IId) Free-form answers regarding participants' opinion of the drone. |

**Table 4: Description of the coding scheme identified in the thematic analysis of the qualitative data from the free-form answers. Mentions of specific elements are referred to as "naming". For example, when a participant reported "the eyes are shining", we counted one occurrence of naming the facial feature 'eye' and of an interpretive descriptor 'shining'.**

| Theme | Category | Description |
|---|---|---|
| **Facial Features** (FF) | Facial feature name | Naming of displayed FF: e.g., mouth, eyes, eyebrows, face. |
| | Invented | Naming of imaginary FF: e.g., teeth, tongue, nose. |
| | Descriptive | Purely informative description about the FF appearance (e.g., big, raised). |
| | Interpretive | Interpretations of FF beyond pure form factor (e.g., frowned, teary). |
| **Emotion Naming** | Provided | Emotions provided in the corresponding wheel of emotions. |
| | Invented | Emotions not provided in the corresponding wheel of emotions. |
| **Drone's State** | Internal | No external factors are mentioned when explaining the drone's state. |
| | External | At least one external factor is mentioned when explaining the drone's state. |
| | *– Effect –* | |
| | Drone to Environment | The drone affects its environment. |
| | Drone to Human | The drone affects a person. |
| | *– Cause –* | |
| | Environment to Drone | The drone's state is caused by its environment. |
| | Human to Drone | The drone's state is caused by a person. |

**Table 5: Description of the coding scheme identified in the thematic analysis of the qualitative data from the free-form answers regarding participants' opinion of the drone**

| Theme | Category | Description |
|---|---|---|
| **Drone & Participants** | Empathy | Mentions of either emotion contagion or empathy. |
| | Motivation to Interact | Mentions of wanting to (prosocially) interact with the drone. |
| | Expectations | Mentions of expectations or preferences around the drone. |
| **Emotions** | Positive | Mentions of positive emotions caused by the drone. |
| | Negative | Mentions of negative emotions caused by the drone. |

considering the role and interpretation of facial features, emotion and intensity recognition, and the understanding of the drone's emotional state. We then report how participants were emotionally affected by the perceived drone emotions.

## 5.1 Role & Interpretation of Facial Features △ ◻

This section describes the role of individual facial features in the recognition of drone emotions. When asked to explain their reasoning in choosing emotions, participants most often referred to specific facial features. Our results reveal the role of these features in making sense of emotions. Figure 4 illustrates absolute frequencies of facial feature naming for static △ (left) and dynamic ◻ (right) stimuli. Interestingly, participants did not only name existing facial features. They also invented new ones, such as teeth, tears, and lips. Invented features were thus included in the analysis for static stimuli △. This was not the case for the dynamic stimuli ◻ where invented features were only named 7 times overall. We describe below the role of each facial feature and their nature of mentioning: **descriptive** vs. **interpretive** (see Table 4). Effect estimates stem from linear mixed-effect models with Poisson link function. They indicate fixed effects, thus non-standardized regression weights $b$ quantifying count differences of naming a given facial feature as **compared to other features within an emotion category** or count differences of naming **a given facial feature across emotion categories**. All reported $b$s are statistically significant at an $\alpha$ level .05. Positive values indicate more counts for the respective facial feature.

*5.1.1 Role of Facial Features Within Emotions △ ◻.* We first describe the frequency pattern of named facial features within each emotion. We found for all emotions that some features were named more

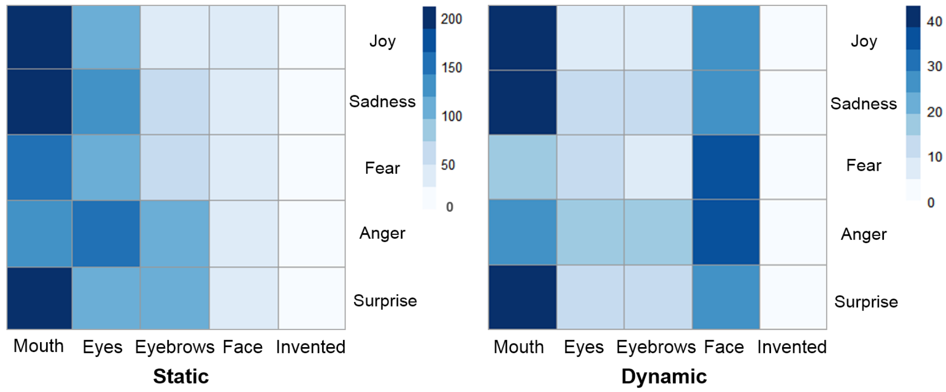



**Figure 4: Absolute frequency of facial feature naming for static (Left) and dynamic (Right) stimuli.**

often than others (△, □). For example, with the static stimulus △ of *Joy*, the mouth was named most frequently compared to all other facial features.

We found that the sequence pattern of facial feature naming frequency for static stimuli △ was: *mouth* > *eyes* > *eyebrows* > *face* > *invented*. Interestingly, the sequence was different when the stimuli were dynamic – with the face significantly more often named than all other facial features with some exceptions when compared to the mouth. The general sequence pattern for the dynamic stimuli □ was: *face* = *mouth* > *eyes* = *eyebrows* > *invented*.

- **Mouth** was the most frequently named facial feature for all emotions in static stimuli △ ($b$ = 0.281 - $b$ = 3.272) except for *Anger* in which we found no significant differences between mouth and eyes naming. In dynamic stimuli □, the mouth was the most frequently named facial feature in *Sadness* only ($b$ = 0.503 - $b$ = 1.363), but it did not differ from face naming for *Joy, Anger,* and *Surprise*. Furthermore, face naming was more frequent than mouth naming for *Fear* ($b$ = 0.693).
- **Eyes** were the second most frequently named facial feature in *Joy*, *Sadness* and *Fear* ($b$ = 0.534 - $b$ = 2.683) for static stimuli △. In the case of *Surprise*, we found no significant differences between eyes and eyebrows naming. For dynamic stimuli □, we did not find any significant differences in eyes and eyebrows naming.
- **Eyebrows** were the third most frequently named facial feature in static stimuli △ for all emotions ($b$ = 0.496 - $b$ = 3.94), except for *Joy*, in which the face was named more often. For dynamic stimuli □, eyebrows were never named significantly more often than any other facial feature.
- **Face** naming was more frequent for static stimuli △ than eyebrows naming for the emotion *Joy* ($b$ = 0.613) and than invented features for the emotions *Sadness* ($b$ = 0.56), *Fear* ($b$ = 1.539), and *Surprise* ($b$ = 2.862). In contrast, face naming for dynamic stimuli □ became more prevalent and was more often named than eyes and eyebrows for all emotions ($b$ = 0.693 - $b$ = 1.846), and than mouth for *Fear* ($b$ = 0.693).
- **Invented Features** were the least named in both static and dynamic stimuli, with the exception of *Anger* in static stimuli △, in which no significant differences were found between face and invented features naming.

*5.1.2 Descriptions and Interpretations of Facial Features Within Emotions △ □.* We further coded whether facial features were mentioned with additional information of either **descriptive** or **interpretive** nature (see Table 4). In static stimuli △, interpretations of facial features were more prevalent than descriptions of facial features within the emotion *Joy* (Figure 5), possibly due to the coding scheme, where "smile" was coded as interpretive. Other emotions had significantly more descriptive than interpretive facial feature naming – except *Sadness*, for which we did not find significant differences.

In dynamic stimuli □, the above found differences in descriptive and interpretive facial feature naming disappeared. Descriptive facial features were more frequently mentioned over interpretive facial features for *Surprise* only (Figure 5). We found more interpretive facial features over descriptive ones for *Sadness* – which was the only emotion for static stimuli △ where descriptive facial features were not mentioned predominantly. These findings suggest that facial features become proportionally more interpretive for dynamic compared to static stimuli.

*5.1.3 Role of Facial Features Across Emotions △ □.* Overall, we found less differences in facial feature naming across emotions in dynamic □ compared to static stimuli △. We describe below the role of each facial feature across emotions.

- **Mouth** naming for *Joy*, *Sadness* and *Surprise* was more prevalent as compared to both *Anger* and *Fear* (*Joy*: $b$ = 0.41, $b$ = 0.308, *Sadness*: $b$ = 0.35, $b$ = 0.25, *Surprise*: $b$ = 0.386, $b$ = 0.284) for static stimuli △. This was also the case for dynamic stimuli □ for *Fear* only (*Joy*: $b$ = 0.817, *Sadness*: $b$ = 0.817, *Surprise*: $b$ = 0.744).
- **Eyes** naming frequency was found to be similar (no statistical differences) across all emotions for both static △ and dynamic stimuli □.
- **Eyebrows** were divided in three subgroups for static stimuli △: *Surprise* = *Anger* > *Sadness* = *Fear* > *Joy*. They were more frequently named in *Surprise* ($b$ = 0.476 - $b$ = 1.377) and *Anger* ($b$ = 0.351 - $b$ = 1.327) compared to all emotions, followed by *Fear* ($b$ = 0.976) and *Sadness* ($b$ = 0.901). For dynamic stimuli □, we found no statistical differences for the frequency of eyebrows naming across emotions.

- **Face** naming, in static stimuli △, showed one significant difference. Face was more frequently named for *Joy* compared to *Anger* ($b = 0.539$). We did not find any significant difference in face naming across emotion categories for dynamic stimuli □.
- **Invented Features** were significantly more frequently named in *Sadness* ($b = 0.916$, $b = 0.799$, $b = 2.303$) and *Anger* ($b = 1.012$, $b = 0.894$, $b = 2.4$) compared to *Joy*, *Fear*, and *Surprise* for static stimuli △.

## 5.2 Emotion Recognition △ □

This section describes results on emotion recognition rates and on the frequency of emotion naming in both studies.

*5.2.1 Emotion Recognition Rates.* Table 6 shows two confusion matrices illustrating the absolute frequencies of emotion category selections and the proportions of correct selections for each study. It shows, for example, that for static stimuli, *Joy* images (i.e., serenity, joy, and ecstasy for low, medium, and high intensities) were correctly recognized 95% of the time. In both studies, *Joy*, *Surprise*, *Sadness*, and *Anger* stimuli △ □ were recognized with high accuracy (above 70% and up to 99%). Interestingly, for *Joy*, *Fear*, and *Surprise*, the recognition accuracy was higher for static △ than for dynamic □ stimuli, while the opposite was true for *Sadness* and *Anger*. While *Fear* was recognized above average accuracy (62%) in static stimuli △, its recognition rate dropped for dynamic stimuli □, where we observed significant confusions with *Sadness* and *Disgust*. We further found that *Disgust* did not perform as well as the other emotion categories, with only 29% accuracy in static stimuli △. The confusion matrix shows that participants selected *Sadness* significantly more often than *Disgust*. As such, *Disgust* was removed from further analysis and was not used as emotion category in Study II.

**Table 6: Confusion matrices illustrating correct emotion recognition rates for static and dynamic stimuli.**

**Static Stimuli**

| | Joy | Sadness | Fear | Anger | Surprise | Disgust |
|---|---|---|---|---|---|---|
| Joy | 95 | | | 0 | 5 | |
| Sadness | | 83 | 9 | 0 | 2 | 5 |
| Fear | 1 | 13 | 62 | 9 | 5 | 11 |
| Anger | 4 | 3 | 2 | 71 | 1 | 19 |
| Surprise | 0 | 0 | 7 | | 92 | |
| Disgust | | 51 | 4 | 15 | | 29 |

**Dynamic Stimuli**

| | Joy | Sadness | Fear | Anger | Surprise | Disgust | Trust | Anticipation |
|---|---|---|---|---|---|---|---|---|
| Joy | 81 | | | | 13 | | 4 | 2 |
| Sadness | | 99 | | | | | | 1 |
| Fear | | 24 | 43 | 10 | | 22 | | |
| Anger | | 3 | 7 | 78 | | 12 | | |
| Surprise | 2 | 1 | 5 | 1 | 87 | | | 4 |

*Notes.* Values are rounded to the nearest integer with entries $< 0.5$ rounded to 0. $p$-values resulting from a binomial test were adjusted using Benjamini-Hochberg (BH) correction. Proportions correct with $p$-value $<$ .05 are highlighted in green (correct choice). Grey indicates confusion frequencies above random choice. Rows correspond to emotion category of the stimuli and columns to the best-choice emotion. **Top.** In Study I, the emotion category is an aggregate across low, medium, and high intensity labels within the same emotion category. **Bottom.** In Study II, stimuli and labels directly correspond to the applied emotion categories.

*5.2.2 Emotion Naming △ □.* Our results demonstrate that in both studies, participants used significantly more **provided** emotional words over **invented** ones (see Figure 5). This was the case for all emotions, except for *Joy* for which the drone was mostly described as "happy" (invented emotion). In the case of *Fear*, we did not find significant differences between provided and invented emotional words in static stimuli △.

## 5.3 Recognition of Intensities △

This section describes the recognition rate for emotion intensities within the correct emotion group and the validation that participants could discriminate between the different emotion intensities.

*5.3.1 Intensities Recognition Rate △.* Table 7 presents the confusion matrix of recognition rates of emotion intensities from Study I (static stimuli). We found that for most emotions, participants tended to use the medium intensity label within the correct emotion category regardless of the intensity of the stimuli. For example, participants significantly chose the label sadness (medium intensity) for images of pensiveness and grief (resp. low and high intensity). We found that for the majority of stimuli, the correct intensity label was chosen significantly above random choice (green cells on the table). As mentioned in the emotion recognition results section, intensities belonging to the *Disgust* emotion category were significantly mistaken for emotions of *Sadness*. Pensiveness, fear, and rage were the only emotion intensities (out of 18) to be significantly recognized above random choice in a wrong emotion category. To summarize, we found a tendency towards labeling images with medium intensity labels within emotion categories, and we only observed few confusions across emotion categories.

*5.3.2 Validity of Intensity Recognition △.* Task 2 of Study I was intended to validate discriminations between intensities within emotion categories. We found that each of the 18 intensity images were labeled with the correct intensity significantly more often than with a wrong intensity label. This result holds for both within and across emotion categories. Furthermore, confusions only happened between emotion categories for pensiveness, annoyance, and loathing; yet their intensity levels were still chosen correctly. For instance, the loathing (*Disgust* high intensity) image was incorrectly labeled as grief (*Sadness* high intensity) and as rage (*Anger* high intensity) in a significant amount of cases (14% and 15% respectively). Our results indicated that the recognition rates of all low and high intensities were significantly higher when participants could perceive all images at once (Task 2 compared to Task 1), except for the *Anger* emotion category. This is evidence that individuals can recognize and distinguish between emotion intensities based on facial expressions on a drone.

## 5.4 Drone's State △ □

This section reports on the free-form answers related to how participants described the drone's emotional state (Tables 4 and 5). We conclude with the described drone's capabilities, behaviors, and expectations.

*5.4.1 Internal and External △ □.* Internal states were mentioned significantly more often within all displayed emotions (Figure 5). Yet, participants attributed emotional states to external factors 23%

**Table 7: Confusion matrix illustrating emotion recognition rates for each of the 18 presented emotion labels and their intensities in Study I Task 1.**

| | | Joy | | | Sadness | | | Fear | | | Anger | | | Surprise | | | Disgust | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Emotion Category** | **Images** | **S** | **J** | **E** | **P** | **S** | **G** | **A** | **F** | **T** | **An** | **Ang** | **R** | **D** | **Su** | **Am** | **B** | **Di** | **L** | **% Correct** |
| Joy | Serenity | 40 | 56 | 2 | | | | | | | | | 1 | 1 | | | | | | 40 |
| | Joy | 9 | 78 | 12 | | | | | | | | | | | | 1 | | | | 78 |
| | Ecstasy | 1 | 53 | 33 | | | | | | | | | | | 2 | 11 | | | | 33 |
| Sadness | Pensiveness | | | | 12 | 43 | 3 | 20 | 2 | | | | | 3 | 2 | 1 | 12 | | 1 | 12 |
| | Sadness | | | | 1 | 83 | 9 | 3 | 2 | | | 1 | | | | | 1 | | | 83 |
| | Grief | | | | | 74 | 24 | | | | | | | | | | 1 | | | 24 |
| Fear | Apprehension | | 1 | | 2 | 5 | 3 | 27 | 31 | 4 | 8 | 6 | 1 | 2 | 2 | | 1 | 5 | 2 | 27 |
| | Fear | | | 1 | | 4 | 11 | 4 | 44 | 12 | 1 | 2 | | | 5 | 1 | 1 | 13 | | 44 |
| | Terror | | | | | 4 | 9 | 4 | 20 | 39 | 3 | 1 | 5 | | 3 | 1 | | 9 | 1 | 39 |
| Anger | Annoyance | 7 | 3 | 1 | 5 | | 1 | 5 | | | 43 | 12 | 1 | 2 | 1 | 1 | 7 | 2 | 8 | 43 |
| | Anger | | | | 1 | | | 1 | | | 11 | 59 | 14 | | | | | 5 | 8 | 59 |
| | Rage | | | | 1 | | 1 | 1 | | | 7 | 20 | 44 | | | | | 22 | 3 | 44 |
| Surprise | Distraction | 1 | | | | | | 1 | 4 | 1 | | | | 1 | 67 | 24 | | | | 1 |
| | Surprise | | | | 1 | | | 3 | 3 | 2 | | | | 1 | 60 | 30 | | | | 60 |
| | Amazement | | | | | | | | 4 | 4 | | | | | 44 | 48 | | | | 48 |
| Disgust | Boredom | | | | | 23 | 22 | 4 | 2 | | 7 | 10 | 2 | | | | 2 | 16 | 10 | 2 |
| | Disgust | | | | | 26 | 32 | 2 | 2 | | 5 | 2 | 3 | | | | | 21 | 7 | 21 |
| | Loathing | | | | | 10 | 41 | 1 | | 2 | 5 | 1 | 9 | | | | 1 | 21 | 8 | 8 |

*Notes. p*-values resulting from a binomial test were adjusted using Benjamini-Hochberg (BH) correction. Color-coded cells represent proportions with *p*-value <.05. Significantly correct choices of intensity and corresponding emotion category are highlighted in green (yellow if only the corresponding emotion category is correct). Confusion rates above random choice are indicated in grey. Rows correspond to the emotion label of the stimuli and columns to the best-choice emotion label chosen by participants. *Abbreviations*: S = Serenity, J = Joy, E = Ecstasy, P = Pensiveness, Sa = Sadness, G = Grief, A = Apprehension, F = Fear, T = Terror, An = Annoyance, Ang = Anger, R = Rage, D = Distraction, Su = Surprise, Am = Amazement, B = Boredom, Di = Disgust, L = Loathing.

of the time for static △ and 30% of the time for dynamic □ stimuli. The number of mentioned external states differed across emotion categories for static stimuli △ where *Fear, Anger*, and *Surprise* were interpreted by using significantly more external factors than for *Joy* and *Sadness* (*Fear*: $b = 0.956$, $b = 0.74$, *Anger*: $b = 1.03$, $b = 0.815$, *Surprise*: $b = 1.125$, $b = 0.91$). This finding did not replicate for dynamic stimuli □. Results on external factors are surprising given that the drone was the only actor in all visual stimuli. They indicate that participants interpreted the drone's emotions in a context involving the outside world, triggered and dependent upon the emotion displayed on the drone.

*5.4.2 Direction of External Drone States (Cause & Effect) △ □.* We found that there was a direction of the drone's external state either directed towards the drone (Cause) or away from it (Effect).

- **Cause and Effect Within Emotions** We found that *Fear* and *Surprise* evoked significantly more Cause over Effect naming for static stimuli △ (Figure 5); while *Anger* provoked the opposite. Participants mentioned significantly less Cause than Effect naming. We did not find statistically significant differences between Cause and Effect for the emotions *Joy* and *Sadness* in static stimuli △. Interestingly, we found the same pattern for dynamic stimuli □ for *Sadness, Fear*, and *Surprise*. However, there was significantly more Cause than Effect naming in *Joy* for dynamic stimuli □. We did not find significant differences between them in *Anger*.
- **Cause and Effect Across Emotions** We found that *Anger* evoked significantly more Effect naming compared to all other emotions for static stimuli △ ($b = 1.156$ - $b = 2.043$) (Figure 6). This finding was partly aligned with findings for dynamic stimuli □, in which Effect naming in *Anger* differed significantly from *Joy* ($b = 1.041$) and *Surprise* ($b = 1.447$), but not from *Sadness* and *Fear*. The dominant emotions for Cause naming were *Surprise* ($b = 1.253$ - $b = 1.54$) and *Fear* ($b = 0.916$ - $b = 1.204$), which differed significantly from all other emotions for static stimuli △. However, this effect almost disappeared for dynamic stimuli □, as the Cause naming in *Surprise* and *Fear* differed significantly from Cause naming only in *Sadness* (*Surprise*: $b = 0.728$, *Fear*: $b = 0.827$).

*5.4.3 External Factors Across Emotions △ □.* For both static and dynamic stimuli participants mentioned two different types of external factors: environmental or human.

- **Environment to Drone** With respect to both the direction and type of external factor, participants mentioned significantly more often that the drone was affected by the environment for *Fear* ($b = 1.119$ - $b = 1.149$) and *Surprise* ($b = 1.402$ - $b = 1.777$) compared to *Joy, Sadness, and Anger* for static stimuli △. This was partly replicated in dynamic stimuli □ for *Surprise* compared to *Sadness* ($b = 0.832$) and *Anger* ($b = 0.938$) as well as *Fear* compared to *Anger* ($b = 0.799$).

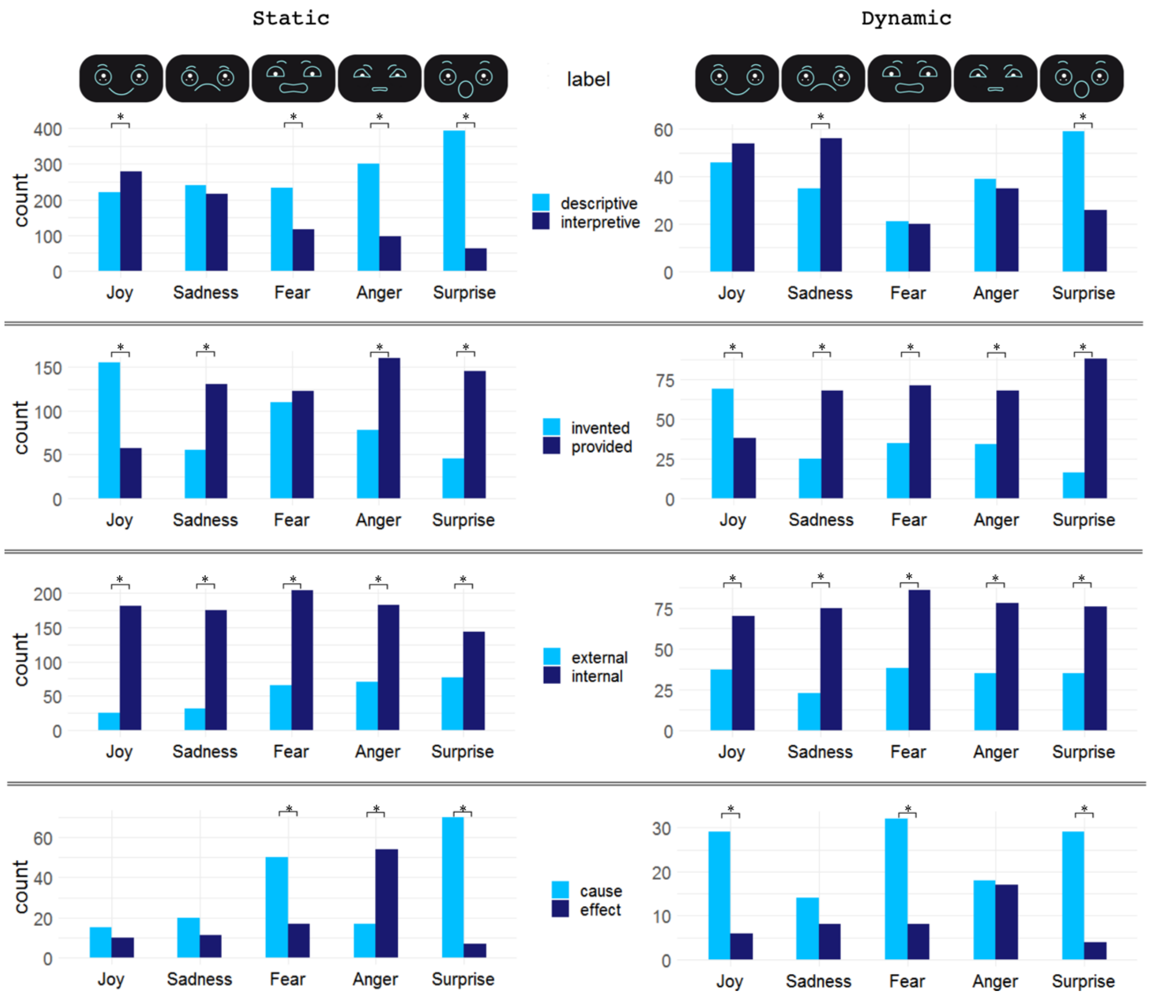



**Figure 5: Barplots illustrating the frequency of coded categories (Table 4) for each of the emotion categories: *Joy, Sadness, Fear, Anger, and Surprise* for static (left) stimuli and dynamic (right) stimuli. Note that the count data for static stimuli are aggregated over the three intensities (low, medium, and high). * indicates $p$-values <.01 for within emotion comparisons, which were provided by linear mixed-effect models with Poisson link function.**

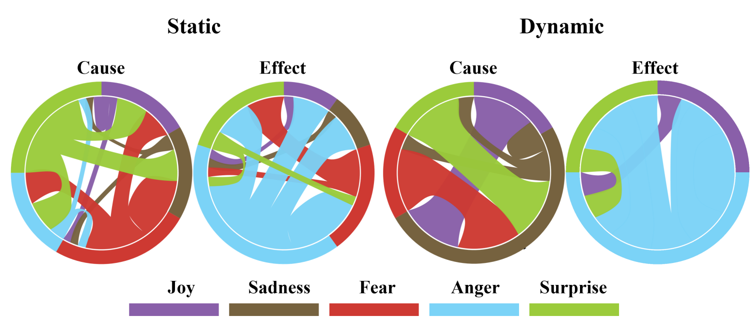



**Figure 6: Representation of comparisons across emotion for Cause and Effect (Table 4) in static (left) and dynamic (right) stimuli. Thick lines indicate that count data for Cause and Effect in the respective reference emotion category (e.g., Surprise) was significantly ($p$-values <.05) higher compared to the emotion categories in which the thick line ends (e.g., Joy, Sadness, Anger). The reverse is shown by thin lines originating from the respective reference emotion category.**

- **Human to Drone** We observed significantly more human as Cause mentioning in *Joy* compared to *Sadness* ($b = 1.321$) for dynamic stimuli □.
- **Drone to Environment** The predominant emotion in which the drone state is directed towards the environment is *Anger*, which was significantly more often mentioned compared to all other emotions for static stimuli △ ($b = 0.783$ - $b = 1.764$). Interestingly, this effect disappeared for dynamic stimuli □.
- **Drone to Human** For external states of the drone directed towards humans, for both static and dynamic stimuli △ □, participants mentioned *Anger* significantly more often compared to all other emotions (△: $b = 1.558$ - $b = 2.94$, □: $b = 0.981$ - $b = 2.08$). This is the only effect that jointly occurs in both static and dynamic stimuli.

*5.4.4 Reasoning Around the Drone State* △ □. Participants discussed several drone capabilities and behaviors that we further describe and classify below.

- **Physical Reactions** △ □ For both static and dynamic stimuli, participants mentioned on several occasions that the drone had physical reactions based on its emotional state. These included *trembling, gasping, cowering in fear, fleeing when it was scared, gritting its teeth in anger, striking out in frustration, and starting a fight to beat someone up*. Participants also acknowledged the flying capability of the drone: *"the drone is feeling fear about something it has to do, and it doesn't want to do it. Maybe doesn't want to fly high"* (S98) and even imagined that the drone could be afraid because *"the flight must have gone wrong"* (D97).
- **Perceived Flying Speed** □ Participants mentioned the flying speed of the drone in dynamic stimuli as related to its emotional state. This was a surprising finding since the displayed speed was constant for all emotions. We found, for instance, that some participants perceived the speed as being slow when the drone looked sad *"He seems to be slow and down hearted"* (D2), *"The expression was negative. It moved very slowly as well"* (D94). In contrast, the flying speed was occasionally perceived as fast, such as when the drone was angry *"Its rotors even seemed to spin faster the madder it got"* (D29), *"It seemed to moving slowly and then pick up speed before landing, as if something just made it [angry]"* (D87).
- **Expectations** □ Participants expressed various expectations around the drone's emotional state *"It seems kind of strange that a drone would be sad"* (D58), *"What would be the point? To feel bad emotions at the hand of an AI system? I think not. I am the one in control in this situation"* (D10). For some, facial expressions on drones were almost surprising and rather positive *"I think the drone's reaction was cute. I wasn't expecting that. I'm curious [about] what the drone find[s] surprising about my presence"* (D78). We also found preferences and concerns around the emotional state of the drone: *"I am not sure I would like a drone that feels fear"* (D45), *"It should always make a happy face by default"* (D13). Interestingly, some people actively recalled that it is a drone they are facing when describing their thoughts: *"He brought out the instinct in me to comfort him. Which is silly cause he is a drone, but his face was so sad"* (D89), *"I loved that it was excited to see me, but yet, I still don't trust it fully because it's a drone"* (D98).
- **Capabilities** △ □ We found a trend towards the drones being described as agents with autonomy, consciousness; as well as cognitive, affective, and behavioral abilities: *"He looks as though someone told him something he didn't want to hear"* (S89). For some, the drone can develop sympathy and antipathy for humans (*"I don't trust it at all. I'm a good person, and he just thinks I'm terrible"* (D98). The drone was described as capable of creating deep bonds with others, feeling complex emotions of love and hate (e.g., *"they may grieving a loss of a loved one"* (S32). However, participants did not uniformly ascribed the same level of capabilities to the drone, leaving some participants wondering what the drone's abilities actually are: *"I felt concerned with what abilities it would have to be able to harm me"* (D74).

Participants reported, both for static and dynamic stimuli △ □, that the drone reacted to or affected them in some way: *"It's clear happiness. In fact, just looking at its happy face made me feel happy for a moment and uplifted"* (S98), *"I feel the drone was curious about me, but then when it reached me, he was suddenly sad about my presence...almost like he/she has met me before and is unhappy with me for some reason"* (D98). In the next section, we analyze how people were emotionally affected by the different drone's emotions.

## 5.5 Affect on Participants' Emotional State □

In Study II, participants were surveyed on their emotional reaction to the drone's video stimuli. We here describe the results of the SAM questionnaire and how participants described being affected by the drone in the free-form answers (see Table 5).

*5.5.1 Emotional Assessment (SAM)* □. The results of the SAM questionnaire expand across three dimensions. Baseline scores (assessed on a 9-point scale) averaged across all participants as follows: Valence 5.55 ($SD = 1.26$), Arousal 3.67 ($SD = 1.81$), and Dominance 5.29 ($SD = 2.05$). We estimated the average difference scores between the baseline and each emotional stimulus category by using DSM (see Section 4.5.2) to measure the affect change elicited for each emotion. Figure 7 illustrates the overall differences in each SAM dimension for each emotion compared to the baseline. Results can be summarized as follows:

- **Valence** was affected according to the emotion, such that it was significantly higher when individuals were exposed to a positive emotion: *Joy*; significantly lower in case of negative emotions: *Sadness, Fear*, and *Anger*; and did not appear to change significantly as compared to the baseline with a neutral emotion, such as *Surprise*.
- **Arousal** was significantly higher for all emotions displayed on the drone.
- **Dominance** was significantly increased for *Joy* and *Surprise*; and significantly decreased for *Fear* as compared to the neutral baseline.

*5.5.2 Participants Emotional Response* □. We found that participants discussed how the drone affected them emotionally in the free-form answer (see Table 5) and report example quotes in Table 8. The emotions *Joy* and *Surprise* significantly triggered participants

**Table 8: Example quotes of how participants were positively or negatively affected after being presented with the drone's emotional state.**

| Drone's Emotion | Participants Emotional Response | |
|---|---|---|
| **Joy** | *I like him. He seems to have a very upbeat cheerful personality* | D89 |
| **Sadness** | *It's depressing and I would want to avoid it. Makes me sad looking at it* | D81 |
| | *I feel protective towards it, like I want to assist it to fix the problem* | D36 |
| | *I wanted to tell it that everything will be okay* | D98 |
| **Fear** | *I didn't like it. I would never want to use something that made an expression like that* | D73 |
| **Anger** | *I didn't like that it challenged me and seemed to threaten my status in the space* | D63 |
| **Surprise** | *The drone was very cute in it's almost childlike expression of excitement*<br>*It created a feeling of cooperation and openness with me* | D48 |

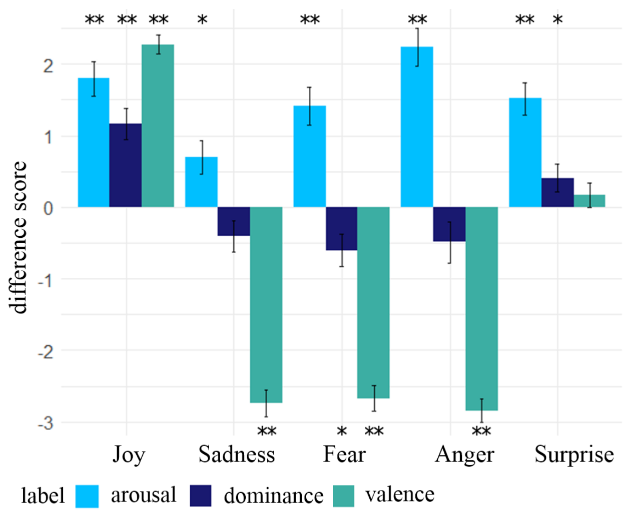



**Figure 7: Results of the SAM questionnaire showing the participants' overall emotional assessment of the drone for each emotion across: Arousal, Dominance, and Valence. The bars represent standard error of mean. Positive values indicate values greater than the corresponding baseline, and negative values indicate values smaller than the corresponding baseline. * indicates $p$-values <.01 and ** $p$-values <.001.**

to mention positive responses, while *Fear* and *Anger* significantly triggered negative emotion mentioning in the free-form answers (see Figure 8). Interestingly, *Sadness* did not appear to lead to significant differences. However, we found much discussion around empathy when participants were exposed to expressions of *Sadness*.

*5.5.3 Empathy and Prosocial Behavior* ◻. Some emotions evoked empathy, such as *Sadness*, which led to significantly higher empathy towards the drone compared to all other emotions ($b$ = 1.299 - $b$ = 3.5), gathering 64% of all empathy quotes. *Fear* and *Joy* also triggered empathy, with 17% and 12% of empathy quotes respectively. We found that empathy was linked to participants' motivation to prosocially interact with the drone. For example, when the drone displayed a sad facial expression, more than a third of participants suggested prosocial interactions (see Table 8).

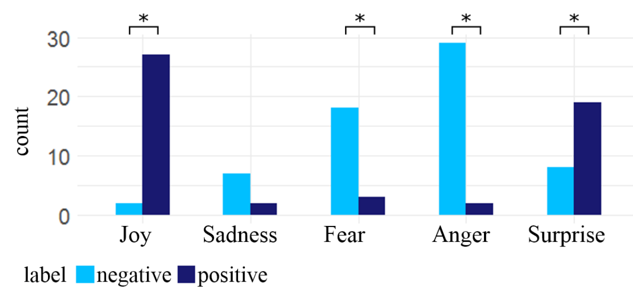



**Figure 8: Compared absolute frequencies of participants reporting on being negatively or positively affected by the drone for each emotion. * indicates $p$-values <.01.**

## 6 DISCUSSION

Our results demonstrate the potential offered by using facial expressions of emotions on drones. Here, we further discuss and highlight particularly interesting aspects of people's perception, recognition, and interpretation of emotional expressions on drones; as well as the use of facial versus bodily expression in HDI.

### 6.1 Perception and Interpretation of Emotions

This section dives into the role of facial features in the emotion recognition process and further contextualizes the ambiguity encountered in this process. We then discuss differences in participants' sensitivity to emotion intensities. Finally, we develop on how the experience of visualizing a drone with emotional features prompts narrative development and storytelling as well as empathetic responses.

***Role of Facial Features.*** Our results show that four facial features: eyes, eyebrows, pupils, and mouth, are sufficient to generate five recognizable emotions on drones. We observed that the facial features were not equally referred to when describing the choice of emotion, and our results point to the fact that some facial features are more powerful than others, with differences both within and across emotions. For example, we observed a tendency towards facial feature naming when the said facial feature was manipulated, so that the mouth, which was the most manipulated feature, was the most frequently named facial feature in both studies. Similarly, we observed that the eyebrows naming occurred significantly more often for emotions in which they were manipulated (e.g., *Anger* and *Surprise* compared to *Joy*) for static stimuli. Prior human-human communication literature demonstrated that some emotions are better recognized using the bottom half of the face (*Joy* and *Disgust*), and others using the top half of the face (*Sadness*, *Fear*, and *Anger*) [16]. Our results suggest that, similarly to human-human communication, **specific facial features may have different contributions to the recognition of specific emotions** in human-drone emotional face recognition. Further investigation is required to establish particular patterns of recognition in human-drone and human-robot communication.

***Ambiguity in Emotion Recognition.*** Our results show high to near perfect recognition rates for four emotions: *Joy* (best in static), *Sadness* (best in dynamic), *Anger*, and *Surprise*. However, *Disgust* showed poor recognition rates in static and was not further investigated in dynamic stimuli. Finally, while *Fear* could be well recognized in static stimuli, its recognition rate dropped by 19% in dynamic stimuli. We found that *Disgust* was more often associated with expressions of *Sadness*; and *Fear* was occasionally associated with *Disgust* or *Sadness* (in dynamic stimuli). We suggest two main factors that could have contributed to this ambiguity. The first one is the **perceived legitimacy** of the emotion in human-drone interaction, where the participants may not have envisioned this emotion as applicable to a drone. For example, participants made comments such as "I think it looks a little weird, seeing a drone with a scared expression", "It would have to be a fake fear as robot's do not feel emotion". Another potential contributing factor is the **design of the facial expressions**. Our choice of facial features did not include a nose into the drone's face, while it is included for *Disgust* in FACS [26] (see Table 1). Similarly, the recognition of *Disgust* as *Sadness* may result from the squeezed eye design that one participant referred to as if the drone had "been crying for a while". This suggests that future face designs should reconsider the shape of the eyes and potentially add a nose in the facial expression of *Disgust*. However, we believe that the constructed facial expression of *Fear* can be used for future designs.

***Sensitivity to Emotion Intensities.*** The accuracy of recognition of emotional intensities seems to be mediated by how it is presented to the participant (individually vs. in a series). In Study I, when stimuli were presented one by one, participants tended to identify each emotion as of a medium intensity. However, when presented in a series (here, all 18 stimuli at once), the intensities were correctly recognized more often than not. These results indicate that the presence of comparison points guides the recognition of intensity, and that **participants tend to interpret facial expression on drones with a focus on the emotion category** rather than on the intensity.

***Narratives and Storytelling.*** When trying to make sense of the drone's emotional facial expression, participants developed narratives in which they integrated external factors that we divided into two categories: people and environment. We found that the drone's state was either explained by a prior external event (e.g., "The drone received some good news") or be preceding a future action of the drone (e.g., "The drone looks angry at me and looks like it will do something to me"). We refer to this notion as the **direction of the external drone's state**, which we found was significantly affected by the displayed emotion. For example, *Anger* (in both static and dynamic stimuli) was often seen as a targeted action towards people - including towards participants. Prior human-human communication works showed that emotions can convey information about the social situation of others [34, 36]. Interestingly, this tendency of making sense of emotions of others and to infer a story behind that emotion seems to hold in HDI.

In addition, we found that participants often included themselves in their stories. For example, they wondered why their presence made the drone feel sad, or felt that the drone recognized them and wanted to greet it back. Our findings further suggest that **participants perceived the drone as an autonomous agent with a range of capabilities**. Put together, these findings show that participants perceived the drone as having its own state, which can be affected by, or affecting, its environment and people within it.

***Taking it Personal.*** The narratives revealed that participants expressed different emotional reactions to the drone's emotions; such as treating the drone's behavior as a reaction to their own actions or presence, or even experiencing empathetic emotions similar to the ones of the drone. The perception of the drone's state as an emotionally charged feedback to participants' actions suggests that some interpersonal mechanisms in human-human communication persist for HDI (e.g., seeking for harmonic reciprocal relationships with drones). These effects can become a powerful mechanism to shape human-drone interactions, where for example, **the drone's expressions of emotions can serve to trigger behavior change, mediate human-human relationships, and be used in the development of novel emotional support systems** [6, 19, 63]. For example, the drone's emotional state could be used to prompt feelings of *Joy* during difficult times or feelings of *Fear* to notify of immediate threats and danger.

## 6.2 Facial versus Bodily Expression

Prior to this work on facial expressions of emotions, the HDI community had focused on bodily expressions of drones in collocated interactions. In the following section, we discuss how our work intersects with this prior literature.

***Emotional States and Perceived Agency.*** Prior works have shown that people can identify a drone's emotional state and personality through its flight path [19, 74]. Our findings extend these works by showing that people can make sense of a drone's emotional state using facial expressions of emotion, at a higher emotional resolution. Both in this work and prior works, participants associated drone's intentions to, respectively, facial expressions and flying behavior. This opens an exciting opportunity to facilitate better emotion recognition and prompt the attribution of intentions and personality to a social drone, by **combining bodily and facial expressions of emotions**. Interestingly, the notion of participants developing narratives to make sense of the drone's state was also found in bodily expressions [74], such as participants mentioning that the drone was happy to see them. This exemplifies the opportunity to use both facial and bodily expressions to create the feeling that people are part of a story that is conveyed through the drone.

***Empathy and Synchrony.*** Our work showed that a drone displaying facial expressions of emotions can evoke empathy and synchrony (i.e., participant reporting feeling a similar emotion). Prior work showed that empathy can be fostered, not only by the facial display of emotion, but also through kinesthesia in human-human interactions [8]. Given the similarities highlighted earlier in this discussion between human-human and human-drone interaction, we ponder on the role of the drone's movement in fostering empathy. Interestingly, recent works investigated collocated HDI using bodily expressions of both people and drones acting in synchrony [28, 49]. In terms of interaction design, this means that empathy and synchrony in HDI can be achieved both with and without elements of anthropomorphism (e.g., facial features). This raises the question of how much interpretation we want to keep open when we interact with social drones, and how our interpretation of the drone is shaped through its bodily and facial expressions. Prior work suggested that a tempered anthropomorphism [72] can be an exciting element for art [28] and can enrich freedom for the interpretation of a drone's bodily expressions. **This opens interesting avenues for future research in understanding the role of anthropomorphism in HDI**.

# 7 RECOMMENDATIONS

To further synthesize and operationalize our findings, we provide a set of design recommendations divided into design and methodological recommendations.

## 7.1 Design Recommendations

Considering the synthesis of our findings from two studies on the recognition and perception of emotional facial expressions in drones, we discuss how they can be applied to the design of social drones and HDI through a set of five design recommendations.

**DR1. Use facial features to convey emotions on drones.** Conveying drone emotions is a challenging task given their traditional non-anthropomorphic form. While prior research had suggested the use of the drone's flight path and behavior to encode personality [19], we propose to add facial expressions onto a drone's body to enable its social perception by humans. We approached it by using minimal facial features. We show that 4 core facial features – namely: eyes, eyebrows, pupils, and mouth – were sufficient to generate 5 recognizable emotions. We further describe how individual facial features are interpreted by participants and show that adding basic anthropomorphic features onto a non-anthropomorphic body dramatically changes how it is being perceived.

**DR2. Design with five basic emotions.** We find that five basic emotions – namely: *Joy, Sadness, Fear, Anger, and Surprise* – are applicable to drones. Emotions can be identified with recognition rates ranging from 62%-95% using static and 43%-99% using dynamic stimuli. *Joy* is the best recognized emotion in static, and *Sadness* in dynamic stimuli. *Disgust* was poorly recognized and as such should be carefully considered in future designs. People correctly interpreted emotions using diverse and nuanced emotional vocabulary in free-form answers, including additional invented facial features, to describe these basic emotions. We found that participants could also discriminate between emotion intensities, although they tended to rate all intensities as medium. More work is still needed to fully appreciate which emotions are appropriate for drones, and in which context they might enrich and enhance human-drone communication.

**DR3. Consider reciprocity in emotional reactions of humans.** Extending evidence from affective robotics [44], our results suggest that drones' emotional expressions tend to provoke an emotional response from humans. This emotional response often mirrors the interpreted emotional state of a drone (e.g., feeling uplifted when the drone is happy, or feeling low when the drone is sad). At times, this empathetic reaction brings in an infantilizing response, such as comparing the drone to a child who needs its mother's comfort. Interestingly, socially undesirable emotional states of a drone, such as *Anger*, can lead to people feeling combative or frightened. Finally, the reciprocal response from humans opens an exciting potential for designing drones for emotional support and for embedding emotional behaviors into the HDI process.

**DR4. Shape prosocial intentions through empathy.** Our results demonstrate that empathetic response to the drone's emotional state can evoke prosocial intentions in observers. For instance, participants who expressed empathy to the sad drone would often express willingness to interact with the drone to make it feel better, comfort it, or give it a hug. Alongside communicative intentions, empathetic responses to the drone seem to also have a potential to evoke broader social behavioral intentions, such as wanting to help the drone fix a problem. These results contribute to a number of direction for applications of emotional displays on drones, e.g., design of drones as pets [19], or for behavior change, such as a physical exercise companion [58].

**DR5. Design to fit the narrative.** The qualitative analysis of the responses revealed the participants' tendency to develop narratives to make sense of the drone's emotional state. In particular, participants would often create scenarios, which would explain why the drone was in the respective emotional state. These narratives appear both when a drone is perceived as a social entity (e.g., "like it just witnessed a tragedy or received horrible news") and as a mechanical entity (e.g., the drone's state is interpreted as a request to change its battery). Additionally, we found that people would often include themselves into the developed narrative, such as interpreting the drone's emotions as a reaction to their actions. This trend is particularly prevalent when the drone is presented dynamically. These results inform the directions for narrative-based design for HDI and open new research directions to identify the factors which affect and direct such sense-making narrative.

## 7.2 Methodological Recommendations

The comparative analysis of data from the static and dynamic stimuli conditions suggest a number of differences in the participants' perception of emotional drones in respective conditions. This, for instance, includes the variations in recognition rates for different emotions, the richness of interpretations of facial features, and the level of participants' inclusion into the sense-making narrative. Correspondingly, we suggest that **the comparison** of the participants' reactions to **static and dynamic stimuli allows to elicit richer and more contextualized data** in investigating the perceived emotional expressions in drones. Furthermore, in dynamic stimuli condition, participants tend to base their reasoning on the overall drone face more often, compared to the static stimuli. Thus, we suggest that **studies focusing on particular facial features might benefit from the use of static stimuli, whereas studies focusing on the holistic perception of drone emotional states might prefer using dynamic stimuli**.

## 8 LIMITATIONS AND FUTURE WORK

This work investigated the emotional perception of facial expression on drones using static and dynamic stimuli. These methodologies are well-established in the literature [11, 40, 81], highly scalable, reproducible, and safe [82]. Yet, they present limitations as there will be additional factors affecting people's perception when exposed to a real drone (e.g., noise and wind generated). As prior work has shown that a drone's movements and behaviors can be used to convey emotions [19, 74], future work is needed to investigate the perception of facial expressions of drones with different flight paths and behaviors, and to fully understand whether some modalities are more persuasive in communicating emotions than others. Finally, future work should further investigate the use of drone facial expressions in context, to fully assess the appropriateness of emotional expressions in different scenarios of use (e.g., delivery vs. law enforcement vs. sports).

## 9 CONCLUSION

This work presented the first systematic exploration of emotional perception using facial expression on drones. We designed a set of rendered robotic faces using a minimal number of facial features to represent basic emotions. In two user studies, we showed that people can recognize five basic emotions: *Joy, Sadness, Fear, Anger, and Surprise* on drones, as well as discriminate between different emotion intensities. We found that beyond recognition, people interpret the drone's emotions and create narratives around the drone's state. In our work, participants were further affected by the drone and displayed different responses, including empathy, depending on the valence of the drone's emotion. We conclude with design and methodological recommendations for future research into social drones. This work contributes to the growing body of work on factors contributing to the acceptability of drones in human spaces.

## ACKNOWLEDGMENTS

We would like to thank Gilad Ostrin for developing the drone video stimuli, Anna Wojciechowska and Chloe Benmussa for their help with the video material, and Dr. Jeremy Frey for his helpful feedback on this work. Many thanks to the other members of the Magic Lab for their support and in particular to Carmel Shavitt.

## REFERENCES


[1] Dante Arroyo, Cesar Lucho, Silvia Julissa Roncal, and Francisco Cuellar. 2014. Daedalus: A sUAV for Human-Robot Interaction. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction* (Bielefeld, Germany) *(HRI '14)*. Association for Computing Machinery, New York, NY, USA, 116–117. https://doi.org/10.1145/2559636.2563709

[2] Mauro Avila Soto, Markus Funk, Matthias Hoppe, Robin Boldt, Katrin Wolf, and Niels Henze. 2017. DroneNavigator: Using Leashed and Free-Floating Quadcopters to Navigate Visually Impaired Travelers. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility* (Baltimore, Maryland, USA) *(ASSETS '17)*. Association for Computing Machinery, New York, NY, USA, 300–304. https://doi.org/10.1145/3132525.3132556

[3] Tadas Baltrušaitis, Laurel D. Riek, and Peter Robinson. 2010. Synthesizing Expressions Using Facial Feature Point Tracking: How Emotion is Conveyed. In *Proceedings of the 3rd International Workshop on Affective Interaction in Natural Environments* (Firenze, Italy) *(AFFINE '10)*. Association for Computing Machinery, New York, NY, USA, 27–32. https://doi.org/10.1145/1877826.1877835

[4] Christoph Bartneck, Juliane Reichenbach, and Albert van Breemen. 2004. In your face, robot! The influence of a character's embodiment on how users perceive its emotional expressions. In *Proceedings of Design and Emotion 2004 Conference* (Ankara, Turkey). 32–51.

[5] Joseph Bates. 1994. The Role of Emotion in Believable Agents. *Commun. ACM* 37, 7 (July 1994), 122–125. https://doi.org/10.1145/176789.176803

[6] Roy Baumeister, Kathleen Vohs, C DeWall, and Liqing Zhang. 2007. How Emotion Shapes Behavior: Feedback, Anticipation, and Reflection, Rather Than Direct Causation. *Personality and Social Psychology Review* 11 (06 2007), 167–203. https://doi.org/10.1177/1088868307301033

[7] Mehmet Aydin Baytas, Damla Çay, Yuchong Zhang, Mohammad Obaid, Asim Evren Yantaç, and Morten Fjeld. 2019. The Design of Social Drones: A Review of Studies on Autonomous Flyers in Inhabited Environments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300480

[8] Andrea Behrends, Sybille Müller, and Isabel Dziobek. 2012. Moving in and out of synchrony: A concept for a new intervention fostering empathy through interactional movement and dance. *The Arts in Psychotherapy* 39, 2 (2012), 107–116.

[9] Margaret M. Bradley and Peter J. Lang. 1994. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry* 25, 1 (1994), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

[10] Cynthia Breazeal. 2001. Affective Interaction between Humans and Robots. In *Proceedings of the 6th European Conference on Advances in Artificial Life* (Prague, Czech Republic) *(ECAL '01)*. Springer-Verlag, Berlin, Heidelberg, 582–591.

[11] Cynthia Breazeal. 2003. Emotion and Sociable Humanoid Robots. *International Journal of Human-Computer Studies* 59, 1–2 (July 2003), 119–155. https://doi.org/10.1016/S1071-5819(03)00018-1

[12] Cynthia Breazeal. 2003. Toward sociable robots. *Robotics and Autonomous Systems* 42, 3-4 (2003), 167–175. https://doi.org/10.1016/S0921-8890(02)00373-1

[13] Elizabeth Broadbent, Vinayak Kumar, Xingyan Li, John Sollers 3rd, Rebecca Q. Stafford, Bruce A. MacDonald, and Daniel M. Wegner. 2013. Robots with display screens: a robot with a more humanlike face display is perceived to have more mind and a better personality. *PloS one* 8, 8 (2013), e72589. https://doi.org/10.1371/journal.pone.0072589

[14] Anke M. Brock, Julia Chatain, Michelle Park, Tommy Fang, Martin Hachet, James A. Landay, and Jessica R. Cauchard. 2018. FlyMap: Interacting with Maps Projected from a Drone. In *Proceedings of the 7th ACM International Symposium on Pervasive Displays* (Munich, Germany) *(PerDis '18)*. Association for Computing Machinery, New York, NY, USA, Article 13, 9 pages. https://doi.org/10.1145/3205873.3205877

[15] A. Bruce, I. Nourbakhsh, and R. Simmons. 2002. The role of expressiveness and attention in human-robot interaction. In *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, Vol. 4. IEEE, 4138–4142. https://doi.org/10.1109/ROBOT.2002.1014396

[16] Andrew J. Calder, Andrew W. Young, Jill Keane, and Michael Dean. 2000. Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance* 26, 2 (2000), 527–551. https://doi.org/10.1037//0096-1523.26.2.527

[17] Lola Cañamero and Jakob Fredslund. 2001. I show you how I like you - can you read it in my face? [robotics]. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 31, 5 (2001), 454–459. https://doi.org/10.1109/3468.952719

[18] Jessica R. Cauchard, Jane L. E, Kevin Y. Zhai, and James A. Landay. 2015. Drone & me: An Exploration into Natural Human-Drone Interaction. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 361–365. https://doi.org/10.1145/2750858.2805823

[19] Jessica R. Cauchard, Kevin Y. Zhai, Marco Spadafora, and James A. Landay. 2016. Emotion Encoding in Human-Drone Interaction. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction* (Christchurch, New Zealand) *(HRI '16)*. IEEE, 263–270.

[20] Ashley Colley, Lasse Virtanen, Pascal Knierim, and Jonna Häkkilä. 2017. Investigating Drone Motion as Pedestrian Guidance. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia* (Stuttgart, Germany) *(MUM '17)*. Association for Computing Machinery, New York, NY, USA, 143–150. https://doi.org/10.1145/3152832.3152837

[21] Juliet M. Corbin and Anselm Strauss. 1990. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology* 13, 1 (1990), 3–21. https://doi.org/10.1007/BF00988593

[22] Frédéric Dehais, Emrah Akin Sisbot, Rachid Alami, and Mickaël Causse. 2011. Physiological and subjective evaluation of a human–robot object hand-over task. *Applied Ergonomics* 42, 6 (2011), 785–791. https://doi.org/10.1016/j.apergo.2010.12.005

[23] Brian R. Duffy. 2003. Anthropomorphism and the social robot. *Robotics and Autonomous Systems* 42, 3-4 (2003), 177–190. https://doi.org/10.1016/s0921-8890(02)00374-3

[24] Jane L. E, Ilene L. E, James A. Landay, and Jessica R. Cauchard. 2017. Drone & Wo: Cultural Influences on Human-Drone Interaction Techniques. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 6794–6799. https://doi.org/10.1145/3025453.3025755

[25] Paul Ekman and Wallace V. Friesen. 1971. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology* 17, 2 (1971), 124–129. https://doi.org/10.1037/h0030377

[26] Paul Ekman, Wallace V. Friesen, and Joseph C. Hager. 2002. *Facial Action Coding System: The Manual on CD ROM*. Salt Lake City, UT, USA.

[27] Hillary Anger Elfenbein and Nalini Ambady. 2003. When familiarity breeds accuracy: Cultural exposure and facial emotion recognition. *Journal of Personality and Social Psychology* 85, 2 (2003), 276–290. https://doi.org/10.1037/0022-3514.85.2.276

[28] Sara Eriksson, Åsa Unander-Scharin, Vincent Trichon, Carl Unander-Scharin, Hedvig Kjellström, and Kristina Höök. 2019. Dancing With Drones: Crafting Novel Artistic Expressions Through Intercorporeality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300847

[29] Friederike Eyssel, Frank Hegel, Gernot Horstmann, and Claudia Wagner. 2010. Anthropomorphic inferences from emotional nonverbal cues: A case study. In *19th International Symposium in Robot and Human Interactive Communication*. IEEE, 646–651. https://doi.org/10.1109/ROMAN.2010.5598687

[30] Julia Fink. 2012. Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction. In *International Conference on Social Robotics*. Springer, 199–208. https://doi.org/10.1007/978-3-642-34103-8_20

[31] Kerstin Fischer, Malte Jung, Lars Christian Jensen, and Maria Vanessa aus der Wieschen. 2019. Emotion Expression in HRI – When and Why. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 29–38. https://doi.org/10.1109/HRI.2019.8673078

[32] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. 2003. A survey of socially interactive robots. *Robotics and Autonomous Systems* 42, 3-4 (2003), 143–166. https://doi.org/10.1016/S0921-8890(02)00372-X

[33] Alan J. Fridlund. 1991. Evolution and facial action in reflex, social motive, and paralanguage. *Biological Psychology* 32, 1 (1991), 3–100. https://doi.org/10.1016/0301-0511(91)90003-y

[34] Nico H. Frijda and Batja Mesquita. 1994. The social roles and functions of emotions. In *Emotion and Culture: Empirical Studies of Mutual Influence*. American Psychological Association, Washington, DC, US, 51–87. https://doi.org/10.1037/10152-002

[35] Chris D. Frith and Uta Frith. 2006. How we predict what other people are going to do. *Brain Research* 1079, 1 (2006), 36–46. https://doi.org/10.1016/j.brainres.2005.12.126

[36] Shlomo Hareli and Anat Rafaeli. 2008. Emotion cycles: On the social influence of emotion in organizations. *Research in Organizational Behavior* 28 (2008), 35–59. https://doi.org/10.1016/j.riob.2008.04.007

[37] Markus Häring, Nikolaus Bee, and Elisabeth André. 2011. Creation and Evaluation of emotion expression with body movement, sound and eye color for humanoid robots. In *2011 RO-MAN*. IEEE, 204–209. https://doi.org/10.1109/ROMAN.2011.6005263

[38] Guy Hoffman and Wendy Ju. 2014. Designing Robots with Movement in Mind. *Journal of Human-Robot Interaction* 3, 1 (Feb. 2014), 91–122. https://doi.org/10.5898/JHRI.3.1.Hoffman

[39] Jihong Hwang, Taezoon Park, and Wonil Hwang. 2013. The effects of overall robot shape on the emotions invoked in users and the perceived personalities of robot. *Applied Ergonomics* 44, 3 (2013), 459–471. https://doi.org/10.1016/j.apergo.2012.10.010

[40] Alisa Kalegina, Grace Schroeder, Aidan Allchin, Keara Berlin, and Maya Cakmak. 2018. Characterizing the Design Space of Rendered Robot Faces. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (Chicago, IL, USA) *(HRI '18)*. Association for Computing Machinery, New York, NY, USA, 96–104. https://doi.org/10.1145/3171221.3171286

[41] Kari Daniel Karjalainen, Anna Elisabeth Sofia Romell, Photchara Ratsamee, Asim Evren Yantac, Morten Fjeld, and Mohammad Obaid. 2017. Social Drone Companion for the Home Environment: A User-Centric Exploration. In *Proceedings of the 5th International Conference on Human Agent Interaction* (Bielefeld, Germany) *(HAI '17)*. Association for Computing Machinery, New York, NY, USA, 89–96. https://doi.org/10.1145/3125739.3125774

[42] L.N. Kendall, Quentin Raffaelli, Alan Kingstone, and Rebecca M. Todd. 2016. Iconic faces are not real faces: enhanced emotion detection and altered neural processing as faces become more iconic. *Cognitive Research: Principles and Implications* 1, 1, Article 19 (2016), 14 pages. https://doi.org/10.1186/s41235-016-0021-8

[43] Bomyeong Kim, Hyun Young Kim, and Jinwoo Kim. 2016. Getting Home Safely with Drone. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* (Heidelberg, Germany) *(UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 117–120. https://doi.org/10.1145/2968219.2971426

[44] Eun Ho Kim, Sonya S. Kwak, and Yoon Keun Kwak. 2009. Can robotic emotional expressions induce a human to empathize with a robot?. In *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 358–362. https://doi.org/10.1109/ROMAN.2009.5326282

[45] Hyun Young Kim, Bomyeong Kim, and Jinwoo Kim. 2016. The Naughty Drone: A Qualitative Research on Drone as Companion Device. In *Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication* (Danang, Viet Nam) *(IMCOM '16)*. Association for Computing Machinery, New York, NY, USA, Article 91, 6 pages. https://doi.org/10.1145/2857546.2857639

[46] Elly A. Konijn and Henriette C. Van Vugt. 2008. Emotions in Mediated Interpersonal Communication: Toward modeling emotion in virtual humans. In *Mediated Interpersonal Communication*. Routledge, 114–144.

[47] Dana Kulic and Elizabeth A. Croft. 2007. Affective State Estimation for Human–Robot Interaction. *IEEE Transactions on Robotics* 23, 5 (2007), 991–1000. https://doi.org/10.1109/TRO.2007.904899

[48] Aleksandra Kupferberg, Stefan Glasauer, Markus Huber, Markus Rickert, Alois Knoll, and Thomas Brandt. 2011. Biological movement increases acceptance of humanoid robots as human partners in motor interaction. *AI & Society* 26, 4 (2011), 339–345. https://doi.org/10.1007/s00146-010-0314-2

[49] Joseph La Delfa, Mehmet Aydin Baytas, Rakesh Patibanda, Hazel Ngari, Rohit Ashok Khot, and Florian 'Floyd' Mueller. 2020. Drone Chi: Somaesthetic Human-Drone Interaction. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376786

[50] Iolanda Leite, André Pereira, Carlos Martinho, and Ana Paiva. 2008. Are emotional robots more fun to play with?. In *RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 77–82. https://doi.org/10.1109/ROMAN.2008.4600646

[51] Jukka M. Leppänen and Jari K. Hietanen. 2004. Positive facial expressions are recognized faster than negative facial expressions, but why? *Psychological Research* 69, 1-2 (2004), 22–29. https://doi.org/10.1007/s00426-003-0157-2

[52] Diana Löffler, Nina Schmidt, and Robert Tscharn. 2018. 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* (Chicago, IL, USA) *(HRI '18)*. Association for Computing Machinery, New York, NY, USA, 334–343. https://doi.org/10.1145/3171221.3171261

[53] Maya B Mathur and David B Reichling. 2016. Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. *Cognition* 146 (2016), 22–32. https://doi.org/10.1016/j.cognition.2015.09.008

[54] Maya B. Mathur, David B. Reichling, Francesca Lunardini, Alice Geminiani, Alberto Antonietti, Peter A. M. Ruijten, Carmel Levitan, Gideon Nave, Dylan Manfredi, Brandy Bessette-Symons, Attila Szuts, and Balazs Aczel. 2020. Uncanny but not confusing: Multisite study of perceptual category confusion in the Uncanny Valley. *Computers in Human Behavior* 103 (2020), 21–30. https://doi.org/10.1016/j.chb.2019.08.029

[55] John J. McArdle. 2009. Latent Variable Modeling of Differences and Changes with Longitudinal Data. *Annual Review of Psychology* 60 (2009), 577–605. https://doi.org/10.1146/annurev.psych.60.110707.163612

[56] Masahiro Mori, Karl F. MacDorman, and Norri Kageki. 2012. The Uncanny Valley [From the Field]. *IEEE Robotics Automation Magazine* 19, 2 (2012), 98–100. https://doi.org/10.1109/MRA.2012.2192811

[57] Florian Mueller and Matthew Muirhead. 2014. Understanding the Design of a Flying Jogging Companion. In *Proceedings of the Adjunct Publication of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST'14 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 81–82. https://doi.org/10.1145/2658779.2658786

[58] Florian Mueller and Matthew Muirhead. 2014. Understanding the Design of a Flying Jogging Companion. In *Proceedings of the Adjunct Publication of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST'14 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 81–82. https://doi.org/10.1145/2658779.2658786

[59] Bilge Mutlu, Fumitaka Yamaoka, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2009. Nonverbal Leakage in Robots: Communication of Intentions through Seemingly Unintentional Behavior. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction* (La Jolla, California, USA) *(HRI '09)*. Association for Computing Machinery, New York, NY, USA, 69–76. https://doi.org/10.1145/1514095.1514110

[60] Jean Newlove and John Dalby. 2004. *Laban for all*. Taylor & Francis US.

[61] Nikolaas N. Oosterhof and Alexander Todorov. 2008. The functional basis of face evaluation. *Proceedings of the National Academy of Sciences* 105, 32 (2008),

11087–11092. https://doi.org/10.1073/pnas.0805664105

[62] Nikolaas N. Oosterhof and Alexander Todorov. 2009. Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion* 9, 1 (2009), 128–133. https://doi.org/10.1037/a0014520

[63] Hannah R. M. Pelikan, Mathias Broth, and Leelo Keevallik. 2020. "Are You Sad, Cozmo?": How Humans Make Sense of a Home Robot's Emotion Displays. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction* (Cambridge, United Kingdom) *(HRI '20)*. Association for Computing Machinery, New York, NY, USA, 461–470. https://doi.org/10.1145/3319502.3374814

[64] Anssi Peräkylä and Johanna Elisabeth Ruusuvuori. 2012. Facial Expression and Interactional Regulation of Emotion. In *Emotion in Interaction*. Oxford University Press, Chapter 4, 64–91. https://doi.org/10.1093/acprof:oso/9780199730735.003.0004

[65] Robert Plutchik. 1980. A general psychoevolutionary theory of emotion. In *Theories of Emotion*. Elsevier, 3–33. https://doi.org/10.1016/B978-0-12-558701-3.50007-7

[66] Mauricio E. Reyes, Ivan V. Meza, and Luis A. Pineda. 2019. Robotics facial expression of anger in collaborative human–robot interaction. *International Journal of Advanced Robotic Systems* 16, 1 (2019), 13. https://doi.org/10.1177/1729881418817972

[67] Tiago Ribeiro and Ana Paiva. 2012. The Illusion of Robotic Life: Principles and Practices of Animation for Robots. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction* (Boston, Massachusetts, USA) *(HRI '12)*. Association for Computing Machinery, New York, NY, USA, 383–390. https://doi.org/10.1145/2157689.2157814

[68] Laurel D. Riek, Tal-Chen Rabinowitch, Bhismadev Chakrabarti, and Peter Robinson. 2009. How Anthropomorphism Affects Empathy toward Robots. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction* (La Jolla, California, USA) *(HRI '09)*. Association for Computing Machinery, New York, NY, USA, 245–246. https://doi.org/10.1145/1514095.1514158

[69] Annie Roy-Charland, Melanie Perron, Olivia Beaudry, and Kaylee Eady. 2014. Confusion of fear and surprise: A test of the perceptual-attentional limitation hypothesis with eye movement monitoring. *Cognition and Emotion* 28, 7 (2014), 1214–1222. https://doi.org/10.1080/02699931.2013.878687

[70] Peter A. M. Ruijten and Raymond H. Cuijpers. 2018. If Drones Could See: Investigating Evaluations of a Drone with Eyes. In *International Conference on Social Robotics*. Springer, 65–74. https://doi.org/10.1007/978-3-030-05204-1_7

[71] James A. Russell and Merry Bullock. 1985. Multidimensional scaling of emotional facial expressions: Similarity from preschoolers to adults. *Journal of Personality and Social Psychology* 48, 5 (1985), 1290–1298. https://doi.org/10.1037/0022-3514.48.5.1290

[72] Eleanor Sandry. 2015. Re-evaluating the form and communication of social robots. *International Journal of Social Robotics* 7, 3 (2015), 335–346. https://doi.org/10.1007/s12369-014-0278-3

[73] Margret Selting. 2010. Affectivity in conversational storytelling: An analysis of displays of anger or indignation in complaint stories. *Pragmatics* 20, 2 (2010), 229–277. https://doi.org/10.1075/prag.20.2.06sel

[74] Megha Sharma, Dale Hildebrandt, Gem Newman, James E. Young, and Rasit Eskicioglu. 2013. Communicating Affect via Flight Path: Exploring Use of the Laban Effort System for Designing Affective Locomotion Paths. In *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction* (Tokyo, Japan) *(HRI '13)*. IEEE, 293–300. https://doi.org/10.1109/HRI.2013.6483602

[75] Takanori Shibata, Kazuyoshi Wada, Tomoko Saito, and Kazuo Tanie. 2005. Human interactive robot for psychological enrichment and therapy. In *Proceedings of the AISB '05 Symposium on Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction* (University of Hertfordshire, Hatfield, UK), Vol. 5. The Society for the Study of Artificial Intelligence and the Simulation of Behaviour (AISB), 98–109.

[76] Stefan Sosnowski, Ansgar Bittermann, Kolja Kuhnlenz, and Martin Buss. 2006. Design and Evaluation of Emotion-Display EDDIE. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Beijing, China). IEEE, 3113–3118. https://doi.org/10.1109/IROS.2006.282330

[77] Lorna H. Stewart, Sara Ajina, Spas Getov, Bahador Bahrami, Alexander Todorov, and Geraint Rees. 2012. Unconscious evaluation of faces on social dimensions. *Journal of Experimental Psychology: General* 141, 4 (2012), 715–727. https://doi.org/10.1037/a0027950

[78] Dante Tezza and Marvin Andujar. 2019. The State-of-the-Art of Human–Drone Interaction: A Survey. *IEEE Access* 7 (2019), 167438–167454. https://doi.org/10.1109/ACCESS.2019.2953900

[79] Tim Treurniet, Lang Bai, Simon à Campo, Xintong Wang, Jun Hu, and Emilia Barakova. 2019. Drones with eyes: expressive Human-Drone Interaction. In *1st International Workshop on Human-Drone Interaction*. Glasgow, United Kingdom, 7. https://hal.archives-ouvertes.fr/hal-02128380

[80] Eva Wiese, Giorgio Metta, and Agnieszka Wykowska. 2017. Robots As Intentional Agents: Using Neuroscientific Methods to Make Robots Appear More Social. *Frontiers in Psychology* 8, Article 1663 (2017), 19 pages. https://doi.org/10.3389/fpsyg.2017.01663

[81] Anna Wojciechowska, Jeremy Frey, Esther Mandelblum, Yair Amichai-Hamburger, and Jessica R. Cauchard. 2019. Designing Drones: Factors and Characteristics Influencing the Perception of Flying Robots. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3, Article 111 (Sept. 2019), 19 pages. https://doi.org/10.1145/3351269

[82] Anna Wojciechowska, Jeremy Frey, Sarit Sass, Roy Shafir, and Jessica R. Cauchard. 2019. Collocated Human-Drone Interaction: Methodology and Approach Strategy. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 172–181. https://doi.org/10.1109/HRI.2019.8673127

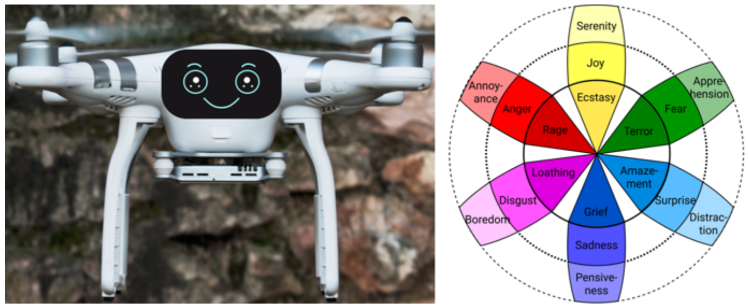
[83] Anna Wojciechowska, Foad Hamidi, Andrés Lucero, and Jessica R. Cauchard. 2020. Chasing Lions: Co-Designing Human-Drone Interaction in Sub-Saharan Africa. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS '20)*. Association for Computing Machinery, New York, NY, USA, 141–152. https://doi.org/10.1145/3357236.3395481

[84] Alexander Yeh, Photchara Ratsamee, Kiyoshi Kiyokawa, Yuki Uranishi, Tomohiro Mashita, Haruo Takemura, Morten Fjeld, and Mohammad Obaid. 2017. Exploring Proxemics for Human-Drone Interaction. In *Proceedings of the 5th International Conference on Human Agent Interaction* (Bielefeld, Germany) *(HAI '17)*. Association for Computing Machinery, New York, NY, USA, 81–88. https://doi.org/10.1145/3125739.3125773

[85] Shen Zhang, Zhiyong Wu, Helen M. Meng, and Lianhong Cai. 2007. Facial Expression Synthesis Using PAD Emotional Parameters for a Chinese Expressive Avatar. In *Proceedings of the 2nd International Conference on Affective Computing and Intelligent Interaction* (Lisbon, Portugal) *(ACII '07)*. Springer-Verlag, Berlin, Heidelberg, 24–35. https://doi.org/10.1007/978-3-540-74889-2_3

[86] Jiayin Zhao, Qi Meng, Licong An, and Yifang Wang. 2019. An event-related potential comparison of facial expression processing between cartoon and real faces. *PLoS ONE* 14, 1, Article e0198868 (2019), 13 pages. https://doi.org/10.1371/journal.pone.0198868

# A APPENDIX




1) Select the emotive word that, in your opinion, best matches the drone's facial expression (your best choice)!

| | | |
|---|---|---|
| ○ Ecstasy | ○ Amazement | ○ Loathing |
| ○ Joy | ○ Surprise | ○ Disgust |
| ○ Serenity | ○ Distraction | ○ Boredom |
| ○ Terror | ○ Grief | ○ Rage |
| ○ Fear | ○ Sadness | ○ Anger |
| ○ Apprehension | ○ Pensiveness | ○ Annoyance |

2) Explain why you chose the emotional word as your best choice

3) How certain are you about your selected emotive word (your best choice)?

| | Not certain at all | | | | | Extremely certain | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Certainty of your best choice | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

4) Rate the intensity of the facial expression of the drone!

| | Not intense at all | | | | | Extremely intense | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Intensity of facial expression | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

5) Select any other emotive word that in your opinion additionally matches the facial expression of the drone (besides your best choice). You can select none, one, or multiple options.

| | | |
|---|---|---|
| ☐ Ecstasy | ☐ Amazement | ☐ Loathing |
| ☐ Joy | ☐ Surprise | ☐ Disgust |
| ☐ Serenity | ☐ Distraction | ☐ Boredom |
| ☐ Terror | ☐ Grief | ☐ Rage |
| ☐ Fear | ☐ Sadness | ☐ Anger |
| ☐ Apprehension | ☐ Pensiveness | ☐ Annoyance |

**Figure A.1: Study I. Screenshot of the study interface used in Task I. Top: Static drone stimulus (here *Joy*) with wheel of six emotions based on Plutchik's theory of emotion [65], followed by Questions 1 to 5 and their possible answers.**

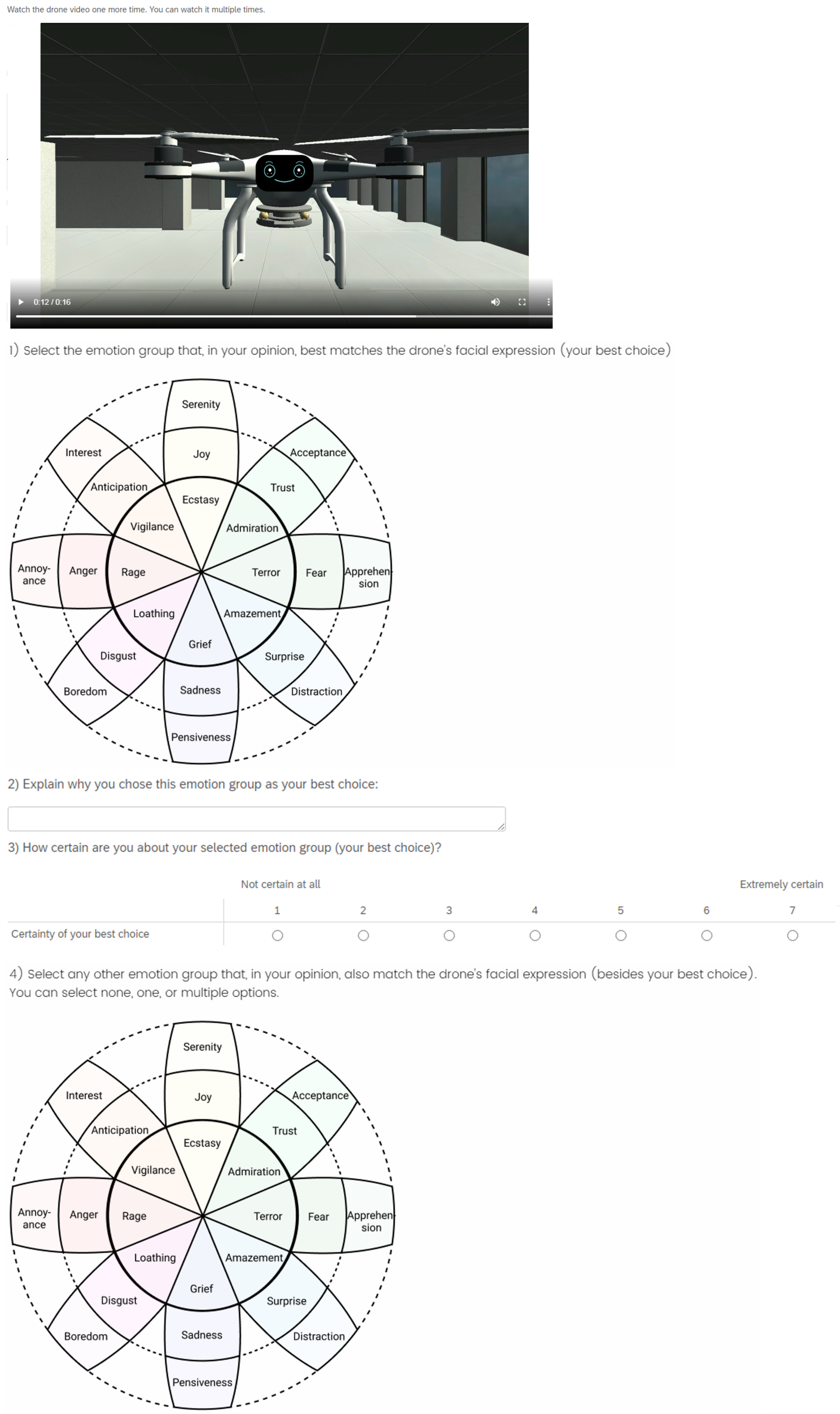



**Figure A.2: Study II. Screenshot of the study interface used in Task II. Top: Dynamic drone stimulus (here *Joy*), followed by Questions 1 to 4 and their possible answers. Note that the Plutchik's wheel of emotions [65] was clickable.**