International Journal of Social Robotics (2024) 16:579–597
https://doi.org/10.1007/s12369-023-01094-x

Emotion Appropriateness in Human–Drone Interaction

Viviane Herdel[1] · Jessica R. Cauchard[1]

Accepted: 13 December 2023 / Published online: 22 January 2024

Abstract

As robotic agents become increasingly pervasive in our lives, recent works have shown promising results on the use of emotions on social drones. However, further research is needed to determine the appropriateness of these emotions in various contexts. We here investigate the appropriateness of seven drone emotional states. In a series of six workshops (N=30) consisting of both individual and group activities, we studied emotion appropriateness for six drone metaphors across the two most popular and radically different domains for Human–Drone Interaction namely: Emergency and Entertainment. Within diverse situations of interactions, participants were able to identify the appropriateness of each emotion. Our results describe how each emotion was found both appropriate and inappropriate depending on context. We provide insights into unique opportunities generated by the perceived emotion appropriateness, depending on different roles of drone emotions in interactions with people. We conclude with design considerations for future social robotic agents, including the importance of using a broad range of emotions, the use of a neutral expression, the temporality of emotions, and novel applications to interaction design. This work contributes to the understanding of the inner workings of emotion appropriateness in drones, providing researchers with a starting point for future work on social flying robots.

Keywords Human–Drone interaction · Affective computing · Emotion appropriateness · Acceptability · Flying robot · Acceptance · UAV · Social robotics

1 Introduction

Recent works have shown promising results on the use of emotions on social drones [1–3]. Yet, little is known about emotion appropriateness when presented on drones in various contexts. *Appropriate* refers to emotions that are *"suitable, acceptable, or correct for the particular circumstances"* [4]. For example, would it be acceptable for a drone accompanying children to school to express sadness when a child is late? What about a drone showing anger when people trespass? This urges the need to explore the appropriateness of emotions across various situations for drones to provide foundations to current and future research on affective drones. This understanding is important as users would most likely reject drones that do not express emotions appropriately. This issue of predicting suitable robotic emotion is wider than drones, as highlighted in a recent survey on emotions in Human–Robot Interaction (HRI) [5]. The authors describe a "rich range of possible theoretical approaches" to emotion appropriateness but a clear need for new theories on HRI to be developed. While the wider research community acknowledges the importance of guidelines for emotional interaction with artificial agents, such guidelines have yet to be developed whether with drones or other agents [6]. We here propose an empirical approach to emotion appropriateness that can support future HRI and Human–Drone Interaction (HDI) theory development using a set of basic emotions [7] widely used in HRI research.

So far, the HDI community has mainly focused on interaction techniques [8], such as input from the human to the drone (e.g., speech [9], gestures [10–12], body postures [13]); and output from the drone to the human [14–16]. Some of these outputs include means of conveying emotional states, such as by using expressive flights [2, 3] and by using facial expressions of emotions on a screen positioned on the front of the drone [1]. While we are now able to convey a range of emotions from a drone, we do not know whether these emo-

✉ Viviane Herdel
herdel@post.bgu.ac.il

Jessica R. Cauchard
jcauchard@acm.org

[1] Magic Lab, Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Ben-Gurion Blvd. 1, 8410501 Beer Sheva, Israel

tions are appropriate to be featured on a drone. For instance, could a drone show fear or anger, if it gets bullied [17]?

Moreover, we posit that the appropriateness of emotions expressed by drones might be moderated by factors such as the drone's role (e.g., delivery or emotional support [18, 19]); the drone's metaphor[1] [20]; the severity of contexts of use [21]; and even the domain in which the drone is being used (e.g., Emergency, Sports, Assistance). To investigate emotion appropriateness in HDI, it was important to collect unique perspectives of potential future users and bystanders who may encounter such drones in their everyday life. We further wanted to ensure that our methodology would enable people to think creatively, beyond current knowledge, and ensure that we limited potential bias due to current technological limitations. As such, we opted for a focus group and design approach with a high level of abstraction, which enabled participants to not only describe their own perspectives, but also, through collaboration, reflect, discuss, and consolidate their insights.

To study a large range of drone roles and applications, with a high level of abstraction, we opted to use drone metaphors, which were recently described as a starting point for researchers to study HDI across different domains [20]. We then ran a series of six workshops (N=30), each exploring one drone metaphor, that were composed of both individual and group activities. We let participants first assess emotion appropriateness on their own, before coming together as a group to share, discuss, and refine their assessments.

This paper first presents a literature review on the use of emotions in both HRI and HDI. We then describe our empirical approach to study emotion appropriateness for six drone metaphors across two distinct domains: Emergency and Entertainment. After this, we present our study results that describe that a wide range of emotions is needed in HDI and that all studied emotions can be both appropriate and inappropriate depending on context. We conclude with a discussion of our findings on the notion of emotion appropriateness, on their contexts of use, and with five design considerations providing foundations for future research.

Our work contributes the following:

- The perceived emotion appropriateness of seven drone emotional states.
- The definition of emotion appropriateness profiles for six drone metaphors.
- An understanding of the role and opportunities of emotions in HDI.

[1] While the term "metaphor" often refers to an idiom or phrase, we here refer to a single descriptor as presented in [20].

2 Related Work

While *emotions* present many definitions, we refer to them as "an episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus" [22]. The five subsystems include different emotion components such as a cognitive component (appraisal) to evaluate objects and events, and a motor expression component (facial and vocal expression) to communicate reaction and behavioral intention. We first present prior work on emotion conveyance in both Human–Robot and Human–Drone Interaction since it is an essential part of emotion appropriateness as people need to recognize a robotic agent's emotions to be able to evaluate and interpret them. We then describe literature on emotion perception and interpretation since expressed emotions inform humans about the robot and drone's behavioral intentions or reactions to its surrounding in a given situation [1]. We finally present prior works on emotion appropriateness.

2.1 Emotion Conveyance

Emotional expressions play a fundamental role in human-human communication [23] for both communicative and social functions [24]. The possibility to convey emotions in ground [25] and flying robots (a.k.a. drones) [1–3] exploits the strong ability of humans to use non-verbal communication to infer intentions and emotional states [26]. Research has further shown that emotional robots can induce people to make sense of their intentions and guide human behaviors [27, 28]. In return, the attribution of intentions to robots can foster feelings of social connection, empathy, and prosociality [29, 30]. For example, it was shown that participants spent a significantly shorter time deciding to collaborate with emotional compared with non-emotional robots [31]. The use of emotions in robotics, therefore, offers exciting opportunities to help humans understand and predict nonhuman agents' behavior [32] and to increase the effectiveness of robots' communication and interaction with a user [33].

In both HDI and HRI, robotic emotions can be expressed through a range of both verbal and non-verbal expressions. Accordingly, a wide range of options has been proposed for conveying emotions, from facial expressions, such as by rendering facial features onto a screen [1] or animating a robot's face [25] to body movements of robots [34]. Prior work also showed combinations of the two, such as the iCat robot [35], which reacts using a combination of facial features and movements, for example, looking happy when a request is understood and sad when it is unclear. Similarly, in animatronics, a dynamic-legged robot—Cassie—performs motions with emotional attributes supported by facial features (eyes) [36], while Cozmo, a fork-lift small-sized robot, features expressive behaviors by leveraging facial expressions, head motion, locomotion, and lift movement [37]. While, in this work, participants' perception seemed to be most influenced by Cozmo's facial expressions [37], both movement and facial features of robots were shown to be effective means of conveying emotions.

In drone research, emotional states have been communicated through expressive flights [2, 3], as well as through facial expressions, which have been shown to accurately convey five drone emotions [1], some emotions even causing people to show empathetic responses. Note that people's willingness to interact with drones was found to be higher on drones with facial features [38, 39]. In both HRI and HDI prior works, people were able to recognize a set of emotions conveyed by ground and aerial robots. We now describe how people perceive and interpret these emotions.

2.2 Emotion Perception and Interpretation

The ability for people to recognize emotions in robotics has been extensively researched in the community, resulting in humans being able to recognize on average 50% of a robot's emotions correctly, with large disparities between emotions, designs, and studies [5]. How emotions are perceived and interpreted when expressed by a robot is important as they can trigger people to ascribe human traits and intentionality to it [28]. Emotional interpretation can then affect the interaction experience with social robots [40].

For example, recent work showed how people involved themselves into narratives that they created to make sense of drone emotions, and that people treated the drone's emotions as a reaction to their own actions or presence [1]. Further, prior work in social robotics described particular aspects of empathy, also designated as emotional contagion, where people rated the responses of the emotional robot more appropriate to the social situation when it mirrored their own emotional state [41]. Similarly, prior work demonstrated that children reacted more positively to a Nao robot when it adaptively expressed itself by reacting to their emotions [42]. Moreover, a study revealed that participants preferred a robot with appropriate emotions in its voice over a monotone voice [43]. In addition, robots expressing emotions oriented towards people were perceived as more caring, likable, trustworthy, and submissive [44]. Finally, prior work showed that a robot's ability to express emotions is an essential characteristic for a robot to be socially acceptable [45].

The effects linked to emotional expressions can be used as a powerful mechanism to shape interactions with robots. Prior work showed that a robot displaying happiness implied that the interaction could continue, while sadness led to a reconsideration of the person's previous actions [46]. Similarly, displays of sadness or anger were found to be effective for robots to enlist help from human collaborators [31]. Prior work also described the role of emotions for trustworthiness and willingness to interact [37]. Interestingly, the kind of emotion displayed at different stages of a collaborative task had an effect on the success rate, as well as the decision time of participants [47]. It is therefore essential to understand what emotions are appropriate and in which conditions.

2.3 Emotion Appropriateness

While the notion of emotion appropriateness in robotics is ill-defined, machine ethics is a well-defined field concerned with the behavior of machines towards humans and other machines [48]. Prior research empathizes that people need to be able to rely on the ability of machines (e.g., robots, drones) to behave ethically and in a socially acceptable manner [48–50]. Consequently research investigated topics such as emotional deception from robots and the emotional attachment that people develop for them [51] with the motivation to address ethical concerns [52]. It is then argued that social robots should have the ability to decide if it is appropriate to express a particular emotion in a given context or if it is better to inhibit it [49]. The need for standardized procedures to establish requirements, boundaries, and the appropriate use of emotions in HRI/HDI is illustrated by different scenarios (e.g., therapeutic or domestic contexts) in which emotional HRI/HDI is applied and causes ethical, legal, and societal implications [6].

The notion of emotion appropriateness has been described in prior HRI/HDI research [1, 53], although it has not been investigated holistically in its own right. Drones are particularly interesting as their design is traditionally non-anthropomorphic in nature [38], and as such, any findings on such machine-like devices would certainly open the space for verifying emotion appropriateness on other types of robotic entities. Interestingly, even though the research community acknowledges the importance of guidelines for emotional interaction with artificial agents, they are yet to be established [6]. Finally, the long-lasting adoption and positive impact of robotic technologies will depend on their acceptability by the general public [54, 55]. Due to that, a possible outcome of using appropriate emotions will be their social acceptability. Our aim is therefore to provide further insights into the emotion appropriateness expressed by drones and contribute a first step towards an understanding of the role of emotional interaction with flying robots.

3 Methodology

We describe the method used to investigate the appropriateness of emotions in drones. We first present the rationale for the design of our user study and then describe our participants, scenarios, study procedure, and data analysis.

This study was approved by the Human Subjects Research Committee of the Ben-Gurion University of the Negev. All methods were carried out in accordance with the relevant guidelines and regulations. Additionally, we obtained informed consent from all participants.

3.1 Rationale

We here investigate the appropriateness of emotions displayed on drones. Our first step was to decide what emotions to investigate as several models of emotions exist and have been used in both Human–Drone and Human–Robot Interaction [7, 56, 57]. Since evidence suggests that there is a set of basic emotions that can be universally recognized by humans [23] regardless of culture, and since these basic emotions are widely used in prior work on emotional perception of robots [25, 58, 59], we opted for Ekman's basic emotion model [7]: Joy, Sadness, Fear, Anger, Surprise, and Disgust. We further included a Neutral emotion as an additional choice, so that participants could indicate that the drone is "feeling indifferent, nothing in particular, and a lack of preference one way or the other" [60]. To explore a large range of situations for drones, we leveraged the concept of drone metaphors [20], which were suggested as means to investigate HDI across domains of use. These metaphors are used to describe and represent rich *sets of applications* for drones in the HDI context. These sets of applications permit testing clusters of related applications together, enabling a holistic perspective on drone research, beyond unique applications compared one by one. The used metaphors include: *Functional*, *Reliable*, *Knowing*, *Helpful*, *Amicable*, and *Sensational* drones[2]. We then decided to explore two domains of use: *Emergency* and *Entertainment* as they correspond to the most popular domains in the field [20]. In addition, these are very distinct domains and we expect that if any differences exist for emotion appropriateness between domains, these two will present a good example. As such, Emergency focuses on a response to an event (e.g., with the potential to save human lives), while Entertainment corresponds to actions with the goal of entertaining people (e.g., street games) [20]. We then formulate the following research questions:

- **RQ1:** Is there a notion of emotion appropriateness for drone metaphors?
- **RQ2:** Is there a difference between emotion appropriateness across domains and drone metaphors?

To explore these research questions, we opted for an exploratory approach. We ran one workshop for each drone metaphor, resulting in a series of six workshops (N=30), which included both individual and group activities, as described below. We chose this approach to enable collecting unique perspectives from participants while getting them to think creatively. To limit bias due to current technological limitations, we worked with textual information using a high level of abstraction. Indeed, high-resolution props, such as videos, or actual drones, could restrict people's imagination while textual descriptions open a wider "manipulation space in the process of translation into visual images" [61]. Therefore, we designed the workshop to explicitly let participants frame their own point of view of the drone metaphor.

3.2 Participants

A total of 30 participants (5 per workshop) were recruited using emails and communications broadcasted through our university, as well as via word of mouth and snowball sampling. Our only recruitment criterion was English proficiency, and we aimed for an equal distribution of male and female participants. In total, 14 female and 16 male participants aged from 23 to 36 y.o. ($\overline{X}$=28, SD=3.3) took part in the study (see Table 1). Participants were compensated \$40 (in local currency) for 3 h of their time.

3.3 Scenarios Design

For each of the six drone metaphors, we created one *Emergency* and one *Entertainment* scenario in form of textual short stories in which drones interact with people and can express a set of emotions. The scenarios were written by the first author and edited in consultancy with a professional storyteller (art and theatre director). This resulted in a total of 12 detailed text descriptions with various drone applications. Importantly, each drone metaphor performs the same applications in both domains (e.g., the Functional drone delivers items in both the Emergency and Entertainment scenarios). Each scenario contains multiple interaction scenes between people and drones, and after each scene, we added a placeholder with an array containing the seven drone emotions (Table 2). The interaction scenes were designed such that a wide range of emotions expressed by the drone could be plausible. We suggested that participants envision the emotions as if displayed through facial features. This *reference* was chosen as prior research demonstrated that facial features can successfully convey drone emotions at higher emotional resolutions compared to other means [1, 3], and as it may be easier to imagine compared to the drone flight for example. Each scenario presented an average of $\overline{X}$=16 (SD=4.3) placeholders.

[2] Note that the *Unusual* drone was not considered as it represents a collection of unique applications, which are less formed in the prior body of work.

Table 1 Demographic Information of participants for each workshop (WS1-6)

Workshop & Metaphors		WS1 Functional	WS2 Reliable	WS3 Knowing	WS4 Helpful	WS5 Amicable	WS6 Sensational	Total all
Participants		5	5	5	5	5	5	30
Gender	Female	2	2	2	3	2	3	14
	Male	3	3	3	2	3	2	16
Age (y.o.)	Mean *(SD)*	26.2 (3)	29.2 (2.8)	29 (2.8)	32 (3.1)	26 (2.5)	26.2 (1.9)	28.1 (3.3)
	Range	23–31	26–33	24–31	28–36	23–29	23–28	23–36

Each workshop focused on a single drone metaphor

Table 2 Examples of two interaction scenes as presented to participants

Amicable drone—entertainment scenario *As the music continues the drones start doing backflips, motivating the crowd to dance* [drone's expressed emotion: Joy, Sadness, Fear, Anger, Surprise, Disgust, Neutral]
Helpful drone—emergency scenario *Soon, the drone finds a man wandering around the debris and asks for their help* [drone's expressed emotion: Joy, Sadness, Fear, Anger, Surprise, Disgust, Neutral]

Each scene is followed by a placeholder with an array of 7 emotional states for participants to assess the appropriateness of each emotion in the given scene and for the given drone metaphor

3.4 Workshop Procedure

All of the six workshops followed the same 6-stage procedure described below. They took place in our research laboratory and each workshop lasted approximately 3 h.

Stage 1: Pre-Workshop Surveys

Upon accepting to participate, participants were asked to fill in an online pre-workshop survey ahead of the workshop. After signing an electronic consent form, they were asked to fill in the following three questionnaires on their own: Values Survey Module (VSM), a 30-item questionnaire comparing culturally influenced values and sentiments [62]; Negative Attitude Towards Robots Scale (NARS), a 14-item self-report inventory measuring attitude towards robots [63] adapted to drones [64]; and a demographics questionnaire.

Stage 2: Welcome and Introduction

Participants were welcomed to the workshop room, introduced to the process and the workshop activity, and asked to sign an additional consent form for their participation in the workshop. Participants then watched an expert presentation on state-of-the-art Human–Drone Interaction research to inspire them and give them a sense of the existing opportunities around social drones. The importance of the participants' involvement with regard to the idea-creation process was emphasized before the start of the activity.

Table 3 Description of Drone Metaphors given to participants as a starting point to the mind-map group activity [20]

Metaphors	Descriptions
Functional	It is intended to be used, practical rather than attractive, working in the expected or necessary way, utilitarian, and unsentimental. It has the following applications: "Deliver", "Make Videos", "Host Projections", "Provide Communication", and "Take Pictures"
Reliable	It can be trusted as it works in the way people expect it to without failing. This drone can deal with critical situations with a focus on prevention, seeks order and clarity. It has the following applications: "Control Crowds/Traffic", "Guard", "Protect", and "Walk Home"
Knowing	It possesses, reflects, and acquires knowledge and information, is highly skilled and always one step ahead. It has the following applications: "Check", "Detect", "Identify", "Prevent", "Surveil/Monitor", and "Track"
Helpful	It is willing to help, kind, supportive, caring, sympathetic, and skilled to help people in various situations. Application groups include: "Ask for Help", "Buy", "Clean", "Inform", "Keep Pace", "Navigate", "Save/Rescue", and "Search"
Amicable	It is friendly, makes people feel comfortable, and behaves in a kind and pleasant way, without a specific goal to fulfill, but with a focus on relationships with people. It has the following applications: "Cheer", "Follow", "Give Company", "Make Conversation", "Motivate", and "Play"
Sensational	It purposely amazes people and attracts their interest with its entertaining, performing, and artistic nature. It loves to impress, is thrilled to be in the spotlight and has the following applications: "Dance", "Draw/Paint", "Perform", and "Race"

The descriptions are adapted from prior work [20]

Stage 3: Mind-mapping the Drone Metaphor

Before diving into the emotion appropriateness assessment, we wanted to ensure that participants had a good understanding of the drone metaphor they were working with. As such, they were asked to create a mind-map for their assigned drone metaphor considering: traits, look and feel, functions and features, and interaction modalities. As a group, participants are then placed around a table and supplied with creative material (paper, pens, and pencils). They are also provided with a colorful paperboard with written information highlighting the main aspects related to their assigned drone metaphor and to the activity itself. The text descriptions of the drone metaphors were adapted from prior work [20] and can be found in Table 3. Note that each workshop group is only provided with the description of the drone metaphor they are allocated to. Participants were instructed to think about their drone's traits, look and feel, functions and features, and interaction modalities. The text on the paperboard is read by the experimenter before the start of the activity. Once the activity started, the instructor remained available to answer clarifications or questions about the activity. Participants were given 25 min for this task and while working on it the experimenter was available for questions.

Stage 4: Emotion Appropriateness Assessment

The next task was for each participant to individually assess the emotion appropriateness of all seven drone emotional states in a range of situations. Before the task, participants were seated separately and given two written scenarios, one for Emergency and one for Entertainment, corresponding to their assigned drone metaphor (see Sect. 3.3). For each placeholder in the text, participants had to highlight with colors whether each of the seven emotions in the array was: appropriate (green color) or inappropriate (red color) to be expressed by a drone given the current interaction scene (Examples in Table 2). If undecided, participants had the option not to color an emotion (i.e., undecided). This task was adapted from prior research used to assess acceptability [65].

Stage 5: Focus Group

Each participant was given three colored cards: green, red, white. They were then asked to sit in a circle. The experimenter acted as moderator and read each scenario out loud. Whenever a placeholder was reached, participants were asked, for every single emotion, to disclose their answer from the individual Emotion Appropriateness Assessment (Stage 4). They did so by placing the corresponding card (green for appropriate, red for inappropriate, white for undecided) on the table in front of them. The moderator then asked that aimed toward understanding participants' choices and underlying factors that contributed to their choice. Questions included: *"What motivated you to pick appropriate/inappropriate/undecided for this emotion in the given interaction scene?", "Can you explain this in more detail?", "What do the others think about it?", "Why do you agree/disagree?"*. The aim of the focus group was to discuss, share ideas, and understand participants' reasoning about emotion appropriateness. The discussion was audio recorded.

Stage 6: Wrap Up

The experimenter compensated and thanked the participants for their participation.

3.5 Data Analysis

We first present our strategy for the data analysis of the mind-maps, then of the individual task, and finally the focus group. Both mind-maps and focus groups were analyzed using a thematic analysis that strives to identify, analyze, and report themes within data [66] that depend on the related data or on the analytical interest of the researcher.

The mind-maps were analyzed following a deductive thematic analysis approach [66] to understand how participants formed their mental model of each drone metaphor. The lead author first copied the written comments on the mind-maps to digital 'post-it' notes and sorted them into the fitting theme out of the four themes that participants were asked to consider when creating the mind-maps, namely: *traits*; *look & feel*; *functions & features*; and *interaction modalities*. If a comment on a 'post-it' note did not fit any of the four themes of interest, it was excluded from the analysis. The final sorting of the comments into the four mind-map themes was then discussed by the lead and second author and if necessary modifications were made. Together, the two authors then grouped related comments into overarching topics within each theme. For example, in the Amicable drones, three comments are identified as belonging to the theme *Functions & Features*: "should play songs"; "ability to entertain"; "show videos" and are grouped in a *Multimedia* topic within the theme. This process was repeated for each of the six drone metaphor mind-maps and is summarized in Fig. 5 showing the 4 mind-map themes of interest and topics within them that create a profile for each drone metaphor.

The analysis of the individual task helped explore both *RQ1: Is there a notion of emotion appropriateness for drones?* and *RQ2: Is there a difference between emotion appropriateness across domains and drone metaphors?* To do so, the lead author counted the number of occurrences (appropriate, inappropriate, undecided) from the colored sheets for each emotion, metaphor, and domain. Note that one participant did not fill out 4 placeholder arrays which resulted in 28 missing data points. We collected a total of 6587 data points overall. We calculated the percentages of how often emotions were appropriate, inappropriate, and undecided for all emotions: overall, per domain, per drone metaphor, and per domain and drone metaphor. We included their means, mean differences, and standard deviations [all in %].

The focus group data was analyzed with the goal of exploring *RQ1: Is there a notion of emotion appropriateness for drones?* and to identify some of the underlying mechanisms. To identify data-driven themes, we coded the content of the entire data sets using a reflexive thematic analysis [67] that takes an inductive approach, to identify codes and themes from the data. The recordings from the focus group were first automatically transcribed. The lead author then read the six workshop transcripts, familiarized themselves with the data, and took notes while reading the transcripts. After the data familiarization phase, the lead author performed the coding manually by taking notes and using highlighters to mark sections of the transcripts. The lead author then considered how different codes may combine and created an initial set of themes. The lead and the second author then used iterative discussions around the codes and the initial set of themes to finally establish three themes: (1) Emotions as a way to create emotions, (2) Emotions as a way to influence behavior, and (3) Emotions as a signal function. The themes provide information on some of the underlying mechanisms of emotion appropriateness for different roles of drone emotions in interactions with humans. The findings from the focus group (Sect. 4.3) are presented according to these themes.

4 Results

In this section, we first describe the results from the mind-map activity that helped us understand the mental models of the drone metaphors. We then continue with the results from the emotion appropriateness assessment, and continue with the findings from the focus group.

4.1 Mental Models of Drone Metaphors

To let participants form a mental model around their drone metaphor, we provided them with the metaphor's description and intended applications from prior work [20]. For each drone metaphor, we report the results from the mind-map activity around 4 topics: *Look & Feel*, *Interaction*, *Functions & Features*, and *Traits*. Look and feel details how each group imagined the appearance, material, design and shape of their drone metaphor. Interaction presents how the drone should interact with people. Functions and Features describe the capabilities envisioned for the drone and traits refer to its personal characteristics. The findings are summarized in Fig. 1.

The **Functional** drone was imagined to be smart, friendly—but not sociable—and reserved. It looks minimalistic, machine-like without anthropomorphic features (e.g., does not have a face), noticeable (e.g., bright colors, light), and round. Its look was described as pleasant and aesthetic. The main functions are its waterproofness and navigation skills ("follow the lights" functionality). It can interact visually and with sound (e.g., through lights and chimes) but its communication with people is minimal. There is no physical contact, but it could have a mechanism like a "*reaching hand to take something*".

The **Reliable** drone was envisioned to be emotionally intelligent and reliable. Its look was described as visible and recognizable, with a rigid metallic body for guarding, protecting and making it "look tough". It also has a face displaying emotions. Its main functions include being prepared (having first aid kits, water, and a siren), it can capture and send data including at night, and can alert, for example, a security center when a person is in danger. Interestingly, in some interaction instances (e.g., with injured people), it was described as being soft when touched, as opposed to its metallic body. It can emotionally connect with people, uses facial expressions to communicate and is quiet.

The **Knowing** drone was described as adaptable, intelligent, reliable, and straight-forward. Its look is not shiny and can adapt to the environment (e.g., nature imitation). It can make sense of its surrounding and decisions in it due to its context-awareness. Moreover, it protects information making it confidential. It provides facts (knowledgeable), and detects all types of information through various sensors. The interaction with people is fast and can be done through audio with voice and visual information as well as with hand gestures and eye-movements from the drone that is anthropomorphized or zoomorphized.

The **Helpful** drone is sociable, vigilant, and a listener. Their look and feel is adaptive, and shape and color-shifting so that they could look "*like cartoon characters or animals—something fun and comforting*" when they interact with children, and have ambulance colors for rescue missions. It is collaborative, can communicate with a server and with other drones. Moreover, it is well prepared and equipped, fast, able to make sense of situations, and flexible. Thus, it can handle emergencies and is described as a flying assistant for differently abled people. It communicates concisely and via voice, interprets people's facial expressions, and can adapt its voice based on whom it is interacting with.

The **Amicable** drone is humorous, intelligent, and conversational, it "*can listen and make good conversation—like an empathetic friend*". Its look is anthropomorphized "*humanoid to be like a friend*" or zoomorphized "*dog-or cat-shaped*", it has warm colors and is soft and warm to hug, "*fluffy like a pet*". It has multi-media functions, such as recommending music and activities based on mood, and showing memories from social media. It is trustworthy, such that it would not communicate what people talked with it about, and provides companionship and care (e.g., calls close friends if needed). It can give hugs and talk.

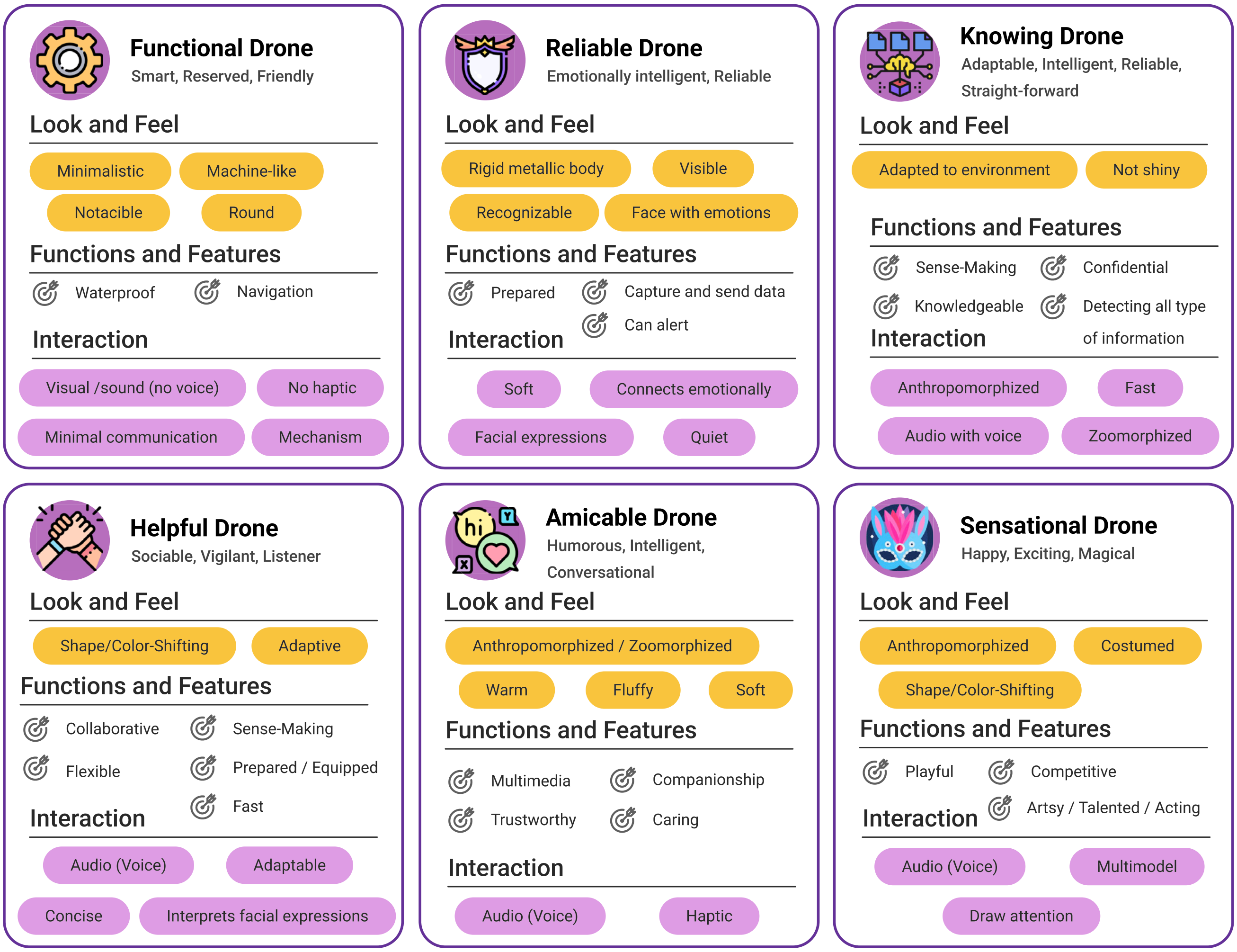

Fig. 1 Metaphor Profiles, i.e., brief description of each metaphors' "Traits", "Look and Feel", "Functions and Features", and "Interaction"

The **Sensational** drone is happy, exciting, and magical. Its look is anthropomorphized, shape/color-shifting, and costumed (e.g., changes clothes depending on the situation). The functions are being playful, competitive, artsy and talented. They include, for example, spraying confetti and water, and competing and fighting each other in shows and races. Moreover, they can be used as substitutes for animals in zoos or races where they take different roles to entertain the audience. They communicate with people via audio, voice, and light, and draw the attention of people.

4.2 Emotion Appropriateness Assessment

This section describes the results from the individual task (Stage 4) that consisted in assessing whether each emotion was appropriate or inappropriate for a range of situations. We first present the overall emotion appropriateness across metaphors and domains, then explain it across domains, and finally present emotion appropriateness profiles for each drone metaphor.

4.2.1 Overall Emotion Appropriateness

Our results indicate that all tested emotions are appropriate and that a broad set of emotions are needed in HDI (Fig. 2). We also found that all emotions have occasions in which they are appropriate ($\overline{X}$=31%, SD=13.6%), and inappropriate ($\overline{X}$=56%, SD=17.2%), and that participants were undecided in 13% (SD=4.1%) of all occurrences. The fact that they were only undecided about the emotion appropriateness in limited instances indicates that they had a good understanding of when emotions are appropriate and when they are not. The emotions most often rated as appropriate are Joy (40%), Surprise (40%), and Neutral (50%). Interestingly, Neutral was chosen as appropriate in only 50% of all situations, indicating the importance of exploring the appropriateness of all emotional states. The three emotions that

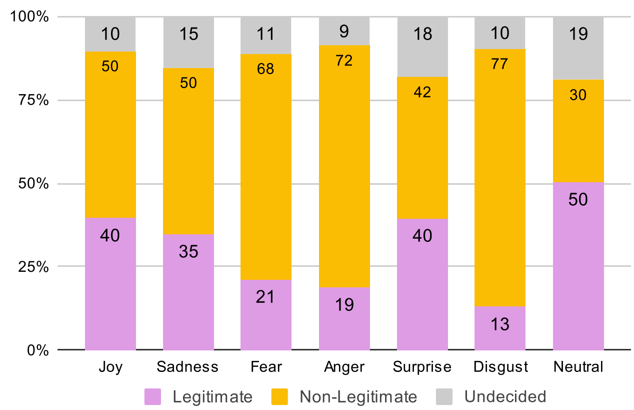

Fig. 2 Overall appropriateness of the seven emotions across drone metaphors and domains of use. Participants assessed each emotion as either appropriate (purple), inappropriate (yellow), or left it undecided (grey). (Color figure online)

were most often rated as inappropriate were Fear, Anger, and Disgust (resp. 68%, 72%, 78%). Yet, they were rated as appropriate resp. 21%, 19%, and 13% of the time. We then describe how the results are affected by domain.

4.2.2 Influence of Domains on Emotion Appropriateness

Figure 3 describes how each emotion was found both appropriate and inappropriate in each domain. Across all metaphors, the emotions most often appropriate were Neutral (51%/50%, Eme/Ent)[3], followed by Surprise (43%, Ent), Joy and Sadness (39%/39%, Eme). The most inappropriate emotions were Disgust (78%/76%), Anger (72%/73%), and Fear (66%/69%). Participants were most undecided about Neutral (21%/17%) and Surprise (21%/16%), and most decisive for Anger (90%/92%) and Disgust (90%/90%). Across domains, the mean difference in percentage for appropriate emotions is 3.4% (SD=3%), for inappropriate emotions 3.9% (SD=3.8%), and for undecided emotions 2.6% (SD=1.7%). We then find that the emotion appropriateness is almost the same across domains. To understand this phenomenon, we analyze how *appropriate* emotions differ for each metaphor across domains.

Figure 4 presents the overall percentage of the emotions chosen as appropriate. When looking into the details of how often emotions are appropriate for each metaphor, we observe small differences across domains, with mean differences: Functional: $\overline{X}$=12%, SD=5%; Reliable: $\overline{X}$=5%, SD=4%; Knowing: $\overline{X}$=16%, SD=11%; Helpful: $\overline{X}$=14%, SD=5%; Amicable: $\overline{X}$=10%, SD=6%; Sensational: $\overline{X}$=10%, SD=7%. We then only present differences that are $\geq$ 20 percentage points to highlight the stronger differences. For, instance, with the Knowing drone, we noticed the appropriateness of Sadness and Fear was 27—resp. 20—percentage points higher in the Emergency than in the Entertainment domain. *Surprise* was rated more often appropriate in Entertainment (Diff=28%). For the Sensational drone, Neutral received more appropriate ratings in Entertainment (Diff=20%). While for the Functional drone, Anger received no appropriate ratings (0%) in Entertainment, this changed in Emergency (Diff=20%). We then report the merged emotion appropriateness (across the two domains) for each metaphor in the next section.

[3] Note: In the results section, *Eme* and *Ent* respectively refer to the Emergency and Entertainment domains.

4.2.3 Influence of Drone Metaphors on Emotion Appropriateness

We found that the drone metaphors differed in the amount of how often emotions were perceived as appropriate (in mean [%]). We also identified the dispersion of the data in relation to the mean (in SD [%]) in the perceived emotion appropriateness within a metaphor. For example, the Reliable drone metaphor has the highest amount of appropriate emotions ($\overline{X}$=42%) with a standard deviation (SD=7%) showing that all emotions are perceived as appropriate in a similar way. In contrast, the Functional drone has the lowest amount of appropriate emotions ($\overline{X}$=23%) with higher standard deviation (SD=21%), showing more pronounced differences between emotions, such as Neutral that is perceived as appropriate 66% of the time and Disgust only 4%. Other drone metaphors present similarities in their appropriate emotions: Knowing ($\overline{X}$=31%, SD=17%); Helpful ($\overline{X}$=28%, SD=17%); Amicable ($\overline{X}$=28%, SD=17%); and Sensational ($\overline{X}$=35%, SD=13%). These results indicate that some drone metaphors enable more emotional expression than others and we present distinct emotion appropriateness profiles for each metaphor in Fig. 5.

We here describe differences across the profiles. Some emotions are often appropriate for some metaphors while rarely appropriate for others. Anger, for example, is the second most appropriate emotion for Reliable (48%) and one of the least for both Helpful (5%) and Amicable (9%) drones. For Functional, Fear was rarely perceived as appropriate (4%) while being appropriate in $\overline{X}$=24% (SD=5.8) of the cases for the other metaphors. Lastly, Disgust was hardly rated as appropriate for the majority of drone metaphors (4%-8%) but was often appropriate for Reliable (41%) and Sensational (15%) drones. Therefore, some metaphors, like Reliable and Sensational, have more often emotions that are appropriate compared to other metaphors that only present a subset of emotions that are appropriate > 10% of the cases (e.g., Helpful and Amicable). Although metaphors like Amicable and Sensational show some similarities (e.g., Joy, Sadness), they differ in other emotions, such as Fear, Anger, and Disgust.

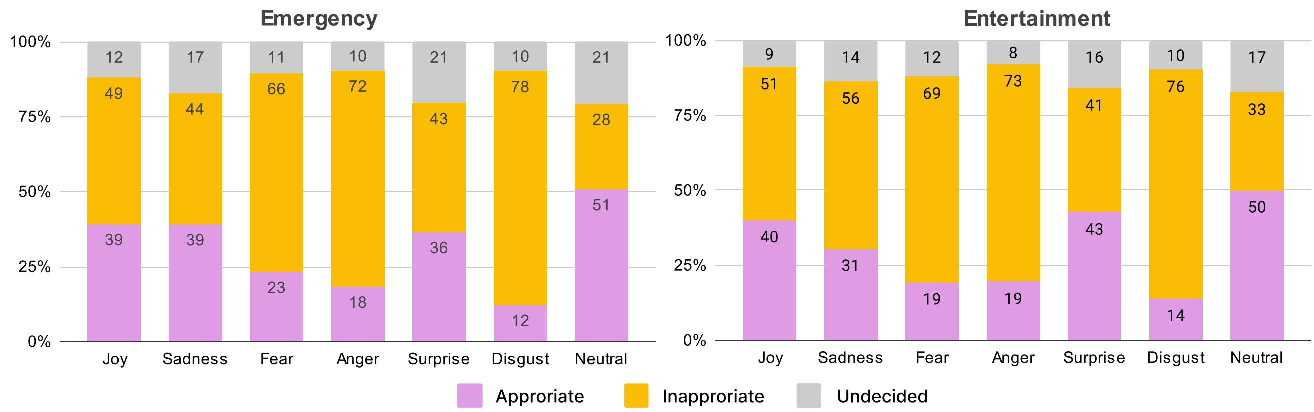

Fig. 3 Appropriateness across drone metaphors for each emotion. Each stacked bar represents how often the given emotion was selected as appropriate (purple), inappropriate (yellow), or left undecided (grey) for the Emergency and Entertainment domains.

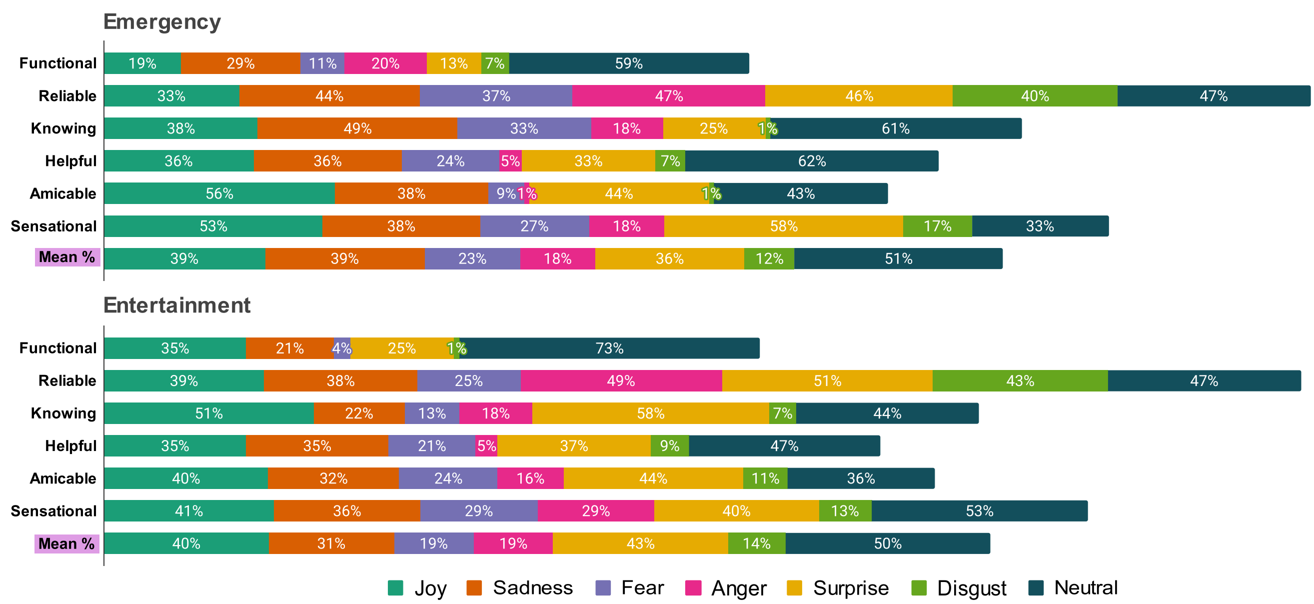

Fig. 4 Emotions selected as appropriate for each drone metaphor for the Emergency and Entertainment domains (presented as cumulative percentages). For each metaphor, the stacked bar represents the percentage of how often each emotion (one per color) was selected as *appropriate*. For instance, in the Emergency domain, for the Functional drone, Joy was selected as appropriate in 19% of all occurrences. The cumulative percentages go over 100% as participants mostly chose more than one emotion as appropriate for each placeholder. Note that *Mean %* corresponds to the percentages shown as appropriate emotions in Fig. 3 (in purple)

4.3 Focus Group

We present our resulting framework from the reflexive thematic analysis comprising three themes (Table 4) identified from the focus group (Stage 5) that aimed to explore the reasoning used by participants to justify why they perceived an emotion as appropriate or inappropriate for a given situation (Stage 4). This exploration into the *reasoning of participants* helps to contextualize their understanding of emotion appropriateness. The first theme addresses how drone emotions can induce emotions in people and affect them either *positively* or *negatively*. We found this to be often linked to perceived emotion appropriateness. The second theme describes how drone emotions can be used to influence the behavior of people and their role in perceived emotion appropriateness. We first describe identified *general strategies*, as well as how participants suggested using drone emotions to deter harmful behaviors from humans, whether *intentional* or *unintentional*. Finally, the last theme addresses the appropriate and inappropriate use of drone emotions as a signal function to show *drone obedience*, *communicate in urgent situations*, and *reassure people*.

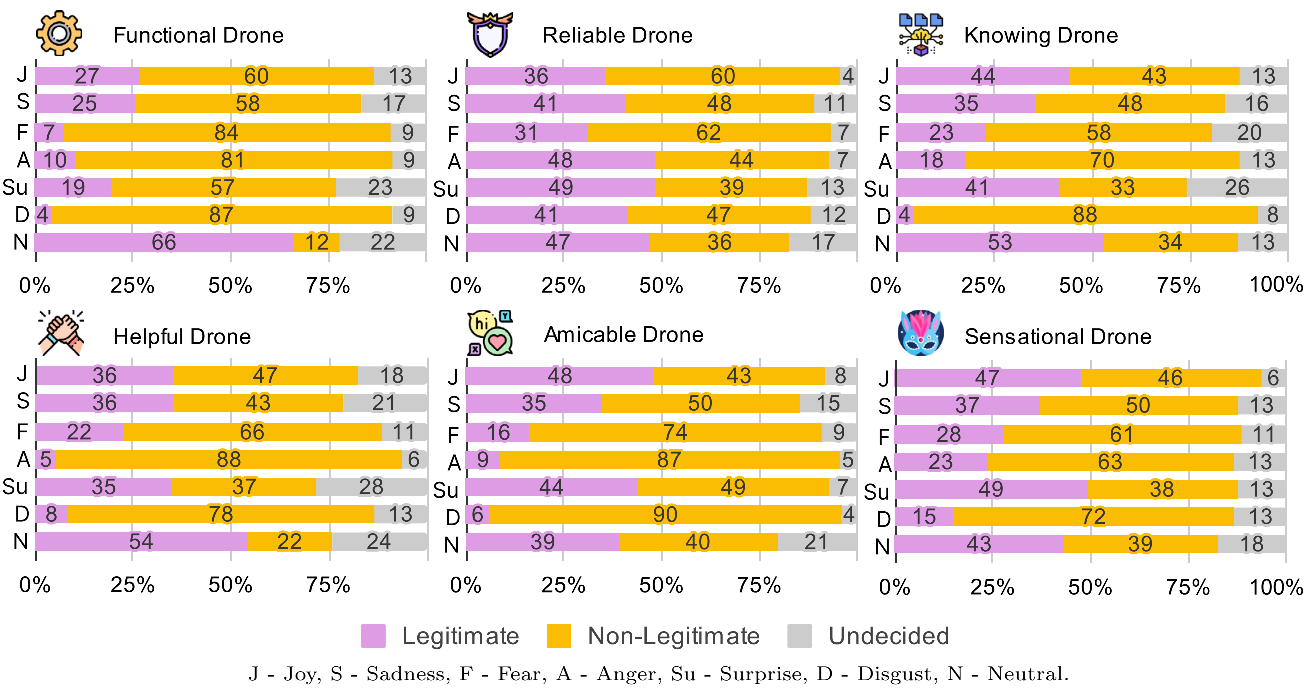

Fig. 5 Emotion appropriateness profiles describing appropriate (purple), inappropriate (yellow), and undecided (grey) emotions for each drone metaphor. The values are averaged across the Emergency and Entertainment domains. (Color figure online)

Table 4 Proposed framework capturing the reasoning used by participants to justify why they perceived an emotion as appropriate or inappropriate for a given situation

Emotions as a...		
1. Way to create emotions	2. Way to influence behavior	3. Signal function
Emotions that affect people positively	General strategies to influence behavior	Obedience and submissiveness
Emotions that affect people negatively	Deterring intentional harmful behaviors	Communicating in urgent situations
	Deterring unintentional but harmful behaviors	Reassuring people

This reasoning helps to contextualize participants' understanding of emotion appropriateness. Themes are displayed in bold with sub-themes described below them

4.3.1 Theme 1: Emotions as a Way to Create Emotions

In the following, we will describe the two identified sub-themes for theme 1, namely: *Emotions that Affect People Positively*, and *Emotions that Affect People Negatively*. We describe how drone emotions could affect participants both positively (e.g., motivating and energizing) and negatively (e.g., discouraging, unprofessional, and judgemental). In many instances, people mentioned that if the emotion was to affect them positively, it would be appropriate. The reverse was true for emotions that would affect them negatively that would be inappropriate. Below, we describe both instances.

Emotions that affect people positively Situations in which participants perceived that the emotions displayed on the drone could affect them positively included when the drone's emotion would energize, cheer people up, and even motivate them. Participants stated that the *Sensational* drone should make people happy with an expression of Joy that can "*affect your own feelings*". In Entertainment scenarios, positive energy could be conveyed from "*confident and happy*" drones, further describing more complex emotional states of the drone. This was also mentioned in other contexts, such as happily surprised expressions (i.e., a mix of Joy and Surprise) that were perceived as encouraging, and "*might even help to accept the drone*". The drone could further encourage people by staying positive in a difficult situation (e.g., expressing Joy in a first-response situation), therefore becoming a positive figure. Participants went further and stated that without the drone's positivity, the seriousness of the situation could feel even worse. In addition, Sadness was described as appropriate to signal sympathy, such as towards a person who got injured. An unexpected example of an emotion that could affect people positively was Anger which was perceived as appropriate when used as a "*motivating aggressiveness*" similar to a sports' coach. Anger was also described as appropriate to use as an element of action in Entertainment "*it is like a movie with violence and they can shout and be in beast-mode*" which adds to the energy level and excitement of the audience.

Emotions that affect people negatively We found that when participants felt that the drone emotion would affect them in a negative manner, they interpreted this emotion as inappropriate. Participants mentioned that a drone expressing *Sadness* could make people sadder, potentially taking their hope away "*when they are supposed to rescue*" (in Emergency). Sadness could also induce fear, as one might think that the drone is feeling hopeless "*already scared [...] there's no need [...] to show them a sad face*". Moreover, when people are already negatively charged, an angry or disgusted drone could aggravate the situation, potentially inducing aggression and violence in people. This could for example happen with drone acting as security guards. Moreover, participants mentioned that an inappropriate drone emotion could lead them to feel bad. For instance, if a person breaks their nose and the drone shows Joy when a "*broken nose [is] not very joyful*", then the person who expected empathy would get upset. Similarly, participants revealed not wanting the drone's emotions to appear judgmental. This was often discussed for Disgust but also occurred for Joy, Anger, Fear, and Surprise.

4.3.2 Theme 2: Emotions as a Way to Influence Behavior

In this section, we describe the three sub-themes that we identified for theme 2, namely: *General Strategies to Influence Behavior*, *Deterring Intentional Harmful Behaviors*, and *Deterring Unintentional but Harmful Behaviors*. We describe how drone emotions were perceived as appropriate and inappropriate when used to influence the behavior of people, such as when used in humorous, encouraging, or discouraging ways. In the following, we dive into the three sub-themes.

General Strategies to Influence Behavior Participants had multiple approaches of how the drone could use its emotions appropriately to motivate people to change behavior. For example, Joy could be used as an encouraging approach in which the drone functions as a positive figure to motivate people, such as children that do not want to leave a playground. Joy could be used as a kind reminder, for example, when people are at a concert without having purchased an entry ticket. Nevertheless, when the situation requires people to stop rather than start an action, Joy would be inappropriate as it could be perceived as encouraging bad behavior instead of stopping it. For severe situations, Joy would sometimes be perceived as inappropriate as it could deliver a wrong message ("*keep doing*"). Sadness could show disappointment or raise empathy as a way to appropriately motivate behavior change. Fear was considered effective and appropriate by some of our participants as it has an alerting function that can convey the importance of required action. However, participants described that it should only be used if there are negative consequences of a behavior, as Fear could unnecessarily stress people. It was perceived as useless and inappropriate in cases where the drone needs to enforce rules and be authoritarian. Anger should carefully be used according to participants as it could be perceived as pushing people and as beyond the role of a drone: "*not its job to be angry but to be entertaining*". Yet, Anger was mentioned as appropriate in mild intensity "*should be a little bit angry*" or to be used in high severity scenarios to quicken somebody's decision to act. Surprise was described as a promising emotion for behavior change. While its nature is alerting like Fear, it was not mentioned in a negative context of emotion contagion, unlike Fear and Anger. It could be used as a naive approach to appropriately encourage and convince people who are not cooperating "*surprising that they are not behaving, not cooperating... might be a way to convince*". Moreover, Surprise could be used appropriately as an initial emotional expression in unclear situations (e.g., a person does not hold a valid ticket: "*I can't believe you just sneaked in*") and provide a positive approach. Even if the situation unfolds as a false alarm, the expression was mentioned to be appropriate. *Disgust* was considered the least appropriate emotion for behavior change, yet, occasionally participants talked about its value in having people reconsider decisions. It was thought of as a last chance if nothing else helped in high severity scenarios, or appropriate to be used in a humorous way. While some participants wanted a professional and neutral expression as a first step, others thought that this expression is not helpful and inappropriate "*put a dead face will not do anything*" and might even be interpreted as a permission to continue as it "*will not reflect anything*". Moreover, they argued that people could connect more to the drone's role when it shows emotions rather than when it remains neutral.

Deterring Intentional Harmful Behaviors

Emotions were found to be appropriate when they helped the drone protect people, or goods (e.g., food and medicine). Anger and Disgust were mentioned to be appropriate to prevent people's actions through intimidation "*[Anger] might discourage the thief*". Specifically, for the Reliable Drone *Disgust* was sometimes mentioned to be a sarcastic and appropriate way to prevent people from doing something. Anger was sometimes perceived as appropriate in high-severity contexts (e.g., a person is in danger and a drone protects them) or if the potential damage of goods can cause harm to people (e.g., medical delivery). Participants compared the drone's emotional reaction to the reaction of a person. However, they also discussed that Anger would only be appropriate if it is the final "*non-verbal warning to not do this again*" after other emotions proved unsuccessful. In addition, we noticed that some emotions were described as inappropriate as they could cause the reverse behavior of what is needed. For example, expressing Joy could show that the drone is not aware of the ongoing situation and could encourage bad behaviors (e.g., harassment). Similarly, *Fear* could encourage bad behavior if the drone backs out and will not do what is necessary to protect people or goods, and was perceived as inappropriate in these cases.

Further, participants discussed strategies of how drones could emotionally react to violence against it in an appropriate manner. Drones could, for instance, show a neutral expression as if they would not care to not empower bad behavior. Participants gave the example of a child that might get bored for receiving no reaction. Others thought that it could take too long and that a response is needed "*to not teach bad behavior*". *Sadness* and *Fear* were mentioned as appropriate to prevent people to harm or bully drones by inducing empathy "*so that people understand that it is wrong*". Both emotions were thought to make people more sensitive and connected to the drone and thus would discourage bad behavior. Participants also discussed that *Fear* could be more an alternative to *Sadness*, as they tended to believe that *Sadness* would be more powerful in inducing empathy. Others expressed that "*stronger*" emotions would be appropriate (e.g., Anger) and needed since *Sadness* and *Fear* might be too "*weak and will not stop them*". This was occasionally mentioned for interactions in which the drone could take serious damage and would need to be authoritarian and set boundaries.

Deterring Unintentional but Harmful Behaviors In some situations, people might be unaware of their damaging behavior towards the drone or might damage the drone by mistake. While some participants thought that the drone should take a positive approach and can take it with humor (Joy), others thought that negative emotions are appropriate. Participants discussed the appropriateness of *Sadness* as a substitute of *Anger* to indicate that a behavior is harmful. Also, *Fear* could be used appropriately as a signal of the drone's fear of getting damaged. *Surprise* was discussed as the most human way to react to unexpected and unintentional but potentially harmful behavior towards the drone and was perceived as appropriate. It was also mentioned that *Disgust* can be appropriate, but only when used in a funny way "*disgusted from the paint, not the person*". On the other side, if directed to people, participants felt that it is too negative and thus inappropriate. They mentioned that the drone is still a machine and in service of people and should not make them feel bad. While some participants thought that these situations do not require an emotional reaction (Neutral) others said that an emotional reaction is appropriate and should be expressed to signal that the behavior can harm the drone "*please be careful*".

4.3.3 Theme 3: Emotion as a Signal Function

For theme 3, we identified three sub-themes, namely: *Obedience and Submissiveness*, *Communicating in Urgent Situations*, and *Reassuring People* in which emotions can be perceived as appropriate and inappropriate such as by reflecting the drone's confidence for a mission or unprofessionalism. We describe the sub-themes below.

Obedience and Submissiveness Participants felt that in some situations, drones would need to signal their submissiveness and obedience. They discussed that it might not be appropriate for a drone to be angry as they are there to help "*a product of technology should be oriented towards submissive behavior, being angry is dominant*"; "*I don't want a machine to be angry at me, that would be weird*". When a drone makes a mistake that affects people, participants mentioned that it would be appropriate to express Sadness as a signal of its obedience and for people to easier accept its mistake and feel forgiving. Similarly, the drone could express Fear which would show an apology note and its submissiveness. Similarly, Sadness, was mentioned to be a way to appropriately show that the drone's behavior was truly a mistake and be used as a message towards the person saying: "*don't judge me*". The expression of *Disgust* could be perceived as appropriate when it would be directed towards the drone itself "*disgusted with myself*".

Communicating in Urgent Situations In a matter of urgency that requires people to react fast the drone could signal alert with different emotions. Anger was often perceived as appropriate as it would draw attention, feels authoritarian, and could stress upon an emergency, pushing people to act. However, participants mentioned that this emotion might only be appropriate for "*very urgent manners*". Some participants felt that Anger would be more efficient than Surprise in drawing attention. Yet, Surprise could be appropriately used as a "*signal for others to come and help*" in urgent situations. Sadness was sometimes mentioned to be appropriate, even though participants were concerned that it could be unproductive and even slow down the process which might be due to its low arousal [57]. Also, Fear could be used appropriately as a signaling function that people would take seriously, however, participants mentioned the concern that even though it would not slow down the process, more direct information could be needed.

Reassuring People In situations where emotions could be reassuring and communicating that everything goes as planned, they were mentioned to be appropriate. For example, Joy was perceived as appropriate as it is a positive and optimistic emotion that communicates: "*everything goes in the best way, people*". Also, a joyful Surprise (i.e., a mix of Joy and Surprise) was mentioned to be reassuring and showing confidence for a mission. Participants also talked about the drone's professionalism and its confidence expressed through different emotions. Anger for example, would be appropriate in conflict situations to signal that the drone is intelligent and is aware of what people are doing. Moreover, a "*stony face [Neutral] can also identify that it can tackle any situation*" and that it would be ready for any escalations that might happen. It was also mentioned as a way to react calm and appropriately, and give people time to understand that they are not succeeding with an undesired action. Fear on the other side, was seen as encouraging people to do whatever they want, would not show confidence, but weakness, and was hence seen as unprofessional, and inappropriate.

5 Discussion

Based on the results from our six workshops used to investigate the appropriateness of emotions in drones, we discuss how our two research questions were answered.

5.1 Notion of Emotion Appropriateness

We found that all of the seven investigated drone emotions are possible and needed for drones. However, emotions are not appropriate at all times and we found situational differences for when emotions are appropriate ($\overline{X}$=31% of the cases) and when they are not ($\overline{X}$=56% of the cases). We, therefore, conclude that, *in response to our first research question, there is a notion of emotion appropriateness for drones*.

In the given scenarios, participants encountered different situations in which they felt that a drone should express emotions. These were, for instance, situations in which participants wished that the drone used its emotions to indicate that it is aware of an ongoing situation. This was important for participants as they wanted to be reassured and able to understand if they could trust future actions of the drone. These emotions could then be used as a signal function to alert and warn people when needed. Generally, participants felt that in social situations, it would be easier to connect to the drone that shows emotions. Furthermore, our results described many situations where a drone emotion was actually required, such as in situations that require empathy (e.g., a person encounters a difficult situation). Participants described that it would be strange, possibly annoying if the drone was to stay in a neutral emotional state. People suggested that this could be interpreted as the drone not caring about them. Similarly, our findings highlight that issues may arise when drones express inappropriate emotions in a given situation. This is critical as emotions were assessed as inappropriate on average in $\overline{X}$=56% of the cases. People mentioned contexts where they could feel misunderstood, judged, or criticized by the drone if its emotional state was inappropriate. This could affect the desirability to use drones and impair their acceptance in human spaces.

We also encountered situations in which participants thought that it would be better if the drone stayed neutral. These were, for example, critical situations in which the drone needs to be "professional" and was compared to a first responder that would just do their job even though it might be scary. This would signal that the drone can handle the situation and that it is prepared for it. Participants further compared such behavior to a doctor who would keep composure, and not show many emotions, when establishing a diagnosis. These were reflected in situations when participants wanted the drone to act as a tool (i.e., utilitarian [68]) and therefore not emotional or social. Some further mentioned that it would be "creepy" for the drone to smile all the time, indicating that emotions should be dynamic and adaptive to the situation in real-time.

We found that some emotions initially carried a negative connotation and were described as 'bad" emotions. These appeared to correspond to negatively charged emotions such as Fear, Anger, and Disgust, and were mentioned in the context of "evil killer machines". While these emotions are intrinsically human (e.g., fear of failing, being angry at a kid for not tidying their room), they may inspire distrust in machines. Nonetheless, participants still chose negative emotions as appropriate for drones in specific scenarios. For example, when the drone's role was to protect a person, it became appropriate to show Anger to a potential aggressor. People described some emotions through a range of intensities (e.g., annoyance, anger, mild aggression, rage, etc.), so that mild Anger becomes positive when it is used to motivate people, as a sports coach would. We found similar comments on Disgust, that could be expressed in a comical way ("ewww!!"). We observe a dissonance between the initial worry of using negative emotions on drones and their selection and justifications of negative emotions as appropriate in contexts.

5.2 Similarities and Differences across Domains and Drone Metaphors

We explored emotion appropriateness for the two most popular domains of HDI (Emergency & Entertainment) with the rationale that if differences exist in domains, they should be revealed across these two radically different ones. We expected to find domain differences as Emergency involves a more serious context than Entertainment and since prior research showed that people become more favorable towards drone capabilities for contexts of higher severity [21]. Moreover, the acceptance of drone technology was shown to depend on its perceived relevance to its context of use [69, 70]. So that the context of deployment matters in people's acceptance of drones, with different usage types (e.g., hobby, commercial, emergency) result in different requirements for the drone design [54]. It was further demonstrated that people's support for public safety and scientific drones were significantly higher than people's support for commercial and hobby usage [71]. Yet, our results present similar patterns in emotion appropriateness throughout—meaning that the appropriateness of all tested emotions was overall unaffected by the domain. Indeed, we found mean differences for appropriate, inappropriate, and undecided emotions across domains were 3.4%, 3.9%, and 2.6%, respectively. Our study also did not reveal that the emotion appropriateness of drones could be affected by the domain of use. While our work did not provide indications that other domains would reveal differences, additional research is needed to fully understand the role that domains may play on emotional appropriateness.

We also investigated the influence of drone metaphors, which correspond to sets of drone applications, onto emotion appropriateness. At first, we found that participants could make sense of the drone metaphors, including the drone's role, characteristics, look and feel, and specific set of applications. They were then able to imagine situations based on the defined mental model. When analyzing the emotion appropriateness depending on drone metaphors, we found that the emotion appropriateness varied across, and we proposed emotion appropriateness profiles for each metaphor. We can therefore *partially affirm our second research question, as we found evidence that emotion appropriateness differs across drone metaphors*. Our results then indicate that the emotion appropriateness might be more influenced by drone applications than by the domain in which it is being used. Specifically, we found differences in how often emotions were assessed to be appropriate across drone metaphors. This also included that some emotions were appropriate for specific metaphors but hardly appropriate for others. We then wonder whether some metaphors might have a wider range of situations in which emotions are being perceived to be appropriate. One explanation for this could be due to the different sets of applications and situations that they face. Another explanation could be due to some metaphors representing more 'social' drones, while others are more 'utilitarian'. We suggest that due to the metaphors' nature and the people's mental models that ensue, these drones would be expected to act differently. The drone metaphors provided us with the appropriate foundation to research various types of drones, as suggested in [20]. Moreover, the metaphors allowed us to evaluate the emotion appropriateness taking a bird's eye view on the matter rather than application-by-application as they encompass sets of applications.

5.3 Design Considerations

In this section, we present design considerations (DC) that were gained from this research.

DC1: Broad spectrum of emotions Our results showed that all tested emotions had situations in which they were perceived as appropriate and others in which they were perceived as inappropriate. A broad spectrum of emotions could then support a large set of use cases and enable fine-tuning interactions. A narrow selection of emotional expressions (e.g., Joy and Sadness) however, could result in people misinterpreting the intended conveyed information. In this work, although some participants ascribed a negative connotation to certain emotions (e.g., Anger), they still selected these emotions as appropriate in some circumstances. This is a critical insight as emotions like Fear, Anger, and Disgust tend to be excluded from HRI/HDI research, for reasons such as their lack of social qualities [72] or the risk they could pose [2]. Yet, our work proves that they should be studied in their own right. We then suggest designing robotics affective states with a broad spectrum of emotions.

DC2: Neutral expression as a form of communication. Participants' reasoning about whether the Neutral emotion was appropriate in a given situation revealed that they understood it as a means to convey information, beyond the drone being indifferent to a situation. For instance, participants referred to the drone as staying calm, not getting provoked, and even being in control of a situation. They described a drone expressing a Neutral emotion as showing strength and professionalism. Another interpretation was that this Neutral expression would be a sign that everything is going to plan. Interestingly, the Neutral emotion was described as inappropriate in situations where an empathetic response was expected from the drone. These findings are in line with recent work showing that a Neutral emotion conveys information and should be considered as a "part of the affective realm rather than a non-affective control condition" [60]. We suggest that future research includes a Neutral emotion, especially since it may be a main state, for example, when transitioning between emotions.

DC3: Timing, duration, and sequence between emotions In the focus group, participants mentioned the importance of: *timing*, i.e., when an emotion is expressed during an interaction; *duration*, i.e., how long the emotion is expressed for; and *sequence of emotions*, i.e., how an emotion follows upon a previous emotion. Participants described that emotions expressed in the wrong moment by the drone (*timing*) could be irritating. In terms of *duration*, participants mentioned that an emotion being expressed for too long could become "creepy" and thus inappropriate. Regarding *sequence of emotions*, participants said that if an expression of emotion does not provoke the desired effect, then it would be appropriate to change the emotion. Also, it was mentioned that some situations might first require acknowledging that something happened with an empathetic expression but then moving the situation forward with a different emotion. As such, the drone could first be acknowledging a sad situation with a sad expression but then continue with a cheering emotion like joy. We, hence, suggest considering the effects of timing, duration, and sequence between emotions as these could be additional factors contributing to the perceived appropriateness of emotions expressed by drones.

DC4: Applications over domains of use In this work, we took a holistic approach to the type of drones being studied by using drone metaphors and not focusing exclusively on a specific application, but rather a set of applications. Prior work described that applications can be domain-dependent, (e.g., "Rescuing lost people" that only occurs in Emergency settings), while other applications are domain-independent (e.g., "Taking pictures" which can occur in Entertainment, Communication, Emergency, and others) [20]. They further described that domain-independent applications are particularly hard to study as the domain may affect the interaction. Our results suggest that in the case of emotion appropriateness, the current drone application—which were embedded into the scenarios—was the prime reason for decision-making. However, our two studied domains led to similar results across emotions. We then suggest that the need for emotional expression across applications may be more nuanced than across domains of use.

DC5: Emotions as mean for transparency Participants expressed that emotions could be a way for the drone to signal its awareness of a situation. An emotional reaction to, for example, a change in situation was mentioned to be reassuring because it indicates that the drone is attentive to its surrounding. These findings are related to previous research proposing the Situation Awareness (SA)-based Agent Transparency (SAT) model that describes agent's transparency as a means to support peoples' situational awareness of an environment involving the agent [73]. The agent's transparency was defined as the "descriptive quality of an interface pertaining to its abilities to afford an operator's comprehension about an intelligent agent's intent, performance, future plans, and reasoning process" [73]. Prior research showed that operator performance, trust, and perceived usability increased as a function of transparency level [74]. We, therefore, suggest that emotions could be used as part of an interface to increase the transparency level of robotic agents which may then lead to increased trust.

6 Limitations and Future Work

While our chosen approach led to important insights, some limitations need to be considered. For instance, while our small participant sample size with 5 participants per metaphor (total $N=30$) did not allow for comparative statistics, our results provide a starting point for larger emotion appropriateness research in robotics. We suggest future work could compare emotions at a larger scale in a more quantitative manner.

In addition, in our workshops, participants were not exposed to an actual physical drone. While this might affect the ecological validity of our results, we made this choice to avoid bias from the drone's appearance onto the perception of its emotion appropriateness. Future work could investigate whether and how the appearance of a drone might play a role in its emotion appropriateness and whether they might differ from our findings. For instance, would a fluffy anthropomorphized drone have wider emotion appropriateness over a mechanical-looking drone? Beyond anthropomorphism, we suggest exploring factors that might influence the appropriateness of machine emotions, whether internal to the machine (e.g., design, interaction modality) or external (e.g., context).

Furthermore, this research piece focused on two domains of use: Emergency and Entertainment. It is then possible that other domains would yield differences in emotion appropriateness that we did not uncover in this work (e.g., Assistance, Law Enforcement [20]). Similarly, future research could explore the usefulness of emotions across domains where drones can be imagined more as social or antisocial robots. This could involve assessing the impact of displayed emotions on a range of factors, such as engagement and information transmission.

7 Conclusion

This work presented the first exploration of emotion appropriateness in Human–Drone Interaction. In a series of six workshops ($N=30$) consisting of both individual and group activities, we studied the appropriateness of seven emotional drone states, namely: Joy, Sadness, Fear, Anger, Surprise, Disgust, and Neutral for six drone metaphors across two domains: Emergency and Entertainment. We showed that within diverse situations of interactions, people were able to identify the appropriateness of each emotion. Further, we describe that—depending on the given context—each emotion had times where it was appropriate, and times where it was inappropriate. Our results imply that a wide range of emotions is needed in HDI. We further found differences across drone metaphors [20] and built emotion appropriateness profiles for each metaphor. Our findings reveal that emotion appropriateness seems more influenced by drone applications than by the domain in which it is being used. The results of our focus group further highlight the wide variety of roles for emotions in HDI. We contribute to the understanding of the role and opportunities of emotions in HDI, to the definition of emotion appropriateness profiles for six drone metaphors, and to the perceived emotion appropriateness of seven drone emotional states. Further, we provide design considerations for the research community. This work opens the door to further research into the appropriateness of emotions in HDI and in robotics more generally.

Acknowledgements We would like to thank Prof. Yisrael Parmet for his support and Zachary McKendrick for his consultancy and help in editing the scenarios. This research was supported in part by the Helmsley Charitable Trust through the Agricultural, Biological and Cognitive Robotics Initiative and by the Marcus Endowment Fund both at Ben-Gurion University of the Negev.

Author Contributions All authors contributed to the study conception and design. Material preparation and data collection were performed by Viviane Herdel, and analysis by Viviane Herdel and Jessica R. Cauchard. The first draft of the manuscript was written by Viviane Herdel and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Partial financial support was received from the Helmsley Charitable Trust through the Agricultural, Biological and Cognitive Robotics Initiative and by the Marcus Endowment Fund both at Ben-Gurion University of the Negev.

Data Availability The data that support the findings of this study are available upon reasonable request from the corresponding author [VH]. The data are not publicly available due to ethical restrictions.

Code Availability Not applicable.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Ethical Approval The questionnaires and methodology for this study was approved by the Human Research Ethics committee of Ben-Gurion University of the Negev.

Consent to Participate Informed consent was obtained from all individual participants included in the study.

Consent for Publication Patients signed informed consent regarding publishing their data.

References

1. Herdel V, Kuzminykh A, Hildebrandt A, et al (2021) Drone in love: Emotional perception of facial expressions on flying robots. In: Proceedings of the 2021 CHI conference on human factors in computing systems. ACM, New York, NY, USA, CHI '21, pp 1–20. https://doi.org/10.1145/3411764.3445495
2. Cauchard JR, Zhai KY, Spadafora M et al (2016) Emotion encoding in human-drone interaction. In: The 11th ACM/IEEE international conference on human robot interaction. IEEE, HRI, vol 16, pp 263–270. https://doi.org/10.1109/HRI.2016.7451761
3. Sharma M, Hildebrandt D, Newman G, et al (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. IEEE, HRI 13, pp 293–300. https://doi.org/10.1109/HRI.2013.6483602
4. Oxford Learner's Dictionaries (2022) appropriate. https://www.oxfordlearnersdictionaries.com/definition/english/appropriate_1. Last Accessed 02 Feb 2023
5. Stock-Homburg R (2021) Survey of emotions in human–robot interactions: perspectives from robotic psychology on 20 years of research. Int J Soc Robot. https://doi.org/10.1007/s12369-021-00778-6
6. Fosch Villaronga E (2019) I Love You, said the robot: boundaries of the use of emotions in human–robot interactions, Springer, Cham, pp 93–110. https://doi.org/10.1007/978-3-319-96722-6_6
7. Ekman P (1999) Basic emotions. In: Handbook of cognition and emotion, vol 98, Nos 45–60, p 16. https://doi.org/10.1002/0470013494.ch3
8. Funk M (2018) Human–drone interaction: let's get ready for flying user interfaces! Interactions 25(3):78–81. https://doi.org/10.1145/3194317
9. Suarez-Fernandez RA, Sanchez-Lopez JL, Sampedro C, et al (2016) Natural user interfaces for human-drone multi-modal interaction. In: International conference on unmanned aircraft systems (ICUAS). IEEE, pp 1013–1022. https://doi.org/10.1109/ICUAS.2016.7502665
10. Cauchard JR, E JL, Zhai KY, et al (2015) Drone & me: An exploration into natural human-drone interaction. In: Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing. ACM, New York, NY, USA, UbiComp '15, pp 361–365. https://doi.org/10.1145/2750858.2805823
11. Obaid M, Kistler F, Kasparavičiūtundefined G, et al (2016) How would you gesture navigate a drone? a user-centered approach to control a drone. In: Proceedings of the 20th international academic mindtrek conference. ACM, New York, NY, USA, AcademicMindtrek '16, pp 113–121. https://doi.org/10.1145/2994310.2994348
12. Alon O, Rabinovich S, Fyodorov C, et al (2021) Drones in firefighting: A user-centered design perspective. In: Proceedings of the 23rd international conference on mobile human–computer interaction. Association for computing machinery, New York, NY, USA, MobileHCI '21. https://doi.org/10.1145/3447526.3472030
13. Cai C, Yang S, Yan P, et al (2019) Real-time human-posture recognition for human–drone interaction using monocular vision. In: Intelligent robotics and applications. Springer, Cham, pp 203–216. https://doi.org/10.1007/978-3-030-27541-9_18
14. Cauchard JR, Tamkin A, Wang CY, et al (2019) Drone.io: a gestural and visual interface for human–drone interaction. In: 14th ACM/IEEE international conference on human–robot interaction (HRI). IEEE, pp 153–162. https://doi.org/10.1109/HRI.2019.8673011
15. Avila M, Funk M, Henze N (2015) Dronenavigator: using drones for navigating visually impaired persons. In: Proceedings of the 17th international ACM SIGACCESS conference on computers and accessibility. ACM, New York, NY, USA, ASSETS '15, pp 327–328. https://doi.org/10.1145/2700648.2811362
16. Ginosar E, Cauchard JR (2023) At first light: Expressive lights in support of drone-initiated communication. In: Proceedings of the 2023 CHI conference on human factors in computing systems. Association for computing machinery, New York, NY, USA, CHI '23. https://doi.org/10.1145/3544548.3581062
17. Salvini P, Ciaravella G, Yu W, et al (2010) How safe are service robots in urban environments? Bullying a robot. In: 19th international symposium in robot and human interactive communication. IEEE, pp 1–7. https://doi.org/10.1109/ROMAN.2010.5654677
18. Baytas MA, Çay D, Zhang Y, et al (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. Association for computing machinery, New York, NY, USA, CHI '19, pp 1–13. https://doi.org/10.1145/3290605.3300480
19. Fartook O, Oron-Gilad T, Cauchard JR (2023) Designing and prototyping drones for emotional support. In: Companion of the 2023 ACM/IEEE international conference on human–robot interaction. Association for computing machinery, New York, NY, USA, HRI '23, pp 234–237. https://doi.org/10.1145/3568294.3580079

20. Herdel V, Yamin LJ, Cauchard JR (2022) Above and beyond: A scoping review of domains and applications for human-drone interaction. In: Proceedings of the 2022 CHI conference on human factors in computing systems. ACM, New York, NY, USA, CHI '22, pp 1–22. https://doi.org/10.1145/3491102.3501881
21. Herdel V, Yamin LJ, Ginosar E et al (2021) Public drone: attitude towards drone capabilities in various contexts. ACM, New York, NY, USA. https://doi.org/10.1145/3447526.3472053
22. Scherer KR (2005) What are emotions? And how can they be measured? Soc Sci Inf 44(4):695–729. https://doi.org/10.1177/0539018405058216
23. Ekman P, Friesen WV (1971) Constants across cultures in the face and emotion. J Personal Soc Psychol 17(2):124–129. https://doi.org/10.1037/h0030377
24. Baumeister R, Vohs K, DeWall CN et al (2007) How emotion shapes behavior: feedback, anticipation, and reflection, rather than direct causation. Personal Soc Psychol Rev 11:167–203. https://doi.org/10.1177/1088868307301033
25. Breazeal C (2003) Emotion and sociable humanoid robots. Int J Hum Comput Stud 59(1–2):119–155. https://doi.org/10.1016/S1071-5819(03)00018-1
26. Frith CD, Frith U (2006) How we predict what other people are going to do. Brain Res 1079(1):36–46. https://doi.org/10.1016/j.brainres.2005.12.126
27. Reyes ME, Meza IV, Pineda LA (2019) Robotics facial expression of anger in collaborative human–robot interaction. Int J Adv Robot Syst. https://doi.org/10.1177/1729881418817972
28. Eyssel F, Hegel F, Horstmann G, et al (2010) Anthropomorphic inferences from emotional nonverbal cues: a case study. In: 19th international symposium in robot and human interactive communication. IEEE, pp 646–651. https://doi.org/10.1109/ROMAN.2010.5598687
29. Wiese E, Metta G, Wykowska A (2017) Robots as intentional agents: using neuroscientific methods to make robots appear more social. Front Psychol 8:1663. https://doi.org/10.3389/fpsyg.2017.01663
30. Kim EH, Kwak SS, Kwak YK (2009) Can robotic emotional expressions induce a human to empathize with a robot? In: The 18th IEEE international symposium on robot and human interactive communication. IEEE, pp 358–362. https://doi.org/10.1109/ROMAN.2009.5326282
31. Zhou S, Tian L (2020) Would you help a sad robot? Influence of robots' emotional expressions on human-multi-robot collaboration. In: 2020 29th IEEE international conference on robot and human interactive communication, pp 1243–1250. https://doi.org/10.1109/RO-MAN47096.2020.9223524
32. Złotowski J, Proudfoot D, Yogeeswaran K et al (2015) Anthropomorphism: opportunities and challenges in human–robot interaction. Int J Soc Robot 7(3):347–360. https://doi.org/10.1007/s12369-014-0267-6
33. Leite I, Pereira A, Martinho C, et al (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, pp 77–82. https://doi.org/10.1109/ROMAN.2008.4600646
34. Hoffman G, Ju W (2014) Designing robots with movement in mind. J Hum Robot Interact 3(1):91–122. https://doi.org/10.5898/JHRI.3.1.Hoffman
35. van Breemen A (2004) Animation engine for believable interactive user-interface robots. In: 2004 IEEE/RSJ international conference on intelligent robots and systems (IROS), vol 3, pp 2873–2878. https://doi.org/10.1109/IROS.2004.1389845
36. Li Z, Cummings C, Sreenath K (2020) Animated Cassie: a dynamic relatable robotic character. In: 2020 IEEE/RSJ international conference on intelligent robots and systems (IROS), pp 3739–3746. https://doi.org/10.1109/IROS45743.2020.9340894
37. Chan L, Zhang BJ, Fitter NT (2021) Designing and validating expressive cozmo behaviors for accurately conveying emotions. In: 2021 30th IEEE international conference on robot and human interactive communication (RO-MAN), pp 1037–1044. https://doi.org/10.1109/RO-MAN50785.2021.9515425
38. Wojciechowska A, Frey J, Mandelblum E et al (2019) Designing drones: factors and characteristics influencing the perception of flying robots. Proc ACM Interact Mobile Wearable Ubiquitous Technol 3(3):1–19. https://doi.org/10.1145/3351269
39. Ruijten PAM, Cuijpers RH (2018) If drones could see: investigating evaluations of a drone with eyes. In: International conference on social robotics. Springer, pp 65–74. https://doi.org/10.1007/978-3-030-05204-1_7
40. Kirby R, Forlizzi J, Simmons R (2010) Affective social robots. Robot Auton Syst 58(3):322–332. https://doi.org/10.1016/j.robot.2009.09.015
41. Hegel F, Spexard T, Wrede B, et al (2006) Playing a different imitation game: interaction with an empathic android robot. In: 2006 6th IEEE-RAS international conference on humanoid robots, pp 56–61. https://doi.org/10.1109/ICHR.2006.321363
42. Tielman M, Neerincx M, Meyer JJ, et al (2014) Adaptive emotional expression in robot-child interaction. In: Proceedings of the 2014 ACM/IEEE international conference on human–robot interaction. ACM, New York, NY, USA, HRI '14, pp 407–414. https://doi.org/10.1145/2559636.2559663
43. James J, Watson CI, MacDonald B (2018) Artificial empathy in social robots: an analysis of emotions in speech. In: 2018 27th IEEE international symposium on robot and human interactive communication (RO-MAN). IEEE, pp 632–637. https://doi.org/10.1109/ROMAN.2018.8525652
44. Brave S, Nass C, Hutchinson K (2005) Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. Int J Hum Comput Stud 62(2):161–178. https://doi.org/10.1016/j.ijhcs.2004.11.002
45. de Graaf MMA, Ben Allouch S, van Dijk JAGM (2015) What makes robots social?: A user's perspective on characteristics for social human–robot interaction. In: Social Robotics. Springer, Cham, pp 184–193. https://doi.org/10.1007/978-3-319-25554-5_19
46. Pelikan HRM, Broth M, Keevallik L (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. ACM, New York, NY, USA, HRI '20, pp 461–470. https://doi.org/10.1145/3319502.3374814
47. Eastwood JD, Smilek D, Merikle PM (2001) Differential attentional guidance by unattended faces expressing positive and negative emotion. Percept Psychophys 63(6):1004–1013. https://doi.org/10.3758/BF03194519
48. Anderson M, Anderson S, Armen C (2005) Towards machine ethics: implementing two action-based ethical theories. In: AAAI fall symposium—technical report, pp 1–7
49. Ojha S, Williams MA, Johnston B (2018) The essence of ethical reasoning in robot-emotion processing. Int J Soc Robot 10(2):211–223. https://doi.org/10.1007/s12369-017-0459-y
50. Petrak B, Stapels JG, Weitz K, et al (2021) To move or not to move? social acceptability of robot proxemics behavior depending on user emotion. In: 2021 30th IEEE international conference on robot and human interactive communication (RO-MAN), pp 975–982. https://doi.org/10.1109/RO-MAN50785.2021.9515502
51. van Maris A, Zook N, Caleb-Solly P et al (2020) Designing ethical social robots–a longitudinal field study with older adults. Front Robot AI 7:1. https://doi.org/10.3389/frobt.2020.00001
52. Sharkey A, Sharkey N (2012) Granny and the robots: ethical issues in robot care for the elderly. Ethics Inf Technol 14(1):27–40. https://doi.org/10.1007/s10676-010-9234-6

53. Karjalainen KD, Romell AES, Ratsamee P, et al (2017) Social drone companion for the home environment: a user-centric exploration. In: Proceedings of the 5th international conference on human agent interaction. ACM, New York, NY, USA, HAI '17, pp 89–96. https://doi.org/10.1145/3125739.3125774
54. Lidynia C, Philipsen R, Ziefle M (2017) Droning on about drones–acceptance of and perceived barriers to drones in civil usage contexts. In: Advances in human factors in robots and unmanned systems. Springer, Cham, pp 317–329. https://doi.org/10.1007/978-3-319-41959-6_26
55. Salvini P, Laschi C, Dario P (2010) Design for acceptability: improving robots' coexistence in human society. Int J Soc Robot 2(4):451–460. https://doi.org/10.1007/s12369-010-0079-2
56. Plutchik R (1980) A general psychoevolutionary theory of emotion. In: Theories of Emotion. Elsevier, pp 3–33. https://doi.org/10.1016/B978-0-12-558701-3.50007-7
57. Russell JA (1980) A circumplex model of affect. J Personal Soc Psychol 39(6):1161–1178. https://doi.org/10.1037/h0077714
58. Bartneck C, Reichenbach J, van Breemen A (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, pp 32–51. https://doi.org/10.6084/m9.figshare.5160769
59. Cañamero L, Fredslund J (2001) I show you how i like you—can you read it in my face? [robotics]. IEEE Trans Syst Man Cybern Part A Syst Hum 31(5):454–459. https://doi.org/10.1109/3468.952719
60. Gasper K, Spencer LA, Hu D (2019) Does neutral affect exist? How challenging three beliefs about neutral affect can advance affective research. Front Psychol 10:2476. https://doi.org/10.3389/fpsyg.2019.02476
61. Goldschmidt G, Sever AL (2011) Inspiring design ideas with texts. Des Stud 32(2):139–155. https://doi.org/10.1016/j.destud.2010.09.006
62. Hofstede G, Hofstede GJ, Minkov M (2005) Cultures and organizations: software of the mind, vol 2. Mcgraw-Hill, New York
63. Nomura T, Suzuki T, Kanda T et al (2006) Measurement of negative attitudes toward robots. Interact Stud 7(3):437–454. https://doi.org/10.1075/is.7.3.14nom
64. Wojciechowska A, Hamidi F, Lucero A, et al (2020) Chasing lions: Co-designing human-drone interaction in sub-Saharan Africa. In: Proceedings of the 2020 ACM designing interactive systems conference. ACM, New York, NY, USA, DIS '20, pp 141–152. https://doi.org/10.1145/3357236.3395481
65. Harrison C, Faste H (2014) Implications of location and touch for on-body projected interfaces. In: Proceedings of the 2014 conference on designing interactive systems. ACM, New York, NY, USA, DIS '14, pp 543–552. https://doi.org/10.1145/2598510.2598587
66. Braun V, Clarke V (2006) Using thematic analysis in psychology. Qual Res Psychol 3(2):77–101. https://doi.org/10.1191/1478088706qp063oa
67. Terry G, Hayfield N (2020) Reflexive thematic analysis. Edward Elgar Publishing, Cheltenham. https://doi.org/10.4337/9781788977159.00049
68. Cauchard JR, Gover W, Chen W et al (2021) Drones in wonderland—disentangling collocated interaction using radical form. IEEE Robot Autom Lett. https://doi.org/10.1109/LRA.2021.3103653
69. Shapira S, Cauchard JR (2022) Integrating drones in response to public health emergencies: a combined framework to explore technology acceptance. Front Public Health 10:1. https://doi.org/10.3389/fpubh.2022.1019626
70. Rosenfeld A (2019) Are drivers ready for traffic enforcement drones? Accid Anal Prevent 122:199–206. https://doi.org/10.1016/j.aap.2018.10.006
71. Aydin B (2019) Public acceptance of drones: knowledge, attitudes, and practice. Technol Soc 59(101):180. https://doi.org/10.1016/j.techsoc.2019.101180
72. Tsiourti C, Weiss A, Wac K, et al (2017) Designing emotionally expressive robots: A comparative study on the perception of communication modalities. In: Proceedings of the 5th international conference on human agent interaction. ACM, New York, NY, USA, HAI '17, pp 213–222. https://doi.org/10.1145/3125739.3125744
73. Chen JYC, Lakhmani SG, Stowers K et al (2018) Situation awareness-based agent transparency and human-autonomy teaming effectiveness. Theor Issues Ergon Sci 19(3):259–282. https://doi.org/10.1080/1463922X.2017.1315750
74. Mercado JE, Rupp MA, Chen JYC et al (2016) Intelligent agent transparency in human-agent teaming for multi-UxV management. Hum Factors 58(3):401–415. https://doi.org/10.1177/0018720815621206

Viviane Herdel is a Ph.D. candidate at the Magic Lab, Department of Industrial Engineering and Management at the Ben-Gurion University of the Negev, Israel. She received her B.Sc. in Sensors and Cognitive Psychology at the Chemnitz University of Technology, Germany, and her M.Sc. in Neurocognitive Psychology at the University of Oldenburg, Germany. Her research interests include HCI, HRI, UX, Social Robotics, and Affective Computing.

Jessica R. Cauchard is an assistant professor at the Department of Industrial Engineering and Management at the Ben-Gurion University of the Negev, where she founded and heads the Magic Lab. Her research is rooted in the fields of Human-Computer and Human-Robot Interaction with a focus on novel interaction techniques. She received her Ph.D. from the University of Bristol and worked as a postdoctoral scholar at Stanford University and Cornell Tech.