



The Effects of Natural Sounds and Proxemic Distances on the Perception of a Noisy Domestic Flying Robot

Downloaded from: <https://research.chalmers.se>, 2025-12-04 11:02 UTC

Citation for the original published paper (version of record):

Wang, Z., Hu, Z., Rohles, B. et al (2023). The Effects of Natural Sounds and Proxemic Distances on the Perception of a Noisy Domestic Flying Robot. *ACM Transactions on Human-Robot Interaction*, 12(4): 1-32.
<http://dx.doi.org/10.1145/3579859>

N.B. When citing this work, cite the original published paper.

The Effects of Natural Sounds and Proxemic Distances on the Perception of a Noisy Domestic Flying Robot

ZIMING WANG, Chalmers University of Technology, Sweden and University of Luxembourg, Luxembourg

ZIYI HU, Chalmers University of Technology, Sweden

BJÖRN ROHLES, University of Luxembourg, Luxembourg

SARA LJUNGBLAD, University of Gothenburg, Sweden and Chalmers University of Technology, Sweden

VINCENT KOENIG, University of Luxembourg, Luxembourg

MORTEN FJELD, Chalmers University of Technology, Sweden and University of Bergen, Norway

When flying robots are used in close-range interaction with humans, the noise they generate, also called consequential sound, is a critical parameter for user acceptance. We conjecture that there is a benefit in adding natural sounds to noisy domestic drones. To test our hypothesis experimentally, we carried out a mixed-methods research study ($N = 56$) on reported user perception of a sonified domestic flying robot with three sound conditions at three distances. The natural sounds studied were, respectively, added to the robot's inherent noises during flying; namely, a *birdsong* and a *rain* sound, plus a control condition of no added sound. The distances studied were set according to proxemics; namely, *near*, *middle*, and *far*. Our results show that adding *birdsong* or *rain* sound affects the participants' perceptions, and the proxemic distances play a nonnegligible role. For instance, we found that participants liked the *bird* condition the most when the drone was at *far*, while they disliked the same sound the most when at *near*. We also found that participants' perceptions strongly depended on their associations and interpretations deriving from previous experience. We derived six concrete design recommendations.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; **Auditory feedback**;

Additional Key Words and Phrases: Close-range human-drone interaction, natural sounds, domestic flying robots, consequential sound

ACM Reference format:

Ziming Wang, Ziyi Hu, Björn Rohles, Sara Ljungblad, Vincent Koenig, and Morten Fjeld. 2023. The Effects of Natural Sounds and Proxemic Distances on the Perception of a Noisy Domestic Flying Robot. *ACM Trans. Hum.-Robot Interact.* 12, 4, Article 50 (December 2023), 32 pages.
<https://doi.org/10.1145/3579859>

This research is funded by the Marianne and Marcus Wallenberg Foundation.

Authors' addresses: Z. Wang, Chalmers University of Technology, Gothenburg, Sweden and University of Luxembourg, Esch-sur-Alzette, Luxembourg; email: ziming@chalmers.se; Z. Hu, Chalmers University of Technology, Gothenburg, Sweden; email: hziyi@student.chalmers.se; B. Rohles and V. Koenig, University of Luxembourg, Esch-sur-Alzette, Luxembourg; emails: bjorn.rohles@uni.lu, vincent.koenig@uni.lu; S. Ljungblad, University of Gothenburg, Gothenburg, Sweden and Chalmers University of Technology, Gothenburg, Sweden; email: sara.ljungblad@chalmers.se; M. Fjeld, Chalmers University of Technology, Gothenburg, Sweden and University of Bergen, Bergen, Norway; email: fjeld@chalmers.se.



This work is licensed under a Creative Commons Attribution International 4.0 License.

© 2023 Copyright held by the owner/author(s).

2573-9522/2023/12-ART50

<https://doi.org/10.1145/3579859>

1 INTRODUCTION

Flying robots, or **unmanned aerial vehicles (UAVs)**, are commonly known as drones. In this article, we use the definition of “Robot” in ISO 8373:2021(en) [32], which indicates that only drones with a certain degree of (semi- to full-scale) autonomous capability count as flying robots, not fully piloted ones. We use the term “flying robot” whenever we wish to explicitly emphasize this autonomous characteristic. However, for convenience’s sake, the term “drone” will be used interchangeably in this article, especially when discussing the existing literature. Drones are already in frequent use for different purposes. There is a growing interest in making drone applications more ubiquitous [1, 48], and domestic drone applications are likewise gaining increased interest [49].

As the mechanical functioning of robots tends to generate consequential sounds, robot noise is generally inevitable in real life. The noise of a flying drone is particularly salient, as it requires a continuous lifting force from high-frequency turbulent airflows generated by propellers, thus creating loud consequential sounds. Such noise is even more intense in close-range human-drone interaction. With domestic drone applications becoming increasingly popular [49, 61], noise has become a critical issue for user acceptance of drones that interact with humans in close proximity [11, 12, 34].

Many strategies to solve the drone noise issue have been tried, but due to limitations of size, cost, and weight, no single strategy has yet achieved satisfying results for domestic drones (see Section 2.1). In this article, we investigate adding natural sounds to mask or mitigate noises. Literature has shown that nature exposure and listening to natural sounds may not only lower stress and annoyance but also improve health and create positive affect [8, 10, 68]. Noise masking by adding natural sounds has been proposed and studied in various areas of research and commercial applications (see Section 2.3). Following this line of research, we propose adding natural sounds to a noisy domestic flying robot and conjecture that this could have positive effects on people’s perceptions. We designed a mixed-methods empirical study with human participants ($N = 56$) to examine this strategy. Through both quantitative and qualitative approaches, we acquired interesting insights that may be helpful for the design of enjoyable human-robot interactions (see Sections 6 and 7).

The contributions of this article are: (i) Investigating the idea of adding natural sounds to flying robots to alter human perception and make close-range interaction with domestic drones more acceptable. (ii) Sorting out the correlation between sound conditions and proxemics of flying robots, namely, investigating the changes in people’s reported perceptions when adding different natural sounds at different proxemic distances. (iii) Presenting an original empirical study exposing participants to a real flying robot in a realistic and controlled environment, offering a full sensory experience with high realism. (iv) Offering empirical findings, especially qualitative data, that supports earlier models on how perceptions and experiences are formed. We present and discuss a visual summary with the potential to explain why an identical stimulus might lead to diverse and even contradictory individual interpretations. (v) Deriving design recommendations for domestic drones and future work.

2 BACKGROUND AND RELATED WORK

This section will discuss the background and related work concerning various aspects, including currently existing solutions to drone noise, inspiring ideas and related research regarding consequential sounds in **human-robot interaction (HRI)**, adding natural sounds to change the perception of a sound event, proxemics and HRI which includes close-range human-drone interaction, and the importance of realism for sensory experiences. We also identify gaps in the existing literature, which motivate our approach.

2.1 Existing Solutions to Drone Noise

The noise of interactive domestic drones, which is a major focus of our research, is a consequential sound (see Section 2.2) that is undesirable and considered an obvious disadvantage [11, 14]. Many methods have been tried to reduce the impact of drone noise on humans. Miljković [43] systematically reviewed these methods and proposed multiple concepts for attenuating drone noise, including both passive and active solutions. In the passive noise reduction method, optimizing the shape of propellers to reduce drone noise has been successfully commercialized. A kind of low-noise propeller is currently being sold by DJI as a modified part for their drones [65]. However, this method only reduces the noise level by a few dB [43]. In user comments on the online shopping platform Amazon, many customers have pointed out that the noise reduction effect from the low-noise propeller was not obvious [16]. The **active noise canceling (ANC)** solution, which emits a same-amplitude sound wave with an inverted phase, could achieve a significant noise canceling effect [47]. However, due to the complexity of the ANC system for drones, it is still in the laboratory research stage and far from being introduced in practice. Also, the need for an additional energy source to power an ANC system limits the scope of this solution to large drones [43]. For small domestic drones, which is the scope of this study, La Delfa et al. commented that these drones' small size already minimizes propeller noise and airflow, thereby tackling the main pain points in close-range human-drone interaction [34]. However, this is debatable—although smaller drones do emit fewer consequential sounds, making them relatively inoffensive, small drones are by no means silent, and their consequential sound is still noisy and problematic when they are flying indoors, especially in close proximity to humans (by analogy, similar to an insect like a mosquito or a bee flying next to your ear). Therefore, while we agree that smaller flying robots are more suitable for close-range human-drone interaction, their consequential sound still needs masking. Finally, another recent solution is to completely remove propellers. Yamada et al. present an indoor helium blimp drone that uses piezo elements for propulsion [76]. While using a balloon makes the drone move with a low noise level, its poor mobility makes it impractical for domestic applications. In summary, currently existing solutions to the problem of noise in flying robots are insufficient.

2.2 Consequential Sound in HRI: Robot Noise and Altering Sound Profile

As one core modality robots can use to communicate with humans, sound may be used to create rich and engaging **human-robot interactions (HRI)** [55]. Thus, a wide variety of different types of robot sounds have been considered in HRI [52, 55]. While in the past, sound-related research in the HRI field mainly focused on speech and semantic-free utterances, implicit communication in the form of robots' non-verbal sounds, including consequential sounds, is currently attracting growing attention [56, 57]. A review of consequential sound in HRI can be found in Robinson et al. [57]. Consequential sounds are sounds radiated by machines during their motions as a consequence of their operation and construction [35, 78] that are perceived by people as noise [35]. Negative effects of uncontrolled consequential sound on users' perceptions of interacting with robots have already been reported [30, 46, 66]. Although past researchers encountered considerable challenges connecting objective auditory characteristics to the subjective perception of consequential robot sounds, past results demonstrate that altering the sound profile of robots might improve user experiences [78]. For instance, masking robots' undesirable consequential sound with musical sounds has been proposed and validated as effective for mitigating unpleasant feelings [23, 67, 79]; Zhang et al. explored consequential robot sound and confirmed that quieter robots were perceived as less discomforting, while higher-pitched sounds were preferred for a UR5e robot arm [78]; Robinson et al. found that overlaying movement sound designs to the same video of a robot both increased and decreased perceived movement quality compared to the silent control condition [57]; and so forth (e.g., References [45, 46, 64]). Nevertheless, we believe an obvious limitation

of the aforementioned past work is that most tested only video or audio recordings with participants rather than actual robots, thereby neglecting the impact of real-world dependent factors (e.g., multisensory inputs and realistic experience—see further discussion in Section 2.5) and potentially leading to poor ecological validity. In addition, consequential sounds of robots are underexplored, and little is known about how they could influence human-robot interaction [37]. Therefore, more research is needed to explore consequential sounds in HRI, especially research using real robots in a real-world HRI environment.

2.3 Changing the Perception of a Sound Event by Adding Natural Sounds

Mitigating the problematic consequential sounds of domestic drones by incorporating natural sounds is an area where we see potential benefits. Recent work has shown the benefits of exposure to nature in terms of behavioral and psychophysiological responses, decreasing negative affect and increasing positive affect and subjective well-being [8]. A systematic literature review and meta-analysis indicated aggregate evidence for the health benefits of natural sounds [10]. Adding natural sounds to mask noises or soothe users has been proposed in both research and commercial applications: (i) De Coensel et al. conducted a listening experiment on sound quality combining road traffic noise with either bird or fountain sounds, concluding that both of these natural sounds had a positive impact on auditory perception, but that bird sound had a significantly better impact [13]. (ii) The sound blocker “otohime,” which adds a water sound to mask embarrassing noise in toilets, has been commonly used in Japan since the 1980s [62, 75]. (iii) According to the developers of a sound box called “Zwitscher Box,” which generates the sound of birds chirping when people pass by, the sound “eases our minds and softens our gaze, we relax intuitively” [53]. De Coensel et al. also mentioned that the soothing effect of adding natural sound is stronger when the noise has low temporal variability [13]. Compared with traffic noise, flying robot noise has smaller fluctuations in sound level and frequency. This implies that we might achieve good results by adding natural sounds to flying robot noise. A white noise like rain always has a masking effect that may reduce noise [40] and soothe people who are exposed to it [58]; thus, it stands to reason to add a rain sound to drones’ consequential sound. Similarly, it is intuitive to associate a flying robot with a flying animal, leading us to expect a positive effect of adding the sound of birdsong to the drone.

2.4 Proxemics and HRI

Although adding natural sounds seems like a promising approach to masking the sounds of drones, as outlined above, it is vital to consider that human experiences of interactions also depend on the distance involved, usually referred to as proxemics. The term “proxemics” was coined by Edward Hall in 1966 and “identifies the culturally dependent ways in which people use interpersonal distance to understand and mediate their interactions with other people” [26]. He defined four proxemic zones for how people interpret interpersonal distance: intimate (less than 45 cm), personal (about 45 cm to 122 cm), social (about 122 cm to 366 cm), and public (about 366 cm to 762 cm). People adjust these distances to solidify their defense mechanisms when others invade these zones. Proxemics interactions have been investigated from several perspectives in the area of HRI, e.g., in industry settings [36], in experimental settings [50], in shape-shifting display settings [63], and in domestic settings [70]. The definitions of proxemic zones are considered the most relevant aspect for HRI [11, 18]. A number of studies have focused on investigating the relationship between sounds and proxemics for HRI (excluding flying robots). For instance, Walters et al. used a mechanical-looking robot with different synthesized voices and found that unhuman-like voices require a larger approach distance between participants and the robot [71]. Trovato et al. investigated the influence of robot noise on proxemics and found an increased effectiveness of a mask for eliminating the bad effects of noise, compared to no mask [67]. In addition, there have

been some studies focusing on proxemics and flying robots, as small domestic drones with aerial capabilities can assist users in collaborative tasks, which necessitate interactions with people in close range [61]. For instance, Duncan and Murphy found no conclusive difference in comfort for a small drone approaching a human at different heights [19]. Yeh et al. designed a social-look drone with a welcoming voice and pet-like face, which decreased the minimum acceptable distance significantly [77]. Nevertheless, we found a lack of literature investigating the relationship between sounds and proxemics specifically with respect to flying robots.

2.5 The Importance of Realism for Sensory Experiences

In the HRI field, non-verbal sound does affect human perceptions, but it is not the sole factor in the multi-modal affective system [5]. Therefore, sound design for products and interactions must take the input of other senses into account [7]. Human perceptions in physical interactions are multisensory. Usually one or more of the senses: sight (vision), hearing (audition), touch (somatosensation), smell (olfaction), and taste (gustation), intervene in the same interaction [6, 51]. It has been demonstrated that one modality of sensory information can change how a person perceives another [4]. For some people, noise and airflow combined could even make the flying robot “threatening” [12]. Due to the multisensory and exquisite repertoire of human nature, offering a full sensory experience with high realism is particularly important in research on close-range human-drone interaction [73]. Moreover, even when it comes to a single sense, experiencing something through a recording is different than sensing the real world. For example, regarding vision, our two eyes with overlapping and slightly different viewpoints are delicate enough to determine fine differences in depth [31], such that seeing a photo or watching a video does not provide the same level of realism as seeing things in reality. As already mentioned in Section 2.2, we consider the use of video or audio recordings in HRI studies as an obvious limitation to their ecological validity. Instead, it would be arguably beneficial to use real robots interacting with participants in a real-world HRI environment.

Our work addresses the lack of knowledge related to human perceptions of noisy domestic flying robots by combining research on how to use natural sounds and proxemic distances. Specifically, our work will address the following needs and objectives: (i) A satisfying solution to the small domestic drone noise problem in the field of human-robot interaction is absent; the current method of suppressing noise is not suitable for small drones [43]. (ii) The validity of adding masking sounds to robots and adding natural sound to noisy traffic settings to improve people’s perceptions has been confirmed [13, 23, 67]. To the best of our knowledge, there is no existing literature or similar proposals to study user perceptions when adding natural sound to robots or drones that take proxemics into account. (iii) Previous studies mainly used only audio or video recordings to investigate how users’ perceptions changed when adding sound to a robot or noisy traffic scenario. However, the lack of other sensory stimuli such as somatosensation may lead to imperfect conclusions. Here, we set up a real scene that offers a full sensory experience with high realism. (iv) Furthermore, we extend our perspective by investigating the individual experiential dimension, building on a combination of quantitative and qualitative methods, namely, a mixed-methods approach [17], which we consider valuable, since most of the aforementioned related work took either a quantitative or a qualitative approach exclusively.

3 METHODOLOGY AND EXPERIMENTS

3.1 Ideation and Hypothesis

Adding bird or rain sounds to a drone while it flies may positively affect people’s perception of its presence, especially for close-range interactions. Proxemic distances may impact how people perceive the flying robot within certain sound conditions.

We designed and prototyped a small flying robot with an on-board loudspeaker to play chosen natural sounds (either birdsong or a rain sound; see details in Section 3.3) while flying at different proxemic distances (see Section 3.4). We hypothesized that adding natural sound would improve people's perceptions of the drone. We also hypothesized that closer distance has a negative effect on perception due to the louder noise and higher risk of collisions at a closer distance.

3.2 Experimental Design

Our experimental design has two factors, while each factor has three levels, namely: 3 (sound conditions: *bird*, *rain*, *none*) \times 3 (distance conditions: *near*, *middle*, *far*). The setup was a randomized within-subjects approach, where each participant experienced all nine conditions, which were presented in different orders. For each factor with three levels A, B, C, there are six possible sequences ($3! = 6$), namely: ABC, ACB, BAC, BCA, CBA, CAB. We listed all six possible permutations of conditions for each factor (namely, either sound or distance) accordingly. For practical reasons, we first stipulated the order of the three distance conditions via complete counterbalancing, so each participant received a prearranged sequence of distance conditions; then, at each distance, the order of the three sound conditions was randomly determined by letting each participant roll a six-sided dice.

3.3 Natural Sound Design and Choices

Based on past studies and the particular rationales articulated in Section 2.3, we set up three different sound conditions for our experiment, namely: the original drone noise condition with no additional sound (control condition), the added *birdsong* condition, and added *rain* sound condition. The three sound conditions will be subsequently referred to as the *none* condition, the *bird* condition and the *rain* condition, respectively, for short.

3.3.1 Original Drone Noise Condition (Control Condition). No additional sound was added in the original drone consequential sound condition. The humming noise generated by drones is primarily due to their high-speed running motors and rotating propellers [44], which is a significant pain point for interactive drones [34].

3.3.2 Birdsong Condition. We chose the sound of the great tit (*Parus major*) as the birdsong sample. This bird is widespread throughout Europe, commonly resident in any sort of woodland [20]. We chose this common bird, as we had expected participants to recognize the sound as a local bird. Moreover, the song of the great tit was perceived as clear, lively, and cheerful during pilot tests. Typically, their song consists of roughly 3-second strophes and a 2-second break. The strophes consist of a series of phases composed of one to four different notes (defined as a continuous sound trace on a spectrogram) [25]. The bird sound recording we used is from the open-source bird sounds website xeno-canto [74].

3.3.3 Rain Sound Condition (White Noise). For a rain sample, we chose an ambient sound called *Weather Ambience Heavy Rain Downpour Splatty 01.wav* from Adobe Audition open-source library [2]. This sound of heavy rain is characteristic and loud, aiming to ensure participants recognize it even when the flying robot's noisy motors are running.

3.4 Proxemic Distance Choices

Three takeoff locations were chosen according to the theory of proxemics [26, 73] and the size constraints of the experimental setting. The three takeoff distances were designated to be approximately 45 cm, 115 cm, and 185 cm away from participants, i.e., in the range of intimate space, personal space, and social space [18], respectively. We tried to understand these distances in the

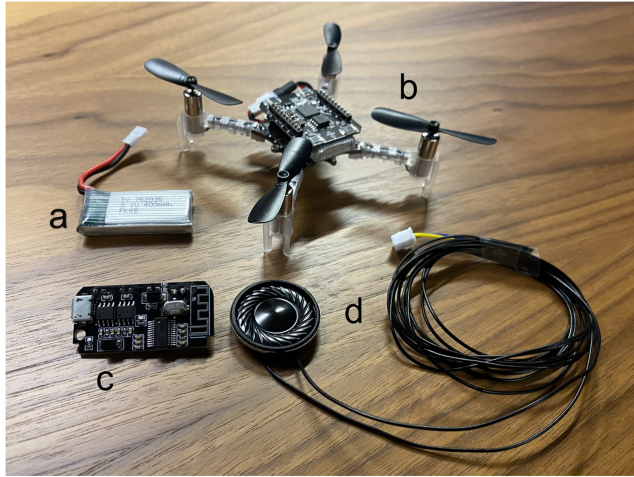


Fig. 1. (a) Li-on battery for Bluetooth board; (b) the “Crazyflie” drone with its on-board battery and lighthouse positioning deck; (c) Bluetooth board; (d) 28 mm loudspeaker and wires.

context of human-robot interaction. These three different locations will be subsequently referred to as the *near*, *middle*, and *far* locations. The three sound conditions were randomly played at each takeoff location.

3.5 Choice of Drone and Engineering the Flying Robot

Crazyflie is a small programmable quadcopter that is designed for indoor flying [15]. Crazyflie provides a wide range of open source Python programs, decks, and components to meet different research demands. In previous work, a mini flying robot with a smooth, stable, and high-precision flight trajectory was considered more acceptable by participants [34]. To achieve stable and precise flying trajectories, we chose the lighthouse positioning system to assist our experiment. With two lighthouse base stations and a lighthouse positioning deck on top of the quadcopter, the flying robots were able to fly with precision under our program’s control. To play natural sounds, we mounted a 28 mm round metal loudspeaker to the bottom side of the drone. The loudspeaker was powered by a 5 W Bluetooth amplifier with an extra Li-on battery through wires. The wires allowed the Bluetooth board and the battery to be installed in a small box and hidden under the desk during the test, which is a temporary solution to save on-board battery by lowering the takeoff weight. The small drone and accessories used in the experiment are shown in Figure 1. Following a **Research through Design (RtD)** approach [80], we decided on the following drone trajectory: The robot would first take off and fly vertically to a height of 40 cm above the table, then stay hovering for 10 seconds, afterward vertically land on the table. The selected natural sounds adapted to the drone flight durations and were played through the on-board loudspeaker.

3.6 Experimental Setup

The experiment was conducted in a soundproof chamber to avoid interference from outside noise. To make the experimental environment closely approximate the household setting where domestic flying robots would be expected, we placed several pieces of furniture as shown in Figure 2. Participants sat in front of the long desk with two lighthouse positioning base stations set behind them out of sight. The desk and chair were pre-located and marked to keep all participants at a similar distance from the flying robot. A long blanket with three position marks was placed in the



Fig. 2. Furniture setting of the experimental scene.

middle of the desk. This blanket was used as an absorber to decrease the reflected sound wave from the desk, and the marks on the blanket were used to show the three different takeoff settings.

3.7 Preliminary Study and Engineering Evaluation

During a preliminary study, we recorded the three chosen sound conditions and measured their sound pressure levels. The total A-weighted SPL of the *bird* condition and the *rain* condition were calibrated to a similar level during the test (i.e., 71 dBA, 66 dBA, 61 dBA from the *near* to the *far* locations, respectively). The frequency spectra of the three sound conditions are shown in Figure 3, which matched our expectations well. The spectrum of flying robot noise is a wide-bandwidth noise mainly concentrated below 1 KHz. The rain sound is more constant and close to white noise, which is known to achieve a decent sound-masking effect when added to a wide-bandwidth noise, as some noise features may be hidden. However, bird sound is usually high-frequency and narrow-bandwidth. For this reason, purely from a spectrum engineering perspective, adding bird sound to a wide-bandwidth noise would have only a very limited sound-masking effect. Nonetheless, previous literature shows that adding bird sound to a similar noise can work very well [13, 27], even better than water sound [13]. This obviously shows that how humans perceive sounds is not purely dependent on the features of mechanical waves, in accordance with the biological principles mentioned later on in the discussion (see Section 6.2).

3.8 Participants and Study Procedures

We recruited participants through multiple ways, including social media, flyers on campus and at student residences, and sending invitations to friends and colleagues (snowball sampling). Each

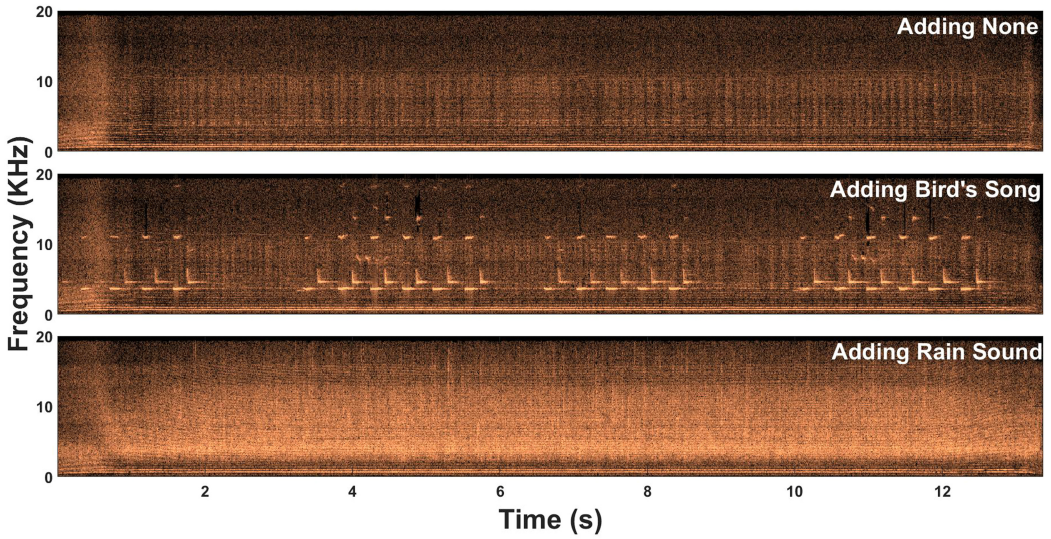


Fig. 3. Frequency spectra of the three sound conditions.

session involved one individual participant, and each participant received a cinema ticket after the test as compensation. As all experiments were carried out entirely in Sweden, we carefully followed the Swedish Ethics Review Authority's guidelines [21] and ensured that the national Ethics Review Act [60] and relevant regulatory requirements were complied with.

3.8.1 Safety Precautions. (i) To avoid physical harm to participants in the case of the flying robot losing control, we did hundreds of tests before the formal experiment and implemented a set of safety precautions: All participants were instructed to protect their bodies with a blanket made available by throwing the unfolded blanket over the flying robot to pull it down if the robot happened to divert from the planned trajectories; (ii) participants who were not already wearing their own glasses were required to wear goggles to protect their eyes; (iii) the robot's battery was exchanged with a fully charged one after every three takeoffs during the manual changes from one takeoff location to another to avoid battery voltage drop and ensure stable operation.

3.8.2 Signing Consent form. Prior to the study, every participant was given the Research Consent Form and enough time to read it. They were then invited to ask any questions before giving their consent to the mentioned procedures, including being observed and audio-recorded, by signing the form agreeing to participate.

3.8.3 Study Phases. There were three phases during each study: (i) In the briefing phase (around 10 min), the researchers introduced the study in detail, including the above-mentioned safety precautions. Participants were told that "We (the researchers) hope that, through your participation, we will learn more about the challenges and opportunities for designing flying robots, especially in terms of sound features and the close-range interactions." To reduce the effect of demand characteristics, we deliberately did not inform about our hypothesis, and we told every participant: "There are no right or wrong answers. We want you to honestly note down your evaluations and later tell us about your feelings and thoughts." (ii) In the experimental phase (around 20 min), participants were exposed to nine performances by a small sonified flying robot, with the order of performances randomized to exclude sequence effects. After each performance, participants were asked to evaluate six features in a questionnaire. Participants were also asked to rate their

preferences among the three sound conditions at each distance. We filmed each experimental condition from the participants' first-person perspective, and the video clips can be accessed via a link.¹ (iii) In the debriefing phase (around 15 min), participants were interviewed regarding their experience, thoughts, and comments on the aforementioned performances.

We ended up having 56 participants, including 31 self-identified males, 24 self-identified females, and 1 person who self-identified as other, leading to a total of 56 experiment sessions. The age range of participants was between 20 and 59 ($M = 28.5$, $SD = 8.63$). Each session took around 40 to 60 minutes, with most differences occurring in the briefing and debriefing phases, as some participants had more things to talk about than others. Eight participants' ratings were removed due to self-reported hearing impairment (2), technical failures during the test (3), and the written notes on their questionnaires indicating that they could not correctly identify all the sound conditions (3). In the end, we included 48 participants' quantitative data from answered questionnaires in the statistical analysis, with the nine experimental conditions counterbalanced via a complete counterbalancing the distance factor (eight times all six possible sequences, $8 \times 6 = 48$) and simple randomization of the sound factor (by rolling a dice) at each distance. Nevertheless, we still considered the interview data from all persons to be very valuable, as it adequately represented the participants' experiences, so we included all 56 participants' qualitative data for the thematic analysis. Please see the following sections.

3.9 Measurements of Reported Perception

Participants were asked to evaluate each performance after its end with respect to six measurements describing the perceived characteristics of the flying robot: "loud," "sharp," "pleasant," "safe," "relaxing," and "attractive," on a scale of 0–10, with 0 representing "not at all" and 10 "extremely." These characteristics were selected based on both existing literature and the focus of this study, as indicated below. After all the performances finished, every participant was also asked to rank their preference for the three sound conditions played at each distance by giving 0 points to the least liked, 1 point to the medium favorite, and 2 points to the most favored.

The measurements of perceived loudness and perceived sharpness were intended to examine how participants would feel about adding natural sounds to the drone noise soundscape. For the rest of the four measurements, participants were explicitly asked to consider their full sensory experience with the demonstrated flying robot performances. Pleasantness and attractiveness are the most commonly used perceptual assessment criteria in previous studies in both user experience [38] and soundscape quality [3, 22, 24]. Safety and causes of stress are further critical parameters for user acceptance of drones used in close proximity; thus, we wanted to examine both perceived safety and perceived absence of stress.

3.10 Post-experiment Interview Questions

The first and second authors conducted all interviews together, with detailed interview notes taken by each author separately. The interviews were primarily conducted in English. However, a number of participants were international students from China newly arrived in Sweden and felt more comfortable communicating in their native language. As both the first and second authors were native Mandarin Chinese speakers, these interviews were conducted in Chinese. Participants from other countries did not indicate a need to switch to another language. The full interviews were audio-recorded.

We used a semi-structured interview guide to elicit information about participants' experiences and perceptions of the different noise conditions. Our questions addressed the participants'

¹<https://youtube.com/playlist?list=PLV514bEMCGdbfbrqVOAGX1y3TV4Kqm5Fa>.

Table 1. Questions during the Interview

<p>Among the nine performances you have experienced:</p> <p>Which performance did you prefer the most? and Why?</p> <p>Which performance did you prefer the least? and Why?</p> <p>Which performance did you have a special feeling about? and Why?</p> <p>How did the <i>bird</i> sound? and Why?</p> <p>Have you heard this specific bird song before?</p> <p>Do you think you are an outdoor person?</p> <p>Where are you from and where did you grow up?</p> <p>How did the <i>rain</i> sound? and Why?</p> <p>Are there any other sounds you think would be suitable to add?</p> <p>Did you feel the airflow when the robot was flying?</p> <p>How did you feel about the airflow?</p> <p>Did you notice the wires when the robot was flying?</p> <p>Was it bothering you?</p> <p>What do you think the wires are used for?</p> <p>Do you have any additional points that we did not discuss?</p>
--

preferences among the performances (and the reasons for these preferences), their impressions of the tested sounds, their personal background, and their impressions of the study setup. Finally, the interviewers asked follow-up questions when appropriate, and the participants had the opportunity to add information they considered important. The specific questions used are listed in Table 1.

4 QUANTITATIVE DATA (MEASUREMENTS OF REPORTED PERCEPTION) ANALYSIS AND RESULTS

In this section, we first describe the overall statistical methods we used for analyzing the quantitative data, with a summary of the results regarding the six measurements' effects on reported perceptions in Table 2. Then, we list the detailed results for each of the six measurements in each of the following subsections, namely: *loudness*, *sharpness*, *pleasantness*, *safety*, *relaxedness*, and *attractiveness*; followed by the last subsection, which discusses the ordinal preference measurement. We provide visualizations for each measurement to support understanding of the data.

4.1 Overall Description of Statistical Methods

Statistical analysis was done using IBM SPSS Statistics (version 28.0.0.0) [29]. For our within-subjects factorial design, we conducted two-way repeated measures ANOVAs on reported perception for each of the six measurements. For each measurement, we checked the significance of the main effects of each of the two factors (sound and distance) and the interaction effect between them. In cases where one factor's main effect was significant, we carried out a post hoc analysis through multiple comparisons with Fisher's **Least Significant Difference (LSD)** test to examine the relationships between the corresponding individual levels. Regardless of whether there was a significant interaction effect, we conducted simple effects tests to compare all pairs of three levels of one factor for each of the three levels of the other factor. The simple effects tests were done with

Table 2. Results from 3 × 3 ANOVAs on Reported Perception

Measurement	Main Effect		Interaction Sound × Distance	Simple Effects					
				Sound under Distance:			Distance under Sound:		
	Sound	Distance	Distance	Near	Middle	Far	None	Bird	Rain
Loudness	Bird > None p < .001	Near > Middle p < .001	p = .008	Bird > None p < .001	Bird > None p < .001	Bird > None p = .012	Near > Middle p = .077(ns)	Near > Middle p < .001	Near > Middle p = .083(ns)
	Bird > Rain p = .004	Near > Far p = .003		Bird > Rain p = .001	Bird > Rain p = .018		Near > Far p < .001	Near > Far p < .001	Near > Far p < .001
	Rain > None p < .001	Middle > Far p < .001		Rain > None p = .033			Middle > Far p = .002	Middle > Far p = .001	Middle > Far p = .016
Sharpness	Bird > None p < .001	Near > Middle p < .001	ns (p = .126)	Bird > None p < .001	Bird > None p < .001	Bird > None p < .001	Near > Far p = .004	Near > Middle p < .001	Near > Middle p = .013
	Bird > Rain p < .001	Near > Far p < .001		Bird > Rain p < .001	Bird > Rain p < .001	Bird > Rain p < .001	Middle > Far p = .096(ns)	Near > Far p < .001	Near > Far p = .037
								Middle > Far p = .035	
Pleasantness	Bird > None p = .003	Middle > Near p < .001	p = .010	Rain > None p = .085(ns)	Bird > None p = .019	Bird > None p < .001	Middle > Near p = .007	Middle > Near p = .004	
	Rain > None p = .045	Far > Near p < .001			Rain > None p = .038	Bird > Rain p = .002	Far > Near p = .002	Far > Near p < .001	Middle > Near p = .002
								Far > Middle p = .003	
Safety	ns	Middle > Near p < .001	ns	ns	ns	ns	Middle > Near p < .001	Middle > Near p < .001	Middle > Near p < .001
		Far > Near p < .001					Far > Near p < .001	Far > Near p < .001	Far > Near p < .001
		Far > Middle p = .011					Far > Middle p = .031	Far > Middle p = .008	Far > Middle p < .001
Relaxedness	Bird > None p = .002	Middle > Near p < .001	ns (p = .090)	Bird > None p = .012	Bird > None p < .001	Bird > None p < .001	Middle > Near p = .004	Middle > Near p < .001	Middle > Near p < .001
	Rain > None p = .057(ns)	Far > Near p < .001		Rain > None p = .037	Rain > None p = .038	Bird > Rain p = .001	Far > Near p < .001	Far > Near p < .001	Far > Near p < .001
Attractiveness	Bird > None p < .001	Middle > Near p < .001	p = .012	Bird > None p = .031	Bird > None p < .001	Bird > None p = .001	Far > Near p = .023	Middle > Near p < .001	Middle > Near p = .003
	Bird > Rain p = .013	Far > Near p < .001		Rain > None p = .009	Rain > None p = .001	Bird > Rain p = .001		Far > Near p < .001	Far > Near p = .027
	Rain > None p = .033							Far > Middle p = .024	

ns = not significant. All effects listed are significant at $p < .05$ or not significant but $.05 < p < .15$.

one-way repeated measures ANOVAs followed by multiple comparisons with the LSD tests. We checked the normality of residuals via the Shapiro-Wilk test and the homogeneity of variance via Levene's Test. We decide to report partial eta squared as the estimate of effect size, denoted as η_p^2 , as it offers a more comparable estimate for factorial designs with multiple independent variables [39]. Table 2 presents a summary of the statistical analysis results.

The preference rating is different from the six parameters mentioned above. Each participant ranked their preferences for the three performances played at each distance by giving 2 points for their most preferred, 1 point for the next preferred, and 0 points for their least favorite. We conducted a one-way repeated measures ANOVA at each distance to compare the effects of the three sound conditions on the participants' preferences. Where the ANOVA revealed a significant difference, we used the LSD test to see the relationships among the three sound conditions at the specific distances.

For the six measurements of reported perceptions, the data were plotted as box-whisker plots with asterisks highlighting the significance level, where * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$. The preference data were plotted as a stacked bar chart. See related figures in the following sections.

4.2 Perceived Loudness

Figure 4 shows the ratings of perceived loudness at the three locations with three sound conditions. The main effect of the sounds ($F(2,94) = 18.44$, $p < .001$, $\eta_p^2 = 0.282$) and the distances ($F(2,94) = 24.87$,

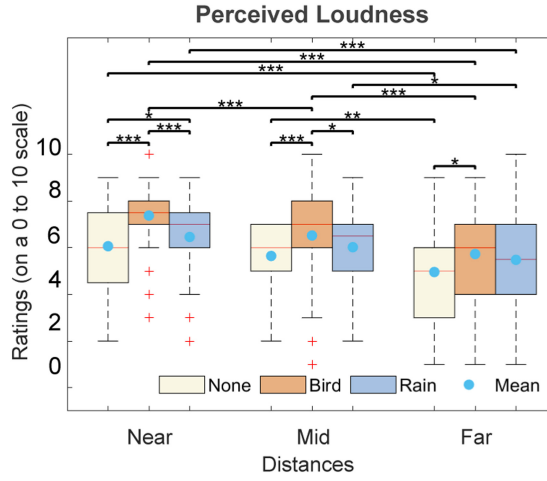


Fig. 4. Box-and-whisker plots of the perceived loudness ratings at the three locations with three sound conditions, where * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$.

$p < .001$, $\eta_p^2 = 0.346$) on perceived loudness were both significant. The interaction between sounds and distances was also significant ($F(4,188) = 3.58$, $p = .008$, $\eta_p^2 = 0.071$).

The simple effects analyses indicated for the three distances: (i) At the *near* location, the mean perceived loudness rating for the *bird* condition ($M = 7.38$, $SD = 1.63$) was significantly higher than both the *none* ($M = 6.06$, $SD = 1.92$) and the *rain* conditions ($M = 6.46$, $SD = 1.66$), and the *rain* condition was also rated significantly louder than the *none* condition. The effect size was $\eta_p^2 = 0.388$. (ii) At the *middle* location, the mean perceived loudness rating for the *bird* condition ($M = 6.52$, $SD = 1.82$) was significantly higher than both the *none* ($M = 5.65$, $SD = 1.39$) and the *rain* conditions ($M = 6.02$, $SD = 1.58$). The effect size was $\eta_p^2 = 0.273$. (iii) At the *far* location, only the mean perceived loudness rating for the *bird* condition ($M = 5.73$, $SD = 1.90$) was significantly higher than the *none* condition ($M = 4.96$, $SD = 1.96$), with an effect size of $\eta_p^2 = 0.140$.

For the three sound conditions: (i) For the *none* condition, the mean perceived loudness ratings at the *near* ($M = 6.06$, $SD = 1.92$) and *middle* locations ($M = 5.65$, $SD = 1.39$) were both significantly higher than the *far* location ($M = 4.96$, $SD = 1.96$). The effect size was $\eta_p^2 = 0.213$. (ii) For the *bird* condition, the mean perceived loudness rating at the *near* location ($M = 7.38$, $SD = 1.63$) was significantly higher than both the *middle* ($M = 6.52$, $SD = 1.82$) and the *far* locations ($M = 5.73$, $SD = 1.90$). The difference between the *middle* and *far* locations was also significant, with an effect size of $\eta_p^2 = 0.487$. (iii) For the *rain* condition, the mean perceived loudness ratings at the *near* ($M = 6.46$, $SD = 1.66$) and *middle* locations ($M = 6.02$, $SD = 1.58$) were both significantly higher than the *far* location ($M = 5.48$, $SD = 2.06$). The effect size was $\eta_p^2 = 0.209$.

4.3 Perceived Sharpness

Figure 5 shows the ratings of perceived sharpness at the three locations with three sound conditions. The main effects of the sounds ($F(2,94) = 27.47$, $p < .001$, $\eta_p^2 = 0.369$) and the distances ($F(2,94) = 13.60$, $p < .001$, $\eta_p^2 = 0.224$) on perceived sharpness were both significant. The interaction between sounds and distances was not significant ($F(4,188) = 1.82$, $p = 0.13$, $\eta_p^2 = 0.037$), but the simple effects analyses nevertheless indicated a possible interaction.

The simple effects analyses indicated for the three distances: (i) The mean perceived sharpness rating of the *bird* condition (*near*: $M = 7.65$, $SD = 1.72$; *middle*: $M = 6.73$, $SD = 2.14$; *far*: $M = 6.29$,

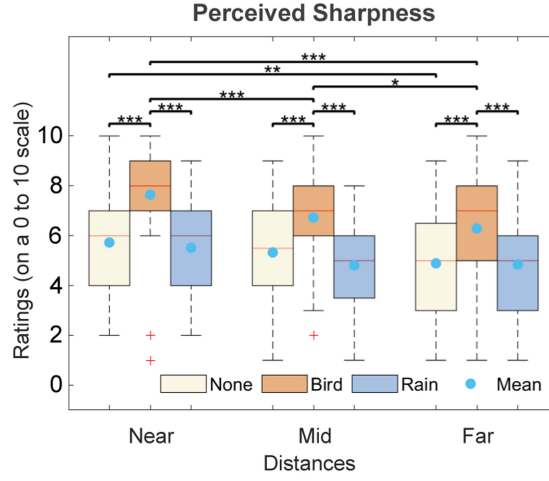


Fig. 5. Box-and-whisker plots of the perceived sharpness ratings at the three locations with three sound conditions, where * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$.

SD = 2.14) was significantly higher than the other two sound conditions. (ii) The mean perceived sharpness ratings for the *none* (*near*: $M = 5.73$, $SD = 2.10$; *middle*: $M = 5.33$, $SD = 1.83$; *far*: $M = 4.90$, $SD = 2.03$) and the *rain* conditions (*near*: $M = 5.52$, $SD = 1.88$; *middle*: $M = 4.81$, $SD = 1.79$; *far*: $M = 4.85$, $SD = 1.99$) were less decisive. The effect sizes for the three locations were $\eta_p^2 = 0.404$ (*near*), $\eta_p^2 = 0.267$ (*middle*), $\eta_p^2 = 0.202$ (*far*).

For the three sound conditions: (i) For the *none* condition, the mean perceived sharpness rating of the *near* location ($M = 5.73$, $SD = 2.10$) was significantly higher than the *far* location ($M = 4.90$, $SD = 2.03$). The effect size was $\eta_p^2 = 0.093$. (ii) For the *bird* condition, the mean perceived sharpness rating at the *near* location ($M = 7.65$, $SD = 1.72$) was significantly higher than both the *middle* ($M = 6.73$, $SD = 2.14$) and the *far* locations ($M = 6.30$, $SD = 2.14$). The difference between the *middle* and the *far* locations was also significant. The effect size was $\eta_p^2 = 0.294$. (iii) For the *rain* condition, the perceived sharpness rating at the *near* location ($M = 5.52$, $SD = 1.88$) was significantly higher than both the *middle* ($M = 4.81$, $SD = 1.79$) and the *far* locations ($M = 4.85$, $SD = 1.99$). There was no significant difference between the *middle* and the *far* locations. The effect size was $\eta_p^2 = 0.294$.

4.4 Perceived Pleasantness

Figure 6 shows the ratings of perceived pleasantness at the three locations with three sound conditions. The main effects of sounds ($F(2,94) = 6.01$, $p = .004$, $\eta_p^2 = 0.113$) and distances on perceived pleasantness ($F(2,94) = 13.08$, $p < .001$) were both significant, as was the interaction between sounds and distances ($F(4,188) = 3.40$, $p = .010$, $\eta_p^2 = 0.068$).

The simple effects analyses indicated for the three locations: (i) The mean perceived pleasantness ratings for the sound conditions did not significantly differ at the *near* location. (ii) At the *middle* location, the mean perceived pleasantness ratings for both the *bird* ($M = 5.48$, $SD = 2.52$) and the *rain* conditions ($M = 5.37$, $SD = 1.96$) were significantly higher than the *none* condition ($M = 4.65$, $SD = 1.72$). There was no significant difference between the *bird* and the *rain* conditions. The effect size was $\eta_p^2 = 0.066$. (iii) At the *far* location, the mean perceived pleasantness rating for the *bird* condition ($M = 6.21$, $SD = 2.31$) was significantly higher than both the *none* ($M = 4.71$, $SD = 2.09$) and the *rain* conditions ($M = 5.08$, $SD = 2.14$), with no significant difference between the *none* and the *rain* conditions. The effect size was $\eta_p^2 = 0.186$. It seems that the sound conditions played an

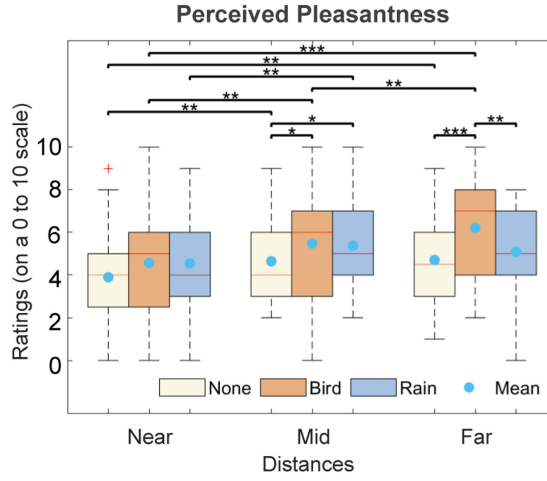


Fig. 6. Box-and-whisker plots of the perceived pleasantness ratings at the three locations with three sound conditions, where * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$.

important role in the perception of pleasantness at the *far* and *middle* distances, but not at the *near* location.

For the three sound conditions: (i) For the *none* condition, the mean perceived pleasantness ratings of both the *middle* ($M = 4.65$, $SD = 1.72$) and the *far* locations ($M = 4.71$, $SD = 2.09$) were significantly higher than the *near* location ($M = 3.90$, $SD = 2.16$), with no significant difference between the *middle* and the *far*. The effect size was $\eta_p^2 = 0.130$. (ii) For the *bird* condition, the mean perceived pleasantness rating at the *far* location ($M = 6.21$, $SD = 2.31$) was significantly higher than both the *near* ($M = 4.56$, $SD = 2.53$) and the *middle* locations ($M = 5.48$, $SD = 2.52$). The difference between the *near* and the *middle* locations was also significant. The effect size was $\eta_p^2 = 0.246$. (iii) For the *rain* condition, only the mean perceived pleasantness rating of the *middle* location ($M = 5.38$, $SD = 1.96$) was significantly higher than the *near* location ($M = 4.54$, $SD = 2.27$). The effect size was $\eta_p^2 = 0.068$. The distances played an important role in the perception of pleasantness for all three sound conditions.

4.5 Perceived Safety

Figure 7 shows the *safety* ratings at the three locations with three sound conditions. The main effects of sounds on perceived safety ($F(2,94) = 1.49$, $p = .23$, $\eta_p^2 = 0.031$) were not significant. However, the main effect of distances on perceived safety ($F(2,94) = 29.68$, $p < .001$) was significant. The sound conditions very likely had no effect on the perception of safety, but the distances did.

The simple effects analyses indicated for the three sound conditions: (i) For both the *none* and the *bird* conditions, the mean perceived safety rating at the *middle* (*none*: $M = 6.79$, $SD = 2.23$; *bird*: $M = 7.02$, $SD = 2.48$) and the *far* locations (*none*: $M = 7.29$, $SD = 2.46$; *bird*: $M = 7.73$, $SD = 2.39$) was significantly higher than the *near* location (*none*: $M = 5.13$, $SD = 2.74$; *bird*: $M = 5.35$, $SD = 2.74$), and the mean perceived safety rating at the *far* location was also significantly higher than at the *middle* location. The effect size was $\eta_p^2 = 0.348$ for the *none* condition and $\eta_p^2 = 0.315$ for the *bird* condition. (ii) For the *rain* condition, the mean perceived safety rating for both the *middle* ($M = 6.96$, $SD = 2.12$) and the *far* conditions ($M = 7.30$, $SD = 2.49$) was significantly higher than the *near* ($M = 5.35$, $SD = 2.62$) condition. The effect size was $\eta_p^2 = 0.313$. The distances played an important role in the perception of safety for all three sound conditions.

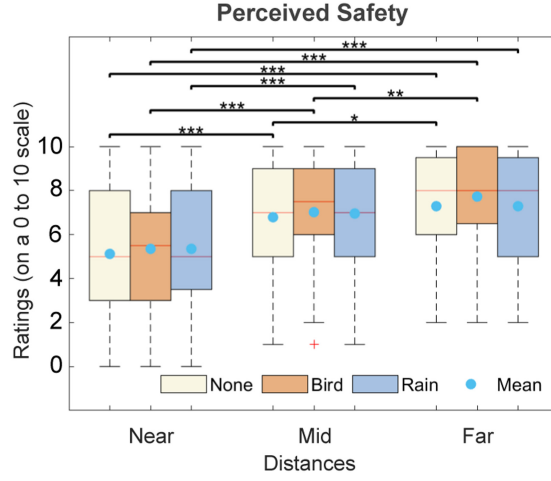


Fig. 7. Box-and-whisker plots of the perceived safety rating at three locations with three sound conditions, where * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$.

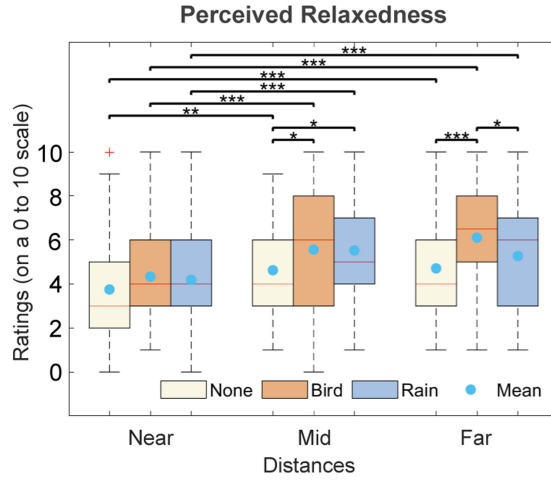


Fig. 8. Box-and-whisker plots of the perceived relaxedness ratings at the three locations with three sound conditions, where * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$.

4.6 Perceived Relaxedness

Figure 8 shows the ratings of relaxedness at the three distances with three sound conditions. The main effect of sounds ($F(2,94) = 5.33$, $p = .006$, $\eta_p^2 = 0.102$) and distances on perceived relaxedness ($F(2,94) = 27.77$, $p < .001$, $\eta_p^2 = 0.371$) were both significant. The interaction between sounds and distances was not significant, ($F(4,188) = 2.05$, $p = 0.09$, $\eta_p^2 = 0.042$), but simple effects analyses indicated the possibility of an interaction.

The simple effects analyses indicated for the three distances: (i) The mean perceived relaxedness ratings between sound conditions did not significantly differ at the *near* location. (ii) At the *middle* location, the mean perceived relaxedness ratings in both the *bird* ($M = 5.56$, $SD = 2.68$) and the *rain* ($M = 5.52$, $SD = 2.26$) conditions were significantly higher than the *none* condition ($M = 4.63$, $SD = 2.10$), with no significant difference between the *bird* and the *rain* conditions. The effect size was $\eta_p^2 = 0.075$. (iii) At the *far* location, the mean perceived relaxedness rating for the *bird* condition

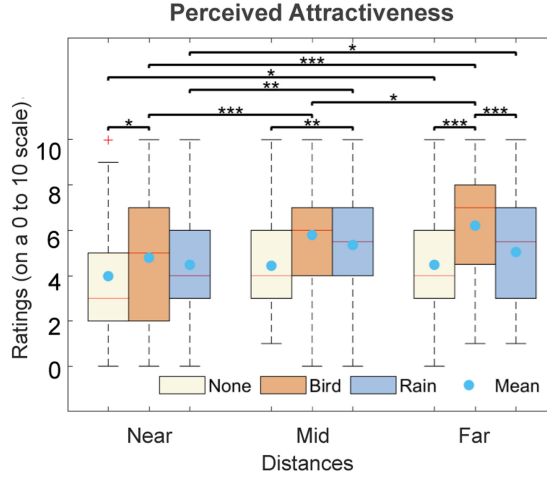


Fig. 9. Box-and-whisker plots of the perceived attractiveness ratings at the three locations with three sound conditions, where * indicates $p < .05$, ** indicates $p < .01$, and *** indicates $p < .001$.

($M = 6.10$, $SD = 2.35$) was significantly higher than both the *none* ($M = 4.71$, $SD = 2.19$) and the *rain* conditions ($M = 5.27$, $SD = 2.52$), with no significant difference between the *none* and the *rain* conditions. The effect size was $\eta_p^2 = 0.251$.

For the three sound conditions, the *none*, the *bird*, and the *rain* conditions: The mean perceived relaxedness ratings at the *middle* (*none*: $M = 4.62$, $SD = 2.01$; *bird*: $M = 5.56$, $SD = 2.68$; *rain*: $M = 5.52$, $SD = 2.26$) and the *far* (*none*: $M = 4.70$, $SD = 2.19$; *bird*: $M = 6.10$, $SD = 2.35$; *rain*: $M = 5.27$, $SD = 2.52$) locations were significantly higher than at the *near* location (*none*: $M = 3.75$, $SD = 2.12$; *bird*: $M = 4.33$, $SD = 2.31$; *rain*: $M = 4.19$, $SD = 2.27$), with no significant difference between the *far* and the *middle* locations. The effect size was $\eta_p^2 = 0.145$ for the *none* condition, $\eta_p^2 = 0.270$ for the *bird* condition, and $\eta_p^2 = 0.226$ for the *rain* condition.

4.7 Perceived Attractiveness

Figure 9 shows the ratings of perceived attractiveness at the three distances with three sound conditions. The main effects of sounds ($F(2,94) = 10.30$, $p < .001$, $\eta_p^2 = 0.180$) and distances on perceived attractiveness ($F(2,94) = 11.38$, $p < .001$, $\eta_p^2 = 0.195$) were both significant. The interaction between sounds and distances was also significant ($F(4,188) = 3.32$, $p = .012$, $\eta_p^2 = 0.066$).

The simple effects analyses indicated for the three distances: (i) At the *near* location, only the mean perceived attractiveness rating of the *bird* condition ($M = 4.79$, $SD = 2.77$) was significantly higher than the *none* condition ($M = 3.98$, $SD = 2.61$), with no significant difference between the other conditions. The effect size was $\eta_p^2 = 0.054$. (ii) At the *middle* location, the mean ratings of perceived attractiveness in both the *bird* ($M = 5.79$, $SD = 2.43$) and the *rain* conditions ($M = 5.35$, $SD = 2.23$) were significantly higher than the *none* condition ($M = 4.43$, $SD = 2.11$), with no significant difference between the *bird* and *rain* conditions. The effect size was $\eta_p^2 = 0.167$. (iii) At the *far* location, the mean perceived attractiveness rating for the *bird* condition ($M = 6.21$, $SD = 2.44$) was significantly higher than both the *none* ($M = 4.48$, $SD = 2.60$) and the *rain* conditions ($M = 5.04$, $SD = 2.34$), with no significant difference between the *none* and the *rain* conditions. The effect size was $\eta_p^2 = 0.351$.

For the three sound conditions: (i) For the *none* condition, only the mean perceived attractiveness rating at the *far* location ($M = 4.48$, $SD = 2.60$) was significantly higher than at the *near*

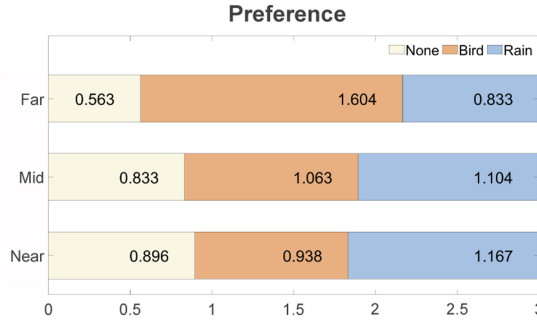


Fig. 10. Stacked bar chart of the participants' mean preference ratings of the three sound conditions at each distance.

location ($M = 3.98$, $SD = 2.61$), with no significant difference between other locations. The effect size was $\eta_p^2 = 0.047$. (ii) For the *bird* condition, the mean perceived attractiveness rating at the *far* location ($M = 6.21$, $SD = 2.44$) was significantly higher than at the *middle* location ($M = 5.79$, $SD = 2.43$), and the *middle* condition was significantly higher than the *near* condition ($M = 4.79$, $SD = 2.77$). The effect size was $\eta_p^2 = 0.237$. (iii) For the *rain* condition, the mean perceived attractiveness ratings at both the *middle* ($M = 5.35$, $SD = 2.23$) and the *far* locations ($M = 5.04$, $SD = 2.34$) were significantly higher than at the *near* location ($M = 4.48$, $SD = 2.63$), with no significant difference between the *far* and the *middle* locations. The effect size was $\eta_p^2 = 0.106$.

4.8 Preference

Figure 10 shows the participants' mean preference ratings among the three sound conditions at each distance. Each participant ranked their preference for the three performances at each distance by giving 2 points to their top favorite, 1 point for the medium favorite, and 0 points for their least favorite.

At both the *near* and the *middle* locations, the *rain* sound was most preferred, followed by the *bird* sound, and the *none* condition was the least preferred. However, the gap between the *rain* and the *bird* sound at the *middle* distance was smaller than at the *near* location. The mean preference rating did not significantly differ between sound conditions at both the *near* and the *middle* locations.

By contrast, at the *far* location, the *bird* condition was ranked the highest, the *rain* condition second, and the *none* condition remained last. Remarkably, the *bird* scored almost double that of the *rain* and triple that of the *none* condition. The main effect of sound on preference was significant at the *far* location ($F(2,94) = 19.39$, $p < .001$, $\eta_p^2 = 0.292$), but not significant at the *near* ($F(2,94) = 1.02$, $p = .364$, $\eta_p^2 = 0.021$) and *middle* ($F(2,94) = 1.02$, $p = .364$, $\eta_p^2 = 0.043$) locations.

For all three locations, adding natural sounds (bird or rain) seemed to have a positive effect on participants' overall preferences; the distance also had an obvious influence on the preference ratings. For instance, ratings of the *bird* condition improved markedly from the *near* to the *far* location, indicating that participants particularly preferred the *bird* sound at the *far*, but not so much at the *middle* or *near* location.

5 QUALITATIVE DATA (INTERVIEWS) ANALYSIS AND RESULTS

The interview responses were analyzed using a thematic analysis, which is a useful and flexible analytic method for identifying themes or patterns from qualitative data in research in and beyond psychology [9]. Data analysis was done in six phases as outlined below, following suggestions by Braun and Clarke [9].

In Phase 1, the qualitative data were analyzed based on the interview notes. We performed three quality checks before taking this decision. First, the first and second authors conducted all interviews together, with detailed interview notes taken by each author separately. Second, the two authors compared their notes before entering the qualitative data into an Excel spreadsheet corresponding to the asked questions to ensure complete and objective data extraction. In the case of disagreement between the two authors' notes, both authors would listen together to the respective audio recording to reach an agreement. Third, the Excel spreadsheet summarizing the notes and audio recordings of the interviews was shared with the third author. The third author transcribed two randomly selected interviews and compared these transcripts to the interview notes, finding that the interview notes adequately captured the interview content (consistency check). In Phase 2, after familiarizing himself with all shared materials, the third author used MaxQDA 2020 [42] to code the qualitative data, while the first author conducted coding via paper-and-pencil. Specifically, we used an inductive coding approach in which we developed coding themes based on the collected data rather than prior theories [9]. During coding, we took care to focus on the explicit, semantic meaning first before moving on to inferred meaning in a later phase [9]. For example, the notes "very artificial, mechanic" (P39), "not real" (P41), and "artificial, metallic" (P50) referring to the rain sounds were coded as "rain sound: artificial." Both authors repeatedly went through their codes and the interview notes as a quality check. However, researchers have an active role in this coding, as coding is never fully independent of interpretation [9]. As a consequence, in Phase 3, the first and third authors collated and discussed their codes to identify initial candidates for overarching themes. Based on these discussions, they performed a refactored analysis of some of the codes in Phase 4, which was reviewed and agreed upon by the second author as an additional quality check. For example, we identified a potential theme that the *sharpness* of the *bird* sound might have created a negative impression but did not sufficiently distinguish between whether participants referred to this sound as "sharp" or "too sharp." As a consequence, we went through the notes again and adapted our codes as appropriate. Afterwards, in Phase 5, the research team created four themes that consolidated important aspects of participants' experiences and thoughts about the experiment. Specifically, we identified the themes of (1) familiarity with the sounds, (2) personal experiences and preferences regarding the sounds, (3) the social dimension of proxemics, and (4) safety associations. These themes were rigorously discussed within the author team, including how they relate to each other and to the sound conditions. Figure 11 provides a visualization of the themes with definitions and sub-themes in the form of an affinity diagram. In Phase 6, we incorporated these identified themes into the present article. We will report on the themes in terms of their impact on the different sound conditions (*bird* and *rain*) in the following section. Finally, we will report findings regarding these themes that were independent of particular sound conditions. All provided quotes were transcribed and some had to be translated into English first (as some interviews were conducted in Mandarin Chinese).

5.1 Themes Identified for the Bird Sounds

Regarding the theme of familiarity, the majority of participants mentioned that the selected bird-song was sharp and implied they had a common understanding of the objective features of the birdsong as a high-frequency sound. However, the theme of familiarity had a strong impact on the participants' attitudes toward the added bird sound, which was closely related to the theme of personal experiences. Since the bird species (great tit) we chose is widespread in Europe, the attitudes of participants who had lived in Europe for at least some years generally tended to be more positive, for they mostly associated the sound with a known bird and nature. P25 stated "I feel relaxed when I hear it. It is a nice melody." P33 said "It sounds very pleasant. The bird reminds me of going to the zoo, like the jungle area in Universeum (a public science center and museum in

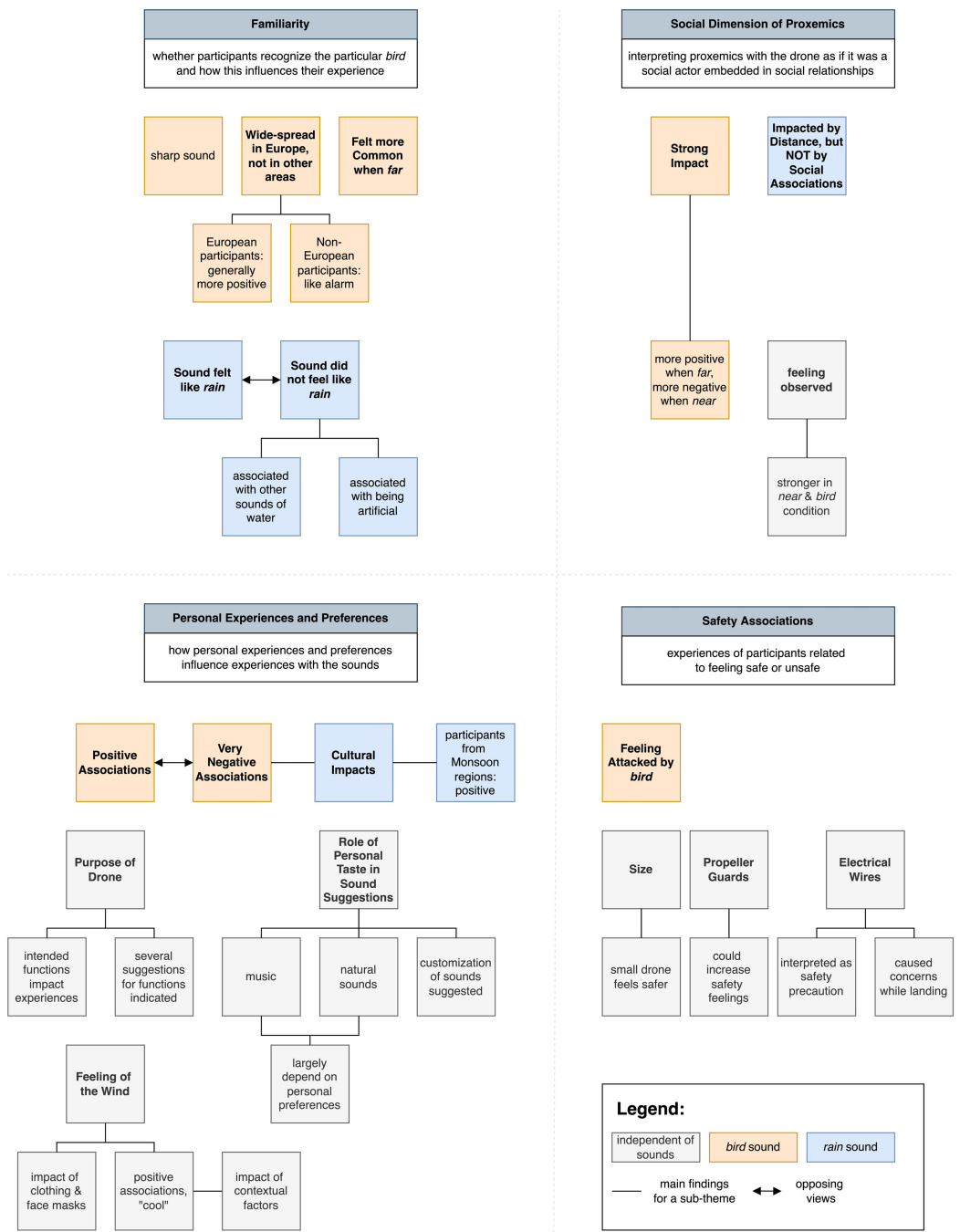


Fig. 11. Affinity diagram with main themes and sub-themes from the qualitative analysis.

Gothenburg).” P31 could even identify the exact bird species from the birdsong. However, among participants newly arrived in Europe (mainly international students), most reported that the sound did not sound like a bird, but rather like an alarm, and for this reason their response tended to be negative. P12 mentioned the sound was sharp and annoying and reminded her of a fire alarm beeping. P18 emphasized “It is a threatening alarm. Feels like it’s coming to attack me!” The attitudes of participants also varied depending on whether they were more outdoors or indoors people, as outdoorsy people were more likely to prefer this birdsong. P32, who identified as an outdoorsy person, stated, “I find it very familiar. I heard this specific bird in spring at our summer house so I recognized it very well! It is really good and comforting.” P34, who mentioned that he was an indoors person, said, “The bird gave me a headache. It was way too loud and high-pitched for me.”

The proxemic distances had a strong impact on the participants’ perceptions of the bird sound, mainly associated with the social dimension of proxemics. Among the nine performances, most participants liked the *bird* condition the most when the drone was at the *far* location, while they disliked the same sound the most at the *near* location—this matches the quantitative data on preference (see Figure 10). Participants reported that at the *far* location, they perceived the *bird* as pleasant, attractive, and comforting, but at the *near* location, they felt it was annoying, stressful, dangerous, and uncomfortable. P04 emphasized “the distances obviously made the *bird* sound very different.” P52 mentioned that the bird’s song had a different effect according to the three distances, “when it was *far*, the *bird* made me feel comfortable and it didn’t sound so sharp, I liked it the most. However, when the same sound was played at *middle* switched from *far*, I felt it became sharper and had an opposite effect which made me feel uncomfortable. I didn’t expect that...When it was *near*, it made me even more uncomfortable.” Participants said it was weird to have a bird so close in a real situation. This finding was closely related to the theme of safety. Some participants mentioned feeling that they were being watched by the drone (see details in the Section 5.3.2), and many felt they might even be attacked by the “bird”: P18: “feels like it is coming to attack me.” In their opinion, it is more common to experience a distant bird rather than a near one, and that might be the reason why they preferred the drone to play the birdsong far away, but not when it is *near*.

Finally, the theme of personal experiences and preferences explained the presence of somewhat extreme opposing attitudes towards the same *bird* condition, especially when the drone was at the *near* location. P23, P34, and P50 seemed to detest and dread the *near-bird* condition. P34 stated: “The *bird* gave me a headache, especially when it was so close to me, I just wanted it to go away! Oh my god! Please! I wanted to kill the bird, almost like ‘get me a rifle.’” P50 said: “when it’s coming so close, I would definitely want to grab it, break the wings and hide it...” In contrast, P31, P32, and P45 were true bird lovers. P31 preferred the *near-bird* condition the most, commenting “I like birds a lot. I’ve seen this bird in real life. It’s a common bird in Europe. I know its name in Swedish, it’s ‘Talgoxe.’” Besides correctly identifying the birdsong as belonging to the talgaxe (great tit), P31 talked about the sound features of the great tit and the change from three syllables to two due to increasing urban noise pollution, which perfectly matches the literature we found (see Section 3.3.2). P32: “I want it (the drone with the birdsong) to come onto my hand,” and P45 stated: “I love it. It made me recall I gave food to wild birds and their babies—they came to the balcony singing.”

5.2 Themes Identified for the Rain Sounds

Regarding the theme of familiarity, participants’ interpretations of the added rain sound fell into two groups. The majority of participants belonged to the first group and claimed the sound was like rain. The second group of around one-third of the participants claimed that it was not like rain, but could be further split into two subgroups. The first subgroup, consisting of P1, P12, P20, P23, P27, P29, P42, P44, P53, associated it with water other than rain, e.g., water leakage, a water

splash, or a waterfall. The other subgroup mentioned that it felt artificial due to the fact that it was together with the drone noise. Participants seemed to have more negative feelings if they associated the sound with something artificial or something wrong (e.g., water leaking). P38 stated “It irritates me. It sounds very artificial, mechanic, and annoying. It is not natural at all.” However, participants tended to feel more positively if they correctly associated the sound with rain. In the case of rain sounds, the theme of familiarity was also closely associated with the theme of personal experiences, as participants from South Asian cultures tended to identify the sound correctly as heavy rain.

The proxemic distances of the rain sounds impacted participants’ perceptions, but in sharp contrast to the bird sound, those perceptions were unaffected by the theme of the social dimension of proxemics. In the case of rain, the impact was related to the blending of the rain sound with the drone noise. Many participants reported that under the *far-rain* condition, they could not distinguish the added rain sound, as it blended into the drone noise, but they could distinguish it at the *near* or *middle* location. P01 stated, “I couldn’t distinguish the rain sound when at *far*, but at *near* it was ok.” P52 mentioned “when *far*, it didn’t sound like rain when it was together with the machine noise—it was not clear. But the closer, the clearer, the more it sounded like rain.” The social dimension of proxemics entirely disappeared in the rain setting, as no sounds from living or artificial social beings were involved.

In the case of rain sounds, the theme of personal experience was mainly associated with cultural factors. We noticed that the passionate rain lovers P45, P47, and P48, who all came from the monsoon region in South Asia, highly rated the *rain* condition. They were able to correctly identify the added sound as a heavy rain even though it was together with drone noise, as P47 said: “The rain sound was suppressed by the machine sound, but I still heard the sound of water, the rain was quite heavy.” And P47 continued: “Especially when at *far*, this rain sound felt special—it reminds me of my home where it has a lot of rain! My school usually reopened during the rainy season, this is exactly the same sound I used to hear when I was a kid sitting in a classroom with heavy rain outside. I could relate to it. It’s a nostalgic feeling.” P45: “This sound is like heavy rain, it reminds me of the rainy season in my home country. I love rain. I have a name that means ‘rain’ in my native language...I recorded the sounds of rain myself...I took shower in the rain—in my country, the raindrops are very big so you can take a shower with them.” P45, P48 associated the rain sound with other sensory experiences as well. P48: “I like rain in my country, it is warm rain [tactile] with the smell of soil [olfactory].” P45: “After raining, it became green [vision], fresh [olfactory], and cool [tactile].”

We did not identify any associations relating to the theme of safety in the case of the *rain* condition, indicating that the sound might be interpreted as a more neutral alternative.

5.3 Themes Independent of Sound Conditions

Several comments by the participants covered the identified themes but did not apply to any particular sound condition. Several participants commented on personal experiences and perceptions independent of particular sound conditions, such as the purpose of the drone, individual suggestions for alternative sounds, and the feeling of the wind. Furthermore, the drone was sometimes experienced as an invasion of privacy, associated with the social dimension of proxemics. The most pronounced theme, however, was safety. Several comments about safety ranged across the different sound conditions, indicating that this is a general concern when interacting with drones.

5.3.1 The Theme of Personal Experiences and Perceptions. Regarding the theme of personal experiences and perceptions, we identified three sub-themes across all sound conditions: (1) the need to discuss the purposes of the drones, (2) individual preferences for alternative sounds, and (3) the experiences of the airflow generated by the drones.

Purposes of drones. Although it was not the intention of this study, some participants mentioned that the purpose of the drone plays a vital role in defining their experiences. The intended function of the domestic drone was neither specified nor discussed. During the interviews, participants P05, P08, P22, P31, P32, P37, and P50 mentioned their considerations or doubts about the intended functionalities of the domestic flying robot. P05 and P22 both asked, “What is it used for?” P8 and P37 both said the choice of sounds to add should depend on the use cases—P37: “if it’s delivering me a drink at a party, it will be very different than if I’m reading a book.” P31 pointed out: “It is more annoying if you don’t know the purpose,” and further explained this with reference to her previous experience encountering a commodity drone: “My neighbor was flying a drone...I first felt annoyed, but later felt better when knowing it’s for advertising (the neighbor wanted to sell his house and used the drone to take photos of the property to showcase).” P50 also said he would have a better feeling and give higher ratings to the drone if he knew the drone was coming to help and accompany him. P32 suggested “this small drone could be a little helper for fun and companionship.” The first and second authors both recalled many other participants casually asking about the intended function/purpose/usage of the small flying robot after the interview during chit-chat (not audio-recorded).

Desired sound depends on personal taste. The participants suggested other sounds they thought might be suitable for adding to the domestic drone besides the bird and rain sounds we used. The most common choices were either music or some other type of natural sound. However, these common choices still varied. For instance, the choice of music ranged from classical Beethoven, country music, or festival music (e.g., Christmas music) to rock and roll, with the suggestions directly related to personal taste—as P44 said: “I like rock music, I want to add rock.” The choices of other natural sounds included ocean waves, a campfire, thunderstorms, and so on. In particular, P22 and P44 asked for customized sounds—P44: “Users should be able to choose which sound and which mode.” In addition, some participants’ choices were special or more personal. P09, a lover of tea culture, wanted to add sounds from the tea ceremony like tea cups being set on the table and the sound of pouring tea into cups. P10 suggested “broadcast, verbal sounds; those containing meaningful messages to bring values.” P30 associated the humming drone to a mosquito and wanted the sound of a croaking frog to prey on the insect. P25 had grown up in Gothenburg and wanted to add Gothenburg-related sounds—he suggested the sounds of strong wind or traffic in the city (Gothenburg is a coastal city that has the largest port in the Nordic countries, with busy traffic and strong winds).

The feeling of the wind. Participants commented on how they experienced the airflow generated by the drone. Nearly all participants agreed that the propellers generated airflow, and it felt the strongest and most obvious at the *near*, weaker at the *middle*, and barely felt at the *far* location. The only exception was three people who were fully covered with thick clothes and face masks and claimed they did not feel airflow. However, how participants perceived the wind was strongly associated with their personal experiences and associations. Half of the participants felt the airflow was a cool and refreshing breeze that made them comfortable and relaxed and thus had a general positive feeling towards it. P28 said, “It reminds me of a summer breeze.” P38 mentioned that the positive feeling from the airflow was even better with the rain sound, “It’s nice and soft, especially with the rain sound, reminds me of soft rain outside on warmer days, gives a cozy feeling.” P31 commented: “The airflow was nice. It made me feel a bit more connected to the drone.”

However, these impressions might depend on contextual factors. P42 mentioned that “it feels pleasant now when it is warm, but might be annoying when it’s cold.” One-fifth of participants were negative about the airflow, as they felt it was cold, uncomfortable, and dangerous, particularly at the *near* position—both P18 and P21 mentioned the wind amplified the presence of the robot and increased fear. P23 explained the negative feeling as arising because “the airflow was surprisingly

strong for such a small robot, much more than I had expected. Very annoying.” However, P23 was more positive towards this surprise: “the airflow was surprisingly strong, and it triggered curiosity.” The rest of the participants felt neutral about the wind. P33, P48, and P54 mentioned that they noticed the airflow also had some visual impact. P33 “visually can see the blanket is moving,” and P54 “even saw the paper was shaking.”

5.3.2 Theme of the Social Dimension of Proxemics. Regarding the social dimensions of proxemics, even though all participants were clearly informed in advance that there was no camera in the experimental environment or on the drone, P08, P30, P37, P38, P50, P55 still mentioned that they felt they were being observed by the drone, especially at the *near* location. This finding was strongly related to the theme of safety and was associated with experiencing the drone as an invasion of privacy. The feeling of being watched by the drone gave participants negative impressions. P30 stated: “I know it didn’t have a camera. But when it’s very close to me, very stable (hovering), it felt like it was staring at me intensively, I didn’t really feel very safe then.” P37 said: “It felt like I was being observed at *near*, just its posture, the way it looks, makes it feel like it’s watching me. I felt it was very invasive.” However, P38 commented that this feeling only arose with the *bird* condition: “with *rain* or *none*, the drone was not so much like a living thing, but with the bird sound, together with the ‘silver thing’ (electronics on the drone), it’s more like an animal—the ‘silver thing’ is like a face, I felt something was looking at me. I didn’t feel safe anymore, I felt in more danger—it was like a mechanical bird.”

5.3.3 Theme of Safety. Regarding the theme of safety, important sub-themes include (1) the small size of the drone, (2) the addition of propeller guards, and (3) the unexpected finding that the electrical wire was perceived as a safety measure.

Small size made the drone feel safer. Many participants acknowledged that the flying robot was small, with some of them pointing out the small size as an advantage, especially in relation to the theme of safety. P03, P08, P18, P31, P32, and P49 mentioned the small size of the drone as a good size that made them feel safer. P31 and P32 thought it was cute at such a small size; as P31 said, “It’s small, it’s cute, it’s like a small animal.”

Adding propeller guards could increase safety. Some participants emphasized that additional propeller guards might be needed to raise feelings of safety. P03 and P49 commented that even though the size is already small, adding a protection frame to each propeller would feel safer. P49: “The size (small) is good, it makes me feel safe. It looks good while flying. However, adding safety protection parts (propeller guards) will be even safer for both human and robot.” Besides P03 and P49, P10, P13, and P48 also suggested adding propeller guards.

Electrical wire was interpreted as a safety precaution. The function of the one-meter-long electrical wires was to connect the Bluetooth board to the loudspeaker, transmitting power and signals. This setting aimed to reduce the drone’s takeoff weight by leaving the Bluetooth board and battery under the desk. It was a temporary solution for prototyping. However, unexpectedly, many participants thought these wires were a safety precaution to restrict the flying area in case the drone got out of control, and it made them feel safer during the experiment. However, some participants worried that the wires would hit the propellers during landing—we noticed that these participants usually had engineering and technical backgrounds.

6 DISCUSSION

Based on the previous two sections, namely, Sections 4 and 5, we discuss here the quantitative and qualitative findings, respectively. It is noteworthy that our quantitative and qualitative analysis results are in line with each other.

6.1 Quantitative Analysis Findings

The statistical analysis of the quantitative data demonstrated a collective pattern among participants regarding their experience encountering a noisy domestic drone with the three sound conditions (*bird*, *rain*, *none*) at three proxemic distances (*far*, *middle*, *near*). The two measurements of the perceived sound characteristics, namely, *loudness* and *sharpness*, met our expectations—natural sound conditions were perceived as louder, and the *bird* in particular was recognized as sharper; the closer the distance, the louder and sharper the sound were perceived. We had hypothesized that it would be beneficial to add natural sounds in terms of people's perceptions of the drone, and we did find support for this. The natural sounds we added, namely, the *bird* and the *rain*, both significantly increased the participants' ratings for the measurements of *pleasantness*, *relaxedness*, and *attractiveness*, but had no effect on perceived safety. Meanwhile, the proxemic distances had significant effects on all measurements, which was in line with our hypothesis that the closer distance would have a more negative effect on the perceptions. Although we had not explicitly considered an assumption for the interaction effects between both factors, in fact, the interaction effects between sound and distance were significant ($p < .05$) for *loudness*, *pleasantness*, and *attractiveness*. The participants significantly preferred the *bird* at the *far* location, but not at the *middle* or *near* location, implying that they started to dislike the *bird* when the distance got close.

6.2 Qualitative Analysis Findings

The qualitative data enabled us to look further into each individual participant's reasons for their feelings and thoughts about the experiment. The findings generalized into four themes: (1) familiarity with the sounds, (2) personal experiences and preferences with the sounds, (3) the social dimension of proxemics, and (4) safety associations. Regarding the *bird* and *rain* sound conditions, we found that participants' sensations were similar with regard to the objective features of the sounds—e.g., no matter whether they liked or disliked the *bird*, they shared the same view that the sound was relatively sharp; regardless of whether they considered the *rain* to be natural or not, they recognized it as somewhat white noise with a water sound, which blended into the drone noise at the *far* location but was more distinct at the *near* location. However, this common understanding of objective features funneled into various subjective associations and interpretations—e.g., some participants claimed the *bird* sounded not like a bird but like an alarm, while others experienced the opposite. This finding emphasized the role that personal experiences and preferences play in the interpretation of sound. Even participants who associated the sound with a bird's song still ranged from interpreting it as a dangerous "bird" that might attack them (especially in the *near* condition, associated with the social dimension of proxemics) to seeing it as a cute and friendly animal. The participants' associations and interpretations of the sound affected their attitude and experience of the flying robot and were dependent on their previous personal experiences, perceptions, and cultural backgrounds. This empirical finding is supported by neurobiology—the number of neurons in the primary auditory cortex, namely, those dedicated to figuring out what sound information means, greatly outnumber those that transform sound into electrical neural signals, resulting in what humans expect to hear plays a great role in what they indeed do hear [41]. In other words, how humans perceive sounds is not purely dependent on actual sonifications, but also and more importantly on their previous experience.

We found further support for this interpretation in the findings for themes across all sound conditions. For instance, participants acknowledged the presence of the drone airflow and reached a consensus that the closer the drone, the more obvious the airflow felt. Then, some associated the airflow with natural weather patterns, while others associated it as an amplification of the perceived presence of the robot. When associated with natural weather, it could be further interpreted as either a comfortable cool breeze or an uncomfortable cold wind. Similarly, the association with

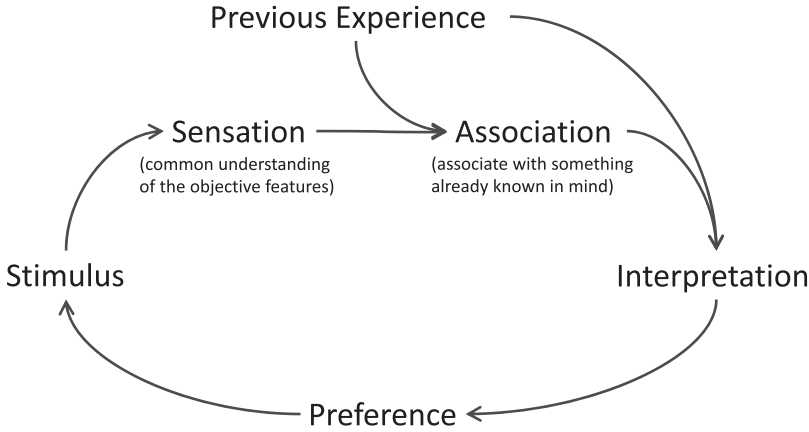


Fig. 12. A visual summary of the relations between stimulus, sensation, association with previous experience, interpretation, and preference.

robot presence amplification could be interpreted as negative feelings of danger or positive feelings of being more connected to the drone. These personal interpretations shaped participants' judgments of the airflow and by extension, the flying robot and the whole experiment. Across all sound conditions, we found support for the vital role that safety perceptions play in HRI, as numerous statements were categorized into the theme of safety.

The thematic analysis of the interview data provided additional information of the participants' qualitative experience beyond the quantitative measurements, which allowed us to extract patterns within participants' perceptions of a noisy flying robot with natural sounds. From the aforementioned themes, we found a generalized pattern within participants' various qualitative experiences of their encounters with the flying robot with added natural sounds in the experiment. This pattern has been consolidated into a visual summary that setting multiple sequential steps in relation with one another (see Figure 12). This visual summary illustrates why an identical stimulus might lead to diverse and even opposite individual judgments. Our visual summary resonates with models from the existing literature in the fields of user experience [28, 33] and neurobiology [41]. Figure 12 shows a sketch of how sounds and other types of stimuli in our experiment are perceived by the participants through a process of sensation and perception, illustrating that perception is a phenomenon of sensemaking. A given stimulus may be first perceived as a physical sensation with certain physical characteristics [54]. This step depends on the functioning of participants' sensory apparatus and usually converges into some common understanding of the objective features of the stimulus (except among people with sensory impairments). This sensation may trigger participants to associate the stimulus with something that they are already familiar with. Then, the association develops into an emotional interpretation. Both of these steps, association and interpretation, highly depend on individuals' previous experience [59], which further determines personal preferences—participants diverge into viewing the same stimulus either favorably or less favorably.

7 DESIGN RECOMMENDATIONS

7.1 Consider Proxemics for Sound Design in HRI

Our study found support for the inclusion of proxemics in the investigation of sound design in HRI. The quantitative results demonstrated statistically significant interaction effects between the studied sound and distance conditions on reported perceptions. Regarding the qualitative results, several participants mentioned that they perceived the drone as an invasion of privacy that was

surveilling them, especially when it was close (despite knowing that there was no camera). These perceptions of a drone as a potential threat were even more pronounced when bird sounds were added. Although participants generally accepted bird sounds when the drone was far away, their perceptions were more diverse when the drone was *near*. Consistent with this interpretation, the rain sounds were not associated with an active social being approaching the participant. It is therefore imperative to consider proxemics for sound design in domestic flying robots. This design recommendation could conceivably apply to sound design in close-range HRI in general.

7.2 Allow Customized Sound

When examining the participants' feelings on the sounds, it is quite interesting to find that the description of natural sound can be contradictory. Participants perceived it as, variously, "annoying" or "soothing," "artificial" or "natural," "noisy" or "relaxing." In general, the added natural sound improved people's feelings about the drone noise, but also introduced the topic of the listener's personal taste [57]. A variety of responses from participants about their own preferred added sounds also revealed individual diversity. We believe that allowing customers to choose their own added sound would be a good design improvement, since it can take personal preference into consideration rather than trying to satisfy everyone with a single common denominator.

7.3 Consider Functionality and Use Cases

Some participants mentioned during interviews that they would be keen to know the intended function or purpose of the small drone, which was not specified in this study. Knowing the intended function of the drone may significantly affect users' perceptions and experience. Other participants informally asked about this point after the interviews. The empirical data thus indicated the necessity of clarifying the purpose of usage. Functionality and use cases need to be considered when designing domestic flying robot interactions.

7.4 Explore the Airflow

The presence of airflow was reported as particularly salient when the drone was flying close to participants, and airflow evoked both positive and negative feelings in participants. Thus, airflow is influential for user acceptance of domestic flying robots, especially in terms of close-range interactions. Coupling the airflow with certain conditions (e.g., sounds, ambient temperatures, vision, smell) may trigger certain effects, and a drone's airflow could be applied to perform some useful functions—for instance, bringing cool breeze on a hot summer day. In general, airflow is a design space that needs exploring.

7.5 Keep Safe Distance

Although the wires under the drone were only used as a temporary solution to attach the loudspeaker and lower the takeoff weight, some participants believed that they worked like a leash to restrict the drone's flying zone, which improved perceived safety. However, in practice, the electric wires may easily become entwined with the propellers and result in danger [72]. Nevertheless, the demand to restrict domestic drones' room to maneuver remains considerable. Inspired by wall barriers for robot vacuums [69], an auxiliary device with an electric fence function that can keep domestic drones a safe distance away from people may be useful, or the flying robot should be capable of autonomously avoiding collisions with humans.

7.6 Add Propeller Guard

Adding a propeller guard or protector was suggested by several participants as an additional safety measure, since the drone flew very close to them during the test. Safety considerations were a

frequent theme in the qualitative analysis across all sound conditions. We also found that the propellers were the most vulnerable parts in a few crashes during commissioning. Thus, adding a propeller guard to a domestic flying robot would better protect both the human user and the robot itself.

8 LIMITATIONS AND FUTURE WORK

Even though the birdsong had a positive influence on participants' perception of the drone and the sound selected was from a common bird, most participants (regardless of whether they recognized it as a bird or not) reported that the sound was too sharp. Some participants suggested choosing another bird species with a lower pitch to achieve better results. Some also mentioned it would be more pleasant to hear a bunch of birds singing in the treetops rather than a single bird singing right in front of you, as a single bird singing such a short distance away from a human does not seem very natural. We conjecture that by making this change, the positive effect of the *bird* condition may become more significant. The fidelity of the loudspeaker was also questioned. Some participants complained about the poor sound quality of our small light loudspeaker. A better loudspeaker with higher fidelity could improve the positive effect of our added sounds. Another issue was the drone's own noise when flying. The same drone sounded different after several takeoffs and landings due to the reduction of the joint gap between the propellers and motors. We tuned the gap to ensure that the noise of the drone would not change a lot during the course of the experiment. However, in actual applications, variations in the noise of the drone may affect participants' evaluations.

Regarding the studied locations, the *near* location in particular needs stricter control. Individual differences in height and body shape obviously introduced additional distance errors between participants and the drone. This error was slight with respect to the distance between the *middle* and the *far* locations relative to the participants' seats, but it became non-negligible regarding the distance from the participant to the *near* location. Even though the experiments were held in a noise-controlled lab, some variables were not controllable, e.g., weather and temperature. Participants mentioned they would be more affected by the *rain* condition on a day with heavy rain. The temperature would also influence the clothing worn and affect the perception of the wind generated by the drone. The design space of domestic drone airflow should be further explored.

Although the functionality of the domestic flying robots was not within the scope of this study, concerns were raised by many participants. This implied that the ambiguity of intended functions might have caused some confusion for the participants and thus to some degree might have biased their judgments. Based on the variety of empirical data we gathered and a thorough consideration of research on close-range human-drone interaction, future work should explore and identify the potential functionalities and usage scenarios of small domestic flying robots. In daily practical usage, the trajectory of a flying robot will be more complex than what we showed in the experiment. How will a complicated flying trajectory that conveys gestural information and fluctuating noise influence people's perceptions of domestic drones? Will the added natural sounds still work in this scenario? Does the added sound need to be updated in real-time in response to the robot's movements? A more comprehensive strategy for adding sound must still be explored.

9 CONCLUSIONS

To the best of our knowledge, this article is the first HRI study to explore the effects of the added natural sound on human participants' reported perceptions of a domestic drone's *loudness*, *sharpness*, *pleasantness*, *safety*, *relaxedness*, and *attractiveness*, while also considering the influence of different proxemic distances. Participants ($N = 56$) were offered a full sensory experience with high realism by being exposed to a real flying robot in a realistic and controlled environment. The quantitative analysis results showed statistically significant evidence that, with the exception of

perceived safety, the two added natural sounds positively influenced participants' perception of the chosen flying robot. Furthermore, the proxemic distance affected the reported perceptions as well. The interaction effects between sound and distance were significant with respect to pleasantness and attractiveness. Moreover, in a qualitative study of post-experiment interviews, we systematically examined the reported thoughts and feelings of all participants. We found that their attitudes toward certain elements of the experimental conditions could be highly affected by their backgrounds and previous experience. Our quantitative and qualitative analysis results support and supplement each other. As a precursor to promoting people's acceptance of close-range interaction with domestic flying robots, our work illustrates the potential of adding natural sounds to improve people's perception of domestic drones. Thus, the combination of sound and proxemic distances should be considered during robot design and development. With further study, this approach may potentially provide an ingenious and effective way of enabling robots like domestic drones to more intimately assimilate into people's daily lives.

ACKNOWLEDGMENTS

We thank all anonymous reviewers and editors for their efforts and valuable inputs. We thank Yiqian Wu for her assistance during the revision process. We acknowledge the Wallenberg AI, Autonomous Systems and Software Program – Humanities and Society (WASP-HS).

REFERENCES

- [1] Parastoo Abtahi, David Y. Zhao, Jane L. E., and James A. Landay. 2017. Drone near me: Exploring touch-based human-drone interaction. *Proc. ACM Interact. Mob. Wear. Ubiq. Technol.* 1, 3 (Sep. 2017). DOI: <https://doi.org/10.1145/3130899>
- [2] Adobe. 2022. Download Sound Effects—Weather (396 MB). (2022). Retrieved from <https://download.adobe.com/pub/adobe/audition/offers/adobeauditiondlcsfx/weather.zip>.
- [3] Östen Axelsson, Mats E. Nilsson, Björn Hellström, and Peter Lundén. 2014. A field experiment on the impact of sounds from a jet-and-basin fountain on soundscape quality in an urban park. *Landsc. Urb. Plan.* 123 (2014), 49–60.
- [4] Christopher C. Berger and H. Henrik Ehrsson. 2013. Mental imagery changes multisensory perception. *Curr. Biol.* 23, 14 (2013), 1367–1372.
- [5] Cindy L. Bethel and Robin R. Murphy. 2006. Auditory and other non-verbal expressions of affect for robots. In *AAAI Fall Symposium: Aurally Informed Performance*. 1–5.
- [6] BetterUp. 2022. You have more than 5 Human Senses. How Are You Using Yours?. Retrieved from <https://www.betterup.com/blog/human-senses>.
- [7] Jens Blauert and Ute Jekosch. 1997. Sound-quality evaluation—A multi-layered problem. *Acta Acust. United Acust.* 83, 5 (Sept./Oct. 1997), 747–753(7). Retrieved from https://www.ingentaconnect.com/content/dav/aaui/1997/00000083/00000005/art00005?utm_source=TrendMD&utm_medium=cpc&utm_campaign=Acta_Acustica_united_with_Acustica_TrendMD_0.
- [8] Gregory N. Bratman, Hector A. Olvera-Alvarez, and James J. Gross. 2021. The affective benefits of nature exposure. *Social and Personality Psychology Compass* 15, 8 (2021), e12630. <https://doi.org/10.1111/spc3.12630>
- [9] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitat. Res. Psychol.* 3, 2 (2006), 77–101.
- [10] Rachel T. Buxton, Amber L. Pearson, Claudia Allou, Kurt Frstrup, and George Wittemyer. 2021. A synthesis of health benefits of natural sounds and their distribution in national parks. *Proc. Nat. Acad. Sci.* 118, 14 (2021), e2013097118. DOI: <https://doi.org/10.1073/pnas.2013097118>
- [11] Jessica R. Cauchard, Jane L. E., Kevin Y. Zhai, and James A. Landay. 2015. Drone & me: An exploration into natural human-drone interaction. In *ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'15)*. Association for Computing Machinery, New York, NY, 361–365. DOI: <https://doi.org/10.1145/2750858.2805823>
- [12] Victoria Chang, Pramod Chundury, and Marshini Chetty. 2017. Spiders in the sky: User perceptions of drones, privacy, and security. In *CHI Conference on Human Factors in Computing Systems (CHI'17)*. Association for Computing Machinery, New York, NY, 6765–6776. DOI: <https://doi.org/10.1145/3025453.3025632>
- [13] Bert De Coensel, Sofie Vanwetswinkel, and Dick Botteldooren. 2011. Effects of natural sounds on the perception of road traffic noise. *J. Acoust. Societ. Amer.* 129, 4 (2011), EL148–EL153.

- [14] Ashley Colley, Lasse Virtanen, Pascal Knierim, and Jonna Häkkinä. 2017. Investigating drone motion as pedestrian guidance. In *16th International Conference on Mobile and Ubiquitous Multimedia (MUM'17)*. Association for Computing Machinery, New York, NY, 143–150. DOI : <https://doi.org/10.1145/3152832.3152837>
- [15] Crazyflie. 2022. Bitcraze. Retrieved from <https://www.bitcraze.io/about/bitcraze>.
- [16] Customer reviews. 2022. DJI Mavic Pro Platinum 8331 Low-Noise Quick-Release Propellers - Gold Tips - 2 Pairs. Retrieved from https://www.amazon.com/-/zh/product-reviews/B076K8CBH9/ref=cm_cr_dp_d_show_all_btm?ie=UTF8&reviewerType=all_reviews&language=en_US¤cy=USD.
- [17] Saraswati Dawadi, Sagun Shrestha, and Ram A. Giri. 2021. Mixed-methods research: A discussion on its types, challenges, and criticisms. *Journal of Practical Studies in Education* 2, 2 (2021), 25–36. <https://doi.org/10.46809/jpse.v2i2.20>
- [18] Rob Diaz-Marino and Miaosen Wang. 2011. Proxemic interactions: The new Ubicomp? *Interactions* XVIII.1, 4 (2011), 42–50. DOI : <https://doi.org/10.1145/1897239.1897250>
- [19] Brittany A. Duncan and Robin R. Murphy. 2013. Comfortable approach distance with small unmanned aerial vehicles. In *IEEE International Symposium on Robot and Human Interactive Communication*. 786–792. DOI : <https://doi.org/10.1109/ROMAN.2013.6628409>
- [20] Marcel Eens, Rianne Pinxten, Rudolf F. Verheyen, Ronny Blust, and Lieven Bervoets. 1999. Great and blue tits as indicators of heavy metal contamination in terrestrial ecosystems. *Ecotoxic. Environ. Safet.* 44, 1 (1999), 81–85.
- [21] Etikprovning myndigheten (Ethics Review Authority). 2022. Vanliga Frågor (Common Questions about Ethical Review). (2022). Retrieved from <https://etikprovning.myndigheten.se/vanliga-fragor/>.
- [22] Hugo Fastl and Eberhard Zwicker. 2007. *Psychoacoustics: Facts and Models* (3rd. ed.). Springer Berlin. DOI : <https://doi.org/10.1007/978-3-540-68888-4>
- [23] Emma Frid, Roberto Bresin, and Simon Alexanderson. 2018. Perception of mechanical sounds inherent to expressive gestures of a NAO robot-implications for movement sonification of humanoids. In *Proceedings of the Sound and Music Computing Conference (SMC)*, 53–59. <https://doi.org/10.5281/zenodo.1422499>
- [24] Laurent Galbrun and Tahir T. Ali. 2013. Acoustical and perceptual assessment of water sounds and their use over road traffic noise. *J. Acoust. Societ. Amer.* 133, 1 (2013), 227–237.
- [25] Wouter Halfwerk and Hans Slabbekoorn. 2009. A behavioural mechanism explaining noise-dependent frequency use in urban birdsong. *Anim. Behav.* 78, 6 (2009), 1301–1307.
- [26] Edward T. Hall. 1966. *The Hidden Dimension*. Doubleday, Garden City, NJ.
- [27] Yiyang Hao, Jian Kang, and Heinrich Wörtche. 2016. Assessment of the masking effects of birdsong on the road traffic noise environment. *J. Acoust. Societ. Amer.* 140, 2 (2016), 978–987.
- [28] Marc Hassenzahl. 2003. The thing and I: Understanding the relationship between user and product. In *Funology*. Springer, 31–42.
- [29] IBM. 2022. IBM SPSS Statistics. Retrieved from <https://www.ibm.com/products/spss-statistics>.
- [30] Kaoru Inoue, Kazuyoshi Wada, and Yuko Ito. 2008. Effective application of Paro: Seal type robots for disabled people in according to ideas of occupational therapists. In *International Conference on Computers for Handicapped Persons*. Springer, 1321–1324.
- [31] Betina Ip and Holly Bridge. 2020. Learning to see in 3D with two eyes: The role of experience, plasticity and neurochemistry. *Biochemist* 42, 5 (09 2020), 12–17. DOI : <https://doi.org/10.1042/BIO20200059>
- [32] ISO 8373:2021(en). 2022. Robotics – Vocabulary. Retrieved from <https://www.iso.org/obp/ui/#iso:std:iso:8373:ed-3:v1:en>.
- [33] Evangelos Karapanos. 2012. *Modeling Users' Experiences with Interactive Systems*. Vol. 436. Springer.
- [34] Joseph La Delfa, Mehmet Aydin Baytas, Rakesh Patibanda, Hazel Ngari, Rohit Ashok Khot, and Florian “Floyd” Mueller. 2020. *Drone Chi: Somaesthetic Human-drone Interaction*. Association for Computing Machinery, New York, NY, 1–13. DOI : <https://doi.org/10.1145/3313831.3376786>
- [35] Lau Langeveld, René van Egmond, Reinier Jansen, and Elif Özcan. 2013. Product sound design: Intentional and consequential sounds. In *Advances in Industrial Design Engineering*. IntechOpen. DOI : <https://doi.org/10.5772/55274>
- [36] Przemysław A. Lasota, Gregory F. Rossano, and Julie A. Shah. 2014. Toward safe close-proximity human-robot interaction with standard industrial robots. In *IEEE International Conference on Automation Science and Engineering (CASE'14)*. 339–344. DOI : <https://doi.org/10.1109/CoASE.2014.6899348>
- [37] Adrian Benigno Latupeirissa and Roberto Bresin. 2020. Understanding non-verbal sound of humanoid robots in films. In *Workshop on Mental Models of Robots at HRI 2020*. Retrieved from <http://kth.diva-portal.org/smash/record.jsf?pid=diva2%3A1454788&dsid=-8534>.
- [38] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group*. Springer, 63–76.
- [39] Timothy R. Levine and Craig R. Hullett. 2002. Eta squared, partial eta squared, and misreporting of effect size in communication research. *Hum. Commun. Res.* 28, 4 (2002), 612–625. DOI : <https://doi.org/10.1111/j.1468-2958.2002.tb00828.x>

- [40] Barry B. Ma, Jeffrey A. Nystuen, and Ren-Chieh Lien. 2005. Prediction of underwater sound levels from rain and wind. *J. Acoust. Societ. Amer.* 117, 6 (2005), 3555–3565.
- [41] Peggy Mason. 2017. *Medical Neurobiology* (2nd. ed.). Oxford University Press. DOI: <https://doi.org/10.1093/med/9780190237493.001.0001>
- [42] MAXQDA. 2022. Introducing MAXQDA 2022. Retrieved from <https://www.maxqda.com/new-maxqda-2022>.
- [43] Dubravko Miljković. 2018. Methods for attenuation of unmanned aerial vehicle noise. In *41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO'18)*. 0914–0919. DOI: <https://doi.org/10.23919/MIPRO.2018.8400169>
- [44] Abdirahman Mohamud and Ashwin Ashok. 2018. Drone noise reduction through audio waveguiding. In *4th ACM Workshop on Micro Aerial Vehicle Networks, Systems, and Applications (DroNet'18)*. Association for Computing Machinery, New York, NY, 92–94. DOI: <https://doi.org/10.1145/3213526.3213543>
- [45] Dylan Moore, Tobias Dahl, Paula Varela, Wendy Ju, Tormod Næs, and Ingunn Berget. 2019. Unintended consonances: Methods to understand robot motor sound perception. In *CHI Conference on Human Factors in Computing Systems (CHI'19)*. Association for Computing Machinery, New York, NY, 1–12. DOI: <https://doi.org/10.1145/3290605.3300730>
- [46] Dylan Moore, Hamish Tennent, Nikolas Martelaro, and Wendy Ju. 2017. Making noise intentional: A study of servo sound perception. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI'17)*. Association for Computing Machinery, New York, NY, 12–21. DOI: <https://doi.org/10.1145/2909824.3020238>
- [47] Michael Narine. 2020. *Active Noise Cancellation of Drone Propeller Noise through Waveform Approximation and Pitch-shifting*. Master's Thesis. Georgia State University, Atlanta, GA. Retrieved from https://scholarworks.gsu.edu/cs_theses/98.
- [48] Eduardo Narváez, Ankit A. Ravankar, Abhijeet Ravankar, Takanori Emaru, and Yukinori Kobayashi. 2021. Autonomous VTOL-UAV docking system for heterogeneous multirobot team. *IEEE Trans. Instrum. Measur.* 70 (2021), 1–18. DOI: <https://doi.org/10.1109/TIM.2020.3039649>
- [49] Mohammad Obaid, Wafa Johal, and Omar Mubin. 2020. Domestic drones: Context of use in research literature. In *8th International Conference on Human-Agent Interaction (HAI'20)*. Association for Computing Machinery, New York, NY, 196–203. DOI: <https://doi.org/10.1145/3406499.3415076>
- [50] Mohammad Obaid, Eduardo B. Sandoval, Jakub Zlotowski, Elena Moltchanova, Christina A. Basedow, and Christoph Bartneck. 2016. Stop! That is close enough. How body postures influence human-robot proximity. In *25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'16)*. 354–361. DOI: <https://doi.org/10.1109/ROMAN.2016.7745155>
- [51] Dinesh K. Pai. 2005. Multisensory interaction: Real and virtual. In *11th International Symposium on Robotics Research*. Springer, 489–498.
- [52] Hannah Pelikan, Frederic Anthony Robinson, Leelo Keevallik, Mari Velonaki, Mathias Broth, and Oliver Bown. 2021. Sound in human-robot interaction. In *ACM/IEEE International Conference on Human-Robot Interaction*. 706–708.
- [53] RELAXOUND. 2022. Treat Yourself with a Relaxing Bird Concert: with the Zwitscherbox. Retrieved from <https://www.relaxound.com/en/zwitscherbox/>.
- [54] Anina N. Rich and Jason B. Mattingley. 2002. Anomalous perception in synaesthesia: A cognitive neuroscience perspective. *Nat. Rev. Neurosci.* 3, 1 (2002), 43–52.
- [55] F. A. Robinson, O. Bown, and M. Velonaki. 2022. Designing sound for social robots: Candidate design principles. *Int. J. Soc. Robot.* 14 (2022), 1507–1525. DOI: <https://doi.org/10.1007/s12369-022-00891-0>
- [56] Frederic Anthony Robinson, Oliver Bown, and Mari Velonaki. 2020. Implicit communication through distributed sound design: Exploring a new modality in human-robot interaction. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI'20)*. Association for Computing Machinery, New York, NY, 597–599. DOI: <https://doi.org/10.1145/3371382.3377431>
- [57] Frederic Anthony Robinson, Mari Velonaki, and Oliver Bown. 2021. Smooth operator: Tuning robot perception through artificial movement sound. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI'21)*. Association for Computing Machinery, New York, NY, 53–62. DOI: <https://doi.org/10.1145/3434073.3444658>
- [58] Vahid Saadatmand, Nahid Rejeh, Majideh Heravi-Karimooi, Sayed Davood Tadrissi, Farid Zayeri, Mojtaba Vaismoradi, and Melanie Jasper. 2013. Effect of nature-based sounds' intervention on agitation, anxiety, and stress in patients under mechanical ventilator support: A randomised controlled trial. *Int. J. Nurs. Stud.* 50, 7 (2013), 895–904.
- [59] Muzafer Sherif. 1935. A study of some social factors in perception. *Archives of Psychology (Columbia University)*. No. 187.
- [60] SVERIGES RIKSDAG (The Swedish Parliament). 2022. Lag (2003:460) om etikprövning av forskning som avser människor (Law (2003: 460) on Ethical Review of Research Concerning People). Retrieved from https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/lag-2003460-om-etikprovning-av-forskning-som_sf-s-2003-460.

- [61] Daniel Szafir, Bilge Mutlu, and Terrence Fong. 2015. Communicating directionality in flying robots. In *10th ACM/IEEE International Conference on Human-Robot Interaction (HRI'15)*. 19–26.
- [62] Marta E. Szczygiel. 2016. From night soil to washlet: The material culture of Japanese toilets. *Electron. J. Contemp. Japan. Stud.* 16, 3 (2016), 12–2016.
- [63] Kazuki Takashima, Takafumi Oyama, Yusuke Asari, Ehud Sharlin, Saul Greenberg, and Yoshifumi Kitamura. 2016. Study and design of a shape-shifting wall display. In *ACM Conference on Designing Interactive Systems*. 796–806.
- [64] Hamish Tennent, Dylan Moore, Malte Jung, and Wendy Ju. 2017. Good vibrations: How consequential sounds affect perception of robotic arms. In *26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'17)*. 928–935. DOI: <https://doi.org/10.1109/ROMAN.2017.8172414>
- [65] The DJI Store. 2022. Mavic 2 Low-Noise Propellers. Retrieved from <https://m.dji.com/se/product/mavic-2-low-noise-propellers>.
- [66] Gabriele Trovato, Martin Do, Ömer Terlemez, Christian Mandery, Hiroyuki Ishii, Nadia Bianchi-Berthouze, Tamim Asfour, and Atsuo Takanishi. 2016. Is hugging a robot weird? Investigating the influence of robot appearance on users' perception of hugging. In *16th International Conference on Humanoid Robots (Humanoids'16)*. IEEE, 318–323.
- [67] Gabriele Trovato, Renato Paredes, Javier Balvin, Francisco Cuellar, Nicolai Bæk Thomsen, Søren Bech, and Zheng-Hua Tan. 2018. The sound or silence: Investigating the influence of robot noise on proxemics. In *27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'18)*. 713–718. DOI: <https://doi.org/10.1109/ROMAN.2018.8525795>
- [68] Timothy Van Renterghem, Kris Vanhecke, Karlo Filipan, Kang Sun, Toon De Pessemier, Bert De Coensel, Wout Joseph, and Dick Botteldooren. 2020. Interactive soundscape augmentation by natural sounds in a noise polluted urban park. *Landsc. Urb. Plan.* 194 (2020), 103705.
- [69] Virtual Wall. 2022. How Do I Use Dual Mode Virtual Wall® Barrier? Retrieved from <https://homesupport.irobot.com/s/article/9053>.
- [70] Michael L. Walters, Mohammedreza A. Oskoei, Dag Sverre Syrdal, and Kerstin Dautenhahn. 2011. A long-term Human-Robot Proxemic study. In *IEEE International Symposium on Robot and Human Interactive Communication*. 137–142. DOI: <https://doi.org/10.1109/ROMAN.2011.6005274>
- [71] M. L. Walters, D. S. Syrdal, K. L. Koay, K. Dautenhahn, and R. te Boekhorst. 2008. Human approach distances to a mechanical-looking robot with different robot voice styles. In *17th IEEE International Symposium on Robot and Human Interactive Communication*. 707–712. DOI: <https://doi.org/10.1109/ROMAN.2008.4600750>
- [72] Ziming Wang, Ziyi Hu, Yemao Man, and Morten Fjeld. 2022. A collaborative system of flying and ground robots with universal physical coupling interface (PCI), and the potential interactive applications. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA'22)*. Association for Computing Machinery, New York, NY. DOI: <https://doi.org/10.1145/3491101.3519766>
- [73] Anna Wojciechowska, Jeremy Frey, Sarit Sass, Roy Shafir, and Jessica R. Cauchard. 2019. Collocated human-drone interaction: Methodology and approach strategy. In *14th ACM/IEEE International Conference on Human-Robot Interaction (HRI'19)*. 172–181. DOI: <https://doi.org/10.1109/HRI.2019.8673127>
- [74] xeno-canto. 2022. Sharing Bird Sounds from Around the World. Retrieved from <https://xeno-canto.org/sounds/uploaded/IIGGEELDDD/XC527981-200215-153900.mp3>.
- [75] YABAI. 2022. Otohime: Using the Sound Princess to Hide Embarrassment. Retrieved from <http://yabai.com/p/4177>.
- [76] Wataru Yamada, Hiroyuki Manabe, and Daizo Ikeda. 2019. ZeRONE: Safety drone with blade-free propulsion. In *CHI Conference on Human Factors in Computing Systems (CHI'19)*. Association for Computing Machinery, New York, NY, 1–8. DOI: <https://doi.org/10.1145/3290605.3300595>
- [77] Alexander Yeh, Photchara Ratsamee, Kiyoshi Kiyokawa, Yuki Uranishi, Tomohiro Mashita, Haruo Takemura, Morten Fjeld, and Mohammad Obaid. 2017. Exploring proxemics for human-drone interaction. In *International Conference on Human-Agent Interaction (HAI'17)*. Association for Computing Machinery, New York, NY, 81–88. DOI: <https://doi.org/10.1145/3125739.3125773>
- [78] Brian J. Zhang, Knut Peterson, Christopher A. Sanchez, and Naomi T. Fitter. 2021. Exploring consequential robot sound: Should we make robots quiet and kawaii-et? In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'21)*. 3056–3062. DOI: <https://doi.org/10.1109/IROS51168.2021.9636365>
- [79] Brian J. Zhang, Nick Stargu, Samuel Brimhall, Lilian Chan, Jason Fick, and Naomi T. Fitter. 2021. Bringing WALL-E out of the silver screen: Understanding how transformative robot sound affects human perception. In *IEEE International Conference on Robotics and Automation (ICRA'21)*. 3801–3807. DOI: <https://doi.org/10.1109/ICRA48506.2021.9562082>
- [80] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. In *SIGCHI Conference on Human Factors in Computing Systems (CHI'07)*. Association for Computing Machinery, New York, NY, 493–502. DOI: <https://doi.org/10.1145/1240624.1240704>

Received 30 April 2022; revised 14 November 2022; accepted 22 December 2022