Robotics: Science and Systems 2024
Delft, Netherlands, July 15-July 19, 2024

The Benefits of Sound Resound: An In-Person Replication of the Ability of Character-Like Robot Sound to Improve Perceived Social Warmth

Nnamdi Nwagwu*, Adeline Schneider*, Ibrahim Syed†, Brian J. Zhang* and Naomi T. Fitter*
*Collaborative Robotics and Intelligent Systems (CoRIS) Institute,
Oregon State University, Corvallis, Oregon, USA
†Music Department, Oregon State University, Corvallis, Oregon, USA

***Abstract*—While robot sound is known to impact perceptions of robots, little research to date has addressed the topic of robot sound. In particular, there are few in-person studies surrounding the topic. To remedy this, we conducted an in-person empirical study with $N = 30$ participants, as a partial replication of an online study in the space. We sought to better understand the effects that character-like and functional sounds have on human teammates' perceptions of a robot during a joint in-person task. Participants rated the robot with character-like sound as more socially warm compared against a no-added-sound condition; this result was akin to insights from our previous work that showed benefits of transformative robot sound for warmth and other factors across multiple robot platforms. Additional evidence newly presented in this work also suggested increased localizability of robots with augmented sonic profiles. The partial replication in the results strengthens past findings on robot sound, especially as related to character-like robot sound's ability to improve perceived robot warmth. This work can help to inform designers and researchers who are interested in enhancing robot interactions via nonverbal robot expression.**

I. Introduction

Currently, there are few robotic platforms that intentionally leverage sound, even though robot sound shows promise as a means for robot communication and augmentation. For example, augmented robot sound profiles have demonstrated potential for displaying emotional and functional information [16, 20, 19], highlighting relative position [3], and improving the perception of robots [22]. In order to expand on related work and encourage more generalized use of robot sound, we looked to replicate the results presented in previous online studies [31, 33] with the addition of further metrics and context informed by a past in-person robot sound study [3]. The work presented in this paper broadly looks to utilize existing methodologies from sound research in robotics to replicate past online and in-person results. Particularly, it revolved around the use of four robot sound types found in our recent review paper [29]: *Consequential sound* is sound generated naturally from the operation of the robot (e.g., motor activity, friction between parts). *Functional sound* is added sound meant explicitly to convey non-emotional informational from the robot. *Transformative sound* is added sound made to change the sonic profile of a robot. *Emotional sound* is added sound explicitly made to convey emotions from the robot. In this work, we combine transformative and emotional sound into a sound type that we call "character-like," as further explained in Section III.

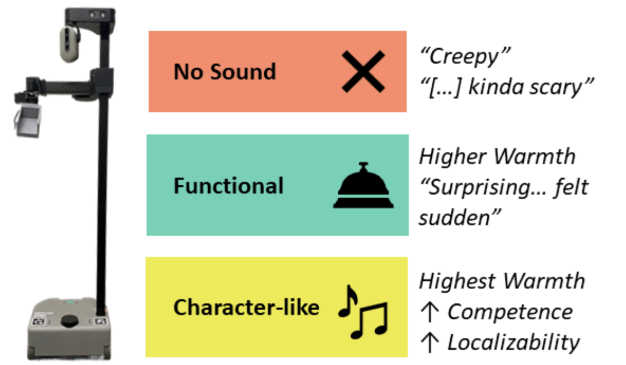

Figure 1: *Left:* The Hello Robot Stretch RE2 used in the study. *Middle:* All studied sound types. *Right:* A summary of key results for each sound condition.

Recent research in robot sound has used design methods to prescribe principles for concrete interactions with an autonomous bus [16] and build design guidelines to support successful collaborations between sound designers and roboticists [34]. Both studies [16, 34] have successfully showcased the benefits of intentionally created robot sound, but are respectively online [34] or apply to a single form factor [16]. The work presented in this paper sought to directly build on past online study-based efforts related to assessing the introduction of transformative sound to robotic sound profiles [31]. We aimed to replicate these past methods in an in-person study, in addition to studying the localizability of robots with added sound (using methods from [3]) and monetary value offered by augmented robot sound profiles (using methods from [33]), as visualized in Fig. 1.

The main goal of this research was to determine whether existing findings stemming from past online robot sound studies can be replicated in in-person work, and to what degree. Our review of related work in Section II provides background on the current state of robot sound research. Based on our previous study results in [31, 33] and other past robot sound explorations in [3], we designed and conducted the replication study described in Section III. The results of participant perceptions of consequential, transformative, and functional robot sound are presented in Section IV. Section V discusses our findings and related design implications for robot sound profiles. Contributions of this work include the extension of robot sound research through replication of existing works and additional perspective regarding the benefits of functional and character-like sounds.

II. Related Work

We reviewed recent human-subjects studies on nonverbal robot sound with a focus on the study methods. Within the past five years, a rise in the number of online, video-based studies has helped enable broad and statistically powerful examinations of the effects of nonverbal sound creation, while in-person experimental studies have continued to advance our understanding of human responses to embodied interactions concerning sound. Studies often focused on one function of sound—consequential, functional, transformative, or emotional—each of which are detailed below. These terms, as well as many of the studies, originate from our review of nonverbal sound in [29].

A. Explorations of Consequential Sound

Consequential robot sound (sound produced as a consequence of the robot's operation, often due to electromechanical components such as motors or interactions with the robot's environment) has seen limited exploration, primarily through online survey-based studies. Tennent et al. compared the sound of a high-end and low-end robot arm overlaid onto a video of the arm's motion, but differences in participants' responses varied based on the motion's social context [25]. Moore et al. delved more finely into comparisons of servo sounds, asking participants to rate and associate sounds with descriptors; however, these ratings did not clearly correlate with objective acoustic properties [12, 13]. Zhang et al., using digitally manipulated consequential sound, found that participants preferred quieter and higher-pitched sounds for robot arms across motions [30], but *consequential sound does not lend itself easily to study, in part due to its similarity to existing robot and robot component sounds.*

In-person work has also sought to identify the role of consequential sound. Trovato et al. found that consequential sound played back through a speaker produced worse subjective ratings compared to no consequential sound [27], while Izui et al. had participants wear earmuffs and earplugs to muffle consequential sound [7]. However, *more detailed explorations and follow-up studies of consequential sound's role remain difficult in in-person experiments.*

B. Explorations of Functional Sound

Functional robot sound (intentionally-produced sound made explicitly to help convey information, such as the robot's state or intent) appears most commonly in the real world, but infrequently in the literature. Some work has delved into the effect of functional sound for particular contexts and purposes. Shrestha et al. used car turning signal sounds in combination with turning signal lights to demonstrate improved subjective metrics for motion legibility [23], while Okimoto et al. used sine wave tones to indicate a robot arm's impending motion and improve participant response time, task completion time, and subjective measures in a collaborative task [14]. While these findings help validate existing trends in nonverbal robot sound work, the field *would benefit from comparisons of functional sound with other sound types, such as character-like sound.*

C. Explorations of Character-Like Sound

Transformative robot sounds (intentionally-produced sound made to complement, blend with, or transform the impression created by a robot's operation and consequential sounds) and emotional robot sounds (intentionally-produced sound made explicitly to help convey emotion, often as part of an affective display) are common features for *robot characters* portrayed in the media. Online survey-based work on transformative sound, the newest area of robot sound investigation, has helped to identify promising transformative sound types; for example, musical designs are recommended over harmonic or mechanical designs [21] and rhythmic is recommended over continuous [5]. Zahray et al. compared and contrasted designs in person, but compared a separate, more discrete sound design set than other works [28]. Currently, *study findings across research groups remain disconnected due to a lack of follow-up in-person experiments.*

Work on emotional sound, the most deeply investigated form of nonverbal robot sound, have employed both online survey-based studies and in-person follow-up studies. Jee et al. were among the first to explore musical motifs to convey emotion through in-person experiments [8], while Read et al. examined the impacts of emotional sounds in various contexts and as parts of multimodal interactions [20, 19, 18]. Studies have primarily focused on the effectiveness of emotional sounds in conveying intended emotions, limiting their usefulness to robots that have no intent to convey emotions. Zhang et al. showed that transformative and emotional sounds (i.e., character-like sound) improved overall ratings of robots in online survey-based studies, particularly social warmth and perceived value [31, 32]; *these findings may be of interest to the broader robotics community, but have yet to be replicated in the real world.*

III. Methods

We conducted an experiment to investigate how different sound profiles inspired by the robot sound literature influenced human perceptions of a robot in a mock collaborative task. All study procedures were approved by our university's Institutional Review Board under protocol #IRB-2020-0592.

A. Study Design

During the study, the human participant worked with a Hello Robot Stretch robot, a robot arm with two prismatic joints and a mobile base, which is intended for personal and service applications.

Figure 2: The study setup. The left image displays the participant workspace and perspective, and the right image shows the robot starting location, as well as a curtain that obscures the view of the researcher.

1) Space Setup: Fig. 2 shows the study room setup. A webcam was located in the human's workspace, along with starting materials for the study task. The participant engaged in a LEGO assembly task while facing away from the robot, and the robot would drop a bag of LEGOs, one in each trial, to support the continued assembly.

2) Central Manipulation: The study conditions involved three different types of robot sound profiles:

- *Consequential (C):* No sound was added
- *Functional (F):* Elevator bell-like sounds played when the robot dropped off a package and as the robot departed
- *Character-like (T+E):* A continuous (transformative sound) melody (taken from [33]) played when the robot's base was moving. Custom emotional sounds played when the robot dropped off a package and as the robot departed.

For the functional sounds, we focused on creating two simple tones to mimic existing technologies that use sound notifications (e.g., elevator chimes, mobile phones). Generally, functional sounds are not focused on conveying emotional information, but rather focus on calling attention in a succinct way. The remaining "custom emotional sounds" leveraged domain knowledge of the authors (roboticists and music technologists) to design appropriate sounds with positive valence. The sounds used are available for reference in the supplemental material, in addition to a video recording of the sounds being used in the context of the study task.

The study was within-subjects, and condition order was randomized for each participant to mitigate ordering effects.

B. Participants

The $N = 30$ participants who completed the study were recruited through Oregon State University email lists and newsletters. Participants were aged 18 to 61 years old ($M = 27$, $SD = 9.7$), with gender demographics consisting of 39% men, 51% women, 6% non-binary persons, and 3% transgender women. 30% of participants had backgrounds in music or sound-related fields, and 53.33% of participants had backgrounds in STEM.

C. Measures

Our post-trial survey collected self-reported measures about robot social attributes, localizability, purchasing interest, and value. We used the Robot Social Attributes Scale (RoSAS) [2] to measure perceived warmth, competence, and discomfort during interactions with the robot. We used a 6-pt Likert scale from "Definitely Not Associated" to "Definitely Associated," omitting a neutral scale option to force decisions. Participants reported robot localizability using their level of agreement with a custom item (i.e., "I could tell where the robot was at all times.") on a 6-pt Likert scale from "Strongly Disagree" to "Strongly Agree." We captured information about value using the Price Sensitivity Meter (PSM) [1]. This inventory collects user purchasing interest on a 5-pt Likert scale from "Not at all interested" to "Extremely interested," in addition to four dollar-valued price points: too cheap, cheap, expensive, and too expensive.

We also conducted a semi-structured interview with participants at the end of the study. Participants answered questions about their impression of the robot and what aspects of the robot influenced their survey responses. The end-of-study interview also included questions about participants' favorite condition (from memory), as well as an opportunity to expound on differences between conditions. Next, the three different conditions were replayed to the participant, accompanied by questions about their favorite condition, opinions of the sounds, information content perceived from the sounds, and whether participants were able to guess the premise of the study before it was revealed.

A final demographic survey recorded participants' age, gender, ethnicity, nationality, hometown, profession, robotics experience, and musical experience.

D. Procedure

Upon entering the study room, the participant was asked to give informed consent before beginning the study. Next, the participant was instructed to construct a LEGO set at a designated workspace in the room and place any bags that the robot dropped on the adjacent table into their workspace. Participants started with the first three bags of the LEGO set already in their workspace. Approximately 2 minutes after the participant started their task, the robot moved to drop off additional LEGOs using the first condition selected for that given trial, and then returned to its starting location. After each condition experience, the researcher (who was otherwise obscured from view) led the participant to a different table to complete the post-trial survey. Lastly, participants completed the semi-structured interview and demographic survey.

E. Hypotheses

Our first hypothesis was formed on the basis of our previous online explorations that considered augmented sonic profiles for multiple robot types [31] and that explored the added value of transformative sound to a mobile robot [33]. Additionally, the concept of localizability was considered based on effects of transformative sound in prior in-person explorations [3]. The resulting hypothesis was as follows:

H1: Adding character-like sound will increase perceived warmth, competence, localizability, purchasing interest, and value compared to the consequential sound condition.

Additionally, we compared functional sound, which had not been explored at length in prior robot sound work, against the other conditions. Based on functional sound benefits in existing devices like elevators, we developed the following hypothesis:

H2: Adding functional sound will increase perceived competence, localizability, purchasing interest, and value compared to consequential sound alone.

Specifically, we expected gains in the less social measures.

F. Analysis

The post-trial questions were evaluated using repeated analysis of variance (rANOVA) tests with an $\alpha = 0.05$ significance level. Concerns such as spherecity violations were considered through use of the Greenhouse-Geiser corrections. In the case of significant main effects, we used Tukey's Honestly Significant Difference (HSD) test to identify significant pairwise differences. For the results of the rANOVA, we reported the F-statistic (F) with the Greenhouse-Geiser corrected degrees of freedom in parentheses, the p-value (p) for determining significance, and the effect size as generalized eta-squared (η_G^2) to improve cross-study comparability [15]. Effect sizes were interpreted relative to Funder and Ozer's updated guidelines [6] after Cohen's original work [4].

In cases of non-normality, a variation of the rANOVA, Friedman's test, was conducted. From the results of Friedman's test, we reported the test-statistic (χ^2) and the p-value (p) respectively.

Financial value responses were further assessed through the PSM analysis methods detailed in prior work [1, 33].

All statistical analyses were conducted using jamovi, a graphical interface for statistics powered by R, as well as open-source modules developed for jamovi [26, 17, 24, 10].

Qualitative data from the semi-structured interviews was thematically coded. We also sorted responses to the conditional preference questions from the interview.

IV. Results

36 participants total enrolled in the study. Due to instances of robot failure or researcher error, six participants were not able to complete the full study; their partial responses were excluded from the data analysis. 30 participants total completed the full study. All data for these participants was successfully recorded other than the ages for four participants who opted not to disclose this information. The following subsections report the results of the rANOVA tests, PSM analyses, thematic coding, and conditional preference sorting.

A. Self-Report Results

1) Social Perceptions Results: The rANOVA tests showed significant effects in responses to two of the RoSAS scales:

Perceived Warmth: ($F(1.81, 52.39) = 14.9$**,** $p < 0.001$**,** $\eta_G^2 = 0.091$**)**

Perceived Competence: ($F(1.91, 55.52) = 4.20$**,** $p = 0.021$**,** $\eta_G^2 = 0.015$**)**

Figure 3 showcases the results of the pairwise comparisons; for warmth, the difference between every condition pairing was significant. The character-like sound ($M = 1.77$) had greater warmth than the consequential sound ($M = 0.961$), as well as the functional sound ($M = 1.38$). The functional sound also had higher warmth than the consequential sound.

Perceived competence was found to have one significant pairwise difference. Character-like sound ($M = 2.99$) had greater perceived competence than consequential sound alone ($M = 2.69$). Functional sound tended to yield a higher rating ($M = 2.93$) than consequential sound, but this difference was not significant.

The rANOVA test on discomfort toward the robot did not yield any significant results ($p = 0.836$). The trending of the data was in the following order for decreasing discomfort: consequential sound ($M = 0.933$), functional sound ($M = 0.878$), and character-like sound ($M = 0.850$).

2) Localization Results: The rANOVA test on localizablity ratings also yielded significance:

Localizability: ($F(1.8, 4.7) = 4.40$, $p = 0.020$, $\eta_G^2 = 0.048$)

Pairwise comparison tests revealed a significant difference specifically between the functional and character-like sounds, where the character-like sound was most localizable ($M = 5.10$, vs. $M = 4.53$ for functional). The consequential sound localizability tended to be rated in the middle ($M = 4.60$), although this difference was not significant from either other condition.

3) PSM Results: Purchasing interest was not found to vary significantly across conditions ($p = 0.141$) from the results of an rANOVA.

The tests on each of the four pricing tiers yielded one significant difference:

Too Cheap: ($\chi^2 = 8.61$, $p = 0.014$)

This result came from a Friedman's test. The Friedman's tests on the other pricing tier input did not yield significant results (all $p > 0.101$).

Pairwise comparisons of the inexpensive price point data with a Holm-Bonferroni correction showed a significant difference between the value of character-like sound ($Median = 100$ dollars) and consequential sound ($Median = 82.5$ dollars). The median value of functional sound was identical to character-like sounds ($Median = 100$ dollars), but was not found to be significantly different.

B. Interview Results

1) Thematic Coding Results: We coded the qualitative interview data; a results summary appears in Table I. Responses were coded according to the facets listed in Table I, which themselves were grouped based on five identified themes from participant responses: 1) personification and other perceptions of the robot, 2) awareness of the robot, 3) robot use context, 4) familiarity, and 5) future considerations.

A review of the facets reveals an association of the character-like condition in particular with positive characteristics ($n =$

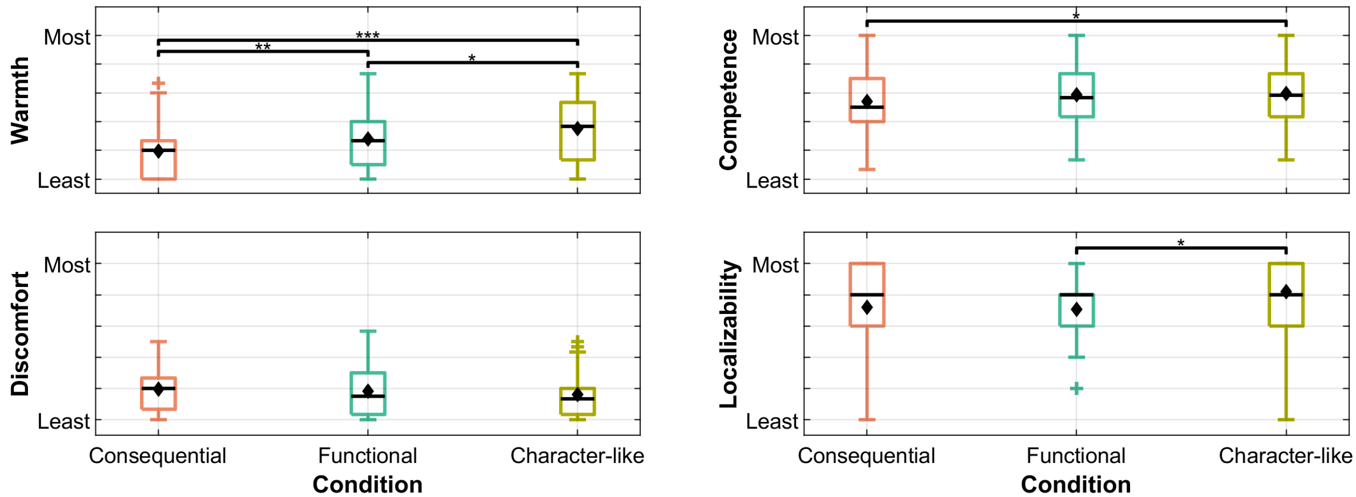

Figure 3: Boxplots for results of the in-person replication study. Black horizontal lines represent the median, diamonds represent the mean, pluses represent outliers, boxes represent the 25th and 75th percentiles, and whiskers cover up to 1.5 times the interquartile range. Brackets represent significant differences. Asterisks above the brackets indicate the level of significance: * for $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$.

19). This feedback contrasted with comments, primarily related to the consequential and functional sound conditions, associating their silence ($n = 9$) or abruptness ($n = 5$) with varying levels of discomfort. The awareness comments involved incorrect labels of what was or was not present in the condition ($n = 7$) and statements about the ability to locate the robot in space, especially for the character-like condition ($n = 9$). Additional comments offered thoughts about cases in which each type of sound may be most appropriate ($n = 7$) and noted similarities between the robot's sounds and common commercial technologies ($n = 2$). The last facet provided opinions on how often the robot should be making sounds ($n = 7$).

2) Conditional Preference Results: When asked about preferred condition, 53.3% participants reported the character-like sound condition to be their preference, 33.3% selected the functional sound condition, 6.67% preferred the consequential sound condition, and 6.67% expressed no preference.

Table I: Results of the coding, including theme numbers, facet (or code) names, example quotes, and frequencies. As in the introduction of the conditions, we use the abbreviations C for consequential, F for functional, and T+E for character-like.

Theme	Facet	Example	Frequency
1	Positive Social Characteristics	"[Character-like] just seemed more friendly and more helpful,... more willing to help."	19 (2F, 17T+E)
1	Discomfort from Silence	"[Silent] one was like, okay, what's going on here? It was kind of scary."	9 (all C)
1	Sharp or Abrupt Interruptions	"[Functional] felt maybe slightly...like surprising, it felt sudden"	5 (all F)
2	Awareness of Conditions	"I didn't even realize during [consequential] that it didn't play a noise."	7 (4C, 3T+E)
2	Localization	"I think it was generally pretty good at making me aware that, like, hey, something is coming."	9 (2C, 7T+E)
3	Underlying Context	"[Character-like] is something you can leave with your kid, you know?"	7 (3F, 4T+E)
4	Familiarity	"it was like the sound that my Apple products make"	2 (all F)
5	Frequency of Sounds	"I could definitely see how [character-like] for a prolonged period of time that could get pretty annoying."	7 (all T+E)

V. Discussion

In this section, we first test the hypotheses, one of which sought to replicate online study findings (from [31], [33]), the other of which offers new robot sound comparisons. We next relate these results to broader application, and we conclude by discussing the strengths and limitations of this work.

A. Key Findings

H1 was partially supported by the results, as evidenced by the significant positive effect of character-like sound on perceived warmth generally, and in one pairwise test (compared to consequential sound) for competence and too cheap pricing. This benefit was echoed in the qualitative results, for example, positive social descriptors of the character-like sound included "friendly", "upbeat", and "more approachable." One participant stretched this association further by describing this robot condition as "more willing to help" and suggesting that the robot made "the environment seem more welcoming [...] because of the tone it played as it went." Participant comments also noted that the character-like robot was "easier to place in space," although a separate set of participants referenced the robot's natural consequential sound as their main point of reference for locating the robot.

These results align well with expectations based on the previous online studies in [31]. The Zhang et al. study gauged perceptions of character-like robot sound when applied to two personal robots (i.e., Cozmo, NAO), one service robot (i.e., TurtleBot 2), and two industrial robots (i.e., Baxter, UR5e). The results showed that the addition of character-like sound led to significant increases in ratings of warmth and competence for most of the investigated robotic platforms. Based on the other motivating study (i.e., [33]), we expected increased price points from the addition of character-like sound. There is partial alignment between the current results and the findings of these past studies in terms of warmth, competence, and the too cheap price point. Taken together, this replication work suggests that added character-like sound can offer strong benefits to warmth and potential benefits to competence and value for robotic systems.

H2 was not supported in the ways we expected. However, we identified trends in competence and localizability that suggest that functional sound may support increases in both compared to consequential sound. Qualitative results surrounding the functional sound condition were mixed, perhaps as a result of participants' existing familiarity with systems that use functional sound. Some participants mentioned the functional sound reminding them of notification sounds from their phones. Although there were no statistically significant results related to discomfort (just trending in the expected order), a subset of participants reported discomfort with the consequential sound condition in the interviews, describing this robot mode as "creepy" or being "weirded out by" it due to the relative silence. One person explained that sound helped "ease the tension" between them and the unfamiliar agent.

While not explicitly tied to a hypothesis, other responses supported the idea that character-like sound can offer advantages over functional sound. Two significant pairwise comparison results, one for warmth and one for localizability, align with this idea. The qualitative results likewise highlight benefits of the character-like sound compared to the functional sound. Multiple participants found the functional sound to appear more abrupt or sharp compared to the emotional component of the character-like sound. Participants also referenced a break in focus as a result of the bell-like sound, in addition to making other associations to the functional sound conjuring "surprise." A number of participants even directly compared the character-like and functional sounds, describing the character-like sounds as more "pleasing." Perhaps the character-like condition's localizability advantage over functional sound arises because the character-like sound condition used a melody while in transit, rather than just during key transition points. The nature of character-like sound affords itself to this type of sound design beyond brief signaling during key transition points, which may in the end have benefits beyond the purely social (e.g., better localizability).

B. Design Implications

The cumulative results suggest that added sound beyond a robot's base consequential sound profile have a measurable benefit. Particularly, the dual application of transformative and emotional sound (as we implemented in the "character-like" condition) may be able to support both successful localization and relative comfort toward a robotic system. The benefits of character-like sound stand out as especially beneficial compared to consequential sound, which, as expressed in the interviews, can be inherently uncomfortable. This qualitative result hints that functional sound (or other simple uses of sound in robotic systems) is preferable to robots with no added sound. These results are consistent with expectations set in other work like the effects of sound on trust of robotic platforms [22].

Currently, many of the robots that use non-verbal sound (e.g., functional, transformative, and emotional sound) are socially-oriented systems. Within this set, systems primarily leverage emotional sound. This design choice is particularly interesting when one considers how common robots are in other scenarios (e.g., industrial manufacturing, food delivery service) in which nearby individuals could actively benefit from the advantages, from better social attributes to improved localization, that transformative sound can offer. We encourage designers to introduce sound in a broader range of robotic systems to take advantage of such benefits and support a more harmonious joint use of space by humans and robots.

C. Strengths & Limitations

This work reproduced selected results from existing academic literature, an important task for ensuring a reliable scientific process. As further discussed in Section V-A, most results were replicated, or at least trended in the expected ways for character-like vs. consequential sound. This confirmation is important, given the replication crisis occurring in psychology (i.e., many classical papers have failed to replicate [11]) and potentially starting to occur in human-robot interaction [9].

Additionally, this effort is one of a small group of robot sound studies that have occurred in person. In-person work affords us a unique position to replicate results in a compelling way, as well as to investigate topics that are highly relevant to current real-life human-robot interaction scenarios (e.g., localizability) but not commonly explored in robot sound work.

Limitations of this work include the relatively small sample size of participants, who themselves are not perfectly representative of our application contexts of interest. The limited instances of significance in the PSM analysis in particular may be due to the smaller sample size in our study compared to typical marketing research. Future robot sound research can use the presented work to perform power analyses and identify fitting sample sizes. Follow-up efforts should also strive to recruit participants who represent the social-cultural perspectives of intended deployment spaces. Further, while this work highlights the benefits of robot sound, it does not prescribe guidance for how these sound profiles should be designed. The specific sounds used can be important; for example, our participants had mixed and conflicting opinions about the right amount of sound to use, a finding echoed in past sound-based robot localization work [3]. Although the sounds created for the presented study leverage existing work and domain knowledge from the area of robot sound, additional sound design options may still make similar or better sense for collaborative tasks. Future work can expand the variety of sounds tested to gain more perspective on how sounds should be designed for collaborative contexts. Robot sound designers may also encounter tensions producing easy-to-hear sound while complying with social norms (or even safety standards) of use domains. Past work from Zhang et al. can inform effective robot sound design collaborations [29] to help navigate these tensions and produce effective nonverbal expression for robots.

VI. Conclusion

In this work, we performed an in-person study to investigate the effects of consequential, functional, and character-like robot sound in a mock human-robot collaboration task. This design allowed us to partially replicate the results of previous online studies, in addition to providing new findings related to functional sound and robot sound's impacts on localizability. Participants perceived the presence of character-like (i.e., transformative and emotional) sound as more positive than the other investigated conditions in all comparisons of warmth, and in some comparisons of competence and localizability. Our results demonstrated that character-like sound helps people to locate a robot without direct line of sight. Additionally, functional sound was perceived as more warm than consequential sound, perhaps due to negative views on complete silence during the mock human-robot teamwork task. This work can help designers understand the potential benefits of adding sound to robot sound profiles to support better social perception and improved robot localizability.

References

[1] Christoph Breidert, Michael Hahsler, and Thomas Reutterer. A Review of Methods for Measuring Willingness-To-Pay. *Innovative Marketing*, 2(4):8–32, 2006.

[2] Colleen M. Carpinella, Alisa B. Wyman, Michael A. Perez, and Steven J. Stroessner. The Robotic Social Attributes Scale (RoSAS): Development and Validation. In *Proc. of the ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI)*, pages 254–262, 2017.

[3] Elizabeth Cha, Naomi T. Fitter, Yunkyung Kim, Terrence Fong, and Maja J. Matarić. Effects of Robot Sound on Auditory Localization in Human-Robot Collaboration. In *Proc. of the ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI)*, pages 434–442, 2018.

[4] Jacob Cohen. *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates, Hillsdale, N.J, 2nd ed. edition, 1988. ISBN 9781134742707.

[5] Emma Frid and Roberto Bresin. Perceptual Evaluation of Blended Sonification of Mechanical Robot Sounds Produced by Emotionally Expressive Gestures: Augmenting Consequential Sounds to Improve Non-verbal Robot Communication. *Int. Journal of Social Robotics*, 2021.

[6] David C. Funder and Daniel J. Ozer. Evaluating Effect Size in Psychological Research: Sense and Nonsense. *Advances in Methods and Practices in Psychological Science*, 2(2):156–168, 2019.

[7] Takamune Izui and Gentiane Venture. Correlation Analysis for Predictive Models of Robot User's Impression: A Study on Visual Medium and Mechanical Noise. *International Journal of Social Robotics*, 12(2):425–439, May 2020. doi: 10.1007/s12369-019-00601-3.

[8] Eun-Sook Jee, Yong-Jeon Jeong, Chong Hui Kim, and Hisato Kobayashi. Sound design for emotion and intention expression of socially interactive robots. *Intelligent Service Robotics*, 3(3):199–206, July 2010. doi: 10.1007/s11370-010-0070-7.

[9] Benedikt Leichtmann, Verena Nitsch, and Martina Mara. Crisis Ahead? Why Human-Robot Interaction User Studies May Have Replicability Problems and Directions for Improvement. *Frontiers in Robotics and AI*, 9:838116, March 2022. ISSN 2296-9144.

[10] Russell Lenth. emmeans: Estimated Marginal Means, aka Least-Squares Means [R package], 2020.

[11] Stephan Lewandowsky and Klaus Oberauer. Low replicability can support robust and efficient science. *Nature Communications*, 11(1):358, January 2020. ISSN 2041-1723.

[12] Dylan Moore, Hamish Tennent, Nikolas Martelaro, and Wendy Ju. Making Noise Intentional: A Study of Servo Sound Perception. In *Proc. of the ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI)*, pages 12–21, 2017.

[13] Dylan Moore, Tobias Dahl, Paula Varela, Wendy Ju, Tormod Næs, and Ingunn Berget. Unintended Consonances: Methods to Understand Robot Motor Sound Perception. In *Proc. of the CHI Conference on Human Factors in*

Computing Systems, pages 1–12, 2019. ISBN 978-1-4503-5970-2.

[14] Jumpei Okimoto and Mihoko Niitsuma. Effects of Auditory Cues on Human-Robot Collaboration. In *Proceedings of the IEEE International Symposium on Industrial Electronics (ISIE)*, pages 1572–1577, Delft, Netherlands, June 2020. IEEE. doi: 10.1109/ISIE45063.2020.9152413.

[15] Stephen Olejnik and James Algina. Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs. *Psychological Methods*, 8(4):434–447, 2003.

[16] Hannah R. M. Pelikan and Malte F. Jung. Designing Robot Sound-In-Interaction: The Case of Autonomous Public Transport Shuttle Buses. In *Proc. of the ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI)*, pages 172–182, 2023. ISBN 978-1-4503-9964-7.

[17] R Core Team. R: A language and environment for statistical computing (Version 4.0) [Computer software], 2021.

[18] Robin Read and Tony Belpaeme. Non-linguistic utterances should be used alongside language, rather than on their own or as a replacement. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 276–277, Bielefeld, Germany, March 2014. ACM/IEEE. doi: 10.1145/2559636.2559836.

[19] Robin Read and Tony Belpaeme. Situational Context Directs How People Affectively Interpret Robotic Non-Linguistic Utterances. In *Proc. of the ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI)*, pages 41–48, 2014.

[20] Robin Read and Tony Belpaeme. People Interpret Robotic Non-linguistic Utterances Categorically. *Int. Journal of Social Robotics*, 8(1):31–50, 2016.

[21] Frederic Anthony Robinson, Mari Velonaki, and Oliver Bown. Smooth Operator: Tuning Robot Perception Through Artificial Movement Sound. In *Proc. of the ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI)*, pages 53–62, 2021.

[22] Richard Savery, Ryan Rose, and Gil Weinberg. Establishing human-robot trust through music-driven robotic emotion prosody and gesture. *Proc. of the IEEE Int. Conf. on Robot and Human Interactive Communication (RO-MAN)*, 2019.

[23] Moondeep C. Shrestha, Ayano Kobayashi, Tomoya Onishi, Hayato Yanagawa, Yuta Yokoyama, Erika Uno, Alexander Schmitz, Mitsuhiro Kamezaki, and Shigeki Sugano. Exploring the use of light and display indicators for communicating directional intent. In *Proceedings of the IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, pages 1651–1656, Banff, AB, Canada, July 2016. IEEE. doi: 10.1109/AIM.2016.7577007.

[24] Henrik Singmann. afex: Analysis of Factorial Experiments [R package], 2018.

[25] Hamish Tennent, Dylan Moore, Malte Jung, and Wendy Ju. Good vibrations: How consequential sounds affect perception of robotic arms. In *Proc. of the IEEE Int. Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 928–935, Lisbon, Spain, 2017.

[26] The jamovi project. jamovi (Version 2.2) [Computer software], 2021.

[27] Gabriele Trovato et al. The Sound or Silence: Investigating the Influence of Robot Noise on Proxemics. In *Proc. of the IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN)*, pages 713–718, 2018. ISBN 978-1-5386-7980-7.

[28] Lisa Zahray, Richard Savery, Liana Syrkett, and Gil Weinberg. Robot Gesture Sonification to Enhance Awareness of Robot Status and Enjoyment of Interaction. In *Proc. of the IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN)*, pages 978–985, 2020.

[29] Brian J Zhang and Naomi T Fitter. Nonverbal Sound in Human-Robot Interaction: A Systematic Review. *ACM Transactions on Human-Robot Interaction*, 1(1): 47, February 2023. doi: 10.1145/3583743.

[30] Brian J. Zhang, Knut Peterson, Christopher A. Sanchez, and Naomi T. Fitter. Exploring Consequential Robot Sound: Should We Make Robots Quiet and Kawaii-et? In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 3056–3062, 2021.

[31] Brian J Zhang, Nick Stargu, Samuel Brimhall, Lilian Chan, Jason Fick, and Naomi T Fitter. Bringing WALL-E out of the Silver Screen: Understanding How Transformative Robot Sound Affects Human Perception. In *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2021.

[32] Brian J Zhang, Christopher A Sanchez, and Naomi T Fitter. Using the Price Sensitivity Meter to Measure the Value of Transformative Robot Sound. In *Proc. of the IEEE Int. Conf. on Robot and Human Interactive Communication (RO-MAN)*. Springer, 2022.

[33] Brian J. Zhang, Christopher A. Sanchez, and Naomi T. Fitter. Using the price sensitivity meter to measure the value of transformative robot sound. In *Proc. of the IEEE Int. Conf. on Robot and Human Interactive Communication (RO-MAN)*, pages 301–307, 2022.

[34] Brian J. Zhang, Bastian Orthmann, Ilaria Torre, Roberto Breslin, Jason Fick, Iolanda Leite, and Naomi T. Fitter. Hearing it out: Guiding robot sound design through design thinking. In *Proc. of the IEEE Int. Conf. on Robot and Human Interactive Communication (RO-MAN)*, 2023.