

Effects of Robotic Companionship on Music Enjoyment and Agent Perception

Guy Hoffman

Media Innovation Lab, School of Communication
IDC Herzliya
P.O.B 167, Herzliya 46150, Israel
Email: hoffman@idc.ac.il

Keinan Vanunu

School of Psychology
IDC Herzliya
P.O.B 167, Herzliya 46150, Israel
Email: kvanunu@idc.ac.il

Abstract—We evaluate the effects of robotic listening companionship on people’s enjoyment of music, and on their perception of the robot. We present a robotic speaker device designed for joint listening and embodied performance of the music played on it. The robot generates smoothed real-time beat-synchronized dance moves, uses nonverbal gestures for common ground, and can make and maintain eye-contact.

In an experimental between-subject study ($n=67$), participants listened to songs played on the speaker device, with the robot either moving in sync with the beat, moving off-beat, or not moving at all. We found that while the robot’s beat precision was not consciously detected by Ps, an on-beat robot positively affected song liking. There was no effect on overall experience enjoyment. In addition, the robot’s response caused Ps to attribute more positive human-like traits to the robot, as well as rate the robot as more similar to themselves. Notably, personal listening habits (solitary vs. social) affected agent attributions.

This work points to a larger question, namely how a robot’s perceived response to an event might affect a human’s perception of the same event.

I. INTRODUCTION

Robotic companions are expected to play both a functional and a social role in human environments, such as homes, schools, offices, nursing homes, and cars. There, robots will be designed to respond to events and stimuli around them, in part to fulfill a task, but also to affect a social response in their human counterparts. A key question that arises is: how might a robot’s perceived response to an event affect a human’s perception of the same event?

We know that humans look to the interpretation of others to form their own perception of events and experiences, a phenomenon called “social referencing” [1]. To what extent does this phenomenon translate to human-robot shared experiences? How does it affect their perception of the robot? Can we design robot responses to third-party experiences so as to affect a particular human response in a desired way? Answering these questions could have consequences on a host of robotic companion areas, from media experiences, through healthcare (imagine making an unpleasant but necessary medical procedure more positively perceived), to work environments, or even just for making standing in traffic less burdensome.

In this work we examine the influence of robotic experience companionship in the realm of music listening, examining how the sense that a robot is experiencing music with a person

affects that person’s enjoyment of the listening experience, their liking of the songs played, and their impression of the robot. This is the first in a series of studies dealing specifically with joint experiences of media, also suggesting the idea of robotic companions as a novel addition to media playback technologies.

A. Social Aspects of Music Listening

Music listening is central to human culture, serving functions from entertainment to mood management and emotional support [2]. As music playback technology evolves, so does the way we consume music. For example, the introduction of affordable portable devices has made listening in the late 20th century increasingly solitary [3]. This trend has recently reversed, perhaps due to the proliferation of playback opportunities and online music sharing. A recent study found that today only 26% of music listening happens alone, compared to 69% in the 1980s [4].

While apparent that music listening is often a social phenomenon (e.g. parties and public concerts), the social aspects of music listening have not been widely explored. A recent book on shared consumption of music deals mostly with online sharing of music and not with physically co-located listening [5]. And while North *et al.* found that people enjoy music less when they are with others, their finding could not be separated from unintended public listening, where participants did not control the music they heard [4]. They found, in contrast, that participants paid more attention to music when listening with their boyfriend or girlfriend, or even with “others”, than alone. In other work, it has been found that people move more vigorously to music when listening to it with others [6], also illustrating a social aspect of music listening.

Can robots provide the kind of social presence that might support a joint music listening experience? We know that social presence can be a mediating factor in human-robot social responses [7]. More specifically, a robot was perceived as more engaging, credible, and informative than an animated character due to its physical embodiment [8]. Another robot’s physical presence has been shown to affect its social presence in relation to personal space, trust, and respect [9]. In an additional study, a robot’s movement to music influenced children’s proclivity to dance to the music [10].

It thus makes sense to investigate how a robotic listening companion may affect people’s music listening experience.

To this end, we developed *Travis*, a robotic speaker dock and music listening companion, (Fig. 1) aimed to enhance a human's music experience by providing listening companionship, and by embodying the music played on the device as a performance. In its initial application, the robot performs beat-synchronized moves to accompany the music played on the device, maintains eye-contact with the user, and uses gesturing for common ground.

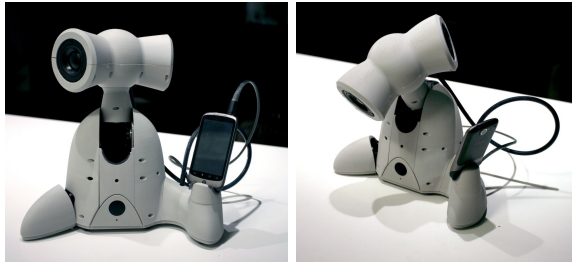


Fig. 1. Robotic speaker dock and music listening companion.

B. Musical Robots and Physical Gestures

Travis also builds on the tradition of musical robots. Robotic musicianship extends other kinds of computer music by adding a physical aspect to computer-generated and interactive musical systems [11]. It provides humans with physical cues, that can help them coordinate their joint playing with the robot. But, importantly, they also create a more engaging experience for the audience by adding a visual element to the sound.

Virtually all robotic musicianship research deals with music production and improvisation [12], [13], with little research on the effect of musical robots for audiences, or the effect of performance on music listening. In human music listening, it has been shown that adding a video channel to a music performance alters audience perception in terms of the affective interpretation of sound features [14]. Musical robots, too, have been shown to positively affect audience appreciation of joint improvisation [13]. This finding, however, was not separated from the other musician's ability to see the robot's gestures as it was playing.

Travis is designed to serve as a research platform to isolate and explore the effects of the performative aspect of robotic musicianship on human's music listening.

There were a number of commercial robots which included music amplification capabilities (e.g. miuro [15] and Rolly [16]). However, their design as non-anthropomorphic mobile robots, and their purpose for entertainment only, were different than the robot presented here.

Several previous works have dealt with rhythmic human-robot interaction (HRI), as well as with dancing robots. The Keepon robot explored dance as an element of social interaction [10], [17], in particular between children and robots. Other work has explored robotic drumming, rhythmic HRI, and human-robot musical synchronization [18], [19], [13].

Most recently, Avrunin *et al.* examined the effects of robotic beat synchronization to music on people's perception of the robot's dance abilities and lifelikeness [20]. We extend

their work in several ways: first, Avrunin *et al.* evaluated participants' responses to video sequences of the robot, while this study uses a physical robot co-located with participants. Furthermore, we do not directly ask for people's evaluation of the quality of the robot's movement, but instead the indirect effect it has on people's perception of the quality of the music and the robot. Finally, we expand on the general notion of the robot's "lifelikeness" by examining specific perceived character traits of the robot and people's perceived similarity to the robot.

II. ROBOTIC PLATFORM

The robot *Travis* is a research platform to examine HRI as it relates to media consumption, nonverbal behavior, timing, and physical presence. This section describes the robot's appearance, hardware, and software design.

A. Appearance

Travis's appearance was designed with a number of guidelines in mind: first, the robot's main application is to deliver music, and to move expressively to the music. Its morphology therefore emphasizes audio amplification, and its DoF placement supports expressive movement to musical content.

Second, the robot needs to be capable of basic nonverbal communication, such as turn-taking, attention, and affect display. Travis is therefore sized and shaped so that, when placed on a desk, its head is roughly in line with a seated person's head in front of it, and is capable of basic gesturing.

The robot's appearance should also elicit a sense of companionship with the human user. Its body is thus designed to evoke a pet-like relation, with a size comparable to a small animal, and an organic, but not humanoid form.

Finally, we have decided not to make the mobile device part of the robot's body, but instead to create the impression that the robot "holds" the device. This is intended to create a sense of identification ("like-me") and empathy with the robot, as Travis relates to the device similarly to the way a human would. Moreover, this setup allows for the device to serve as an object of common ground [21] and joint attention [22] between the human and the robot, setting the stage for nonverbal dialog.

B. Hardware

Travis's hardware platform is a five degree-of-freedom robot. Each DoF is controlled via direct-drive using a Robotis Dynamixel MX-28 servo motor. The robot has two speakers, acting as a stereo pair, in the sides of its head, and one sub-woofer speaker pointing downwards in the base. In addition, the robot contains an Android Accessory Development Kit (ADK) control board [23], and a digital amplifier with an audio crossover circuit (Fig. 2).

All of the software runs on an Android smartphone, communicating to the ADK board using a low-latency variable-interval position-velocity packet protocol. The board runs bridge firmware translating the ADB interface into the MX-28 network protocol. Each motor maintains its own position and velocity control through the servo firmware of the motor unit.

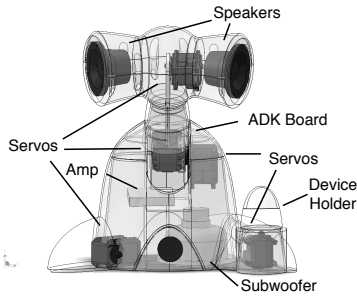


Fig. 2. *Travis* mechanical structure.

C. Software

In its initial application, *Travis* plays songs from the mobile device's music library and responds to them by generating dance moves based on the song's beat, segment, and genre. The segmentation and classification of songs is beyond the scope of our work, as there is a large body of work concerned with methods to automatically track beats in musical audio (e.g. [24], [25]), as well as for splitting musical audio into segments (for a review, see: [26]). More recently, network-based services offer identification and classification of musical audio based on short audio samples. Some of these services provide beat and segmentation information, as well [27].

Our work therefore focused on the expressive gesture and animation system given a song's accurate genre and beat segmentation. Fig. 3 shows an overview of the robot's system software.

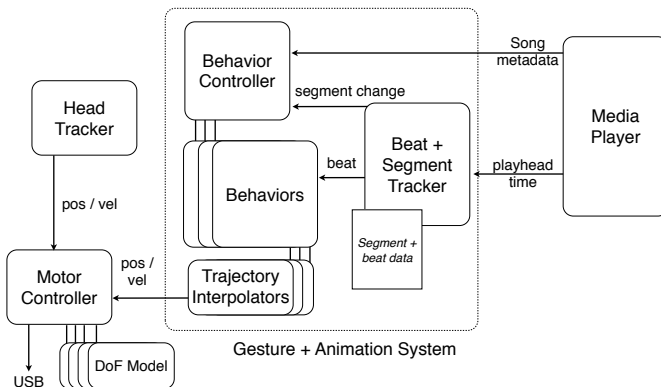


Fig. 3. *Travis* software diagram.

The “Behavior Controller” tracks the song’s metadata to trigger “Behaviors” based on the current segment and genre of the song. In parallel, the “Beat and Segment Tracker” follows beat segmentation and sends events to the active Behavior.

In case of a beat trigger, the Behavior executes one of two beat responses: (a) a repetitive beat gesture involving one or more DoFs; or (b) a probabilistic adjustment gesture, adding variability to the repetitive motion. Each motion is then split by DoF and sent to a trajectory interpolator associated with the DoF. The interpolator simulates ease-in/ease-out through a high-frequency interpolator, inspired by the animation arbitration system used in [28], and similar to the one we used in a previous robot [13].

Finally, as gaze behavior is central to interaction both between humans [29] and between humans and robots [30], the “Head Tracker” maintains eye-contact by using the camera of the mobile device. We make use of existing face detection software on the phone to track and follow the user’s head.

Our face tracking uses an active perception approach [31], [32]. Since the phone is mounted on a pan DoF, linear compensation feedback can keep the head centered in the camera view. However, as the camera is coupled to the robot’s hand, gaze behavior requires an additional transformation between the neck pan DoF angle θ' and the active perception result angle θ . The robot compensates for parallax induced by the disparity d between the two DoF centers by $\theta' = \arctan(\tan\theta - \frac{d}{h})$, with h being the estimated frontal distance of the human’s head.

Motor requests from the Behavior system and the Head Tracker are arbitrated by the “Motor Controller” module, which then uses specific DoF models to send position and velocity commands to the robot’s hardware.

A detailed description of the robot’s design, hardware, and software modules is provided in a separate publication [33].

III. RESEARCH QUESTIONS AND HYPOTHESES

We are interested in examining the effects of a robotic listening companion on music appreciation and impressions of the robot. In particular, this work evaluates whether an embodied artificial agent listening to music with a person can affect that person’s liking of a song, or their enjoyment of the music listening experience overall. How does it affect their appreciation of the robot’s human-like traits? To what extent does the the robot’s response affect users’ sense of the robot as being similar to them?

In addition, we are curious how sensitive people were to the robot’s accuracy at representing the beat of the songs it is reacting to? Also, do people’s listening habits affect their appreciation of a robotic listening companion? To evaluate the above, we tested the following hypotheses:

Hypothesis 1 (Song Liking) — A robotic companion listening to a song with a person will cause them to like the song better.

Hypothesis 2 (Experience Enjoyment) — A robotic companion listening to a song with a person will cause them to enjoy listening to music more.

Hypothesis 3 (Agent Impression) — A robotic companion listening to a song with a person will cause them to attach positive character traits to the robot.

Hypothesis 4 (Human-robot Similarity) — A robotic companion listening to a song with a person will cause them to consider the robot to be similar to themselves.

Hypothesis 5 (Listening Habits) — Listening habits act as a mediating factor to the above-mentioned effects, such that persons who usually listen to music with others will show stronger responses to the measured variables.

To estimate these variables, we manipulated the robot’s behavior to provide listening companionship. We also manipulated the robot’s response to the music to gain insight on the importance of precise beat synchronization, as suggested by Avrunin *et al.* [20].

IV. METHOD

We conducted a controlled laboratory experiment, in which participants listened to three songs through the robot's speakers, with the robot situated on a table next to them. All participants listened to the same three songs, taken from different genres and periods. To prevent order effects, the order of the songs was randomized for each participant.

A. Design

To minimize interfering variables, this experiment evaluated only a single beat-tracking behavior of the robot, which was kept constant throughout each song and without regard for the songs' genre. We also excluded any other nonverbal communication, as well as eye contact behaviors. In other words, the robot's only behavior was that, when a song was played, the robot moved in a constant repetitive motion according to the beat, and then stopped moving when the song was stopped.

We manipulated one between-subject variable, the robot's movement response to the music. In the *ON-BEAT* condition, the robot performed accurate on-beat movements to the song. In the *OFF-BEAT* condition, the robot's movements were at a similar tempo, but consistently off-beat to the music. Some beats were skipped completely, and some were time-shifted by a random amount of ± 100 -500ms. In the *STATIC* condition, the robot did not move at all, although the music still came out of the robot's speakers, as in the other two conditions.

In addition, we measured two additional between-subject variables: one was music listening habits, *solitary* or *social*, by asking two questions ("I usually listen to music alone" (reverse scale), "I usually listen to music with friends"). The other was a scale variable of experience with AI and robotics.

At any point in the song, participants could decide to stop playback and move on to the next song. All song control was done from a computer monitor situated next to the robot.

We recorded the length of time the participant decided to listen to the song, as well as the participant's self-reported liking of the song ("Song Liking"). After all three songs were played and evaluated, a questionnaire measured the participant's overall enjoyment; whether the robot was "listening" to the song with them; whether the robot was "playing" the song for them; and how similar it was to them. We also collected a number of impressions of the robot's positive human-like characteristics ("Impression of Agent"), and the participants' relationship to the robot.

Through video recording, we captured the participant's physical reaction to the songs to evaluate the amount of rhythmic head-nodding and foot-tapping movements. These were not used in the currently presented results.

B. Procedure

The experiment was conducted in a small office room with controlled lighting and no outside distractions. The room contained a desk on which a computer monitor and mouse were placed, as well as the robotic speaker device (Fig. 4).

Each participant entered the room individually with the experimenter. They were asked to sit in a chair at the end of

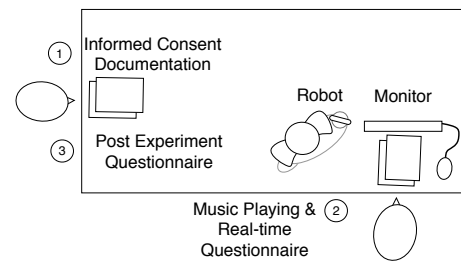


Fig. 4. Experimental setup diagram.

the desk that was furthest from the computer monitor and the robot, facing the back of the robot. The experimenter explained the experiment guidelines, and received informed consent from the participant while they were seated.

The participant then moved to the other end of the desk to sit by the monitor. In front of the monitor was the real-time "song liking" questionnaire with a pen, allowing the participant to fill it out while listening to each individual song.

Each participant was told that they would be listening to three songs through a prototype speaker device. The robot was situated slightly to their left, about 30 centimeters away. At no point was the device referred to as a "robot", but only as a "speaker device" in order to not prime subjects in the *STATIC* group to expect movement from the device.

The participant was shown briefly how to use the computer program to switch between songs. They were told to answer the three questions measuring song liking before moving on to the next song. Participants were also instructed that they could skip to the next song at any point, and were not required to listen to each song to the end. Finally, they were told that it was supposed to be a fun experience, and were encouraged to enjoy and have a good time.

The participant was then left alone in the office to complete the experiment, which started by clicking a button on the screen, triggering the first song. To stop the current song they would click the same button on the screen, fill out the three relevant enjoyment questions, and click the same button to start the next song.

Upon completion of the experiment they would click another button labeled "Call Experimenter" and the experimenter would re-enter the room. The experimenter would then shut off the robot in front of the participant and ask them to return to the chair at the end of the desk in order to fill out the post-experiment questionnaire. This in order for the participants to not be influenced by the robot's movement or lack thereof while completing their evaluation of the robot.

C. Participants

A total of 67 people participated in the experiment. The subjects were multi-national, and all communication was done in English. They were recruited from the International School of Communication at IDC Herzliya, in return for class credit.

V. MEASURES

All measures are on a 7-point Likert scale (from "strongly disagree" to "strongly agree") unless otherwise noted.

A. Song Liking

We estimate the real-time liking of the played songs by taking the mean of the response to three questions answered during or immediately after listening to each song, “I enjoyed this song”, “I believe others would enjoy this song”, and “I would like to listen to this song again in the future”. The total Song Liking score is the mean of the score for all three songs (Cronbach’s $\alpha = 0.75$). This measure estimates participants’ rating of the songs themselves.

B. Experience Enjoyment

We estimate the overall enjoyment of the experience (presumably including the songs and the robot’s response to them) by taking the mean of the response to two questions answered in the post-experiment questionnaire: “My overall experience was enjoyable”, “My overall experience was boring” (reverse scale). Cronbach’s α for this measure was 0.74.

C. Impression of Agent

Participants rated their impression of the robotic agent as a composite measure of five items: the robot’s perceived friendliness, confidence, warmth, cooperativeness, and sociability. This measure was validated in previous studies [34], [35], and Cronbach’s α in our data was 0.79.

D. Human-Robot Similarity

Finally, we asked participants to rate on a 7-point scale to what extent the robot was similar to them.

VI. RESULTS

Our experiment contained three experimental conditions (*ON-BEAT*, *OFF-BEAT*, and *STATIC*). We tested for group differences between the three conditions. No significant differences were found between conditions regarding gender, age, familiarity with artificial intelligence and past experience working with robots.

A. Sensitivity to Beat Precision

Somewhat surprisingly, participants were overall not consciously aware of the robot’s beat precision. When asked whether the device “moved on or off-beat”, 22/23 (96%) participants in the *OFF-BEAT* condition said that the robot moved on-beat. This is on par with the *ON-BEAT* condition, where 21/22 (95%) perceived the robot to move on-beat.

B. Manipulation Checks

Our conditions were intended to manipulate the sense of joint listening. We confirmed our manipulation by a composite measure of two items, asking whether the device “was listening to the song with” the participants, and whether it “enjoyed the songs”. Cronbach’s α for this composite was 0.84. One-way ANOVA confirms that the manipulation had a significant effect on the sense of joint listening ($F(2, 64) = 18.192, p < .001, \eta^2 = .36$). Multiple comparisons, applying the Bonferroni correction, revealed that participants in both the *ON-BEAT* ($M = 5.29, SD = 1.32$) and *OFF-BEAT* ($M = 5.78, SD = 1.15$) conditions showed significantly

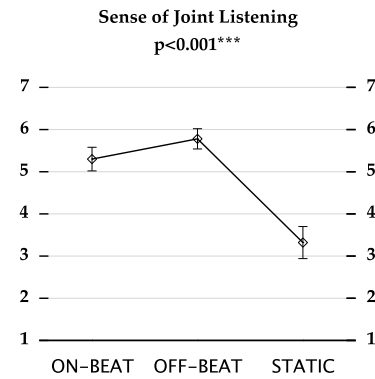


Fig. 5. Means and SEs of perceived sense of joint listening per condition.

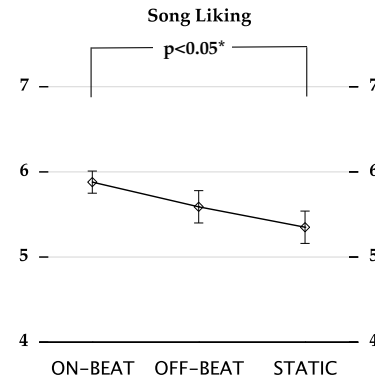


Fig. 6. Means and SEs of experienced song liking per condition.

higher means than the control group ($M = 3.32, SD = 1.8$). No significant differences were found between the *ON-BEAT* and *OFF-BEAT* conditions (Fig. 5).

C. Response to Music

1) *Song Liking*: The composite song liking variable was collected in real-time during or after each song, and estimated how much participants liked the songs they listened to. Song liking for the *ON-BEAT* condition is highest ($M = 5.89, SD = .62$), for *STATIC* is lowest ($M = 5.36, SD = .87$), and for *OFF-BEAT* in-between ($M = 5.6, SD = .92$) (Fig. 6).

A one-way ANOVA did not result in a significant result ($F(2, 64) = 2.202, n.s.$). However, a planned contrast analysis between the *ON-BEAT* and the *STATIC* conditions yielded a significant difference ($t(42) = 2.095, p < .05, d = .7$).

2) *Experience Enjoyment*: The composite experience enjoyment variable was measured at the end of the study, and estimated whether the overall experience was enjoyable to participants. Neither one-way ANOVA nor planned contrasts revealed significant differences between the conditions ($F(2, 64) = .31, n.s.$ and $t(42) = .93, n.s.$).

D. Evaluation of the Robot

1) *Impression of Agent*: The impression of agent composite variable measured positive human traits attributed to the robot. A one-way ANOVA yielded significant results, ($F(2, 64) = 11.18, p < .001, \eta^2 = .26$). Multiple comparisons, applying the Bonferroni correction, showed that both the *ON-BEAT*

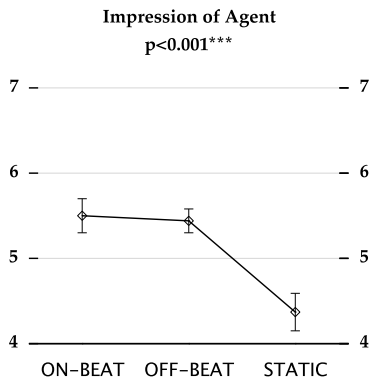


Fig. 7. Means and SEs of perceived impression of agent per condition.

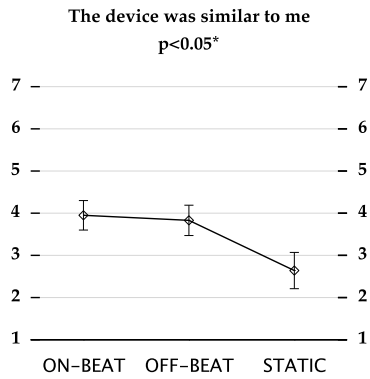


Fig. 8. Means and SEs of perceived human robot similarity per condition.

($M = 5.5$, $SD = .93$) and the *OFF-BEAT* ($M = 5.44$, $SD = .67$) conditions significantly differ from the *STATIC* control condition ($M = 4.37$, $SD = 1.02$) (Fig. 7).

2) *Human-Robot Similarity*: A one-way ANOVA on the human-robot similarity scale yielded significant results ($F(2, 64) = 3.64$, $p < 0.05$, $\eta^2 = .12$). Multiple comparisons, applying the Bonferroni correction, revealed that participants in the *ON-BEAT* condition ($M = 3.95$, $SD = 1.62$) felt that the device was more similar to them when compared to the control condition ($M = 2.64$, $SD = 2.01$). No other significant differences were found (Fig. 8).

E. Effects of Listening Habits

To measure the interaction between individual differences in music listening habits and the above effects, we performed a two-way ANOVA for condition (*ON-BEAT*, *OFF-BEAT*, and *STATIC*), and listening habits (*solitary* vs. *social* as described above).

The reported significant main effect for the “Impression of Agent” variable was qualified by significant disordinal two-way interaction for condition \times listening-habit, ($F(2, 61) = 3.36$, $p < .05$, partial $\eta^2 = .1$). Simple effects analysis revealed that among participants who usually listen to music with others, both the *ON-BEAT* ($M = 5.69$, $SD = .9$) and *OFF-BEAT* ($M = 5.6$, $SD = .67$) conditions rated the robot higher compared to the *STATIC* condition ($M = 4.04$, $SD = .76$). No significant differences were found among participants who

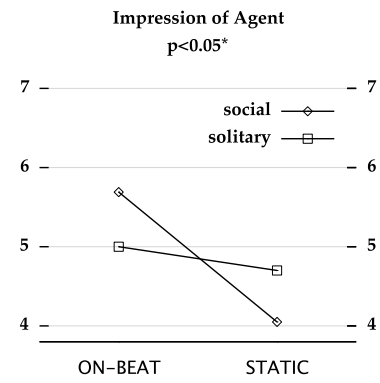


Fig. 9. Interaction between participants' listening habits and experimental condition on impression of agent.

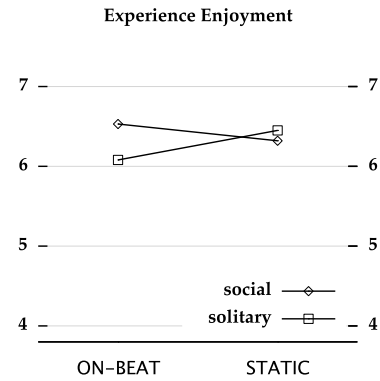


Fig. 10. Interaction between participants' listening habits and experimental condition on experience enjoyment.

usually listen to music alone (Fig. 9)¹. A similar analysis for the “Experience Enjoyment” variable was not found significant ($F(2, 61) = 1.44$, n.s.) (Fig. 10).

VII. DISCUSSION

Our results show support for Hypothesis 1: a robot that responds to music on-beat causes participants to rate the same songs significantly higher than a robot that just plays the song without responding to the music. Interestingly, even though participants did not report noticing the robot's imprecision in the *OFF-BEAT* condition, their song liking measure falls roughly halfway between the two more extreme conditions, albeit not significantly.

We did not find support for Hypothesis 2. Experience enjoyment was not affected by our manipulation. However, we did find that this variable is sensitive to the interaction between listening habits and robot response (see below).

Both Hypothesis 3—participants' impression of the agent—and Hypothesis 4—their sense that the robot is like them—were supported, with significant differences between the robot response conditions and the control condition. We did not find an effect of the robot's beat precision on either of these two variables.

¹Due to the similarity of the *ON-BEAT* and *OFF-BEAT* conditions, figures in this section only show *ON-BEAT* alongside the control condition, for clarity.

Since we believed that our beat precision manipulation was noticeable, we were surprised to not have found much difference between a precisely moving robot and a robot that misses beats, and errs on most other beats relative to the music. Participants did not detect the beat precision consciously, and for all but one measure (song liking), both conditions behaved as if they were the same. One explanation could be the robot's high novelty in both conditions, masking the actual robot's movement effect. Another could be that the manipulation was too subtle, relative to people's sensitivity to choreography precision. Also, participants in our *OFF-BEAT* did not have any point of reference to estimate how well the robot is performing, possibly supporting their sense of this being the robot's "best attempt" at syncing to the music.

Moreover, considering the detected difference in song liking between the on-beat and the static robot, it is interesting that there was no difference in overall experience enjoyment between a completely non-moving speaker device and a robot that responded to songs.

Interaction analysis suggests that individual differences in the form of listening habits could be in play: social listeners significantly rate music-responsive robots more positively. Solitary listeners not only did not show this effect, but actually display a decrease in agent impression when the robot responds to the music. We find a similar trend (albeit not significant) in experience enjoyment (see: Fig. 10), with social listeners enjoying the task more with a responsive robot, and solitary listeners enjoying the task less when the robot responds to the music.

It is worth comparing our findings to previous work, which did find significant differences between an on-beat and an off-beat dancing robot, specifically in terms of their lifelikeness and entertainment value [20]. There are three methodological differences between the two experiments which can shed light on this contrast:

Most importantly, Avrunin *et al.* used a within-subject design in which participants watched videos of an on-beat and an off-beat robot moving simultaneously and side-by-side to the same musical soundtrack. The distinction bias occurring in such joint evaluation may account for the more pronounced difference between the two conditions. The lack of a nonmoving baseline could further contribute to emphasizing the difference between an on-beat and an off-beat moving robot. In contrast, our experiment followed a between-subject design with each participant experiencing only one robotic behavior, either on-beat, off-beat, or static.

Another methodological difference that could account for the contrasting outcome is that Avrunin *et al.* asked participants to judge the purpose of the study, focusing their attention on the robot's dance and dependent variables, and possibly contributing to demand effects. Our design used an indirect approach, asking participants to just listen to songs without drawing their attention to the robot's behavior. This design was aimed to both simulate a more natural listening experience, and explore the implicit effects of the robot's joint listening on people's enjoyment of the music.

Finally, Avrunin *et al.*'s work used on-screen video evaluation, whereas our work placed the robot physically in the participants' space. We know that a robot's physical presence

affects people's impression of the agent [8], [9]. This could also have contributed to their higher rating of our off-beat robot condition than if it was presented to them on a screen.

VIII. CONCLUSION

In this work we explored the idea of a robotic music listening companion. We presented a robotic speaker dock that responds to the music played on it, and is capable of nonverbal behaviors and eye contact. In a between-subjects experiment we evaluated the effects of robotic listening companionship on people's liking of songs, their enjoyment of listening to music, their perception of the robot, and their sense that the companion is similar to them. We also evaluated the interaction of people's listening habits with these variables.

Our support of Hypothesis 1 (Song Liking) shows first evidence of "robotic social referencing"—a robot's perceived enjoyment of an event influencing people's perception of the same event. While we focused on music listening, this effect can have implications beyond media enjoyment, and suggests a novel role for personal robots as contributors to, and possibly amplifiers of, people's own evaluation of external events.

Be it in schools, the workplace, or nursing homes, robotic experience companions could shape people's perception of external occurrences. People's reactions could also be purposefully manipulated by a robot's apparent reaction to them, thus encouraging people to enjoy activities more, or "sweeten a bitter pill" when a negative outcome is unavoidable.

Our support for Hypotheses 3 (Agent Impression) and 4 (Human-robot Similarity) suggest that people rate a robot significantly higher on positive human-like traits, and as significantly more similar to them when the robot responds to the music as they do. This finding points to a straightforward new way to affect people's impressions of robots, increasing empathy by causing the robot to respond similarly to them to an external event.

Notably, we found evidence that individual differences affect the response to robotic experience companions. Social listeners preferred a responsive robot, whereas solitary listeners found it detrimental for the robot to listen to music with them. This finding suggests that robotic experience companions might not be for everyone, but could be more appropriate for people who score higher on the sociability scale.

We also find these results to merit further personality-related studies of robotic companions, as well as in the field of HRI at large, in order to understand what different people look for in a personal robots.

In conclusion, the work presented here is a first step in studying robotic experience companionship: how does a human's sense that a robot experiences something change that person's perceptions of the experience. We showed evidence in the realm of music listening as one example of a joint experience. However, looking at HRI more broadly, robotic experience companionship can affect non-media realms, such as healthcare, workplace robotics, tourism, sports, transportation, and elderly care.

Our results suggest that the idea of robotic experience companionship could be seen both as a novel form of human-robot

interaction, and as a tool to achieve human-robot empathy, but also that the user's personality should be taken into account.

REFERENCES

- [1] S. Feinman, "Social referencing in infancy," *Merrill-Palmer Quarterly*, vol. 28, pp. 445–470, 1982.
- [2] T. F. M. Ter Bogt, J. Mulder, Q. a. W. Raaijmakers, and S. Nic Gabhainn, "Moved by music: A typology of music listeners," *Psychology of Music*, vol. 39, no. 2, pp. 147–163, Aug. 2010.
- [3] R. Larson and R. Kubey, "Television and Music: Contrasting Media in Adolescent Life," *Youth & Society*, vol. 15, no. 1, pp. 13–31, Sep. 1983.
- [4] A. C. North, D. J. Hargreaves, and J. J. Hargreaves, "Uses of Music in Everyday Life," *Music Perception*, vol. 22, no. 1, pp. 41–77, 2004.
- [5] K. OHara and B. Brown, Eds., *Consuming Music Together*, ser. Computer Supported Cooperative Work. Berlin/Heidelberg: Springer-Verlag, 2006, vol. 35.
- [6] L. D. Bruyn, M. Leman, and D. Moelants, "Does Social Interaction Activate Music Listeners?" in *CMMR 2008*, S. Ystad, R. Kronland-Marinet, and K. Jensen, Eds. Springer-Verlag Berlin Heidelberg, 2009, pp. 93–106.
- [7] K. M. Lee, W. Peng, S.-A. Jin, and C. Yan, "Can Robots Manifest Personality?: An Empirical Test of Personality Recognition, Social Responses, and Social Presence in Human-Robot Interaction," *Journal of Communication*, vol. 56, no. 4, pp. 754–772, Dec. 2006.
- [8] C. Kidd and C. Breazeal, "Effect of a robot on user perceptions," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2004)*, 2004.
- [9] W. Bainbridge, J. Hart, E. Kim, and B. Scassellati, "The effect of presence on human-robot interaction," in *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008)*, 2008.
- [10] M. P. Michalowski, S. Sabanovic, and H. Kozima, "A Dancing Robot for Rhythmic Social Interaction," in *HRI '07: Proceeding of the ACM/IEEE international conference on Human-robot interaction*, Arlington, Virginia, USA, Mar. 2007, pp. 89–96.
- [11] G. Weinberg and S. Driscoll, "Toward Robotic Musicianship," *Computer Music Journal*, vol. 30, no. 4, pp. 28–45, 2006.
- [12] K. Petersen, J. Solis, and A. Takanishi, "Toward enabling a natural interaction between human musicians and musical performance robots: Implementation of a real-time gestural interface," in *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008)*, 2008.
- [13] G. Hoffman and G. Weinberg, "Interactive improvisation with a robotic marimba player," *Autonomous Robots*, vol. 31, no. 2-3, pp. 133–153, Jun. 2011.
- [14] W. F. Thompson, P. Graham, and F. A. Russo, "Seeing music performance : Visual influences on perception and experience," *Semiotica*, pp. 203–227, 2005.
- [15] J. Aucouturier, Y. Ogaï, and T. Ikegami, "Making a robot dance to music using chaotic itinerancy in a network of fitzhugh-nagumo neurons," *Neural Information Processing*, 2008.
- [16] J. Kim, S. Kwak, and M. Kim, "Entertainment robot personality design based on basic factors on motions: A case study with Rolly," in *18th IEEE International Symposium on Robot and Human Interactive Communication, (RO-MAN '09)*, 2009, pp. 803–808.
- [17] H. Kozima, M. P. Michalowski, and C. Nakagawa, "Keep on: A playful robot for research, therapy, and entertainment," *International Journal of Social Robotics*, vol. 1, no. 1, pp. 3–18, Nov. 2009.
- [18] C. Crick and B. Scassellati, "Synchronization in social tasks: Robotic drumming," in *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Reading, UK, 2006.
- [19] G. Weinberg and B. Blosser, "A Leader-Follower Turn-taking Model Incorporating Beat Detection in Musical Human-Robot Interaction," in *HRI '09: Proceeding of the ACM/IEEE international conference on Human-robot interaction*, 2009.
- [20] E. Avrunin, J. Hart, A. Douglas, and B. Scassellati, "Effects related to synchrony and repertoire in perceptions of robot dance," in *HRI '11: Proceeding of the ACM/IEEE international conference on Human-robot interaction*. New York, New York, USA: ACM Press, 2011, p. 93.
- [21] H. H. Clark, *Using Language*. Cambridge, UK: Cambridge University Press, 1996.
- [22] C. Breazeal, A. Brooks, and D. Chilongo, "Working collaboratively with humanoid robots," *Humanoid Robots, 2004 4th IEEE/RAS International Conference on*, 2004.
- [23] Google, "Android Open Accessory Development Kit." [Online]. Available: <http://developer.android.com/guide/topics/usb/adk.html>
- [24] M. Goto, "An Audio-based Real-time Beat Tracking System for Music With or Without Drum-sounds," *Journal of New Music Research*, vol. 30, no. 2, pp. 159–171, 2001.
- [25] M. E. P. Davies, S. Member, M. D. Plumbley, and A. P. Art, "Context-Dependent Beat Tracking of Musical Audio," *Language*, vol. 15, no. 3, pp. 1009–1020, 2007.
- [26] E. Peiszer, T. Lidy, and A. Rauber, "Automatic Audio Segmentation : Segment Boundary and Structure Detection in Popular Music," in *Proceedings of LSAS*, 2008.
- [27] T. Bertin-mahieux, D. P. W. Ellis, B. Whitman, and P. Lamere, "The million song dataset," in *Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR 2011)*, 2011.
- [28] J. Gray, G. Hoffman, S. O. Adalgeirsson, M. Berlin, and C. Breazeal, "Expressive, interactive robots: Tools, techniques, and insights based on collaborations," in *HRI 2010 Workshop: What do collaborations with the arts have to say about HRI?*, 2010.
- [29] M. Argyle, R. Ingham, and M. McCallin, "The different functions of gaze," *Semiotica*, vol. 7, no. 1, pp. 19–32, 1973.
- [30] Y. Yoshikawa, K. Shinozawa, H. Ishiguro, N. Hagita, and T. Miyamoto, "Responsive robot gaze to interaction partner," in *Proceedings of robotics: Science and Systems*, 2006.
- [31] R. Bajcsy, "Active Perception," *Proceedings of the IEEE*, vol. 76, pp. 996–1005, 1988.
- [32] K. Daniilidis, C. Krauss, and M. Hansen, "Real-time tracking of moving objects with an active camera," *Real Time Imaging*, vol. 4, no. 1, pp. 3–20, Feb. 1998.
- [33] G. Hoffman, "Dumb Robots , Smart Phones : a Case Study of Music Listening Companionship," in *RO-MAN 2012 - The IEEE International Symposium on Robot and Human Interactive Communication*, 2012.
- [34] M. Slater, A. Sadagic, M. Usuh, and R. Schroeder, "Small Group Behaviour in a Virtual and Real Environment : A Comparative Study," *PRESENCE: Teleoperators and Virtual Environments*, vol. 9, pp. 37–51, 2000.
- [35] R. E. Guadagno and R. B. Cialdini, "Online persuasion: An examination of gender differences in computer-mediated interpersonal influence," *Group Dynamics: Theory, Research, and Practice*, vol. 6, no. 1, pp. 38–51, 2002.