"Improvisation $\neq$ Randomness": a Study on Playful Rule-Based Human-Robot Interactions

Irene Alcubilla Troughton[a], Hendrik von Kentzinsky[b], Maaike Bleeker[c], and Kim Baraka[d]

***Abstract*— To develop and sustain rich social interactions between humans and robots, previous research has mostly looked at task-oriented performance metrics or the ability for a robot to adequately express messages, emotions, or intents. In contrast, our research starts from the premise that movement, as a nonverbal modality of social interaction, can cover other essential aspects of social interaction that do not have to do with the expression of messages, inner states, or drives but that nonetheless contribute to improving the quality of interaction. These aspects have to do with interaction dynamics and highly depend on appropriate action choice. Drawing inspiration from rule-based improvisation, this paper seeks to show that there exists implicit expert knowledge that can be used to inform these movement action choices, contributing to rich, playful, and non goal-oriented interactions between humans and robots. We present an experimental study conducted at a performing arts festival, in which participants interacted with a robot in three simple rule-based movement games, in two conditions: one where the robot was fully controlled by an improvisation expert (*Improv Timing/Improv Action*) and one where the timing of the actions was controlled by the expert but the robot's action choices were drawn randomly (*Improv Timing/Random Action*). This was done in order to focus on action choice, beyond the timing of a response. Our results show that the Improv Timing/Improv Action condition not only performs better in terms of anthropomorphism and animacy, but also increases the interest of people in interacting with the robot for longer periods of time. These results serve as preliminary evidence of how improvisational knowledge in this context contributes to improving the quality of an interaction, and point at the value of further work in this field.**

***Index Terms*— Robot improvisation; Performing arts; Action selection; User studies.**

I. Introduction

An open challenge for the field of Human-Robot Interaction (HRI) is to create and sustain rich and effective social interactions with robots. This is not merely a matter of goal-oriented task performance and efficient communication. It also involves aspects of communication that are not (first and foremost) directed towards communicating a message, expressing an intention or emotion, or performing a task, but about sustaining interest, attention, and a sense of connection via shared involvement. In a recent study, [15] argue that precisely these aspects of communication might be what makes interaction social. Even if the authors still focus on the informational content of the interaction, their observations also indicate the relevance of a broader understanding of social interaction beyond effective communication of messages and intentions. This is not to deny that actual social situations will usually also involve the communication of content and the expression of intentions, yet approaching this from an alternative perspective directs attention to different aspects of social interactions that are also essential for developing HRI.

This embodied knowledge about sustaining interactions beyond the communication of a message, emotion, or intention, is essential for dancers and other performing artists. Experts at improvisational dance hold implicit knowledge about how to engage with each other through the exchange of bodily cues. These cues, though affective and communicative inasmuch as they propose ways of moving and engaging, do not express a pre-determined, clear-cut message or idea but rather work as offers or affordances in an interaction. We hypothesize that tapping into this implicit knowledge will serve the development of richer robot behavior that is able to enhance the quality of interaction with humans, including engagement, flow, responsiveness, etc. Specifically, we focus on *rule-based improvisation*, a set of improvisational methods in which dancers rely on rule-based structures to make choices in relation to the environment, such as the objects surrounding them, other bodies, previous choices, etc. Improvisation in this case is understood not as free form, but as a *choice-making practice within a set of rules of arbitrary complexity*. The goal of these rules is to have the dancer react pre-reflexively to specific input, thus selecting movement choices while minimizing their reliance on internal states, such as drives or intentions. Rule-based dance improvisation provides then a method for choice making that contributes to rich interaction dynamics while moving beyond the informational content of the communication at hand.

With the broad vision of bringing the expertise of these types of improvisers into the HRI space, we tackle the development of under-explored types of social interactions with robots (namely, open-ended, playful interactions). Particularly, we are interested in the choice-making of improvisers in rule-based settings in order to inform future development of decision-making algorithms for a robot autonomously engaging in open-ended interactions with a human in an improvised manner. This work is a first step towards this goal, and serves as preliminary evidence that the implicit knowledge of improvisers indeed improves the

[a]Department of Media and Culture, Utrecht University (corresponding author: +31625571402, i.alcubillatroughton@uu.nl)

[b] Department of Computer Science, Free University of Amsterdam (hvkentzinsky@gmail.com)

[c]Department of Media and Culture, Utrecht University (M.A.Bleeker@uu.nl)

[d]Department of Computer Science, Free University of Amsterdam (k.baraka@vu.nl)

quality of interaction in those contexts. It builds on our previous exploratory work [22] where we examined rule-based improvisations in an semi-structured workshop with dancers trained in Forsythe improvisational technologies [5].

This paper presents a controlled experiment we conducted as part of a Performing Arts Festival, in which a robot interacted with human participants in three rule-based tasks/games covering aspects of multi-modal communication, namely gaze, proxemics and body movement. Each participant underwent two conditions: one in which the robot's choices and timing were controlled by an improviser through a Wizard of Oz interface (Improv Timing/Improv Action condition, from now on IT/IA condition), and another one in which the timing is chosen by the improviser but the responses are selected randomly (Improv Timing/Random Action condition, from now on IT/RA condition). We compared these two conditions on several subjective measures, including the Godspeed questionnaire (Animacy, Likeability, Perceived Intelligence) and measures of interaction quality. With this setup, we show that the way in which an improviser engages with choice-making is essential for creating and sustaining social interactions of good quality.

The rest of the paper is organized as follows. Section II discusses related work on performing arts-inspired approaches to social interactions with robots. Section III describes our study design, whose results are reported in Section IV and discussed in Section V. Section VII concludes and outlines next research steps.

II. Related Work

We briefly discuss existing work in HRI and social robotics drawing from improvisational practices in the performing arts.

A. Improvisational Robots in Performance

Robots have been used in dance performances in several ways. One of them is as a way of aiding humans in improvisation exercises or performances. For instance, [16] used robots to develop Viewpoint exercises for humans (a structured technique for dance improvisation and composition). Another way is to make use of improvisation exercises with robots as a means of gaining inspiration for a, later on, choreographed dance performance, like [4].

Other experiments have tried to imbue robots with improvisational capabilities in order to have them dance with a human performer, such as [24] and [10]. The authors of [24] explored how to program a 6-legged robot to achieve interactive improvisational behavior. According to them, this can take place if the human feels that the robot is in a state of "offer" and if it responds to the human's movement. So far, only one of the limbs has been programmed using Laban Movement Analysis to capture and analyze the human's movement. The authors also programmed three modes (copy, oppose and innovate) that can be organized in "follow" or "random". [10] has also determined a system of structured improvisation for a humanoid robot. In this last case, similarly to us, the author opts for rule-based improvisation techniques and develops an automatized system in which a robot can select from a series of motion primitives based on rules of sequencing. Motion quality is also determined through optimization-based modules that draw from Laban's theory of Effort.

These projects generally aim to improve the performing capabilities of robots on stage (and they are focused on how to inform the humans' improvisations), whereas our research looks for improvisation tools that can be applied in playful and open-ended human-robot social interactions. We follow a similar path to [24] and [10], inasmuch as we rely on structured or rule-based improvisation, but we differ from their approach in our use of the implicit knowledge of improvisers instead of pre-programming meta-rules of choice-making.

B. Improvisation for Better Social Robot Capabilities

Dance improvisation has also been employed as a way of making a robot social in its interactions with humans. Regarding this, two approaches can be found: first, improvisation is used as an inspiration for the human to design robot motion and behavior; and second, the insights of an improviser are accessed in Wizard of Oz settings and later used to program robots.

1) Inspiration for Motion Design: The authors of [6], [7] have developed their approach "Performative Body Mapping" where they use dancers from the Body Weather Technique to improvise interactions with robot prototypes in order to gain inspiration for future robotic motion and behavior. [11] work with LBMS and improvisational practices like Contact Improvisation, as a form of inspiration for robot designers. Finally, [8] experiment consisted of having one mobile robot move with three improvisers from different backgrounds (breakdance, physical theatre, and modern dance). Based on the iterations with dancers, motion algorithms were created.

2) Wizard of Oz Studies: Several experiments tap into the knowledge of improvisers with Wizard of Oz techniques in order to develop social characteristics in robots. One early experiment is [18] and it consists on an improv game called "Relativity", where a robot (a mobile cylinder) reacts showing relative interest of disinterest to a point in space. Their goal is to explore whether a robot can be a creative partner to a human being, and to gather information about how people tend to interpret simple motions. Other projects developed by Knight and colleagues focus on the expressive capabilities of chairbots, and on how improvisers are able to make a chair effectively communicate in live interactions with humans [2]. In further research, the authors automatize those behaviors and conduct more experiments [1], [23].

Wendy Ju has similarly worked on this field for many years and has established an array of collaborations especially with regards to robotic furniture. An early work of hers [20], operationalises her theory of implicit interactions with two studies on how automatic door gestures can convey approachability to users. Ju's theory of implicit interactions [9] claims that people use gestures and patterns of interaction to communicate offers, responses, and feedback that give subtle and often non-conscious information about how to engage in interaction and cooperation. According to her, this embodied knowledge can be gathered and applied to the design of interactive and automated devices without the need of linguistic communication. She has tried to gather this implicit knowledge of improvisers in several projects exploring the motion of robotic furniture with WOZ prototyping [17], [19], [25]. One instance is [14] where the authors investigate patterns of dominant behavior on a robotic footstool by using Keith Johnstone's Improv on how to communicate a "master" and "servant" intent through movement.

This set of projects research how the implicit knowledge of improvisers help in creating meaningful social interactions. In similarity with these projects, we opt for a Wizard of Oz technique to tap into the implicit embodied knowledge of improvisers with regards to social interaction. Having these improvisers then directly controlling the robot allows us to analyze their data of choice-making. In this sense, we build on the previous projects that aim to analyze this information through similar techniques. However, the projects using WOZ to access implicit knowledge mostly focus on expressing a particular communicative intent, whereas we move away from that perspective. Instead, we explore how movement can work as an improvisational cue that creates and responds to offers, that provides affordances for interaction without the need of expressing and deciphering pre-determined inner states.

III. Study Design

In order to evaluate whether the knowledge of rule-based improvisers contributes to creating and sustaining the quality of the interactions among humans and robots, we created a within-subject study with two conditions:

- **Improv Timing/Improv Action Condition**: the expert improviser selects an action for the robot to perform in a Wizard of Oz set up, choosing when and how to respond to the human participant.
- **Improv Timing/Random Action Condition** (baseline): the expert improviser only decides when to respond to the human, but the action is selected at random.

There is evidence from the literature that randomness can go a long way in fostering engagement and surprise, especially in short-term interactions [12], [13], [21]. An example of this can also be found in robotic toys that use randomized responses, such as Furby. That is why, in this short-term setting (3x 3mn), we view a random baseline as an appropriate way of testing our hypothesis. Moreover, there is a long tradition of understanding improvisation as free-form, unscripted unconscious flow, and therefore often perceived as "random", so we also seized the opportunity to contribute to debunking this idea.

Ideally, we would have added a third condition (e.g., "naive improv", in which the robot purely imitates the participant's behavior), but this would have made the sessions too long and caused participant fatigue. The goal of the experiment is therefore to compare expert choice-making to

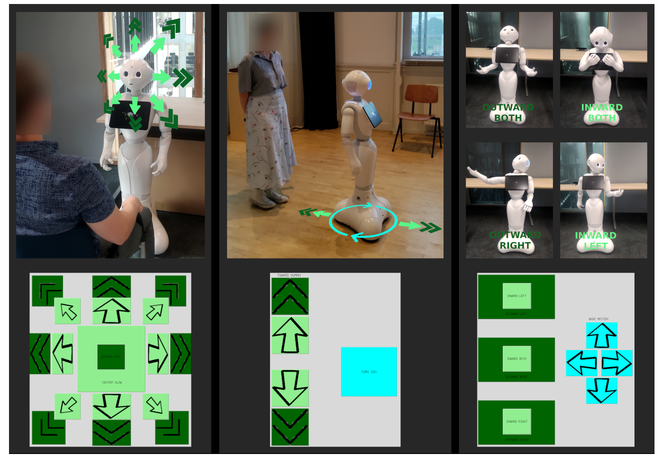

Fig. 1. GUIs of the three tasks and illustrations of selected actions

a "naive" algorithmic solution for choice making. We were interested in using people's knowledge about interaction dynamics and found that improv experts were great wizards when it came to that and provided us with valuable data that could be used in the future to make the robot improvise autonomously, e.g., through imitation learning.

In a nutshell, this study design aims to show that factors besides timing (namely, the improviser's actual action choices) contribute to perceived robot responsiveness, attunement, etc. The decision of which action to trigger based on the history of robot actions and human actions puts to test improvisers' implicit knowledge on how to create interaction dynamics in this playful, open-ended context (including imitating, contrasting, initiating, following…).

Each participant interacts with the robot in the two conditions during three rule-based tasks or games: the first one involves gaze; the second, proxemics; and the third is an open-ended task (see Section III-B). We expect the Improv Timing/Improv Action condition to perform better on most of our subjective measures (see Section III-E), indicating a better quality of interaction.

A. Experimental setup

The experiment took place in a performing arts venue over the course of three days, where participants interacted with a Pepper robot. The Wizard of Oz setup was achieved by segmenting the experiment room with a partition, in front of which there was the robot and behind of which there were the wizard controlling the robot and two members of the research team. To see what was unfolding in front of the partition, the wizard had access to a screen that showed the live-stream of a camera recording the interaction of the participant with the robot. Said webcam was placed on top of the partition. To control the robot, the wizard made use of a tablet showing a GUI consisting of buttons (see Fig. 1). By pressing those buttons, a script that connects to the robot via a Wi-Fi router was executed, which in turn would start a respective action in the robot. The Python code for the GUI and connecting to the robot is publicly available [1]. Both conditions had the same interface, and, in the IT/RA, the wizard was instructed to pretend controlling the robot with the GUI and select a fitting action, despite knowing that each button would execute an action at random. This way both conditions would have similar reaction times and only the selected action would differ.

B. Task description

The experiment consisted of three tasks that were repeated in each condition (see Fig. 1). The tasks were organized from more to less structured, the last one being an open-ended interaction in which the human was not subjected to rules. This last one was an attempt to replicate daily encounters with robots, and to research how improvisation cues could still create and sustain an interaction even if the human is not aware of following rules. For all tasks, pressing a new action overrode any running action in order for the transitions in actions to be quick and easy to spot and therefore making the robot appear more responsive. The Choregraphe project, which contains the robot's behaviors, is publicly available[1].

Task 1 was a turn-taking game concerning head movement. The participant was sitting on a chair, facing the robot roughly at eye level. The participant was instructed to move their head vertically, horizontally, diagonally or to the center, then hold the position and wait for the robot to respond, and then repeat by carrying out another action. The wizard could move the robot's head in the same directions, but in two different speed settings: normal or fast.

Task 2 was a turn-taking game involving proxemics. The participant was facing the robot, this time standing up, and had four possible actions to choose from: one step towards other, one step away from other, turn away 180°, and turn back 180°. Similarly to the previous task, the robot could perform the same movements as the human, and in two different speed settings (normal or fast). For turning away, the robot would look over its shoulder to keep the connection with the participant. The participant was instructed to do the same while turning away in order to keep watching what the robot was doing.

Task 3 was comprised of an open-ended interaction. The participant was specifically instructed to forgo any rules and turn-taking-restrictions, and just interact freely with the robot. The only limitation was that they had to move within a specifically demarcated space, so that the camera could still record them. For moving in space, the wizard could choose from eight actions to move in cardinal directions in a slower or faster speed. For upper body movement, the wizard had six categories to choose from: An inward/outward oriented motion directed towards the center/left/right. Each category came with two different actions, to afford more variety, such that e.g. upon pressing the "Inward left" button one out of two actions would be selected. These actions were previously choreographed by a professional puppeteer and were intended to show an open/inviting gesture vs. a closed/self-directed gesture. We intentionally decided to authorize movements that would not be interpreted in clear social cues (e.g., waving a hand) or that would not inspire a particular emotional expression (e.g., being sad), as we want to analyze how the robot's actions invite movements and contributes to the interaction without relying on the expression and deciphering of intentions and feelings. Being an open-ended interaction, we needed more creative and varied choices than gaze and proxemics to offer the participants enough to react to, reason why we decided to authorize more movement choices. In case of moving in space, an action was programmed as moving 0.5m in the respective direction at uniform speed. For the upper body movement, each action was programmed as a single set of joint angles.

The actions generated in both conditions, therefore, are predefined motions: either simple orientation motions or more complex movements created with the help of an expert puppeteer. It is important to emphasize that the knowledge that we were interested in capturing did not depend on action complexity or number. As such, we aim to draw the participants to the dynamics of the interaction, not to the movement generation of the robot.

C. Procedure

Each participant was allocated a 30-minute slot, which began by them being welcomed by a research assistant at the entrance of a room. Participants signed an informed consent form giving the researchers permission to use the data gathered during the interaction with the option to refuse and/or withdraw at any point without consequences (the study was approved by the university's ethical committee before the experiment took place). If they did not consent, they could still experience the robot on the IT/IA condition. The researcher then explained the procedure: the participant interacts with a Pepper robot during three tasks, fills a short questionnaire assessing their experience and perception of the robot on a table in the corner of the room, and then comes back to repeat those three tasks with the same robot, but this time with a different programming. Lastly, the participant fills a new questionnaire outside of the room. At the end of the experiment, they were debriefed by the assistant, told that the robot was controlled by an out-of-sight operator, and asked again to give their consent.

During each task, the researcher explained in detail the rules of the game, and then held a short demonstration where the participants could practice the movements and the robot would imitate them, in order to familiarize themselves with the movements and the robot's capabilities. After the demonstration, the participants were told that each task would last three minutes, during which the researcher and technical staff would be behind a partition and only intervene if there was a technical failure. it was also emphasized that during the task the robot would not necessarily imitate them as in the demonstration, but that it would react to their behavior. During the procedure, a go-pro camera and a webcam recorded the session with the participants from

[1] https://www.dropbox.com/sh/oxpt20ml2qcltjr/AAAzMajUTg-F76zPRE-c5F0Ha?dl=0

two different angles. This data, along with the questionnaires (see Section III-E), will be used for dissemination purposes. Finally, the puppeteers were prompted on how to manipulate the interface, and how to puppeteer the robot, before the session and in a previous try-out (see puppeteers' manual in the supplementary material linked at the end of the paper).

D. Participants

The experiment was advertised as an interactive installation within the frame of a performing arts festival. Participants could sign up for their slot through a scheduling link in the page of the festival that announced the installation. Anyone above 18 years old could participate, with no restrictions, however we expected most of the interested participants to be related to the field of performing arts. 23 people signed up in total: 10 on the first day; 7, on the second and 6 on the third. One of the participant had to be excluded, as he didn't follow the given instructions, which resulted in data for 22 participants. Among those 22 participants, 55% and 41% stated they were female and male respectively; 64% and 32% stated they were of Dutch or non-Dutch nationality respectively. The average age was 37.6, ranging from 20 to 74. 41% stated that they had interacted with a robot beforehand. 73% and 82% were characterized as having experience in watching or acting in performing arts, respectively.

We decided to conduct a within-subject, as we considered that the participants would be better equipped to judge a novel situation when faced with a comparison. Moreover, we tested this experiment on non-experts, as we aim to use this knowledge to further develop open-ended interactions between robots and humans in daily contexts. Our study thus analyzes how improvisation experts are able to provide offers and respond to movement choices in ways that increase the overall quality of the interaction and that provide more interesting and engaging options for non-experts.

Regarding the wizards, all three wizards came with improvisational experience. Wizard 1 and 2 were professional dancers and wizard 3 was a professional puppeteer. We selected more than one improviser to address issues of generalizability. Ideally, we would have liked to include more wizards but due to practical constraints, we opted for three.

E. Measures

We logged each action the puppeteer took, i.e. pressing a button on the tablet, with a respective timestamp. Thereby we could compute further features, e.g. actions per minute for each puppeteer and action selection distribution. For measuring the participant's perception of the robot, we seize on the established Godspeed Questionnaire Series [3], which rates contrasting word-pairs on a five-point-scale regarding Anthropomorphism, Animacy, Likeability, Perceived Intelligence and Perceived Safety (e.g. fake/natural or dead/alive). As Likeability and Perceived Safety were not relevant in this context, we constrained our questionnaire to the 16 items regarding Anthropomorphism, Animacy and Perceived Intelligence. Additionally, the questionnaire contained a series of statements to be rated on a five-point scale as well. Furthermore, we added four open-ended questions at the end of the questionnaire, asking the participants how they would characterize the interaction, how they experienced the robot's demeanor, and about moments/aspects they found more or less engaging/interesting (the questions can be found in the supplementary material).

IV. RESULTS

We report here on the analysis of the questionnaires and the logged data. First, the numerical items of the questionnaire (the Godspeed items and the statements) were analyzed statistically, both across conditions and across wizards. Additionally, the logged data was used to discern further differences among improvisational styles.

A. Comparison of conditions

Fig. 2 shows the average scores on the Godspeed items and the questionnaire statements. Overall, the IT/IA condition ranks higher for all measures. To evaluate whether these differences are statistically significant, a paired-t-test was conducted with a Shapiro-Wilk test as a normality check. The differences on the Godspeed items were found to be normally distributed, but for all the statements, except "Repeated Interaction", the hypothesis of normality was rejected. For those statements where the t-test assumptions did not hold, a Wilcoxon signed-rank test was conducted instead. The t-test showed a significant difference in "Anthropomorphism"($t(21) = 2.81$, $p = 0.01$), "Animacy" ($t(21) = 3.69$, $p = 0.001$) and the statement "Interact again"($t(20) = 2.97$, $p = 0.008$), but no significant difference in "Perceived Intelligence"($t(21) = 0.98$, $p = 0.339$). The Wilcoxon signed-rank test prompted only for "Robot Perception" ($W(21) = 4$, $p = 0.013$) a significant difference, however "Creativity"($W(20) = 10$, $p = 0.052$) and "Engagement"($W(21) = 27.5$, $p = 0.056$) barely missed the 5% mark. "Surprise"($W(21) = 13$, $p = 0.25$) and "Flow"($W(21) = 39$, $p = 0.115$) were found to have the least distinct, non-significant differences.

B. Individual differences among wizards

Fig. 3 shows the average scores on the Godspeed items per puppeteer for the IT/IA condition. Overall, there are little differences in the IT/RA condition and more considerate differences in the IT/IA condition - which was to be expected, as the puppeteers had more control there and therefore more opportunities to express different puppeteering styles. However, to evaluate whether these differences are actually significant, a 2-factor ANOVA was conducted - the two factors (variables) being which condition and which puppeteer. A linear model was fitted, using OLS (ordinary least squares), to explain the target variable, which is respectively one of the Godspeed items or questionnaire statements. Significant differences were found among all Godspeed items("Anthropomorphism"($F(2) = 3.577$, $p = 0.038$), "Animacy"($F(2) = 3.872$, $p = 0.029$), "Perceived

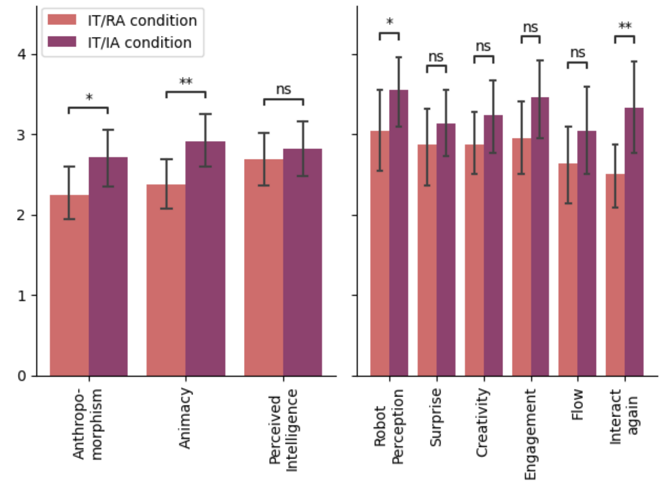

Fig. 2. Mean scores of the Godspeed items (left) and the questionnaire statements (right) per condition, with 95% confidence intervals. Significance shown according to a paired t-test / Wilcoxon signed-rank test: ns: $p > 0.05$; *: $0.01 < p \leq 0.05$, **: $0.001 < p \leq 0.01$. Labels of the questionnaire statements: "Robot Perception": The robot was perceiving what I was doing, "Surprise": The robot was surprising in its choices, "Creativity": The robot was creative in its choices, "Engagement": I felt engaged in the interaction, "Flow": I felt like the interaction was flowing, "Interact again": I would like to interact with the same robot again over a longer period of time.

Intelligence"($F(2) = 3.789$, $p = 0.032$)), whereas all differences among the statements ("Robot Perception"$F(2) = 0.794$, $p = 0.459$, "Surprise"($F(2) = 1.88$, $p = 0.167$), "Creativity"($F(2) = 2.157$, $p = 0.13$), "Engagement" ($F(2) = 1.714$, $p = 0.194$), "Flow"($F(2) = 2.188$, $p = 0.126$), "Interact again"($F(2) = 1.485$, $p = 0.24$)) were found to be non-significant.

We ran a per-wizard t-test, which showed statistically significant differences in 3 out of 9 measures for wizard 3 (puppeteer), and only one measure for wizards 1 and 2 (improvisational dancers). Keeping in mind uneven sample sizes per wizard, this result suggests that the puppeteer may have been more comfortable with the setup and/or had a higher level of expertise.

Fig. 3 shows the average amount of times the wizards would choose an action that is mirroring, contrasting or complementing the last action of their respective participant for task 1 and 2 in the IT/IA condition. The data was collected by one of the authors by going through the video material and manually labeling when the participant started an action. The actions of the participants were then being aligned time-wise with the logged actions that the wizards took. An action was considered to be mirroring, if it was exactly the same action, and contrasting, if it was exactly the opposite action (e.g. move-forwards and move-backwards, look-down and look-up). For task 1 only, actions were categorized as complementing, too. An action was considered to be complementing, if it relates in an 45° angle to the participant's action (e.g. look-down and look-down-left, look-down and look-down-right). Hereby, only the direction of an action was considered, and not the velocity of it. Any action not

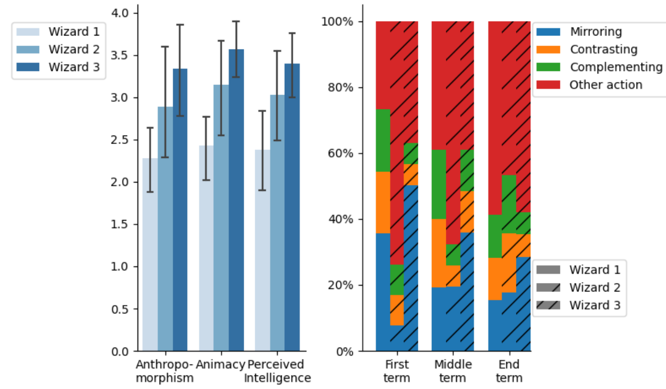

Fig. 3. Left: Mean scores of the Godspeed items per wizard for the IT/IA condition, with 95% confidence intervals. Right: Choice-making per wizard, averages of all sessions, sessions categorized into first/middle/end terms.

considered to be mirroring/contrasting/complementing was lumped into the category "Other actions".

V. Discussion

A. Comparison of conditions

Overall, the IT/IA condition performed better than the IT/RA condition, even though only marginally and non-significant for some of the variables (cf. Fig. 2). For the variables of the Godspeed questionnaire, the IT/IA condition scored significantly higher in anthropomorphism and animacy, which might indicate that the expertise of improvisers help in making the robot be perceived as more responsive, aware, and lifelike. As for the statements that we provided, one significant result was that the participants felt the robot in the IT/IA condition to perceive what they did more than in the IT/RA condition. This indicates the importance of selecting a fitting action as response and not only react in a timely manner, as the timing was similar in both conditions. In engagement, surprise, creativity and flow, the IT/IA condition did not score significantly higher.

Finally, the most standing out difference of the questionnaire statements is that participants rather would interact with the IT/IA condition again - complementing the condition's overall better performance. To sum up, our hypothesis that the IT/IA condition would perform better in our selected variables is partially confirmed, as it scored higher in Anthropomorphism and Animacy in the Godspeed Questionnaire, and people wanted to interact again and longer with the robot in that condition.

B. Individual differences among wizards

The results found that the wizards scored differently, and statistically differently, on the Godspeed items (cf. Fig. 3). Together with different styles in choosing mirroring/contrasting/complementing actions (cf. Fig. 3), this is indicating how different control styles lead to different results. Nonetheless, the wizard with the best scores had the lowest sample size, thereby a higher variability and less of an influence on the average scores on the Godspeed items and the questionnaire statements(cf. Fig. 2). However, more research is needed to understand the underlying differences in style, as, for comparing the wizards actions, the participant's actions have to be incorporated in order to understand the patterns of how a wizard is responding.

C. Limitations

The first limitation in our experiment is the differences in styles of puppeteering, which have influenced the ratings that the robots had in each category. Furthermore, even if we tried to avoid a delay in the IT/IA condition by having the puppeteers using the same interface and instructing them to select a fitting response, regardless of the result, it might have influenced the timing. Finally, also other factors influenced the perception of the robots capabilities, especially the weather: on day 1 and 2 the sunlight was troubling the Lidar sensors, leading the safety mechanism of the robot to sense that something would be in the way and therefore not moving occasionally, whereas on day 3 it was mostly rainy and the performance was significantly better; Task 1 was, however, unaffected.

VI. Lessons Learned

This section reports on the qualitative information of the open-ended questions, and the experience of the puppeteers and the members of the team that were observing the interactions. It therefore serves as a reflection on the experiment, as well as a commentary on future steps.

Participants commonly mentioned as engaging when the robot took the leader role in an interaction dynamic and proposed something new to the current movement pattern. This happened in two ways in the IT/IA condition: when the robot actively proposed a new movement, and when it stayed in place or still, thus rejecting the offer made by the human. People also reported feeling inclined to copy the robot's movements, stepping into a follower role. This was also perceived as a tendency by the members of the team and the puppeteers.

The fact that people tended to follow the robot was also mentioned in the IT/RA condition, although comments that would reflect on how the robot actively took on a leader position were less common. A couple of interesting situations took place during this condition in which the robot, because of the random selection of choices, would end up sometimes moving away from the allocated space. This action, which had to be corrected, was encountered with amusement by several participants, and in one case it was reported on in the questionnaire. A similar situation also arose in our previous research with expert improvisers interacting with a Pepper robot (Alcubilla et al., 2022). The improvisers experienced the moments in which the robot was refusing to follow the humans as an intentional "stubbornness" rather than a malfunction, even when it was so. We hypothesize that this has to do with the broader system in which the robot is embedded. That is, when the robot "stepped out of line" and it "had to be corrected" (participant's words), people of the team would intervene and put him back in the right space, apologizing for the inconvenience, which might have contributed to the perception of the robot as being unruly. This points at the importance of the "staging" of an interaction beyond the traditional robot-human dyad, thus paying attention to how things are framed by other humans, the space, and other elements (including random malfunctions and how they are conceptualized by people around it).

With regards to leader-follower dynamics in the third task, there was a clear difference between two types of participants, which was observed by the puppeteers and members of the team: those who had a tendency to be in the leadership position, and those who were on a follower position and more open to co-create something with the robot. In the first case, participants would try to test the robot, measuring the interaction in terms of whether or not the robot responded according to their expectations. More creative moves or initiations on the part of the robot were not highly appreciated. In the second case, people seemed to have less clear expectations, being more open to the robot's offers. This made us hypothesize about the possibility of having different styles of improvising, one more "adventurous" and another more "conservative", depending on the participant's tendencies.

The cases in which the robot was leading, it was perceived as having an intentionality (e.g., the robot was making "its own choices", or "suggesting something"). This intentionality, instead of being linked to an interiority that was being expressed, was mostly understood in terms of an offer. That is, the robot was clearly communicating in certain circumstances a proposal for moving or not moving, and that was perceived as an engaging aspect of the interaction. In relation to this, participants mentioned the importance of understanding how the robot was responding to them, and also the offers that the robot gave them. Creativity could therefore be appreciated once there was a base of legibility. Participants did not need to know exactly which patterns the robot was choosing (and, up to an extent, those choices needed to remain surprising), but they did have to perceive a sense of coherent responsiveness in order to remain engaged; otherwise they reported feelings of uneasiness and unpleasantness. With regards to this, the puppeteers admitted starting the interaction with some mimicking or clear opposition of the participant's behavior, to create a legible and clear connection before turning to more creative choice-making.

Finally, feeling connected to the robot was a factor that increased the interest in the interaction. Moreover, physical and affective proximity was almost always interlinked, as participants mentioned feeling more connected to the robot whenever it would come close to them. Being conscious of this affective aspect of proxemics will also be essential in further research.

VII. Conclusion

Our research contributes to the improvement of social capabilities of robots in improvised scenarios by tapping into the implicit embodied knowledge of rule-based improvisers.

Our study provides a first step into researching the patterns, or meta-rules of sequencing in interactions that deal with movement beyond the expression of messages, intentions or drives. To this end, we developed an experiment to prove that the implicit knowledge of improvisers with regards to choice-making improves the quality of social HRI. Our results partially confirm our hypothesis, showing that people perceive the robot as more anthropomorphic and animate in the IT/IA condition, and that they would like to interact with it longer. Moreover, the open-ended questions gave us essential information as to which elements were contributing to a sense of engagement in the interaction. Namely, a sense of initiative and intentionality whenever the robot would propose an offer, creative choice-making once a base of legibility was established, playing with modes of physical and affective connection, and being conscious of the "staging" of the interaction.

Future work will include analyzing and annotating the collected video material in order to develop a decision-making algorithm for the robot. By annotating which actions the participants made, in conjunction with the participant's perceptions of the robot and the puppeteer's actions, we plan to extract patterns of choice-making, as well as instances in which the patterns are broken. More research is also needed to understand the underlying differences in styles in puppeteering and how they affect the interaction. We believe our results can aid in informing algorithms for robot choice-making in improvised, open-ended, playful situations. Even if we are aware that this knowledge cannot be applied in every HRI scenario, we do believe that our experiment shows preliminary evidence of how improvisational knowledge in this context contributes to improving the quality of the interaction, and points at the value of further work in this field.

VIII. Acknowledgements

This research is part of the project "Acting like a Robot. Theatre as Testbed for the Robot Revolution" (funded by NWO – Creative Industries: Smart Culture – Arts and Culture, GW.000885.2

References

[1] Abhijeet Agnihotri and Heather Knight. Persuasive chairbots: a (mostly) robot-recruited experiment. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 1–7. IEEE, 2019.

[2] Abhijeet Agnihotri, Alison Shutterly, Abrar Fallatah, Brian Layng, and Heather Knight. Chairbot café: Personality-based expressive motion. In *adjunct Proceedings of the 13th Annual ACM/IEEE International Conference on Human-Robot Interaction. ACM*, 2018.

[3] Christoph Bartneck, Elizabeth Croft, and Dana Kulic. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, 1(1):71–81, 2009.

[4] Catie Cuan, Ishaan Pakrasi, and Amy LaViers. Time to compile: An interactive art installation. In *16th Biennial Symposium on Arts & Technology*, volume 51, page 19, 2018.

[5] William Forsythe and Christine Bürckle. William forsythe, improvisation technologies: A tool for the analytical eye. *Germany: ZKM, digital arts editions*, 1999.

[6] Petra Gemeinboeck and Rob Saunders. Towards a performative body mapping approach. In *The AISB50 Online Conference*, 2014.

[7] Petra Gemeinboeck and Rob Saunders. Towards socializing non-anthropomorphic robots by harnessing dancers' kinesthetic awareness. In *International Workshop on Cultural Robotics*, pages 85–97. Springer, 2015.

[8] Elizabeth Jochum and Jeroen Derks. Tonight we improvise! real-time tracking for human-robot improvisational dance. In *Proceedings of the 6th International Conference on Movement and Computing*, pages 1–11, 2019.

[9] Wendy Ju. The design of implicit interactions. *Synthesis Lectures on Human-Centered Informatics*, 8(2):1–93, 2015.

[10] Amy LaViers. Programmed improvisation inspired from autonomous humanoids. 2019.

[11] Amy LaViers, Catie Cuan, Catherine Maguire, Karen Bradley, Kim Brooks Mata, Alexandra Nilles, Ilya Vidrin, Novoneel Chakraborty, Madison Heimerdinger, Umer Huzaifa, et al. Choreographic and somatic approaches for the development of expressive robotic systems. In *Arts*, volume 7, page 11. MDPI, 2018.

[12] Tuck Leong, Steve Howard, and Frank Vetere. Feature take a chance on me: using randomness for the design of digital devices. *Interactions*, 15(3):16–19, 2008.

[13] Tuck Wah Leong, Frank Vetere, and Steve Howard. Randomness as a resource for design. In *Proceedings of the 6th conference on Designing Interactive systems*, pages 132–139, 2006.

[14] Jamy Li, Andrea Cuadra, Brian Mok, Byron Reeves, Jofish Kaye, and Wendy Ju. Communicating dominance in a nonanthropomorphic robot using locomotion. *ACM Transactions on Human-Robot Interaction (THRI)*, 8(1):1–14, 2019.

[15] Mike EU Ligthart, Mark A Neerincx, and Koen V Hindriks. Core elements of social interaction for constructive human-robot interaction. *arXiv preprint arXiv:2110.04054*, 2021.

[16] David V Lu, Chris Wilson, Annamaria Pileggi, and William D Smart. A robot acting partner. In *ICRA Workshop on Robots and Art*, 2011.

[17] Brian K Mok, Stephen Yang, David Sirkin, and Wendy Ju. Empathy: interactions with emotive robotic drawers. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pages 250–251, 2014.

[18] Jesse Rond, Alan Sanchez, Jaden Berger, and Heather Knight. Improv with robots: creativity, inspiration, co-performance. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 1–8. IEEE, 2019.

[19] David Sirkin, Brian Mok, Stephen Yang, and Wendy Ju. Mechanical ottoman: how robotic furniture offers and withdraws support. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, pages 11–18, 2015.

[20] Leila Takayama, Doug Dooley, and Wendy Ju. Expressing thought: improving robot readability with animation principles. In *Proceedings of the 6th international conference on Human-robot interaction*, pages 69–76, 2011.

[21] Daina Teranishi and Masahiro Araki. Improving user engagement with dialogue systems through meaningful response generation. In *Proc. 1st RobotDial Workshop on Dialogue Models for Human-Robot Interaction*, pages 2–8.

[22] Irene Alcubilla Troughton, Kim Baraka, Koen Hindriks, and Maaike Bleeker. Robotic improvisers: Rule-based improvisation and emergent behaviour in hri. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 561–569. IEEE, 2022.

[23] Jeremy Urann, Abrar Fallatah, and Heather Knight. Dancing with chairbots. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 364–364. IEEE, 2019.

[24] Mick Wallis, Sita Popat, Joslin McKinney, John Bryden, and David C Hogg. Embodied conversations: performance and the design of a robotic dancing partner. *Design Studies*, 31(2):99–117, 2010.

[25] Stephen Yang, Brian Ka-Jun Mok, David Sirkin, Hillary Page Ive, Rohan Maheshwari, Kerstin Fischer, and Wendy Ju. Experiences developing socially acceptable interactions for a robotic trash barrel. In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 277–284. IEEE, 2015.

IX. Supplementary material

Supplementary material containing more details about the study can be found at [https://bit.ly/supplementarymaterial].