Perception of Control in Artificial and Human Systems: A Study of Embodied Performance Interactions

Catie Cuan, Ishaan Pakrasi, and Amy LaViers

Mechanical Science and Engineering Department,
University of Illinois Urbana-Champaign,
Urbana, IL 61801, USA
crcuan@gmail.com, pakrasi2@illinois.edu, alaviers@illinois.edu

Abstract. Robots in human facing environments will move alongside human beings. This movement has both functional and expressive meaning and plays a crucial role in human perception of robots. Secondarily, how the robot is controlled – through methods like movement or programming and drivers like oneself or an algorithm – factors into human perceptions. This paper outlines the use of an embodied movement installation, “The Loop”, to understand perceptions generated between humans and various technological agents, including a NAO robot and a virtual avatar. Participants were questioned about their perceptions of control in the various agents. Initial results with human subjects demonstrate an increased likelihood to rate a robot and a robotic shadow as algorithmically controlled, versus a human performer and a human-shaped VR avatar which were more likely rated as human actor controlled or split between algorithm/human actor. Participants also showed a tendency to rate their own performance in the exercise as needing improvement. Qualitative data, collected in the form of text and drawings, was open-ended and abstract. Drawings of humans and geometric shapes frequently appeared, as did the words “mirror”, “movement”, and variations on the word “awareness”.

Keywords: Human Robot Interaction · Robot Performance · Embodied Learning · Virtual Reality

1 Introduction

As humans continuously welcome new forms of technology into their quotidian experience, the mysterious intricacies of how those technologies affect their lived, embodied experience, and resulting actions, is a question for researchers and artists. In many cases, perceptions of robots varied based on cultural context as well as previous exposures to robots in media [3, 4]. As that media-based, disembodied experience is supplanted with daily life experience alongside robots, perceptions will change. Moreover, an understanding of how humans perceive robots requires multiple contexts of interaction, social influence from others, and knowledge of humans' own perception of self.

The SoftBank NAO, a widely-used humanoid [35], has numerous applications in education and socialization for children [12], including in autism research [36]. Further research has examined the effects of NAO's motion [14] and conversational capacity [19] on human interactions with the robot, showing that non-verbal cues enhance interaction with the NAO.

Previous researchers have used a Kinect in conjunction with a robot to track human movement in an interactive environment. This has occurred in research detecting elderly adults for elder care [26] as well as in robot teleoperation [11, 32]. Researchers have also employed virtual reality and mixed reality systems to teleoperate robots [42, 34]. The overarching methodology in these teleoperation systems is to map ranges and poses of human movement onto robot actions, outlining rules of translation, limits, and parameters (such as speed).

There will be an estimated 82 million virtual reality (VR) headset users worldwide by 2020 [21]. In contrast, there will be an estimated 1.68 billion TV users by 2021 [37]. Therefore virtual reality remains a relatively novel technology for most people. Virtual reality has applications in stroke rehabilitation [23] and K-12 instruction [28] as well as entertainment and gaming. Many of these applications involve virtual characters, or facsimiles of humans in VR, that play pivotal roles in the content and scenarios of the simulation, thus defining the user experience. Earlier work towards building expressive virtual characters has remarked upon the importance of emotion, non-verbal communication, and personality in virtual characters, not only speech-based communication [41]. Trust of virtual characters may vary based on previous exposures to VR and other forms of virtual characters. Additionally, theory that people generally treat computers as social actors [33] supports the notion that VR users may themselves be emotionally affected by the VR environments they engage with [7].

Embodied interaction with various technologies like VR is also highlighted as a pedagogical strategy. Immersive simulations were used to teach students about science and understand their misconceptions about the field [25]. An immersive VR environment was used to teach physics - the sensation of presence and gesture control (increased embodiment in learning) led to better retention of certain types of knowledge when used during the encoding phase [18, 16, 17]. Several embodied learning studies show that the body plays a critical role in thinking, reasoning, and retention [1].

Embodied dyadic interactions which share similar behavior patterning have been shown to increase positive arousal and valence in different contexts, such as a mother and infant [38], structured argument pairings [22], childhood playgroups [5] and opposite-sex romantic dates [13]. Tendency towards similarity may be conscious or conscious as seen in [8]. Embodied dyadic interactions frequently appear in dance and dance therapy. In one study, posture sharing, or intentioned dyadic movement mirroring, was found to increase positive assessments of self-performance and pair performance [30, 2]. The "mirror game" is a form of an improvised dyadic interaction between two movers, often dancers, without a leader. The two parties stand facing one another and attempt to remain synchronized in their motion. Mirror games between dancers, improvisers, and musicians without leaders have been shown to create more complex, novel motion than those with a clearly defined leader [31].

Humans have a demonstrated tendency towards behaving in a manner which increases or maintains positive self-evaluation [40]. Additional studies have shown that this effect increases when in proximity to others (social comparison) and when comparing ones own performance in one field or task towards another (dimensional comparison) [39].

In [27] and [24], we see that human interaction with technology can often lead to frustration. [6] outlines how this frustration develops based on prior experiences, psychological characteristics, and social systems, and often occurs when user expectation does not match interaction outcome. In the case of moving robots, an intentional design of movement that takes into consideration the end user's past characteristics could lead to an interactive experience that matches user expectation, and thus is less frustrating.

Many dance historians see the study of dance as a critical aspect of human anthropology [20]. Dance frequently appears in social interaction across cultures, as a method of signaling adulthood (for example, during quinceaneras - [9]), passing down folktales (a frequent practice for Native American tribes [29]), and selecting a romantic partner (in the case of high school students [15]. The socialness of dance is further evidenced by the common pairing of dance and live music, creating interactions between the dancers and the musicians (whether in performance or at recreational events).

An early version of "The Loop", an interactive, embodied installation, was outlined in [10]. The instantiation in this paper modifies and builds upon that earlier work towards the question on understanding how humans perceive robots and how moving tasks with robots versus computer-based tasks with robots alter human perceptions of the robots in question. Section 2 discusses the artistic themes and research questions explored; Section 3 explains the user study design; Section 4 presents the results; and Section 5 summarizes the work.

2 Artistic and Research Motivation

Artistic themes and research motivations were initially outlined in [10] and are extended for context below. "The Loop" is a subsection of "Time to Compile", a collaboration between an artist and a robotics lab. The three central themes of "Time to Compile" are the following:

- *Are humans becoming more robot-like?* Humans' rich subjectivity and susceptibility to change renders them more likely to alter their behaviors and interactions when working with machines. Machines are less dynamic in this capacity and therefore, less likely to change quickly.
- *The hidden human network.* While various novel technologies possess an inherent mysticism for the uninitiated user, all machines and technologies are built by human beings. Unmasking this network is central to the piece.

– *Time to Compile.* The vastly different compile times for a short dance versus a computer program reveal the human experience of working across disciplines and bodies, as well as the frustration that may occur when relying on a tool, such as a machine, rather than one's own bodily functionality.

In addition to these three central themes, "The Loop" was designed for this instantiation with additional research questions in mind: How does a movement-based task with a robot differ than a programming-based task with a robot? What are the perceptions of control for each? Finally, how does self-evaluation change based on the degree to which a task is embodied? In posing these questions, the installation was a symbiotic research and artistic initiative, allowing each to inform the textures and experiences of the other.

The design elements of "Time to Compile" include live dance, text, music, projection, film, and props. "The Loop" experience shared similar elements and was designed to feel abstract, continuous, meditative, and welcoming. The music, ushers, lighting, and verbal instructions purposely supported these sentiments. For example, the usher noted "There is no right or wrong way to do this. Simply make choices when you are unsure of what to do."

3 Study Structure

The study structure was as follows: Study participants were told that an artistic installation and experiment was open for a fixed time block. They were invited to arrive at their leisure and participate. Once study participants arrived, they filled out a consent form and were instructed to wait in an area outside the installation. One by one, they were guided by an usher to five stations, where they were prompted to "Follow the motion of the agent you see in front of you." They interacted at each station for 50 seconds, prompted by the beginning and ending of an amplified musical track. In some cases, they could see the next stations, and in other cases, the station was hidden from view. A minimum of one and a maximum of five participants could be in the installation at one time. The agents at each station are listed below with the controlling element described in italics, Figure 1 shows the structure.

– Station 1: A NAO robot. *The NAO robot performed a repeated loop of pre-programmed movements through the software program Choregraphe.*
– Station 2: A video projection of a NAO robot, at larger scale. The screen was roughly 7 feet high, making the projected shadow roughly six feet tall. *The video projection was a pre-recorded video of the same movement sequence from Station 1.*
– Station 3: An HTC VIVE headset broadcasting a virtual avatar in the shape of a person. *The VR avatar's movement was controlled by a Microsoft Kinect positioned to capture the Station 4 participant's movement.*
– Station 4: A moving human performer ("Performer 1"). *Performer 1 watched a live Skype feed from the moving participant at Station 2 and attempted to copy the Station 2 participant as closely as possible. Performer 1 occasionally improvised movement when the Skype connection was poor.*

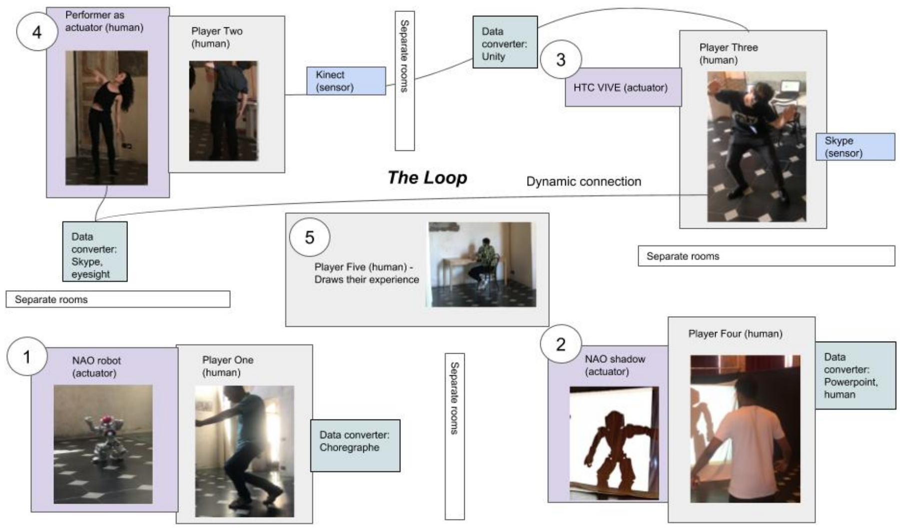

Fig. 1: "The Loop" station agents and their relevant software. The installation was spread across three rooms. In order, each station was separated by a wall, though Stations 1 and 3 were in the same room. The solid connecting line shows a dynamic, live connection between Stations 3 and 4.

Station 5 differed from the other stations: participants were invited to draw a picture of their experience. At the end of all five stations, participants then filled out a survey with the following question for each station: "At the (station number) station, what do you think was motivating the movement of the (agent in motion - for example, at Station 1 it would have been "robot")? (Select more than one if you would like.)"

4 Results

There were 19 total participants. The experience lasted roughly 10 minutes for each participant. Each participant completed a survey and a drawing; the survey and drawing were not paired together with the participant number. Over half the time, participants rated the NAO and shadow of the NAO at Stations 1 and 2 as being algorithmically controlled. Performer 1 was most frequently rated as being controlled by a human actor. The VR avatar was split between human and algorithm in terms of control.

It is noteworthy that at all of the stations, human actor and automated algorithm appeared in the highest and second highest positions. Additional work is necessary to understand if this distribution occurred because participants believe the human actor may be working in conjunction with or as the creator of the algorithm, or if participants were undecided or mystified between the two. Figure 2 shows the distribution of responses at each station. In actuality, the

NAO and shadow of the NAO at Stations 1 and 2 were controlled by a prerecorded algorithm designed and programmed by an artist. The VR avatar and human performer in Stations 3 and 4 were live elements controlled by the other participants as well as Performer 1's movement interpretation and influence.

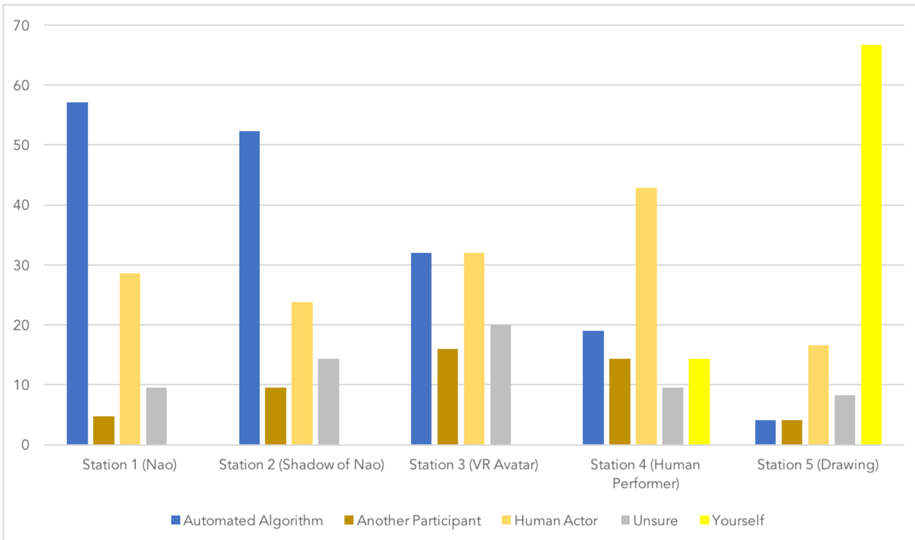

Fig. 2: Perceptions of control for the agents at each station. Each participant selected an average of 1.17 choices at each station. Station 3 (VR avatar) and Station 5 (drawing) each garnered 24 responses, all other stations garnered 21. The values shown in this graph are on a 100 point scale - computed by dividing the number of times a particular response appeared by the total number of checked responses for that station. These choices were discerned by participants who did not know the structure of the installation but could interpret based on what they were seeing at each station.

Samples of the drawings from Station 5 are highlighted in Figure 3. The drawings provided a opportunity for creative, qualitative feedback that may provide further insight into the subconscious aspects of the participant's experience. Human shapes appeared in 16 of the 19 drawings. 3 of the 19 drawings featured arrows pointing between different human shapes, alluding to the connections between the different stations. An additional 3 drawings showed humans with smiling expressions, while 4 drawings featured humans with confused or frowning expressions. This may be linked to participants' feedback that the exercise was challenging, or that the "movements were sometimes hard to follow". Figure 4 highlights attributes of the drawings.

In written responses to the prompt "Write a few sentences about your experience", 11 of the 19 participants used words like "difficult" and "challenging".

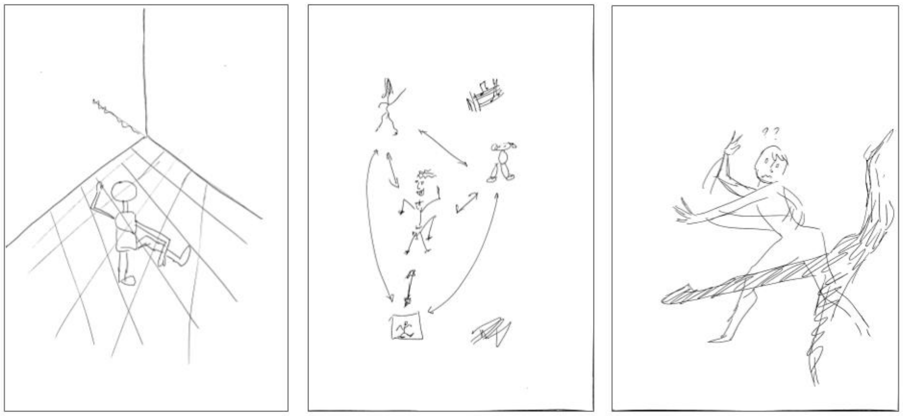

Fig. 3: A sampling of participants' drawings from Station 5. On the left, a human figure is suspended in an grid of varying lines. In the center, pathways are drawn between moving humans and others. On the right, a human mover is shown with a disembodied shadow of themselves or another. As a group, they all illustrate recognition of connectedness and relationship the external world.

Many of them also compared the stations to one another or tried to rank them in terms of difficulty. This monitoring of self-performance was unexpected and provokes additional questions about frustration when interacting with technology in general, whether through computer-based or movement-based exercises. The social aspects of the exercise – such as occasionally being able to see other participants moving, presence of ushers, and interaction with other participants as mediated through the loop-like structure – may have led to additional feelings of self-evaluation or wanting to appear skillful in front of others.

5 Conclusions

In this paper, we presented an instantiation of "The Loop", an embodied interactive art installation, with a formalized experiment structure. Participants were more likely to rate the NAO robot and the shadow of the NAO robot as being controlled by algorithms, whereas they were more likely to rate Performer 1 being controlled by a human actor and the VR avatar as a split between both. Participants drew detailed pictures of humans, connections, and similar postures in their pictures. Self-assessment was frequent during the exercise, despite the open, non-competitive prompting.

In future work, we will conduct more experiments asking about the perceived difficulty of each station to understand why participants frequently see an ab-

Feature in drawing	***Number of occurrences***
Humans	16
Arrows pointing between humans	3
Robots	3
Abstract, geometrical shapes (no humans in picture)	3
Circles enveloping human bodies	3
Windows or mirrors	3

Fig. 4: This table shows frequently illustrated features of the participants' drawings. Humans appeared most often, while elements like arrows and windows illustrate a desire to understand the underlying structure of the installation and connections between participants.

stract movement task through the lens of success or failure. We also aim to test whether other types of robots, with more or less humanoid features as the NAO, will result in similar perceptions of control. Further experiments will probe how embodied exercises with robots generate feelings of empowerment or belittlement for participants, as contrasted with a computer or programming-based task.

Acknowledgement

This work was conducted under IRB #17427 and supported by a National Science Foundation (NSF) grant #1528036. The authors would like to thank the organizing committee of the 5th Annual Conference on Movement and Computing, which was the site of this experiment and provided production support. Novoneel Chakraborty contributed to the creation of the installation. The authors wish to thank Alexandra Bacula, Roshni Kaushik, and Cameron Scoggins for their assistance in running the experiment on site.

References

1. Abrahamson, D., Black, J., DeLiema, D., Enyedy, N., Hoyer, D., Fadjo, C., Trninic, D.: Youre it! body, action, and object in stem learning. In: Proceedings of the International Conference of the Learning Sciences: Future of learning (ICLS 2012). vol. 1, pp. 283–290 (2012)
2. Ashenfelter, K.T., Boker, S.M., Waddell, J.R., Vitanov, N.: Spatiotemporal symmetry and multifractal structure of head movements during dyadic conversation. Journal of Experimental Psychology: Human Perception and Performance **35**(4), 1072 (2009)
3. Bartneck, C., Nomura, T., Kanda, T., Suzuki, T., Kato, K.: Cultural differences in attitudes towards robots. In: Proc. Symposium on Robot Companions (SSAISB 2005 Convention). pp. 1–4 (2005)

4. Bartneck, C., Suzuki, T., Kanda, T., Nomura, T.: The influence of people's culture and prior experiences with aibo on their attitude towards robots. Ai & Society **21**(1-2), 217–230 (2007)
5. Benenson, J.F., Apostoleris, N.H., Parnass, J.: Age and sex differences in dyadic and group interaction. Developmental psychology **33**(3), 538 (1997)
6. Bessiere, K., Ceaparu, I., Lazar, J., Robinson, J., Shneiderman, B.: Social and psychological influences on computer user frustration. Media access: Social and psychological dimensions of new technology use pp. 169–192 (2004)
7. Bradley, M.M., Lang, P.J.: Emotion and motivation. Handbook of psychophysiology **2**, 602–642 (2000)
8. Burgoon, J.K., Stern, L.A., Dillman, L.: Interpersonal adaptation: Dyadic interaction patterns. Cambridge University Press (2007)
9. Cantú, N.E.: La quinceañera: Towards an ethnographic analysis of a life-cycle ritual. Southern Folklore **56**(1), 73 (1999)
10. Cuan, C., Pakrasi, I., LaViers, A.: Time to compile: An interactive art installation. In: 16th Biennial Symposium on Arts & Technology. vol. 51, p. 19 (2018)
11. Du, G., Zhang, P., Mai, J., Li, Z.: Markerless kinect-based hand tracking for robot teleoperation. International Journal of Advanced Robotic Systems **9**(2), 36 (2012)
12. Gelin, R., d'Alessandro, C., Le, Q.A., Deroo, O., Doukhan, D., Martin, J.C., Pelachaud, C., Rilliard, A., Rosset, S.: Towards a storytelling humanoid robot. In: AAAI Fall Symposium: Dialog with Robots. Arlington (2010)
13. Grammer, K., Kruck, K.B., Magnusson, M.S.: The courtship dance: Patterns of nonverbal synchronization in opposite-sex encounters. Journal of Nonverbal behavior **22**(1), 3–29 (1998)
14. Han, J., Campbell, N., Jokinen, K., Wilcock, G.: Investigating the use of non-verbal cues in human-robot interaction with a nao robot. In: Cognitive Infocommunications (CogInfoCom), 2012 IEEE 3rd International Conference on. pp. 679–683. IEEE (2012)
15. Hansen, S.L.: Dating choices of high school students. Family Coordinator pp. 133–138 (1977)
16. Johnson-Glenberg, M.: Immersive vr and education: Embodied design principles that include gesture and hand controls. Frontiers in Robotics and AI **5**, 81 (2018)
17. Johnson-Glenberg, M.C., Megowan-Romanowicz, C.: Embodied science and mixed reality: How gesture and motion capture affect physics education. Cognitive research: principles and implications **2**(1), 24 (2017)
18. Johnson-Glenberg, M.C., Megowan-Romanowicz, C., Birchfield, D.A., Savio-Ramos, C.: Effects of embodied learning and digital platform on the retention of physics content: Centripetal force. Frontiers in psychology **7**, 1819 (2016)
19. Jokinen, K., Wilcock, G.: Multimodal open-domain conversations with the nao robot. In: Natural Interaction with Robots, Knowbots and Smartphones, pp. 213–224. Springer (2014)
20. Kaeppler, A.L.: Ii. dance ethnology and the anthropology of dance. Dance Research Journal **32**(1), 116–125 (2000)
21. Kendal, R.: Infographic: Virtual reality stats everyone should see. Tech. rep., BOSS Magazine, https://thebossmagazine.com/virtual-reality-statistics-infographic/ (01 August 2018)
22. Kuhn, D., Shaw, V., Felton, M.: Effects of dyadic interaction on argumentive reasoning. Cognition and instruction **15**(3), 287–315 (1997)
23. Laver, K.E., George, S., Thomas, S., Deutsch, J.E., Crotty, M.: Virtual reality for stroke rehabilitation. Cochrane Database of Systematic Reviews (2) (2015)

24. Lazar, J., Jones, A., Shneiderman, B.: Workplace user frustration with computers: An exploratory investigation of the causes and severity. Behaviour & Information Technology **25**(03), 239–251 (2006)
25. Lindgren, R., Tscholl, M.: Enacted misconceptions: Using embodied interactive simulations to examine emerging understandings of science concepts. Boulder, CO: International Society of the Learning Sciences (2014)
26. Machida, E., Cao, M., Murao, T., Hashimoto, H.: Human motion tracking of mobile robot with kinect 3d sensor. In: SICE Annual Conference (SICE), 2012 Proceedings of. pp. 2207–2211. IEEE (2012)
27. McCarthy, J., Wright, P.: Technology as experience. interactions **11**(5), 42–43 (2004)
28. Merchant, Z., Goetz, E.T., Cifuentes, L., Keeney-Kennicutt, W., Davis, T.J.: Effectiveness of virtual reality-based instruction on students' learning outcomes in k-12 and higher education: A meta-analysis. Computers & Education **70**, 29–40 (2014)
29. Morris, R., Wander, P.: Native american rhetoric: Dancing in the shadows of the ghost dance. Quarterly Journal of Speech **76**(2), 164–191 (1990)
30. Navarre, D.: Posture sharing in dyadic interaction. American Journal of Dance Therapy **5**(1), 28–42 (1982)
31. Noy, L., Dekel, E., Alon, U.: The mirror game as a paradigm for studying the dynamics of two people improvising motion together. Proceedings of the National Academy of Sciences **108**(52), 20947–20952 (2011)
32. Qian, K., Niu, J., Yang, H.: Developing a gesture based remote human-robot interaction system using kinect. International Journal of Smart Home **7**(4) (2013)
33. Reeves, B., Nass, C.: The Media equation: how people treat computers, television, and new media. Cambridge University Press (1997)
34. Rosen, E., Whitney, D., Phillips, E., Chien, G., Tompkin, J., Konidaris, G., Tellex, S.: Communicating robot arm motion intent through mixed reality head-mounted displays. arXiv preprint arXiv:1708.03655 (2017)
35. Shamsuddin, S., Ismail, L.I., Yussof, H., Zahari, N.I., Bahari, S., Hashim, H., Jaffar, A.: Humanoid robot nao: Review of control and motion exploration. In: Control System, Computing and Engineering (ICCSCE), 2011 IEEE International Conference on. pp. 511–516. IEEE (2011)
36. Shamsuddin, S., Yussof, H., Ismail, L., Hanapiah, F.A., Mohamed, S., Piah, H.A., Zahari, N.I.: Initial response of autistic children in human-robot interaction therapy with humanoid robot nao. In: Signal Processing and its Applications (CSPA), 2012 IEEE 8th International Colloquium on. pp. 188–193. IEEE (2012)
37. Statista: Number of tv households worldwide from 2010 to 2021. Tech. rep., Statista, https://www.statista.com/statistics/268695/number-of-tv-households-worldwide/ (01 January 2018)
38. Stern, D.N.: Mother and infant at play: The dyadic interaction involving facial, vocal, and gaze behaviors. (1974)
39. Strickhouser, J.E., Zell, E.: Self-evaluative effects of dimensional and social comparison. Journal of Experimental Social Psychology **59**, 60–66 (2015)
40. Tesser, A.: Toward a self-evaluation maintenance model of social behavior. In: Advances in experimental social psychology, vol. 21, pp. 181–227. Elsevier (1988)
41. Vinayagamoorthy, V., Gillies, M., Steed, A., Tanguy, E., Pan, X., Loscos, C., Slater, M., et al.: Building expression into virtual characters (2006)
42. Whitney, D., Rosen, E., Phillips, E., Konidaris, G., Tellex, S.: Comparing robot grasping teleoperation across desktop and virtual reality with ros reality. In: Proceedings of the International Symposium on Robotics Research (2017)