On the Flexibility of Robot Social Identity Performance: Benefits, Ethical Risks and Open Research Questions for HRI

Katie Winkle
KTH Royal Institute of Technology
Stockholm, Sweden
winkle@kth.se

Ryan Blake Jackson
MIRRORLab
Colorado School of Mines
Golden, Colorado, USA
rbjackso@mines.edu

Alexandra Bejarano
MIRRORLab
Colorado School of Mines
Golden, Colorado, USA
abejarano@mines.edu

Tom Williams
MIRRORLab
Colorado School of Mines
Golden, Colorado, USA
twilliams@mines.edu

1 INTRODUCTION

Trends in Human Robot Interaction (HRI) research, development, and commercialisation suggest that the first generation of pervasive social robots are likely to perform carefully designed *social identities*, which fundamentally influence (and are typically designed to maximise) those robots' "effectiveness". A significant amount of social HRI work is concerned with understanding how human-like identity cues (such as gender [5, 8], race [2, 32], and personality [25, 35]) can be used to influence typical HRI "effectiveness" measures such as trust, likeability, persuasiveness etc. However, while this previous work has typically examined the potential *benefits* of emulating humanlike identity cues, little work has investigated the potential *risks*. Moreover, little work has explored the unique potential robots offer for abrupt (distinctly non-humanlike) changes in *social identity performance*.

In this article, we reflect on this concept of (robot) social identity performance and its importance in HRI, identifying a number of discussion points and further research questions relating to these under-explored topics. We further question whether we HRI practitioners can go beyond risk avoidance and instead actively leverage robot identity performance to challenge existing norms.

Although many different socially constructed facets of human identity may be regularly projected onto robots (e.g., gender, race, and class), we use *gender* as a typical and pertinent example throughout this paper. Research has demonstrated that humans have a strong tendency to ascribe gender to artificial social actors, even those that are non-anthropomorphic with minimal gender cues [23]. While some designers have attempted to avoid gendering their artificial entities (e.g., the ostensibly genderless voice "Q" [22]), it remains to be seen whether robot designers can (or should) prevent ascriptions of gender to social robots. In addition, perceptions of robot gender seemingly affects user behavior and judgements. Research has shown that robot gendering influences measures of robot friendliness [23], knowledge [23], agenticness [8], communality [8], suitability for certain tasks/roles [5, 8, 33], persuasive capacity [31], trustworthiness [31], credibility [31], and likeability [14, 31]. Despite taking gender as an example, we are cognisant of the fact that gender is only one of the many socially constructed facets of identity that are important to social interaction. However, we suggest that our critical consideration of robot gendering and related risks/opportunities could be similarly applied to other social identity traits.

1.1 Defining Social Identity (Performance)

To understand the importance of robot social identity performance in HRI, we must first understand the socially constructed nature of robot social identity. Social constructivist perspectives of human identity suggest that concepts such as race and gender are not essential categories but are instead human creations, and that the gendering (for example) of humans results from the confluence of one's performance of gender and the way in which that performance is perceived and understood through the collective societal understanding of gender at the time of perception.

Because human perception of robot gender is mediated by the same socially constructed system of gender stereotypes and norms that govern human gendering, robot gender can similarly be understood as a socially constructed phenomenon. However, the robot's gender performance represents a realisation of design intentions held by the robot's developers, rather than the agentic assertion of one's identity in the world. Accordingly, at the time of interaction, robot gender is predominantly a matter of user perception of whatever cues are given off by the robot, intentionally or otherwise.

As such, we use *robot social identity* to refer to users' perceptions of a robot's social identity as mediated by the social constructs created and maintained by society; perceptions that we can expect to be heavily influenced (but not fully determined by) the robot's social identity *performance*, i.e. the summation of those identity cues (morphological, behavioural or otherwise) that are manipulated by robot designers and developers. Specific cues and their role in social identity performance are discussed further in Section 2.

1.2 The Importance of Social Identity in HRI

We previously noted a number of works demonstrating how the manipulation of individual identity traits such as gender and personality can impact HRI in a variety of settings. However, here we draw particular attention to interactions with *socially assistive robots* [9] where the role of the robot is typically to prompt and/or encourage particular user behaviour(s) (e.g. encouraging rehabilitation exercises [10, 18, 41]). From studies of equivalent human-human interactions (e.g. therapist-patient interactions), it is clear that the social relationship between a therapist and their client massively affects client engagement [19, 34]. Our own work with therapists and other domain experts further emphasizes the extent to which they intentionally tailor their social identity performance (e.g. purporting to share an interest in the clients' hobbies, exaggerating particular personality traits) in order to get maximum engagement from the client [39, 44]. We have previously demonstrated that such socially persuasive strategies do indeed seem to map to HRI,

with users undertaking more exercise (our chosen measure of HRI "effectiveness") when a social robot indicates, for example, that it shares their opinions with respect to exercise preferences [42]. The implication then, is that identifying the "best" robot social identity for the user, and using that identity in long term interactions, fostering relationship development, is likely crucial to achieving both instantaneous effectiveness and long term engagement.

Customisable Social Identity

The first implication of designing personalised identities for long-term engagement is the potential for robots with customisable social identity cues. The Furhat robot, for example, enables different faces intended to cue different races and genders. Robot identity cues can fluidly and instantaneously change in a way that has no equivalent in humans, potentially allowing a robot to change its social identity (or aspects thereof) to better suit its context or goals as the robot interacts with different people, or even within single dyadic interactions.

This raises questions as to how users perceive and make sense of changes in these cues, and how much change can be made before users perceive a robot to have taken on a new, distinct identity. Moreover, it raises questions as to who gets to control *if* and *when* robot identity performance can be changed, and what identity *options* are available to particular users. The answers to these questions may have significant impact on the course of individual human-robot interactions, the nature of human-robot relationships, and the impact of robots on society. Furthermore, this potential requires us to reflect on what makes up robot social identity, what morphological and behavioural cues can be leveraged to evoke particular identity ascriptions, and how we should approach designing robot social identities for HRI "effectiveness". We will discuss these considerations in Section 2.

Body-Agnostic Social Identity

The second implication of designing personalized identities for long-term engagement is the potential for one personalized social (robot) identity to be moved between bodies and/or across the various interactive technologies and devices with which a user interacts. Researchers have long explored the possibility of *agent migration*, in which robots "hop" between bodies [11, 12, 21, 24], with an emphasis on the ability for those agents to maintain and project consistent and coherent identities [15, 16, 27], with some recent work also examining the potential for using techniques specifically within the context of robot personalisation [28].

This potential raises questions regarding whether users will recognise "their" personalized social robot companion if it is ported onto a mobile phone and limited to voice-only interactions, and how users would perceive and make sense of different robot identities with distinct social identity cues in multi-embodiment contexts (i.e. when multiple identities actually or ostensibly embody the same robotic hardware at the same time).

2 DESIGNING IDENTITY PERFORMANCE

Robot designers can leverage many cues to prompt human ascription of social identity factors. Some of these cues are unavoidable in robots with certain capabilities (e.g. speech), so designers should be deliberate in how these cues manifest rather than simply defaulting to an arbitrary choice. Importantly, not all cues that inform social identity are those intentionally leveraged by the designer. We might also expect, for example, that user perception of a robot's manufacturer may also be important. Here however, we consider only those cues within the power of the HRI designer to change. We discuss four categories of such cues surrounding robot gender, but similar cues apply to other identity factors.

Naming – Giving a robot a name facilitates interaction by allowing people to more easily refer to and address the robot, but naming is intimately tied to identity [7] and is an important design choice. Certain names are culturally associated with certain genders, and using one of these names for a robot induces perceptions of the corresponding gender (e.g. Amazon's Alexa). HRI researchers often use names to gender robots in experiments, for example using "John" for a male robot and "Joan" for a female robot [33].

Appearance – Robot morphology and appearance provide visual cues that help to construct its identity. For example, researchers have used hairstyle (or plastic resembling hair) to evoke perceptions of certain genders in robots (with a short-haired robot gendered as male and a long-haired robot as female) [8]. Body shape can also elicit perceptions of gender. For example, researchers have found that waist-to-hip ratio and shoulder width are effective gender cues in anthropomorphic robots [3]. Dressing robots in stereotypically gendered clothing provides yet another visual gender cue [5].

Speech – In language-capable robots, voice, tone, word choice, and dialect are cues for many facets of identity. Voice is a sufficiently powerful gender cue that some researchers effectively manipulate robot gender presentation by changing only the robot's voice [23, 31]. Various pragmatic aspects of speech are also gender cues. Many linguistic resources that index power also indirectly index masculinity [20]. Past feminist research often cited women as using "powerless" speech (e.g., indirectness, deference, hesitation, etc.) [17], and, though it is now clear that this stereotype was based on white middle-class women and that not all women use this type of speech, it remains indexing of femininity for many communities [20].

Behavior – Certain roles/tasks are culturally associated with certain genders, and these associations extend to gendered robots [8, 33]. For example, female presenting robots are typically viewed as more suitable for healthcare roles [33], and this association is likely bidirectional (i.e., robots in healthcare roles being perceived as female barring overriding cues).

The typical paradigm in the HRI works referenced thus far is for participants to interact with a robot that has these cues pre-set and fixed to generate a single, consistent identity performance. This is reflected in (most) commercial digital assistants which have e.g. a default gender performance (typically female) on device start-up, as recently critiqued in UNESCO's ''I'd Blush if I Could" report [36]. In fact, one recommendation of that report is simply to ask users to choose the gender of their digital assistant on the first device start-up. Therefore, future work might also consider how giving users *choice* over their robot's identity performance impacts on resultant interactions.

In summary, social HRI research has always been concerned with understanding how particular robot cues can evoke trust, likeability, credibility etc. We simply suggest that this ought to be properly situated within the broader context of robot social identity (performance). Further, we suggest more work should be

done to understand the potential for/impact of identity *fluidity*, given that we've highlighted the importance of social identity for "effective" HRI. However, we also want to highlight that the very concept of designing social identity performance for maximum HRI "effectiveness" itself raises deeper questions around the ethics of intentionally leveraging human-like identity cues, not least around the potential for misrepresenting, further marginalising, and/or appropriating from people of marginalised identities (whose voices might not be heard during the design process).

3 MINIMISING RISK, CHANGING NORMS

Over the course of the previous sections, we have discussed the unique potential for flexible and customisable performance of robot social identity. However, as we have alluded to throughout this piece, this new and flexible design space introduces significant ethical risks. To consider these risks, we might take inspiration from works considering the risks of anthropomorphism in HRI more broadly (e.g. [30]). The concept of (robot) Ethical Risk Assessment (ERA) has recently been posited as one tool that designers and developers might use to systematically assess and mitigate the ethical risk posed by a particular robot [4, 38]. In [40] we utilise ERA to specifically consider ethical risk versus "effectiveness" when intentionally utilising (varying degrees of) anthropomorphic social behaviour in a socially assistive robot. Combining behavioural and perceptual experimental measures with qualitative data regarding perceptions of acceptability and deception we argue that, for socially assistive robots, anthropomorphism is important to overall robot function and poses relatively low ethical risk. Notably however, acceptability of this anthropomorphic behaviour (and any resultant ethical risk) was closely related to participants agreeing with its intended purpose in the specific application discussed (i.e. getting users to exercise more and/or making them feel more comfortable). Based on these results, we might expect similar for social identity performance, i.e. that users' are comfortable with HRI designers manipulating social identity cues to make a robot more "effective" if they agree with that robot's overall purpose.

However, this work was primarily focused on the user; the potential that *they* might be deceived, and to what extent *they* find the robot to be (un)acceptable. Particularly when leveraging increasingly human-like social identity performance, we must go further to consider the risks posed e.g. to those who share some of those identity traits performed by the robot. If we focus only on "effectiveness" when designing social identity performance, we designers risk (un)intentionally identifying and amplifying (harmful) stereotypes. For example, continuing with our example of robot gender, the previously mentioned UNESCO report provides a comprehensive assessment of exactly this risk [37]. It explains that Amazon and Apple's choice to use a female sounding voice for digital assistants is motivated by the "effectiveness" argument; i.e. what pleases customers and is "backed up" by (cherry-picked) academic texts suggesting people simply prefer female voices, which are "helpful", and therefore best suited to customer service roles, compared to male voices which are "authoritative". In the context of robot identity fluidity, it is easy to imagine an equivalent design thought process that goes something like: *'female voices are perceived as being most helpful, male voices are perceived as being most authoritative, so our robot should utilise a female sounding voice when empathising with the user and a male sounding voice when providing an instruction'*. However, as pointed out in the report, results regarding the perceptions of male and female voices no doubt stem from gender biases which predate the digital era. By leveraging this bias in robot design, we would be propagating harmful norms regarding women being docile and obliging and men being unable to provide emotional support, to achieve our own design goals[1]. Whilst we are not the first to consider and critique robot gendering, we note related works also typically use "effectiveness" based arguments, based on robot gender performance (not) influencing task performance [26] for example, rather than considering these broader ethical risks that such design choices might pose.

Further, we draw attention to the lack of works exploring how (fluid) robot identity performance might actually be able to *go against* these stereotypes, to actively model *alternative* identity norms to reflect an actively feminist, anti-racist and pro-social justice design and development stance, motivated by the potential for HRI to influence human norms [13]. Our initial efforts in this direction [43] demonstrate it might even be possible to make the more typical "effectiveness" argument for doing so, in that we were able to boost the credibility of a female robot which was argumentative rather than subservient. However, that work also showcases the difficulty in designing a feminist robot, and the potential for further marginalization of the very group whose identity cues the robot leveraged *and* whose oppression we were trying to stand against.

Moving forward, we suggest that both the minimization of ethical risk and the potential for designing norm-changing robots requires HRI researchers to be cautious of designing robot social identity performance based on "effectiveness" as assessed in the typical HRI experimental paradigm. Mutual shaping [29] stands out as being the more appropriate approach to take, as it requires us to (i) consider the broader impact of robot deployment on society and (ii) utilise participatory design methods to ensure we are in constant dialogue with those whose needs our robots should serve, and whose identity cues our robots might leverage; especially people of marginalized identities who might not be represented in the design team. However, we also suggest that there is space for HRI researchers to explore how feminist approaches can inform our methods, as has been seen in the neighbouring fields of HCI [1] and Data Science [6].

ACKNOWLEDGMENTS

This work was funded in part by Digital Futures, in part by a NASA Early Career Faculty Award and in part by an AFOSR Young Investigator Award.

Dr. Katie Winkle is a Postdoctoral Research Fellow at KTH.
Ryan Blake Jackson is a PhD student at Colorado School of Mines.
Alexandra Bejarano is a PhD student at Colorado School of Mines.
Dr. Tom Williams is an Assistant Professor of Computer Science at Colorado School of Mines.

[1] see also e.g. the black CGI model Shudu Gram created by a white man, as highlighted in Dr. Alex Ketchum's Introduction to Feminist and Social Justice Studies Podcast, Episode 17: https://introtofeministstudies.blogspot.com/2020/07/episode-17-technology.html

REFERENCES

[1] Shaowen Bardzell. 2010. Feminist HCI: Taking Stock and Outlining an Agenda for Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. Association for Computing Machinery, New York, NY, USA, 1301–1310. https://doi.org/10.1145/1753326.1753521

[2] Christoph Bartneck, Kumar Yogeeswaran, Qi Min Ser, Graeme Woodward, Robert Sparrow, Siheng Wang, and Friederike Eyssel. 2018. Robots and racism. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 196–204.

[3] Jasmin Bernotat, Friederike Eyssel, and Janik Sachse. 2017. Shape it–the influence of robot body shape on gender perception in robots. In *International Conference on Social Robotics*. Springer, 75–84.

[4] BSI. 2016. BS 8611:2016 Robots and Robotic Devices: Guide to the Ethical Design and Application of Robots and Robotic Systems.

[5] Julie Carpenter, Joan M Davis, Norah Erwin-Stewart, Tiffany R Lee, John D Bransford, and Nancy Vye. 2009. Gender representation and humanoid robots designed for domestic use. *International Journal of Social Robotics* 1, 3 (2009), 261.

[6] Catherine D'Ignazio and Lauren F. Klein. 2020. *Data Feminism.* MIT Press.

[7] Kenneth L Dion. 1983. Names, identity, and self. *Names* 31, 4 (1983), 245–257.

[8] Friederike Eyssel and Frank Hegel. 2012. (s)he's got the look: Gender stereotyping of robots. *Journal of Applied Social Psychology* 42, 9 (2012), 2213–2230.

[9] David Feil-Seifer and Maja J Matarić. 2005. Defining Socially Assistive Robotics. In *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference On*. IEEE, 465–468.

[10] Rachel Gockley and Maja J Mataric. 2006. Encouraging Physical Therapy Compliance with a Hands-off Mobile Robot. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*. ACM, 150–155.

[11] Wan Ching Ho, Kerstin Dautenhahn, Mei Yii Lim, Patricia A Vargas, Ruth Aylett, and Sibylle Enz. 2009. An initial memory model for virtual and robot companions supporting migration and long-term interaction. In *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 277–284.

[12] Michita Imai, Tetsuo Ono, and Tameyuki Etani. 1999. Agent migration: communications between a human and robot. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, Vol. 4. IEEE, 1044–1048.

[13] Ryan Blake Jackson and Tom Williams. 2019. Language-Capable Robots may Inadvertently Weaken Human Moral Norms. In *Companion Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (alt.HRI)*.

[14] Ryan Blake Jackson, Tom Williams, and Nicole Smith. 2020. Exploring the Role of Gender in Perceptions of Robotic Noncompliance. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 559–567.

[15] Kheng Lee Koay, Dag Sverre Syrdal, Michael L Walters, and Kerstin Dautenhahn. 2009. A user study on visualization of agent migration between two companion robots. In *Proceedings of the 13th Conference on Human-Computer Interaction*.

[16] Michael Kriegel, Ruth Aylett, Kheng Lee Koay, KD Casse, Kerstin Dautenhahn, Pedro Cuba, and Krzysztof Arent. 2010. Digital body hopping-migrating artificial companions. *Digital Futures* (2010).

[17] Robin Lakoff. 1973. Language and woman's place. *Language in society* 2, 1 (1973), 45–79.

[18] Norjasween Abdul Malik, Fazah Akhtar Hanapiah, Rabiatul Adawiah Abdul Rahman, and Hanafiah Yussof. 2016. Emergence of Socially Assistive Robotics in Rehabilitation for Children with Cerebral Palsy: A Review. *International Journal of Advanced Robotic Systems* 13, 3 (June 2016), 135. https://doi.org/10.5772/64163

[19] FAOTA Marilyn B. Cole MS, OTR/L and OTR/L FAOTA Valnere McLean MS, BCN. 2003. Therapeutic Relationships Re-Defined. *Occupational Therapy in Mental Health* 19, 2 (March 2003), 33–56. https://doi.org/10.1300/J004v19n02_03

[20] Sara Mills. 2005. Gender and impoliteness.

[21] Pauli Misikangas and Kimmo Raatikainen. 2000. Agent migration between incompatible agent platforms. In *Proceedings 20th IEEE International Conference on Distributed Computing Systems*. IEEE, 4–10.

[22] Dalia Mortada. 2019. Meet Q, The Gender-Neutral Voice Assistant. https://www.npr.org/2019/03/21/705395100/meet-q-the-gender-neutral-voice-assistant

[23] Clifford Nass, Youngme Moon, and Nancy Green. 1997. Are machines gender neutral? Gender-stereotypic responses to computers with voices. *Journal of applied social psychology* 27, 10 (1997), 864–876.

[24] Tetsuo Ono, Michita Imai, and Ryohei Nakatsu. 2000. Reading a robot's mind: A model of utterance understanding based on the theory of mind mechanism. *Advanced Robotics* 14, 4 (2000), 311–326.

[25] Jong-Chan Park, Hyunsoo Song, Seongyong Koo, Young-Min Kim, and Dong-Soo Kwon. 2010. Robot's Behavior Expressions According to the Sentence Types and Emotions with Modification by Personality. In *Advanced Robotics and Its Social Impacts (ARSO), 2010 IEEE Workshop On*. IEEE, 105–110.

[26] Natalia Reich-Stiebert and Friederike Eyssel. 2017. (Ir)Relevance of Gender? On the Influence of Gender Stereotypes on Learning with a Robot. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI '17)*. Association for Computing Machinery, New York, NY, USA, 166–176. https://doi.org/10.1145/2909824.3020242

[27] Samantha Reig, Jodi Forlizzi, and Aaron Steinfeld. 2019. Leveraging robot embodiment to facilitate trust and smoothness. In *Companion Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 742–744.

[28] Samantha Reig, Michal Luria, Janet Z Wang, Danielle Oltman, Elizabeth Jeanne Carter, Aaron Steinfeld, Jodi Forlizzi, and John Zimmerman. 2020. Not Some Random Agent: Multi-person interaction with a personalizing service robot. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 289–297.

[29] Selma Sabanovic. 2010. Robots in Society, Society in Robots. *International Journal of Social Robotics* 2, 4 (Dec. 2010), 439–450. https://doi.org/10.1007/s12369-010-0066-7

[30] Matthias Scheutz. 2011. 13 The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots. In *Robot Ethics: The ethical and Social Implications of Robotics*. MIT Press, 205.

[31] Mikey Siegel, Cynthia Breazeal, and Michael I Norton. 2009. Persuasive robotics: The influence of robot gender on human behavior. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2563–2568.

[32] Megan Strait, Ana Sánchez Ramos, Virginia Contreras, and Noemi Garcia. 2018. Robots Racialized in the Likeness of Marginalized Social Identities are Subject to Greater Dehumanization than those racialized as White. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 452–457.

[33] Benedict Tay, Younbo Jung, and Taezoon Park. 2014. When stereotypes meet robots: the double-edge sword of robot gender and personality in human–robot interaction. *Computers in Human Behavior* 38 (2014), 75–84.

[34] Renee R. Taylor, Sun Wook Lee, Gary Kielhofner, and Manali Ketkar. 2009. Therapeutic Use of Self: A Nationwide Survey of Practitioners' Attitudes and Experiences. *American Journal of Occupational Therapy* 63, 2 (March 2009), 198–207. https://doi.org/10.5014/ajot.63.2.198

[35] Daniel Ullrich. 2017. Robot Personality Insights. Designing Suitable Robot Personalities for Different Domains. *i-com* 16, 1 (April 2017), 57–67. https://doi.org/10.1515/icom-2017-0003

[36] Mark West, Rebecca Kraut, and Han Ei Chew. 2019. I'd blush if I could: closing gender divides in digital skills through education. (2019).

[37] Mark West, Rebecca Kraut, and Han Ei Chew. 2019. *I'd Blush If I Could: Closing Gender Divides in Digital Skills through Education.* Technical Report.

[38] Alan F. T. Winfield and Katie Winkle. 2020. RoboTed: A Case Study in Ethical Risk Assessment. *2020 5th International Conference on Robot Ethics and Standards (ICRES)* (Sept. 2020). arXiv:2007.15864

[39] Katie Winkle. 2020. *Expert-Informed Design and Automation of Persuasive, Socially Assistive Robots.* Ph.D. Dissertation. University of Bristol, Bristol, UK.

[40] K. Winkle, P. Caleb-Solly, U. Leonards, A. Turton, and P. Bremner. 2021. Assessing and Addressing Ethical Risk from Anthropomorphism and Deception in Socially Assistive Robots. In *2021 16th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*.

[41] Katie Winkle, Praminda Caleb-Solly, Ailie Turton, and Paul Bremner. 2018. Social Robots for Engagement in Rehabilitative Therapies: Design Implications from a Study with Therapists. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI '18)*. ACM, New York, NY, USA, 289–297. https://doi.org/10.1145/3171221.3171273

[42] K. Winkle, S. Lemaignan, P. Caleb-Solly, U. Leonards, A. Turton, and P. Bremner. 2019. Effective Persuasion Strategies for Socially Assistive Robots. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 277–285. https://doi.org/10.1109/HRI.2019.8673313

[43] Katie Winkle, Gaspar Isaac Melsión, Donald McMillan, and Iolanda Leite. 2021. Boosting Robot Credibility and Challenging Gender Norms in Responding to Abusive Behaviour: A Case for Feminist Robots. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21)*. ACM, New York, NY, USA.

[44] Katie Winkle, Ailie Turton, Praminda Caleb-Solly, and Paul Bremner. 2018. Patient Engagement with Rehabilitative Therapy Programmes: Therapist Strategies and Affecting Influences. Proceedings of SRR. *Clinical Rehabilitation* 32, 10 (Oct. 2018), 1414–1415. https://doi.org/10.1177/0269215518784346