Examining the State of Robot Identity

Lux Miranda
Department of Information Technology, Uppsala University,
Uppsala, Sweden
lux.miranda@it.uu.se

Ginevra Castellano
Department of Information Technology, Uppsala University,
Uppsala, Sweden
ginevra.castellano@it.uu.se

Katie Winkle
Department of Information Technology, Uppsala University,
Uppsala, Sweden
katie.winkle@it.uu.se

ABSTRACT

Human-robot interaction has the power to influence human norms and culture. While there is potential benefit in using this power to create positive social change, so too is there risk in merely reinforcing existing social biases which uphold systems of oppression. As the most salient forms of oppression arise along lines of social identity, it stands to reason that we must take utmost care in leveraging human-like identity cues when designing social robots and other agentic embodiments. Yet, the understanding of how to do this is not well-developed. Towards forming an ethics of robot identity, we begin by surveying the state of thought on the topic in human-robot interaction. We do this by conducting a structured review of HRI conference proceedings analyzed from a feminist, intersectional perspective. Our initial findings suggest that existing literature has not fully engaged with intersectionality, embodies an alarming pathologization of neurodivergence, and almost wholly neglects the examination of race.

CCS CONCEPTS

• **Social and professional topics** → **User characteristics**; • **Human-centered computing** → HCI theory, concepts and models; • **General and reference** → Surveys and overviews.

KEYWORDS

identity; intersectionality; embodiment

ACM Reference Format:
Lux Miranda, Ginevra Castellano, and Katie Winkle. 2023. Examining the State of Robot Identity. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI '23 Companion), March 13–16, 2023, Stockholm, Sweden.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3568294.3580168

1 INTRODUCTION

Robots, like any technology, are not socioculturally inert. All design—consciously or otherwise—encodes the values, norms, and biases of its designers. And as culture shapes robots in this way, so too do robots shape our cultures [31]. Humans freely see robots as social actors [18], and we thus have the same capacity to have our culture and norms changed from human-robot interaction as we do from human-human interaction. And, as robots and other artificial agents, unlike humans, may be deployed rapidly and at immense scales, the power for them to shape our culture and norms is proportionally immense. Thus, the norms and values which robots' designs reinforce have tremendous potential to be re-perpetuated throughout our societies. Previous work in HRI gives evidence of this mechanism in action, finding that misaligned robot responses to immoral user requests can weaken human application of moral norms [14].

Conversely, it is apparent that benefits may be had in leveraging this mechanism to bring about more positive social changes [29]. Yet, design which is not careful to consider its biases has equal potential to merely reinforce and amplify those biases. Gender bias and stereotyping in the design and application of social robots has drawn particular attention and critique in recent years [20, 24, 27, 29]. As per UNESCO's *I'd Blush If I Could* report [27], consider that digital assistants such as Siri, Alexa, Cortana, and the Google Assistant, by default, use voices that are widely interpreted as feminine. What message does this send, what stereotypes does this propagate, that the docile, serving agents in all of our smartphones are naturally "women?"

We may approach this issue by understanding *social identity*, the aspects of one's identity which are socially constructed such as race, gender, class, etc. [3, 22]. Principally, a person's social identity is not constructed in their own mind, but in the minds of those perceiving them. The identity categories a perceiver ascribes to a person vary depending on how they interpret the person's visual and behavioural characteristics. In this way, both humans and robots may have social identity. Roboticists often attempt to design robots to garner ascriptions of a particular identity from users, but, since the identity is ultimately constructed within the mind of the user away from the direct control of the roboticist, the intended and perceived identity do not always align. For example, several attempts to create a "genderless" voice for AI have resulted in most users merely ascribing a "man" or "woman" gender to the voices with roughly equal probability [25].

Given that the most salient of oppression and discrimination arise upon lines of social identity [6], it stands to reason that we must, as a minimum, take the utmost care in designing and leveraging robot identities that do not reinforce these systems of oppression. Yet, we invite readers to take one step further: Can we explore if and how such robot identities might be designed and leveraged to actually fight these systems? Can we design from a stance that is feminist, anti-racist, oriented towards social justice, and, above all, compassionate towards all different types of people? A starting point is to explore the ways in which robots might be used as norm-breaking tools and to challenge harmful or inappropriate user behaviour [10, 29].

Further, we are witnessing the emergence of new forms of identity enabled by new technology. From cyborgs, to metaverses, to

HRI '23 Companion, March 13–16, 2023, Stockholm, Sweden

ACM ISBN 978-1-4503-9970-8/23/03.
https://doi.org/10.1145/3568294.3580168

language models which pass the Turing test, the distinction between human and machine continues to blur [4], and the ability to easily distinguish wholly artificial agents becomes ever more difficult.

Thus: We need practical ethical guidelines to inform how we think about, construct, and leverage different human (and human-like) identities, not only in robots but in all agentic embodiments. As part of a broader project of developing these, we must first begin by understanding the HRI community's current strengths and weaknesses in thinking about and practicing the ethics of robot identity. We must also understand what is known about how robot identity intersects with user identity, as inherently alluded to in works on, for example. robot personalization and adaption [10]. This will allow us to direct ethical development where it is needed most.

This late-breaking report provides initial results on a review of HRI literature to map out the state of knowledge on the topic of robot identity. Using intersectionality as an analytical framework [6], we inductively establish fourteen "axes" of human-like identity, detailed in the following section, to guide our review. As a starting point, we begin with a structured search of all ACM/IEEE HRI conference proceedings from 2006 to 2022 as a small representative sample of HRI literature in attempt to locate papers which are explicitly aware of their manipulation of identity. We use our axes of human-like identity as labels to tease out which types of identity have been covered in this sense. Then, we perform an additional search of each of these labels to search for papers where authors may not have strictly acknowledged their manipulation of identity.

We conclude with our findings, limited to HRI conference proceedings in this first instance of study. Our findings suggest that extant HRI conference literature, in general: (1) has not fully engaged with intersectionality in identity studies, (2) embodies an alarming pathologization of neurodivergence, and (3) almost wholly neglects the examination of race.

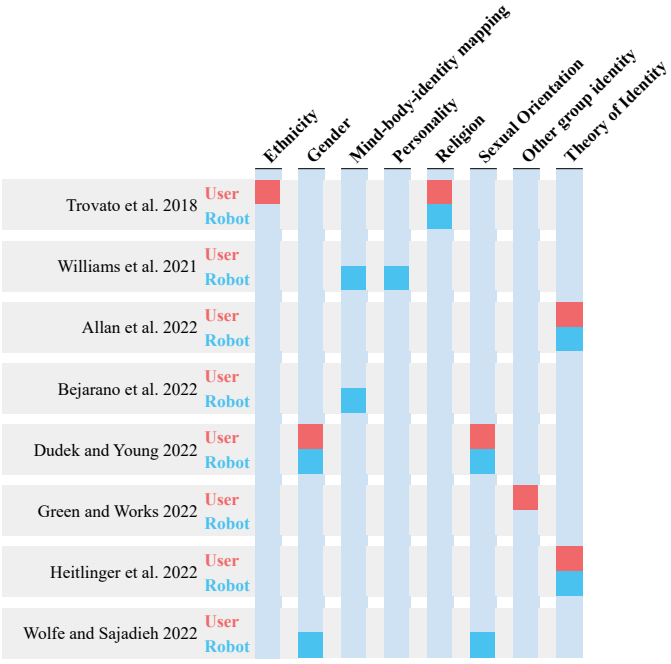
1.1 Research questions

- RQ1: How do HRI papers consider identity for robots vs users?
- RQ2: Which axes of identity have [not] been covered in HRI proceedings?
- RQ3: How has identity been operationalized? That is, what features do authors manipulate to achieve different identity ascriptions?

2 AXES OF HUMAN-LIKE IDENTITY

To guide our review, we use intersectionality [6] as a framework to inductively create a set of human-like identity axes for use as a label coding system. Intersectionality, whilst a widely accepted concept today, originates from Black feminist schools of thought [5] and was developed through the work of Black feminist scholars and activists dealing with interlocking and overlapping systems of oppression [7, 8, 13, 17]. It allows us to grapple with the notion that identity is complex and multi-faceted. Indeed, we cannot capture all of identity in a list of axes alone. But, we can at least be cognizant of the ways in which it is the intersection of one's various identity traits that generally informs how one experiences the world. In Western societies, for example, Black women experience sexist oppression differently to white women. This is on account of the ways in which sexist oppression interweaves with racist oppression under a patriarchal social structure that values whiteness.

Figure 1: The identity traits discussed within our identified works, with an indication of whether they discuss user and/or robot identity.



Thus, we are particularly concerned with modes of identity which are: (1) historically or contemporarily subject to forms of oppression, and/or (2) emerging modes of identity enabled by robots and cybernetic technology. We choose to think of the latter in terms of the different possible configurations[1] of mind, body, identity, and environment [15, 21], as, for our purposes, this seems to offer a reasonably expansive paradigm for conceptualizing these emerging forms of identity [4].

We use the term "axes" to recognize that an individual may be partially, fully, or not at all perceived to be along any one particular axis. For example, an American born and raised by one White parent and one Latinx parent may be perceived as White, Latinx, both, or neither, depending on the social and cultural context they are in (and this is to say nothing of how they personally identify themself).

It is important to note that all axes are socially constructed, and thus their construction can vary significantly from culture to culture. This is neither an attempt to create an exhaustive list of hard, definite categories of identity; We recognize that each axis may mutually overlap with every other axis in one way or another.

[1] It is apparent that robotic systems can take on many such configurations, such as a single system having one "mind" and "identity" but multiple bodies, or perhaps having one body that switches between minds and identities. Western systems of thinking often assume a 1-1-1 mind-body-identity mapping for human beings, but this is not necessarily the case; For example, people who experience multiplicity may have multiple identities within the same body and mind [19]. Or, a group of friends in a metaverse may conceive of each others' virtual avatars as their "bodies" within that space.

We merely attempt to cast as wide of a net as possible in our search and so capture information in a way that might be useful.

Without further ado, our axes of human-like identity are:

- **Age**, as in: infant, child, adolescent, adult, etc.
- **Fat/thin**, as cultural labels pertaining to body type.
- **Class/caste**, as in socioeconomic status or caste membership.
- **Disability and/or health condition**, coded separately from conditions considered cognitive disabilities.
- **Ethnicity**, broadly conceived in the anthropological sense of deriving group identity from common traditions, culture, society, heritage, etc., but coded separately from ethnic identity constructed via religion or nationality.
- **Gender**, including gender identity, performance, and expression.
- **Mind-body-identity mapping**, as in the quantity, location, and temporality of mind, body, and identity relative to each other.
- **Nationality**, or national origin.
- **Neurodiversity**, such as being on the autism spectrum or experiencing variable attention to stimulus.
- **Personality**, coding specifically for if the paper utilizes a formal taxonomy of personality such as the Big Five model.
- **Race**, regardless of the racial categorization framework used.
- **Religion** or spirituality.
- **Sexual orientation**, as well as romantic orientation.
- **Other group identity**, such as membership of a particular team or interest group.

Lastly, we introduce one additional coding label which is not specifically an axis of identity: We use the label **theory of identity** to code for papers which are concerned with identity in general or fundamental theoretical work on human or robot identity.

Note that, in some cases, certain facets that might be considered part of identity are, under the framework which we indicate here, in fact means by which other axes of identity are being operationalized. For example, while the language one speaks is often considered a facet of one's identity, we do not explicitly include it in our framework. We note that when speakers of a certain language, dialect, or manner of speech are subject to social differentiation, it is likely not due to the language itself, but because the language is being used as a cue to construct impressions of ethnicity, class, nationality, etc. It is oppression along these lines which we are concerned with.

3 METHODOLOGY

We begin with a working definition of robot and human identity, for the purposes of identifying literature of interest:

Papers on "Robot Identity:" Papers concerned with (a) user / participant ascription of human-like attributes to robots (e.g. gender, race) and/or (b) roboticist intent to influence such ascriptions.

Papers on "User Identity:" Papers concerned with user/other persons'/robots' (self-)ascription of these same attributes in the context of HRI.

Table 1: Search queries for each identity axis

Label	Query
Age	"age"
Fat	"fat" OR "thin"
Class/caste	"class" OR "caste"
Disability and/or health condition	"disability"
Ethnicity	"ethnicity" OR "ethnic" OR "culture" OR "cultural"
Gender	"gender"
Mind-body-identity mapping	("mind" AND "body") OR "re-embody"
Nationality	"nationality"
Neurodiversity	"neurodiverse" OR "neurodiversity" OR "neurodivergent" OR "neurodivergence"
Pathologized neurodiversity	"ADHD" OR "attention deficit" OR "autism" OR "autistic" OR "obsessive-compulsive" OR "OCD" OR "dyslexia" OR "dysgraphia" OR "dyspraxia" OR "dyscalculia" OR "dysnomia" OR "Tourette" OR "schizophrenia" OR "schizophrenic" OR "bipolar"
Personality	"personality"
Race	"race" OR "racial"
Religion	"religion" OR "religious" OR "spiritual"
Sexual orientation	"sex" OR "sexual" OR "romance" OR "romantic"

3.1 Search criteria

For our first structured search, we include papers if authors have engaged in explicit discussion to investigate identity. We do **not** include or code for papers using post-hoc judgement on our part about which identities certain design decisions may represent. For example, we would not post-hoc label a particular robot embodiment as fat or thin, rather only identify papers which discuss designing for a particular robot or user body type.

Using the tools available through the ACM Digital Library, we search all research articles and short papers within the HRI conference proceedings to find any abstracts containing the word "identity," in quotes.

For our second search, we search the same corpus but using individual queries for terms within abstracts specific to each axis of identity, conducting a simple count of the number of results. These queries are detailed in table 1.

4 FINDINGS

For our first search, looking for the word "identity" in abstracts in attempt to find papers where authors explicitly recognize their manipulation of identity, we found 12 results. Of these, one was a workshop description and another was an HRI Pioneers workshop submission for which corresponding published work exists. Dropping these two, we examined the remaining 10 papers. We found

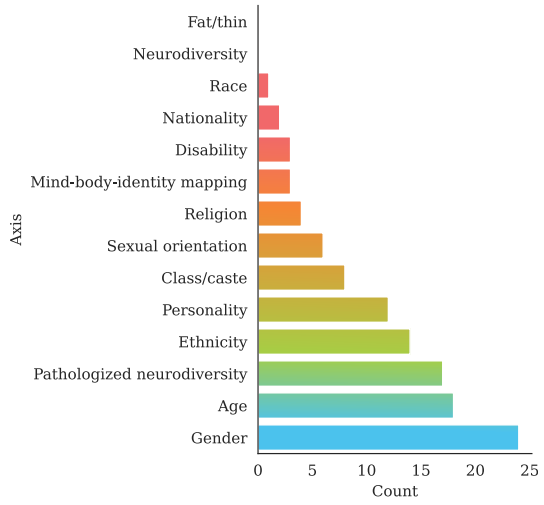


Figure 2: Number of HRI publications found per axis of identity.

that two of them used the word "identity" in their abstract, but did not actually deal with ascriptions of human-like identity to humans or robots. The way in which we coded the remaining eight papers is indicated in Figure 1.

Despite HRI's seventeen-year tenure, six of the results were published in 2022, one was published in 2021, and one was published in 2018. This suggests a very recent interest in robot identity, possibly in part due to the Robo-Identity workshop in 2021 [16] influencing research decisions.

We briefly summarize how each of these eight papers touches on and operationalizes identity: Tovato et al. [26] explores robots in religion through so-called "theomorphic" robots, which are made to represent various beings from religious tradition. They particularly consider the perception of these and their relation to user ethnicity (how, e.g. someone raised in a predominantly Judeo-Christian culture is likely to recognize a robot representing an angel, even if the individual does not personally subscribe to a Judeo-Christian religion). Williams et al.[28] and Bejarano et al.[2] exist as part of the same body of research examining robot mind-body-identity mapping from a fundamental theoretic standpoint. Dudek and Young [9], in their treatment of queer sex robots, are very thorough in their examination of gender and sexual orientation of both robots and users. Green and Works [11] briefly consider university membership as a facet of identity that is operationalized in attempt to appeal to the user to take a certain action. Wolfe and Sajadieh [30] details a theatrical art exhibit featuring a "female" robot created through campily exaggerated breasts, a skirt, simulated lipstick, and long eyelashes. The robot wanders around a room and aggressively and indiscriminately flirts with any encountered human, attempting to get them to give their phone number to the robot. It is debatable whether the users' gender is meant to be operationalized, as although the authors very transparently hint that the piece was primarily targeted towards getting men to understand an experience uncommon to those of their gender, the authors are careful to avoid this exact wording. Thus, we exclude the user identity coding, but note the limitation of doing so. Lastly, Allan et al. [1] and Heitlinger et al. [12] both use social identity as a framing to discuss how people conceive of humans as an ingroup and robots as an "other" or outgroup. Due to their concern with human and robot identity in the broadest sense, we code these as "theory of identity" papers.

The sparsity of the matrix in Figure 1 suggests that, although some authors were very thorough in their treatment of a particular identity axis, none of them are taking an especially broad intersectional approach.

For our second search, we conducted a simple count of the number of results when searching abstracts for terms corresponding to each axis of identity (Table 1). The counts are reported in Figure 2.

Gender is by far the most common topic. Age is the second most common, as, we hypothesize, user age is an often manipulated axis in various studies targeting youth or older adults. Criticially, searching for neurodiversity and related terms yielded zero results. This came as a surprise given our own intuitions on how often HRI research tends to focus on neurodivgent users. Thus, we decided to instead try searching for the most common pathologizations of neurodiversity. In technology literature, neurodivergence is commonly medicalized and seen as a problem to be treated in the individual—rather than seeing it as a natural, valid state of being that society, not the individual, should be changed to accommodate [23]. Including these terms revealed the third largest results category in this search, suggesting that pathologization of neurodiversity in HRI conference proceedings is widespread.

In terms of the limitations of this search: We note that distilling the search for works on mind-body-identity mapping down to a few keywords is difficult, as there is not yet commonly agreed-upon terminology for the concept. There are likely more papers on this topic which simply describe it using different vocabulary.

We further found no papers manipulating robot or user fatness/thinness. This may be due to the stigma of using the word "fat" to describe a human bodies, possibly leading authors to use other terminology which does not lend itself well to a targeted search of robotics literature, such as "weight" or "body type."

The last finding which we shall note: Searching "race" or "racial" reveals a single paper. Given the many critical issues surrounding race in nations such as the United States, this seems to be a glaring omission.

4.1 Summary and future work

Following two structured searches of HRI conference proceedings, our findings suggest that the existing body of proceedings (1) has not yet fully reckoned with intersectionality as a concept or research framework, (2) overwhelmingly pathologizes neurodivergence, and (3) almost entirely neglects race as a pertinent type of identity.

Continuing with our research goal of understanding the state of thinking on robot identity across the broader human-robot interaction field, there is clear room to expand the scope of this review to corpuses other than the HRI proceedings to further elucidate our findings and garner new insights. We invite researchers to take on a participatory role in developing thought on the ethics of robot identity, as we do not think that any single individual or lab has the diversity of insight necessary to do so in a robust way. Good ethics should, at their core, serve us all.

REFERENCES

[1] D. D. Allan, Andrew J. Vonasch, and Christoph Bartneck. 2022. Better than Us: The Role of Implicit Self-Theories in Determining Perceived Threat Responses in HRI. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, Sapporo, Japan, 215–224. https://doi.org/10.1109/HRI53351.2022.9889520

[2] Alexandra Bejarano, Samantha Reig, Priyanka Senapati, and Tom Williams. 2022. *You Had Me at Hello: The Impact of Robot Group Presentation Strategies on Mental Model Formation.*

[3] Judith Butler. 2002. *Gender trouble.* Routledge.

[4] Wesley P Clawson and Michael Levin. 2022. Endless forms most beautiful 2.0: teleonomy and the bioengineering of chimaeric and synthetic organisms. *Biological Journal of the Linnean Society* (July 2022), blac073. https://doi.org/10.1093/biolinnean/blac073

[5] Patricia Hill Collins. 2002. *Black feminist thought: Knowledge, consciousness, and the politics of empowerment.* routledge.

[6] Patricia Hill Collins and Sirma Bilge. 2020. *Intersectionality.* John Wiley & Sons. Google-Books-ID: fyrfDwAAQBAJ.

[7] Kimberle Crenshaw. 1997. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. In *Feminist Legal Theories.* Routledge. Num Pages: 29.

[8] Kimberlé Williams Crenshaw. 1994. Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color. In *The Public Nature of Private Violence.* Routledge. Num Pages: 26.

[9] Skyla Y. Dudek and James E. Young. 2022. Fluid Sex Robots: Looking to the 2LGBTQIA+ Community to Shape the Future of Sex Robots. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction (HRI '22)*. IEEE Press, Sapporo, Hokkaido, Japan, 746–749.

[10] Hideki Garcia Goo, Katie Winkle, Tom Williams, and Megan Strait. 2022. Robots Need the Ability to Navigate Abusive Interactions. https://www.researchgate.net/profile/Tom-Williams-16/publication/359176686_Robots_Need_the_Ability_to_Navigate_Abusive_Interactions/links/622bd153a39db062db95d50a/Robots-Need-the-Ability-to-Navigate-Abusive-Interactions.pdf

[11] Nathan Green and Karen Works. 2022. Measuring Users' Attitudinal and Behavioral Responses to Persuasive Communication Techniques in Human Robot Interaction. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, Sapporo, Japan, 778–782. https://doi.org/10.1109/HRI53351.2022.9889306

[12] Lea Heitlinger, Ruth Stock-Homburg, and Franziska Doris Wolf. 2022. You Got the Job! Understanding Hiring Decisions for Robots as Organizational Members. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, Sapporo, Japan, 530–540. https://doi.org/10.1109/HRI53351.2022.9889444

[13] Bell Hooks. 2000. *Feminist Theory: From Margin to Center.* Pluto Press. Google-Books-ID: uvIQbop4cdsC.

[14] R. B. Jackson and T. Williams. 2019. Language-Capable Robots may Inadvertently Weaken Human Moral Norms. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 401–410. https://doi.org/10.1109/HRI.2019.8673123 ISSN: 2167-2148.

[15] Ryan Blake Jackson, Katie Winkle, Alexandra Bejarano, and Tom Williams. 2021. Design, Performance, and Perception of Robot Identity.

[16] Minha Lee, Dimosthenis Kontogiorgos, Ilaria Torre, Michal Luria, Ravi Tejwani, Matthew J. Dennis, and Andre Pereira. 2021. Robo-Identity: Exploring Artificial Identity and Multi-Embodiment. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction* (Boulder, CO, USA) *(HRI '21 Companion)*. Association for Computing Machinery, New York, NY, USA, 718–720. https://doi.org/10.1145/3434074.3444878

[17] Jennifer C. Nash. 2018. *Black Feminism Reimagined: After Intersectionality.* Duke University Press. Google-Books-ID: QbB8DwAAQBAJ.

[18] 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. https://doi.org/10.1111/j.1559-1816.1997.tb00275.x _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1559-1816.1997.tb00275.x.

[19] Sarah Parry, Zarah Eve, and Gemma Myers. 2022. Exploring the Utility and Personal Relevance of Co-Produced Multiplicity Resources with Young People. *Journal of Child & Adolescent Trauma* 15, 2 (June 2022), 427–439. https://doi.org/10.1007/s40653-021-00377-7

[20] Giulia Perugia, Stefano Guidi, Margherita Bicchi, and Oronzo Parlangeli. 2022. The Shape of Our Bias: Perceived Age and Gender in the Humanoid Robots of the ABOT Database. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction (HRI '22)*. IEEE Press, Sapporo, Hokkaido, Japan, 110–119.

[21] Samantha Reig, Michal Luria, Elsa Forberger, Isabel Won, Aaron Steinfeld, Jodi Forlizzi, and John Zimmerman. 2021. Social Robots in Service Contexts: Exploring the Rewards and Risks of Personalization and Re-embodiment. In *Designing Interactive Systems Conference 2021*. ACM, Virtual Event USA, 1390–1402. https://doi.org/10.1145/3461778.3462036

[22] Audrey Smedley and Brian D Smedley. 2005. Race as biology is fiction, racism as a social problem is real: Anthropological and historical perspectives on the social construction of race. *American psychologist* 60, 1 (2005), 16.

[23] Katta Spiel, Eva Hornecker, Rua Mae Williams, and Judith Good. 2022. ADHD and Technology Research – Investigated by Neurodivergent Readers. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–21. https://doi.org/10.1145/3491102.3517592

[24] Yolande Strengers and Jenny Kennedy. 2020. *The Smart Wife: Why Siri, Alexa, and Other Smart Home Devices Need a Feminist Reboot.* MIT Press. Google-Books-ID: 9s7tDwAAQBAJ.

[25] Selina Jeanne Sutton. 2020. Gender Ambiguous, not Genderless: Designing Gender in Voice User Interfaces (VUIs) with Sensitivity. In *Proceedings of the 2nd Conference on Conversational User Interfaces*. ACM, Bilbao Spain, 1–8. https://doi.org/10.1145/3405755.3406123

[26] Gabriele Trovato, Loys De Saint Chamas, Masao Nishimura, Renato Paredes, Cesar Lucho, Alexander Huerta-Mercado, and Francisco Cuellar. 2021. Religion and Robots: Towards the Synthesis of Two Extremes. *International Journal of Social Robotics* 13, 4 (July 2021), 539–556. https://doi.org/10.1007/s12369-019-00553-8

[27] 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. https://unesdoc.unesco.org/ark:/48223/pf0000367416.page=1

[28] Tom Williams, Daniel Ayers, Camille Kaufman, Jon Serrano, and Sayanti Roy. 2021. Deconstructed Trustee Theory: Disentangling Trust in Body and Identity in Multi-Robot Distributed Systems. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, Boulder CO USA, 262–271. https://doi.org/10.1145/3434073.3444644

[29] 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*. ACM, Boulder CO USA, 29–37. https://doi.org/10.1145/3434074.3446910

[30] Hannen E. Wolfe and Sahar Sajadieh. 2022. A Robo-Pickup Artist Breaking Gender Norms. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, Sapporo, Japan, 1210–1211. https://doi.org/10.1109/HRI53351.2022.9889610

[31] Selma Šabanović. 2010. Robots in Society, Society in Robots: Mutual Shaping of Society and Technology as a Framework for Social Robot Design. *International Journal of Social Robotics* 2, 4 (Dec. 2010), 439–450. https://doi.org/10.1007/s12369-010-0066-7