You're Wigging Me Out! Is Personalization of Telepresence **Robots Strictly Positive?**

Naomi T. Fitter Oregon State University Corvallis, OR, USA naomi.fitter@oregonstate.edu

Megan Strait University of Texas Rio Grande Valley Edinburg, TX, USA megan.strait@utrgv.edu

Eloise Bisbee **Tufts University** Medford, MA, USA eloise.bisbee@tufts.edu

Maja J. Matarić University of Southern California Los Angeles, CA, USA mataric@usc.edu

Leila Takayama University of California, Santa Cruz Santa Cruz, CA, USA takayama@ucsc.edu

ABSTRACT

With their ability to embody users in physically distant spaces, telepresence robots have gained popularity in environments including hospitals, schools, and offices. However, with platforms lacking in individuation and social presence, users often personalize telepresence robots with clothing and accessories to increase their recognizability and sense of embodiment. Toward understanding personalization preferences, as well as perceptions of personalized platforms, we conducted a series of five studies that investigate patterns in personalization of a telepresence robot and evaluate the impacts of common personalizations along five dimensions (robot uniqueness, humanness, pleasantness/unpleasantness, and people's willingness to interact with it). Finding a strong preference for the use of clothing and headwear in Studies 1–2 (N = 52), we systematically manipulated a robot's appearance using these items and evaluated the qualitative and quantitative impacts on observer perceptions in Studies 3–4 (N = 160). Observing that personalization increased perceptions of uniqueness and humanness, but also decreased positive responding, we then investigated the associations between personalization preferences and perceptions via a fifth study (N = 100). Across the five studies, tensions emerged between operators' interest in using wigs and interlocutors' dislike of wigs. This result highlights a need to consider both operator and interlocutor perspectives when personalizing telepresence robots.

CCS CONCEPTS

• Computer systems organization → Robotics; • Human-cen $tered\ computing o Collaborative\ and\ social\ computing,\ Interac$ tion paradigms.

KEYWORDS

telepresence robots, personalization, human-like robot attributes

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

HRI '21, March 23-26, 2021, Boulder, CO, USA © 2021 Association for Computing Machinery. ACM ISBN 978-1-4503-8289-2/21/03...\$15.00 https://doi.org/10.1145/3434073.3444675

ACM Reference Format:

Naomi T. Fitter, Megan Strait, Eloise Bisbee, Maja J. Matarić, and Leila Takayama. 2021. You're Wigging Me Out! Is Personalization of Telepresence Robots Strictly Positive?. In Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21), March 8-11, 2021, Boulder, CO, USA. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/ 3434073.3444675

1 INTRODUCTION

Although people know that robots are not human, we still respond to them in social ways [20] - forming relationships with robots as simple as vacuums [28] - and any individuation (e.g., via personalization) seems to amplify our sociality towards them. For example, personalization of the Roomba robot using decals has been found to improve people's social perceptions of it [27]. People also appear to have a natural interest in robot personalization [5, 9, 14–16, 26]. For example, in a hospital-based deployment of a mobile delivery robot, nurses and other staff spontaneously personalized their robot with stickers [17, 21]. Similarly, in a manufacturing-based deployment of the Baxter robot, co-located workers - without prompting - outfitted the robot with a wig and hat to humanize it and give it more personality [22]. Even corporate entities have been observed to customize their robotic platforms. For example, Kuka robots [1] deployed in Tesla factories [3] are painted Tesla red instead of the standard Kuka orange. Similarly, the Relay delivery robot platform [2] has a standard design, but hotels that deploy it use vinyl decals to add corporate colors and visual styling to the robot to communicate their corporate identity to hotel guests.

Related work on telepresence robots: Considering the emergence of telepresence robots in social settings such as classrooms, homes, hospitals, and offices [8] and the goal of embodying remote individuals, personalization of these platforms seems especially appropriate (e.g., to increase one's social presence). However, is personalization of telepresence robots strictly positive? While the existing literature on robot personalization reflects consistent, positive impacts, findings specific to the personalization of telepresence robots are more mixed. For example, personalization of telepresence robots using decals and other humanizing items including clothing and headwear did not yield social improvements compared to a non-personalized robot in terms of remote presence feeling and remote learning outcomes [5, 6]. Some personalizations have even

| Study | pop. | N | %men | ages | M_{age} | %STEM | M_{fam} . |
|---------------|-------|-----|------|---------|-------------|-------|--------------|
| 1 (in-person) | USC | 20 | 50% | NR | NR | 100% | NR |
| 2 (in-person) | USC | 64 | 42% | 18 - 24 | 20 ± 2 | 50% | $1.73\pm.99$ |
| 3 (in-person) | OSU | 30 | 46% | 18 - 54 | 23 ± 9 | 83% | $2.67\pm.88$ |
| 4 (online) | MTurk | 130 | 60% | 20-71 | 36 ± 10 | 28% | $2.10\pm.77$ |
| 5 (online) | OSU | 100 | 25% | 18 - 45 | 22 ± 1 | 47% | $1.88\pm.08$ |

Table 1: Participant demographics including source population, total participants, gender ratio, age range, mean age $(\pm SD)$, proportion with STEM backgrounds, and mean robotics familiarity reported $(\pm SD)$. NR denotes "not recorded".

yielded the opposite, with people liking the remote interactant with personalization less than they did without personalization [19]. At the same, personalization has value in embodied telepresence. For example, clear communication of remote coworkers' identities (whether through robot personalization or otherwise) is critical to achieving effective collaborations across physically distant spaces [18]. Moreover, personalization of even non-robotic technological systems appears to generally benefit human-machine interactions [4, 12]. This complexity in findings may stem from the fact that, unlike non-telepresence robots and non-robotic technological systems, a telepresence robot has two concurrent social roles: to the remote operator, it is a human proxy; to co-located interlocutors, it is also a robot.

Present work: The multitude of perspectives involved in embodied telepresence renders it difficult to know whether general principles of computer-mediated communication or human-robot interaction norms apply to telepresence robots. Towards better understanding the benefits and downsides of telepresence robot personalization, we carried out a series of five exploratory studies to identify ways in which prospective operators prefer to personalize a telepresence robot (Studies 1-2) and to measure prospective interlocutors' perceptions of common personalizations (Studies 3-5); see Table 1 for participant demographics. Based on the personalization preferences specified by participants in Studies 1-2, we created a set of 18 photographs depicting a personalized and non-personalized Ohmni robot. Within the set, we systematically manipulated the robot's personalization using the items most preferred by participants (shorts, shirts, hats, and wigs). In Studies 3-5, we then investigated people's perceptions of the personalizations along five constructs - uniqueness, humanness, pleasantness, unpleasantness, and approachability of the robot - derived from thematic analysis of the comments by participants in Study 2.

By investigating personalization from the perspective of prospective operators (Studies 1–2) and interlocutors (Studies 3–5), as well as the associations between personal preferences for and evaluation of telepresence robot personalizations (Study 5), this work contributes to theoretical understanding of the impacts of personalization in embodied telepresence. The set of studies additionally provides preliminary evidence of the reproducibility of our findings.

2 STUDIES 1-2

To investigate the ways that people prefer to personalize a telepresence robot, we conducted two studies in which we provided 20 individuals (Study 1) and 32 dyads (Study 2) with various items

that they could use to personalize an Ohmni telepresence robot and observed the frequency at which they selected each item. All study procedures were approved by the USC Institutional Review Board (IRB) under protocol #UP-17-00639. We then used the observed frequencies to select *personalization items* for further, systematic evaluation (Studies 3–5) and identified five constructs of interest via thematic analysis of the transcribed Study 2 discussions.

2.1 Study 1 (Individual Personalizations)

In this pilot study, we recruited 20 university students to test-operate an Ohmni telepresence robot and optionally customize it using the variety of *clothing* (skirts, shorts, shirts, dresses, jackets, and scarves), *headwear* (hats, wigs, and stick-on facial hair), and *miscellaneous items* (name tags, stickers, and other personal items) available to them. Participants were asked to personalize the robot to represent themselves, using as many or few items as they would like. Figures 1–2 depict personalizations created by participants in Study 1. In observing the frequency at which participants selected the available personalization items, we found all participants opted to personalize the robot with 2 – 8 items (M = 5, SD = 2) and gravitated toward clothing (shirts and shorts), headwear (hats and wigs), and name tags to personalize the robot; see Table 2.

2.2 Study 2 (Dyadic Personalizations)

To investigate the reproducibility of our Study 1 observations regarding operator preferences for robot personalization, as well as the potential impacts of these personalizations on telepresent interactions with interlocutors, we recruited 64 participants to work as pairs – with one participant randomly assigned to serve as the operator of an Ohmni robot and the other as interlocutor (N=32). They were asked to customize the robot to represent the operator and, subsequently, drive around the room with the interlocutor's help via telepresence (see [6] for more details). All agents (the robot and both participants) were co-located throughout the personalization process and we again offered the items that were provided to participants in Study 1. Item selection frequencies here mirrored those observed in our pilot investigation: participants selected three to nine items (M=5, SD=1) and again gravitated toward shorts, shirts, wigs, hats, and name tags, as well as stickers.



Figure 1: Example operators and their robot personalization.



Figure 2: Example personalizations created in Study 1.

| | N | skirt | shorts | shirt | dress | jacket | scarf | facial hair | wig | hat | name tag | sticker | personal item | no items |
|---------------------|-----|-------|--------|-------------|-------|--------|-------|-------------|-------------|-------------|----------|---------|---------------|----------|
| Study 1 (in-person) | 20 | 5% | 60% | 80% | 10% | 15% | 50% | 15% | 60% | 65% | 65% | 35% | 20% | 0% |
| Study 2 (in-person) | 32 | 0% | 66% | 75 % | 19% | 19% | 19% | 28% | 75 % | 59 % | 47% | 50% | 47% | 0% |
| Study 5 (online) | 100 | 2% | 3% | 14% | 4% | 12% | 3% | 0% | 9% | 10% | 59% | 40% | 2% | 25% |

Table 2: Frequency of each item's selection by prospective operators to personalize the Ohmni telepresence robot in Studies 1, 2, and 5. Bolding denotes selection by the majority of participants (> 50%), while items selected by less than 25% are grayed.

In transcribing each pair's discussion during the collaborative driving portion (a four-minute, free-form interaction), two recurring themes about operators' self-identification with telepresence hardware and interlocutors' reactions to personalization emerged:

- Humanization: Nine operators clearly identified with the robot's embodiment in some way (e.g., "I'll come see myself", "I'm stuck", and "I'm not blocking the door") and one interlocutor even noted that the resemblance of the robot to the remote robot operator was striking. It additionally seemed natural for the interlocutor to think of the robot anatomy as human anatomy with many participants referring to the robot's wheeled base as "feet" that could "walk".
- Uncanniness: 13 dyads used one or more of "creepy", "weird",
 "Black Mirror"-like, and "uncanny" to describe their experience, reflecting a discomforting effect of personalization
 potentially related to and evocative of an uncanny valley.

2.3 Summary of Findings

With all participants electing to personalize the Ohmni telepresence robot, Studies 1–2 indicate a clear preference of prospective operators for personalization. Item selection frequencies further indicate a particular preference for clothing (namely, shorts and a shirt) and headwear (a wig and/or a hat), with the majority of participants utilizing these items in their personalizations. Moreover, thematic analysis of the Study 2 discourse indicates that operators identified as the robot (whether independent of or facilitated by personalization) and, with interlocutors noting its humanizing effects, personalization may further benefit telepresent interactions. However, given many comments reflecting interlocutor discomfort, the analysis also suggests potential downsides to personalization.

3 STUDIES 3-4

To further investigate (prospective) interlocutors' *perceptions* of robot personalization, we designed an experiment wherein we systematically manipulated the Ohmni robot's appearance based on the personalizations most frequently observed in Studies 1–2 and asked participants to evaluate specific attributes of each depiction. We first ran this experiment in-lab (**Study 3**). Observing that the participant demographics were relatively homogeneous (young, technically-trained students at Oregon State University), we also ran the experiment online to engage a broader and larger participant pool (**Study 4**). All procedures were approved by the OSU Institutional Review Board under protocol #IRB-2019-0172.

3.1 Method

To investigate the effects of **personalization** (*none* vs. *personalized*), we created a set of 18 photographs depicting the nonpersonalized Ohmni robot (2 photos) and the robot outfitted with personalizing items (16 photos); see Figure 3. Within the set of 16 personalized depictions, we manipulated the robot's appearance based on the items most frequently employed by participants in Studies 1–2. Specifically, we manipulated the robot's **headwear** (*none*, hat^2 , or wig) and, using a shirt or shirt-and-shorts combo, **outfit completeness** (*partial* vs. *full*). Given the range of clothing

¹To help inform participants' understanding of telepresence robots, they were first given a brief synopsis ("[here] you will complete a survey consisting of questions about several images. These images will contain telepresence robots, which are used for two-way audio and video conferencing, and navigating around a distant space.") and demonstrative video of an Ohmni robot being piloted around (see the supplementary materials). ²In Study 3, we utilized two hat types: a baseball cap and wide-brimmed hat. However, lacking a significant difference in perceptions of the two hat styles, we collapsed their ratings (resulting in a total of three levels) and we correspondingly retained only one hat style (the baseball cap) in the Study 4 stimulus set.



Figure 3: Stimuli used in Studies 3–5. The non-personalized robot (left), followed by all depictions resulting from the headwear (none, wig, cap, and hat) \times outfit completeness (partial vs. full) \times gender typicality (masculine vs. feminine) manipulations. Note that we subsequently excluded depictions involving a hat (Studies 4–5) as well as partially-dressed depictions (Study 5).

| | | Study 3 (<i>df</i> = 29) | | | | | Study 4 $(df = 129)$ | | | | | Study 5 $(df = 99)$ | | | | |
|-----------------|-------|----------------------------------|--------|--------|------|-------|-----------------------------|--------|--------|------|-------|----------------------------|--------|--------|------|--|
| | M_d | SE | t | p | d | M_d | SE | t | p | d | M_d | SE | t | p | d | |
| uniqueness | .646 | .200 | 3.228 | .003 | .589 | 1.114 | .163 | 6.852 | < .001 | .601 | .235 | .166 | 1.418 | .159 | .142 | |
| humanness | 1.438 | .265 | 5.427 | < .001 | .991 | .981 | .120 | 8.149 | < .001 | .715 | 1.397 | .153 | 9.129 | < .001 | .913 | |
| pleasantness | .483 | .221 | 2.186 | .037 | .399 | 036 | .131 | 273 | .785 | 024 | .293 | .167 | 1.752 | .083 | .175 | |
| unpleasantness | .210 | .175 | 1.195 | .242 | .218 | .907 | .146 | 6.222 | < .001 | .546 | .317 | .148 | 2.134 | .035 | .213 | |
| approachability | 140 | .127 | -1.096 | .282 | 200 | 512 | .124 | -4.118 | < .001 | 361 | 155 | .139 | -1.119 | .266 | 112 | |

Table 3: Main effects of personalization (Student's t statistic, p-value, and Cohen's d) on ratings of the robot depictions, observed in Studies 3, 4, and 5. M_d denotes mean difference (personalized – non-personalized) and SE denotes standard error.

selections in Studies 1–2, we also manipulated the robot's **gender typicality** using a green t-shirt and a white shirt/red cardigan combo (stereotyped as relatively *masculine* vs. *feminine*). In total, we generated 18 photos encompassing four within-subjects manipulations of the robot's appearance: an overall manipulation of personalization – *none* versus *personalized* – and three sub-level manipulations comprising the 16 *personalized* depictions: headwear (three levels²) × clothing (two levels) × gendering (two levels).

Participants: 30 people, recruited from Oregon State University and the surrounding community, participated in Study 3. In Study 4, 130 participants were recruited via Amazon Mechanical Turk (MTurk), with location restriction (to people in the US), > 97% prior task approval rate, and > 5000 previously approved tasks requirements imposed. Participation incentives were \$10 and \$3.50 USD respectively, and across both studies, participants indicated little familiarity with robots using a 7-point Likert scale from 1 (no experience) to 7 (expert-level); see Table 1.

Procedure (Study 3): Upon arrival at the testing location, participants who consented to participate were asked to complete an electronic **survey** consisting of demographic questions, questions about preconceived notions of robots, and quantitative and qualitative evaluations of each of the 18 robot depictions. The *quantitative* evaluation consisted of seven Likert-type scales and the *qualitative* evaluation consisted of a prompt to verbally explain to the on-site researcher what aspects of the given image most influenced their evaluation. The order of presentation of photos and questions was randomized by participant. Following the survey, a researcher additionally administered a brief, **semi-structured interview** that queried participants about which features of the robot depictions they liked and disliked, how they themselves would personalize a telepresence robot, and other thoughts they wished to share.

Procedure (Study 4): Upon opening the task, participants completed the same consenting process and **survey** as did Study 3 participants, with three minor modifications to adapt the procedure to its fully online execution: exclusion of (conceptually) duplicate photos (caps vs. hats) reduced the stimuli to 13 depictions; a free-response field prompting for written description of the features that most influenced their perceptions replaced the verbal, qualitative evaluation; and the post-survey interview was omitted.

Measures: We assessed perceptions of the uniqueness, humanness, pleasantness, unpleasantness, and approachability of each depiction via participants' agreement with seven statements regarding how commonplace, humanlike, pleasant, eerie, strange, and unnerving the robot appeared, as well as how willing they

were to interact with the given robot. "Unpleasantness" was constructed from responses to eerie, strange, and unnerving (Cronbach's $\alpha_{S3}=.91$ and $\alpha_{S4}=.78$); the other four measures were single-item constructs with "uniqueness" inversely inferred from responses to the commonplace item, "humanness" and "pleasantness" interpreted directly (from the humanlike and pleasant items), and "approachability" inferred from participants' willingness to interact. These constructs were derived from related work on general attitudes towards robots [29], perceptions of anthropomorphized robots [24, 25], and human affect [11]. All statements used a 7-point Likert scale with labeled anchors from 1 (strongly disagree) to 7 (strongly agree).

3.2 Results

We analyzed participants' ratings of the robot depictions along the five dependent variables using Student's *t*-tests and analyses of variance (ANOVAs). To evaluate the effects of personalization (*none vs. personalized*), we used two-tailed, paired *t*-tests comparing ratings of the non-personalized robot and ratings of the personalized robots (collapsed across all cells comprising the personalization manipulations). To evaluate the effects of each personalizing factor, we used three-way repeated measures ANOVAs with *headwear* (none, hat, or wig), *outfit completeness* (partial vs. full), and *gender typicality* (green vs. red top) included as independent variables. Significant effects were further analyzed for pairwise differences using post hoc Bonferroni-corrected paired *t*-tests.

Table 3 gives the test statistics for the main effects of personalization, and the main effects of headwear and pairwise differences therein are reported in Tables 4 and 5. Reported in the Appendix are all descriptive and test statistics for the remaining effects (Tables 6 and 7), test statistics from the analyses of variance (Table 8), descriptive statistics for personalization and headwear (Tables 9 and 10), and paired comparisons of *hat* vs. *no headwear* (Table 11).

We analyzed the qualitative data (participants' free responses and interviews) using thematic analysis to identify patterns across respondents. Due to recording errors, qualitative data were only available for 23 participants in Study 3. In addition, an initial readthrough of the free-response data in Study 4 revealed three responses unrelated to the questions asked, which were then discarded. The adjusted sample sizes in the qualitative analyses were thus N=23 (Study 3) and N=127 (Study 4). Below, we discuss the significant findings and most salient themes that emerged.

Effects of personalization: Across Studies 3–4, we observed a significant main effect of personalization (none vs. personalized) on *uniqueness* and *humanness*, with ratings reflecting the perception

| | Study 3 | | | | | Stud | ly 4 | Study 5 | | | | |
|-----------------|-------------|-----------------|-----------------|----------------|-------------|-----------------|-----------------|------------|-------------|-----------------|-----------------|----------------|
| | F statistic | ϵ_{GG} | <i>p</i> -value | η_{p}^{2} | F statistic | ϵ_{GG} | <i>p</i> -value | η_p^2 | F statistic | ϵ_{GG} | <i>p</i> -value | η_{p}^{2} |
| uniqueness | 6.970 | .783 | .004 | .194 | 28.609 | .774 | < .001 | .182 | 7.836 | | < .001 | .073 |
| humanness | 1.304 | | .279 | .043 | .260 | .801 | .721 | .002 | 1.228 | | .295 | .012 |
| pleasantness | 10.230 | .808 | < .001 | .261 | 28.858 | .723 | < .001 | .183 | 10.926 | | < .001 | .099 |
| unpleasantness | 20.970 | .591 | < .001 | .420 | 56.095 | .634 | < .001 | .303 | 35.357 | .862 | < .001 | .263 |
| approachability | 7.093 | .605 | .008 | .197 | 17.498 | .666 | < .001 | .119 | 15.892 | .886 | < .001 | .138 |

Table 4: Main effects of headwear (none, hat, or wig) observed in Study 3, 4, and 5. Baseline degrees of freedom (df_n, df_d) in testing were (2, 58), (2, 258), and (2, 198) respectively, and in cases where the assumption of sphericity was violated, a Greenhouse-Geisser adjustment (ϵ_{GG}) was applied to the degrees of freedom and the p-value was adjusted accordingly.

that the **personalized robots are more unique, as well as more humanlike,** than the non-personalized robot. We also observed a significant main effect of personalization on *pleasantness* (Study 3), as well as *unpleasantness* and *approachability* (Study 4) – the directionality of which was mixed. Compared to the non-personalized robot, Study 3 participants rated the personalized depictions as more pleasant; whereas Study 4 participants rated the personalized depictions as more unpleasant and less approachable. See Table 3.

Qualitative themes: Although many participants commented on the humanizing effect of personalization, their comments also reflected the sentiment that personalization is not necessarily, or partic**ularly, positive**, echoing our quantitative observations. In Study 3, 30% described the personalized robots as 'trying to be human' and, while some comments reflected a positive tone (e.g., "it's trying its best (laughs) to emulate humans, or human nature"), 39% expressed negative sentiments ranging from dismissive (e.g., "it just looks silly seeing clothes on it") to discomforted (e.g., "this robot's definitely more creepy than the other one [...] it's almost trying too much to be a human"). Similarly, in Study 4, 9% described the personalized robots as 'trying to be human', with sentiments again ranging from dismissive (six participants described personalization as 'ridiculous' and five described it as 'silly': e.g., "I preferred the robots without much human clothing as I think it made them look rather silly" and "I feel like when you try to make a robot look 'too human' when it doesn't really look like a real human, it makes it look a little silly") to discomforted (e.g., "the robots that were 'trying too hard' to imitate humans were creepy", "adding human-like hair (in the form of a wig) really turned up the creep factor, as it was like an alien trying to pose as a human, and failing miserably") and even opposed to personalization (e.g., "at the end of the day, I want the robots to be themselves and not try to be human to please me", "when robots try to look and behave like humans, that makes them less appealing to me", and "these are machines and they will never be accepted as humans").

Effects of headwear: Across both studies, participants' responses to the personalized robots (i.e., excluding responses to the non-personalized robot) showed a consistent, significant main effect of headwear on *uniqueness*, *pleasantness*, *unpleasantness*, and *approachability*, with ratings reflecting the perception that, relative to depictions of the robot outfitted with a hat and no headwear, **robots outfitted with a wig are more unique**, **but also less pleasant**, **more unpleasant**, **and less approachable**. See Tables 4 and 5.

Qualitative themes: Participants' free-response feedback similarly fixated on use of the wig and its unpleasant effects. 96% of respondents in Study 3 mentioned the wig, with the majority (73%) commenting that the wig had a negative effect on how they viewed the robot (e.g., "I changed my answer [...] to 'somewhat eerie' mostly 'cause of the wig", "I think wigs are inherently kinda creepy", "The more I see the wig, the more I hate it. It's horrible", "I disliked the wigs a lot", and "the hair makes it really creepy for some reason [...] it seems like something a serial killer would have").

In Study 4, 61% of participants mentioned the wig in their comments. While 18% of the mentions were neutral-to-positive (e.g., "I liked it when the robots had hats or wigs on, since it made it feel like there was more of a head on it"), 82% were strongly negative. For example, three participants compared the effect of the wig to the movie *Psycho* and others noted particular discomfort (e.g., "the wigs generally gave me serial killer vibes", "putting a wig on was the worst thing you could do", and "I think the ones that had the wigs on were the creepiest things I've ever seen").

Effects of clothing variations: We observed two significant main effects of the completeness of the robot's outfit on pleasantness in Study 3 (p = .015) and humanness in Study 4 (p < .001), with ratings reflecting the perception that a complete outfit is more pleasant and more humanlike than a shirt alone. A significant interaction with headwear (p = .037), however, suggests that the humanizing effect of a complete outfit is limited to depictions with no headwear or a hat. See Tables 6, 7, and 8 in the Appendix.

Qualitative themes: Echoing the observed effects of outfit completeness, 52% of Study 3 participants and 9% of Study 4 participants commented on the degree of clothing, with 'complete' outfits (shirt and shorts) perceived more positively than a shirt alone (e.g., "I was influenced by whether the robot was fully clothed in [having] the pants and a shirt and the hat"). However, despite no other significant effects of the clothing manipulations on ratings, participants' commentary reflected further sensitivity to the outfit composition (e.g., "I don't like the pairing between the shirt and this hat. I think the shirt looks like very formal and the hat doesn't", "...kind of works to make a more complete look, as opposed to the prior photo, which just kind of seemed like two items were thrown together at random", and "I like this outfit because I think it matches the best").

Certain item combinations even yielded attribution of social traits (e.g., "I preferred the red shirt to the green. For some reason that combination felt more human-like and warm to me"). Robots outfitted in the white shirt and red cardigan, in particular, were

| | | Study 3 $(df = 29)$ | | | | | Study 4 $(df = 129)$ | | | | | Study 5 $(df = 99)$ | | | | |
|-----------------|-------|---------------------|--------|--------|-------|-------|-----------------------------|--------|--------|------|-------|---------------------|--------|--------|------|--|
| | M_d | SE | t | p | d | M_d | SE | t | p | d | M_d | SE | t | p | d | |
| uniqueness | .292 | .082 | 3.546 | .002 | .647 | .519 | .071 | 7.228 | < .001 | .639 | .255 | .072 | 3.527 | .002 | .353 | |
| uniqueness | .229 | .082 | 2.786 | .022 | .509 | .384 | .071 | 5.398 | < .001 | .473 | .240 | .072 | 3.320 | .003 | .332 | |
| plaasantnass | 483 | .139 | -3.470 | .003 | 634 | 496 | .068 | 7.303 | < .001 | 641 | 350 | .088 | -3.961 | < .001 | 396 | |
| pleasantness | 592 | .139 | -4.248 | < .001 | 776 | 371 | .068 | -5.463 | < .001 | 479 | 365 | .088 | -4.131 | < .001 | 413 | |
| unpleasantness | .739 | .133 | 5.566 | < .001 | 1.016 | .833 | .085 | 9.763 | < .001 | .856 | .652 | .086 | 7.573 | < .001 | .757 | |
| unpieusummess | .750 | .133 | 5.650 | < .001 | 1.032 | .720 | .085 | 8.440 | < .001 | .740 | .598 | .086 | 6.953 | < .001 | .695 | |
| approachability | 300 | .098 | -3.054 | .010 | 558 | 387 | .071 | -5.461 | < .001 | 479 | 460 | .087 | -5.276 | < .001 | 528 | |
| | 338 | .098 | -3.436 | .003 | 627 | 333 | .071 | -4.700 | < .001 | 412 | 380 | .087 | -4.359 | < .001 | 436 | |

Table 5: Post hoc paired comparisons (mean difference, standard error, Students' t-statistic, p-value, and Cohen's d) of participants' ratings of the personalized robots outfitted with wigs vs. robots with no headwear (top) and vs. robots outfitted with hats (bottom) for each significant main effect of headwear. No significant differences between hats and no headwear were found.

described as having a matronly or feminine aura (e.g., "nosy old lady gardening outfit", "clothing [...] an old lady would wear", and "clothes that a grandmother would wear"). While some attributions were negative (e.g., "The red cardigan and white shirt was a turn-off to me. It reminded me of what a supervisor/office snitch I used to work for wore"), comments generally ranged from neutral to positive (e.g., "The same person would wear the hat as would wear this sweater, so it gives the robot a more cohesive personality", "For some reason I liked the more cohesive outfits (shirt/sweater/shorts) the most and found those robots the most friendly", and "I really liked the white shirt and red jacket combo. It's extremely pleasant and reminds me of someone who is going to be nice").

3.3 Summary of Findings

Across Studies 3 and 4, the data – both quantitative (i.e., ratings) and qualitative (i.e., comments) – reflect the perception that personalized telepresence robots are more unique and humanlike than non-personalized robots. At the same time, participants – from the perspective of prospective interlocutors – expressed discomfort with personalization, and ratings by Study 4 participants reflect the perception that personalized depictions are less pleasant, more unpleasant, and less approachable than the non-personalized robot. However, as evidenced by the degree of negative attention wigs received in participants' comments and the perception reflected by participants' ratings that robots outfitted with a wig are less pleasant, more unpleasant, and less approachable than all other depictions, this discomfort seems to stem from wigs in particular.

4 STUDY 5

Having found that prospective telepresence users commonly selected wigs to personalize their robot (Studies 1–2), but also that the wig personalizations were perceived by observers as particularly unpleasant and less appealing to interact with (Studies 3–4), we designed a fifth study to investigate the reproducibility of the findings from both sets of studies, as well as the association between wig selection and wig perception. Specifically, here we asked participants for their preferred personalization of a telepresence robot (from the perspective of the user), as well as their feedback (from

the perspective of an interlocutor) on a subset of the previously-designed robot personalizations. Due to the COVID-19 pandemic, we carried out Study 5 fully online. All procedures were approved by the OSU Institutional Review Board (protocol #IRB-2019-0172).

4.1 Method

Here we assessed what items a participant would prefer to use to personalize a robot for their own use (i.e., perspective of operator) in addition to further evaluating the effects of **personalization** (non-personalized vs. personalized) and **headwear** manipulation via a subset of the personalized depictions used in Studies 3–4.³

Participants & procedure: 100 people, recruited from Oregon State University, participated in Study 5. The incentive for participating was course credit, and like participants in Studies 3–4, Study 5 participants indicated little familiarity with robots (see Table 1). Participants completed the same consenting process and survey as in Studies 3-4, with one major change to incorporate an assessment of participants' personalization preferences (before respondents saw any examples of personalized robots, they were asked to "imagine you were to use this robot to attend your classes remotely" and select what personalization items they would use in this scenario).

Measures: In addition to retention of the measures used in Studies 3–4 to evaluate perceptions of the defined robot personalizations, we also recorded participants' selections from the 12 personalization options offered in Studies 1–2 (see Table 2).

4.2 Results

We again analyzed participants' comments using thematic analysis and ratings of the robot depictions using the same statistical models/methods (i.e., paired *t*-test, repeated measures ANOVA).

Personalization preferences: Although prompted to adopt the operator's perspective, participants' item selections differed substantially from those by participants in Studies 1–2 (see Table 2). Here, name tags and stickers were the most frequently selected

³With benefits of complete outfits reflected by pleasantness ratings (Study 3), humanness ratings (Study 4), and participants' commentary (Studies 3–4), we retained only fully-clothed depictions in Study 5, thus eliminating this factor. We, however, retained the green vs. red top manipulation, lacking sig. differences in perceptions thereof.

items, with few selections of clothing or headwear. Moreover, unlike in Studies 1–2 wherein all participants opted to personalize the robot with *at least two items*, participants here used an average of two items (SD=1; range: 0-7) and 25% selected no items to personalize the robot, reflecting less interest in personalization.

Effects of personalization: Like in Studies 3 and 4, we observed a significant main effect of personalization (none versus personalized) on *humanness*, with ratings again reflecting the perception that, relative to the non-personalized robot, the **personalized robots are more humanlike**; and, like in Study 4, we observed a significant main effect of personalization on *unpleasantness*, with participants again indicating the **personalized robots are more unpleasant** than the non-personalized robot. See Table 3.

Qualitative themes: Overall, reactions to personalization were more mixed than in Studies 3–4. While 41% commented on perceived negatives of personalization (e.g., "it felt patently false"), 38% expressed a receptiveness to the use of clothing (e.g., "I think outfits add something to it") and 19% commented fairly neutrally (16% commented on the humanizing aspects of personalization and three voiced that though they had no negative impression of the personalization, they also perceived no benefit of it).

Effects of headwear: Like in Studies 3–4, we observed a significant main effect of headwear on participants' ratings of the personalized robots' *uniqueness*, *pleasantness*, *unpleasantness*, and *approachability*. Ratings again reflected the perception that, relative to the robots outfitted with a hat and robots with no headwear, **robots outfitted with a wig are more unique**, **but also less pleasant**, **more unpleasant**, **and less approachable**. See Tables 4 and 5.

Qualitative themes: Consistent with participants' feedback in Studies 3–4, the majority of participants (54%) here commented on the wig. Though 9% were relatively neutral (e.g., "I did not expect it") and perhaps even positive (e.g., "the hair makes it seem more real"), 45% were expressly negative (e.g., "the wig was a little creepy").

Preference vs. perception: To investigate associations between preferences for a wig in personalizing a telepresence robot for oneself and perception of our example personalizations involving a wig, we used Spearman's test to compute the correlation between participants' *selection* of the wig (selected vs. not selected) and their evaluations of the robot outfitted with a wig (with ratings collapsed across repetitions), but no significant results emerged (*uniqueness*: $\rho = -.031$, p = .762; *humanness*: $\rho = .062$, p = .539; *pleasantness*: $\rho = -.012$, p = .909; *unpleasantness*: $\rho = .078$, p = .443; and *approachability*: $\rho = -.002$, p = .981).

4.3 Summary of Findings

Like in Studies 3–4, participants' ratings here reflected the perception that personalized robots are more humanlike but also less pleasant than the non-personalized robot. The sentiments reflected in participants' comments, however, were more mixed than in Studies 3–4, with a nearly equal number of positive and negative reactions. Nevertheless, both ratings and comments continued to reflect the perception that while robots outfitted with a wig are more unique, they are also less pleasant, more unpleasant, and

less approachable than all other depictions, and, consistent with these perceptions, the *wig* was not a commonly selected item in participants' personalization preferences as prospective operators.

5 DISCUSSION

The aim of the present work was to investigate the use and effects of personalization in embodied telepresence, from the perspective of both (prospective) operators and interlocutors. In Studies 1–2, we identified common personalization preferences by observing the frequency with which operators elected to use personalizing items such as clothing and headwear. Study 2 also provided insight into the impacts of personalization on interlocutors, underscoring several social dimensions of relevance to their reception. We then systematically manipulated the appearance of an Ohmni telepresence robot based on the personalization strategies observed, and, via Studies 3–5, evaluated perceptions thereof by prospective interlocutors along five constructs (uniqueness, humanness, pleasantness, unpleasantness, and approachability) derived from Study 2.

5.1 Overall Findings

Prospective operators embraced personalization and exhibited particular preference for clothing and headwear. Across Studies 1–2, we observed an average use of five items (SDs = 1-2) to personalize the Ohmni telepresence robot and, in both studies, the majority of operators elected to use *shorts*, *shirts*, *wigs*, and *hats*. Via Study 2's dyadic design, we additionally observed that the operator-specified personalizations have humanizing effects on perceptions by both operators and interlocutors which may facilitate the treatment of the robot as a human proxy. At the same time, Study 2 interlocutors voiced substantial discomfort, with 40% of participants mentioning uncanny valley-like effects such as creepiness, weirdness, and uncanniness of the personalized robot.

Personalization, however, was not entirely positive. Systematic investigation via Studies 3–5 yielded further observations that mirrored those from Studies 1–2. Specifically, participants' ratings and comments (from the perspective of interlocutors) reflect the perception that personalized robots appear more humanlike than a non-personalized robot. At the same time, ratings also reflect the perception that personalizations are more unpleasant (Studies 4–5) and less approachable (Study 4).

Among the personalizing features investigated, **prospective interlocutors disliked** *wigs* **in particular**. Across Studies 3–5, participants' ratings consistently reflect the perception that the robot outfitted with a wig is less pleasant, more unpleasant, and less approachable than all other personalizations investigated, and the most common theme that emerged from participants' verbal and written comments was their discomfort with the wig.

The reason for tension between operators' preferences for and interlocutors' dislike of wigs remains unclear. By prompting participants to adopt the perspective of a prospective operator in addition to that of an interlocutor, Study 5 provided a more in-depth look at how a person's preferences for personalization items might relate to their perceptions of personalized telepresence robots. However, testing for correlations between participants' selection of wigs (selected vs. not selected) and their ratings of the

robot outfitted with a wig did not reveal any significant associations. But, unlike the preferences of participants in Studies 1–2, Study 5 participants selected few personalizing items (M=2,SD=1) and just 9% of participants selected wigs, resulting in limited statistical power available for the correlational analyses. Correspondingly, further investigation is warranted.

5.2 What is wrong with wigs?

Prior research on human-robot interaction has found that people tend to struggle with making sense of whether telepresence robots are people, machines, or something else [9, 30], and robots that are simultaneously too humanlike and not humanlike enough can provoke significant discomfort in observers [7, 10, 24]. Unlike other personalizing items, wigs are made to replace hair, not just cover heads - thus use of a wig in particular may have called too much attention to the fact that the physical telepresence robot is not a living, breathing person. For example, participants described the personalized robots as both trying to be human (e.g., "the hair makes it seem more real", "seems as if the robot is trying to disguise itself as not a robot") and creepy (e.g., "I think the ones that had the wigs on were the creepiest things I've ever seen"). Wigs may also be perceived more like prostheses than they are as clothing or accessories, which can elicit uncanny valley-like aversion [13, 25]. This may also explain why we did not observe significant, negative impacts of other headwear (e.g., hats), as well as why we observed infrequent selection (15%, 28%, and 0% of participants in Studies 1, 2, and 5) of other hair-based items (e.g., facial hair). Careful consideration of hair-based personalizations in relation to these hypotheses would thus likely help illuminate the mechanisms underlying the different reactions to personalization that we observed here.

5.3 Broader Implications

Though participant comments and ratings across Studies 3–5 reflect negative perceptions of personalization, we did not observe *any* significant differences in perceptions of the non-personalized robot and the non-wig-wearing personalized depictions, suggesting that personalization is not strictly negative. For example, prospective operators in Studies 1–2 expressed substantial interest in personalization, and prospective interlocutors in Studies 3–5 expressed some positive reactions to the use of clothing (e.g., "I really liked the white shirt and red jacket combo. It's extremely pleasant and reminds me of someone who is going to be nice"). Moreover, the attribution of social traits elicited by certain clothing choices may benefit operators and interlocutors alike (e.g., automatic projection that the operator is nice may elicit more prosociality from interlocutors which, in turn, may facilitate more effective interactions).

At the same time, operators' natural personalization preferences may not be well received by interlocutors. In particular, while the majority of Study 1–2 operators gravitated toward wigs, wigs provoked significant discomfort in interlocutors (Studies 2–5) suggesting that wigs should likely be avoided in personalization. Moreover, the sensitivity to superficial variation in clothing reflected by interlocutors' comments (e.g., the perception of the white shirt-red cardigan combination as a "nosy old lady gardening outfit") suggests that attention to the remote social environment and co-present interlocutors is important when personalizing a telepresence robot.

5.4 Limitations

While our findings speak to how onlookers may perceive robot personalizations in ad hoc interactions with telepresence robots, there are a number of limitations that highlight potential avenues for further research. In particular, with the use of two distinct setups across the five studies (in-person interactive studies and non-interactive online surveys) it is possible that participants' responses to the personalized robots were amplified or attenuated relative to the specific context. For example, if wigs appear less unpleasant in person, it may explain the differences in wig selection among participants in Studies 1-2 vs. in Study 5. It is also possible that the online modality constrains imagination of the robot's physical presence, which may explain Study 5 participants' limited personalization preferences as prospective operators relative to those of operators in Studies 1-2. Although we attempted to minimize potential differences by including a video demonstrating the Ohmni robot in operation at the start of the online survey, various comments suggest that at least some participants saw the platform as a robot and not a human proxy (e.g., "at the end of the day, I want the robots to be themselves and not try to be human to please me"). Thus, there is more exploration to be done of the effects of the evaluation context (e.g., physically co-present interaction vs. photo-based observation).

Another limitation of the current work is that we used a constrained set of personalization items rather than a larger, comprehensive collection. Thus, while our observations offer guidance for future work (e.g., as no significant differences emerged across outfits of varying gender typicality, it is unlikely that manipulating shirt color will be a fruitful area of further study in understanding perceptions of telepresence robot personalizations, at least in the contexts studied here), broader exploration of robot personalizations is still needed. For example, colors have social meaning in other contexts (e.g., sports teams, political parties). Also, people's perceptions may be influenced by both their gender and the robot's gendering (e.g., [23]), as well as the congruence between the gender of the depicted operator and the gender typicality of the personalizations. Overall, the present work provides preliminary identification of preferences for and impacts of personalization; still, consideration of further social and cultural contexts is warranted.

6 CONCLUSIONS

This research was inspired by field observations we have made of people spontaneously personalizing telepresence robots (e.g., scarfs on Beams) in ways that differ from decoration of other robots (e.g., stickers on Roombas). To systematically explore how personalization impacts embodied telepresence, we carried out two studies of operator personalization preferences and three experiments on interlocutors' perspectives. Together, the results reveal that, while personalization may improve operator presence and individuation, it may not necessarily be well-received by interlocutors.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation (NSF IIS-1528121). We were also assisted in the design of Studies 1–2 by Elizabeth Cha and in data collection by Chris Sanchez.

REFERENCES

- [1] 2019. Kuka Robotics. https://www.kuka.com
- [2] 2019. Savioke. https://www.savioke.com
- [3] 2019. Tesla. https://www.tesla.com
- [4] Jan O Blom and Andrew F Monk. 2003. Theory of personalization of appearance: why users personalize their pcs and mobile phones. *Human-computer interaction* 18, 3 (2003), 193–228.
- [5] Naomi T Fitter, Yasmin Chowdhury, Elizabeth Cha, Leila Takayama, and Maja J Matarić. 2018. Evaluating the Effects of Personalized Appearance on Telepresence Robots for Education. In Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI). 109–110.
- [6] Naomi T Fitter, Luke Rush, Elizabeth Cha, Thomas Groechel, Maja J Matarić, and Leila Takayama. 2020. Closeness is Key over Long Distances: Effects of Interpersonal Closeness on Telepresence Experience. In Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction. 499–507.
- [7] Jari Kätsyri, Klaus Förger, Meeri Mäkäräinen, and Tapio Takala. 2015. A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. Frontiers in psychology 6 (2015), 390.
- [8] Annica Kristoffersson, Silvia Coradeschi, and Amy Loutfi. 2013. A review of mobile robotic telepresence. Advances in Human-Computer Interaction 2013 (2013), 3.
- [9] Min Kyung Lee and Leila Takayama. 2011. Now, I have a body: Uses and social norms for mobile remote presence in the workplace. In *Proceedings of the SIGCHI* conference on human factors in computing systems. 33–42.
- [10] Maya B Mathur and David B Reichling. 2016. Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. *Cognition* 146 (2016), 22–32.
- [11] Albert Mehrabian. 1997. Comparison of the PAD and PANAS as models for describing emotions and for differentiating anxiety from depression. Journal of psychopathology and behavioral assessment 19, 4 (1997), 331–357.
- [12] Louise Mifsud and Anders I Morch. 2007. 'That's my PDA!'The Role of Personalization for Handhelds in the Classroom. In Fifth Annual IEEE International Conference on Pervasive Computing and Communications Workshops (PerComW'07). IEEE. 187–192.
- [13] Masahiro Mori. 1970. Bukimi no tani [the uncanny valley]. Energy 7 (1970), 33–35.
- [14] Carman Neustaedter, John Tang, Samarth Singhal, Rui Pan, Yasamin Heshmat, and Azadeh Forghani. 2017. Telepresence Attendance at the ACM CSCW 2016 Conference. Connections Lab Technical Report, 2017-0419-01, Simon Fraser University (2017)
- [15] Veronica Ahumada Newhart and Judith S Olson. 2017. My student is a robot: How schools manage telepresence experiences for students. In Proceedings of the 2017 CHI conference on human factors in computing systems. ACM, 342–347.
- [16] Veronica Ahumada Newhart, Mark Warschauer, and Leonard Sender. 2016. Virtual inclusion via telepresence robots in the classroom: An exploratory case study. International Journal of Technologies in Learning 23, 4 (2016), 2327–2686.

- [17] Sean O'Brien. 2017. Employees work in a robot-transformative environment. Reading Eagle (Apr 2017). https://www.readingeagle.com/business-weekly/article/employees-work-in-a-robot-transformative-environment
- [18] Gary M Oson and Judith S Olson. 2000. Distance matters. Human-Computer Interaction 15, 2-3 (2000), 139–178.
- [19] Irene Rae, Leila Takayama, and Bilge Mutlu. 2012. One of the gang: supporting in-group behavior for embodied mediated communication. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 3091–3100.
- [20] Byron Reeves and Clifford Ivar Nass. 1996. The media equation: How people treat computers, television, and new media like real people and places. Cambridge University Press.
- [21] Adam Satariano, Dina Bass, and Jack Clark. 2015. This Robot Is Cute, Artificially Intelligent and Employed. Bloomberg (May 2015). https://www.wired.com/2015/ 02/incredible-hospital-robot-saving-lives-also-hate/
- [22] Allison Sauppé and Bilge Mutlu. 2015. The social impact of a robot co-worker in industrial settings. In Proceedings of the 33rd annual ACM conference on human factors in computing systems. 3613–3622.
- [23] Megan Strait, Priscilla Briggs, and Matthias Scheutz. 2015. Gender, more so than age, modulates positive perceptions of language-based human-robot interactions. In 4th international symposium on new frontiers in human robot interaction. 21–22.
- [24] Megan Strait, Lara Vujovic, Victoria Floerke, Matthias Scheutz, and Heather Urry. 2015. Too much humanness for human-robot interaction: exposure to highly humanlike robots elicits aversive responding in observers. In Proceedings of the 33rd annual ACM conference on human factors in computing systems. ACM, 3593–3602.
- [25] Megan K Strait, Victoria A Floerke, Wendy Ju, Keith Maddox, Jessica D Remedios, Malte F Jung, and Heather L Urry. 2017. Understanding the uncanny: both atypical features and category ambiguity provoke aversion toward humanlike robots. Frontiers in psychology 8 (2017), 1366.
- [26] Amanda Sullivan and Marina Umaschi Bers. 2017. Dancing robots: integrating art, music, and robotics in Singapore's early childhood centers. *International Journal of Technology and Design Education* (2017), 1–22.
- [27] Ja-Young Sung, Rebecca E Grinter, and Henrik I Christensen. 2009. Pimp my Roomba: Designing for personalization. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 193–196.
- [28] Ja-Young Sung, Lan Guo, Rebecca E Grinter, and Henrik I Christensen. 2007. "My Roomba is Rambo": Intimate home appliances. In *International Conference on Ubiquitous Computing*. Springer, 145–162.
- [29] Dag Sverre Syrdal, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L Walters. 2009. The negative attitudes towards robots scale and reactions to robot behaviour in a live human-robot interaction study. Adaptive and emergent behaviour and complex systems (2009).
- [30] Leila Takayama and Janet Go. 2012. Mixing metaphors in mobile remote presence. In Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work. 495–504.

A DESCRIPTIVE AND TEST STATISTICS

All main effects of **personalization** (personalized versus nonpersonalized) on ratings of the included robot depictions in terms of *uniqueness*, *humanness*, *pleasantness*, *unpleasantness*, and *approachability* are reported in Section 3.2: Table 3, and the corresponding descriptives are reported in Table 9 below.

Table 8 below reports the test statistics – F statistic, Greenhouse-Geisser correction (where applicable), p-value, and effect size (η_p^2) – for all main and interaction effects of the personalization manipulations in Studies 3–4 (**headwear**, **outfit completeness**, and **gender typicality**). As Study 5 involved only one manipulation of interest (**headwear**) all test statistics for Study 5 analyses of variance are already reported (in Section 3.2: Table 4).

Descriptive statistics corresponding to the main effects of **headwear** (Studies 3–5) are reported in Table 10 below and results of post hoc paired comparisons are reported in Section 4.2: Table 4 (wig vs. no headwear and wig vs. hat contrasts) and Table 11 below (no headwear vs. hat contrasts). Descriptive statistics for and test statistics from post hoc paired comparisons within the main effects of **outfit completeness** (Studies 3–4) and the **headwear** × **outfit completeness** interaction (Study 4) are reported in Tables 6–7. No other main or interaction effects were observed.

| | descri | ptives | contrast | | | | | | | | |
|---|-------------------|-------------------|----------|------|-------|--------|------|--|--|--|--|
| | shirt only | plus shorts | M_d | SE | t | р | d | | | | |
| 1 | 4.339 ± 1.092 | 4.511 ± 1.092 | .172 | .067 | 2.567 | .015 | .470 | | | | |
| 2 | 2.960 ± 1.644 | 3.171 ± 1.821 | .210 | .055 | 3.816 | < .001 | .335 | | | | |

Table 6: Descriptive statistics $(M \pm SD)$ for pleasantness ratings (Study 3; Row 1) and humanness ratings (Study 4; Row 2) of the personalized robots outfitted with a shirt and shorts versus those outfitted with a shirt only, as well as test statistics – mean difference (shirt plus shorts – shirt only), standard error, Student's t statistic, p-value, and Cohen's d – from post hoc comparisons thereof.

| | descri | ptives | contrast | | | | | | | |
|------|-------------------|-------------------|----------|------|-------|--------|------|--|--|--|
| | shirt only | plus shorts | M_d | SE | t | р | d | | | |
| none | 2.904 ± 1.481 | 3.125 ± 1.702 | .313 | .074 | 4.209 | < .001 | .241 | | | |
| hat | 2.977 ± 1.527 | 3.204 ± 1.789 | .227 | .074 | 3.066 | .035 | .288 | | | |
| wig | 3.000 ± 1.681 | 3.092 ± 1.793 | .092 | .074 | 1.247 | > .999 | .111 | | | |

Table 7: Descriptive statistics $(M\pm SD)$ for humanness ratings of the personalized robots outfitted with a shirt and shorts versus those outfitted with a shirt only, by *headwear* (none, hat, or wig), and post hoc, paired comparisons thereof including *mean difference* (shirt plus shorts – shirt only), standard error, Student's t statistic, p-value, and Cohen's t.

| | | | Study 3 | | | | Study 4 | | | |
|-----------------|--|-------------------|---------|-----------------|--------|------------|---------|-----------------|--------|------------|
| | | | F | ϵ_{GG} | р | η_p^2 | F | ϵ_{GG} | p | η_p^2 |
| | headwear | | 6.970 | .783 | .004 | .194 | 28.609 | .774 | < .001 | .182 |
| | outfit completeness | 1 | 2.273 | | .142 | .073 | 1.308 | | .255 | .010 |
| : | 1 1 | gender typicality | .787 | (00 | .382 | .026 | 1.405 | | .238 | .001 |
| uniqueness | headwear × | gender typicality | .015 | .692 | .954 | < .001 | 1.373 | | .255 | .011 |
| | headwear × outfit completeness | | .240 | | .787 | .008 | 2.695 | | .069 | .020 |
| | outfit completeness × | | .576 | | .454 | .019 | .222 | 051 | .638 | .002 |
| | headwear × outfit completeness × | gender typicality | 1.379 | | .260 | .045 | 1.389 | .951 | .251 | .011 |
| | headwear | | 1.304 | | .279 | .043 | .260 | .801 | .721 | .002 |
| | outfit completeness | | 1.231 | | .276 | .041 | 14.559 | | < .001 | .101 |
| | | gender typicality | 2.448 | | .129 | .078 | .261 | | .611 | .002 |
| humanness | headwear × | gender typicality | 1.415 | | .251 | .047 | 1.043 | .938 | .350 | .008 |
| | headwear × outfit completeness | | .912 | | .408 | .030 | 3.338 | | .037 | .025 |
| | outfit completeness × | | .144 | | .707 | .005 | .270 | | .604 | .002 |
| | headwear × outfit completeness × | gender typicality | .069 | .836 | .905 | .002 | 2.913 | | .056 | .022 |
| | headwear | | 10.230 | .808 | < .001 | .261 | 28.858 | .723 | < .001 | .183 |
| | outfit completeness | | 6.637 | | .015 | .186 | .036 | | .849 | < .001 |
| | | gender typicality | .738 | | .397 | .025 | 3.124 | | .079 | .024 |
| pleasantness | headwear × | gender typicality | 2.513 | | .090 | .080 | .959 | .917 | .378 | .007 |
| | $headwear \times outfit completeness$ | | .008 | | .992 | < .001 | .077 | .955 | .919 | < .001 |
| | outfit completeness × | gender typicality | .187 | | .668 | .006 | .066 | | .798 | < .001 |
| | headwear × outfit completeness × | gender typicality | 1.180 | .805 | .308 | .039 | .404 | | .668 | .003 |
| | headwear | | 20.970 | .591 | < .001 | .420 | 56.095 | .634 | < .001 | .303 |
| | outfit completeness | | 1.694 | | .203 | .055 | 1.437 | | .233 | .011 |
| | | gender typicality | .067 | | .797 | .002 | 1.437 | | .233 | .011 |
| unpleasantness | headwear × | gender typicality | 2.149 | .787 | .138 | .069 | .109 | .946 | .887 | < .001 |
| | $headwear \times outfit completeness$ | | .268 | .819 | .722 | .009 | 1.248 | | .289 | .010 |
| | outfit completeness $	imes$ | gender typicality | .870 | | .359 | .029 | 1.195 | | .276 | .009 |
| | headwear × outfit completeness × | gender typicality | .355 | | .703 | .012 | .079 | .891 | .906 | < .001 |
| | headwear | | 7.093 | .605 | .008 | .197 | 17.498 | .666 | < .001 | .119 |
| | outfit completeness | | 1.260 | | .271 | .042 | .748 | | .389 | .006 |
| | | gender typicality | .069 | | .795 | .002 | .495 | | .483 | .004 |
| approachability | $headwear \times$ | gender typicality | 3.240 | .759 | .061 | .101 | .639 | | .528 | .005 |
| | $headwear \times outfit\ completeness$ | .033 | | .967 | .001 | .563 | | .570 | .004 | |
| | outfit completeness $	imes$ | gender typicality | .693 | | .412 | .023 | .064 | | .801 | < .001 |
| | $headwear \times outfit \ completeness \times$ | gender typicality | .044 | .685 | .902 | .002 | 1.574 | | .209 | .012 |

Table 8: Main and interaction effects of headwear (none, hat, or wig), outfit completeness (shirt only vs. shirt-plus-shorts), and gender typicality (green vs. red top) in Studies 3–4. Baseline degrees of freedom for all main effects of and interactions with headwear were (2,58) in Study 3 and (2,258) in Study 4, and, for all other effects, (1,29) and (1,129) respectively. In cases where assumptions of sphericity were violated, a Greenhouse-Geisser adjustment (ϵ_{GG}) was applied to the degrees of freedom and p-values were adjusted accordingly.

| | Study | 7 3 | Study | 4 | Study 5 | | | |
|-----------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|--|--|
| | non-personalized | personalized | non-personalized | personalized | non-personalized | personalized | | |
| uniqueness | 3.750 ± 1.165 | $4.396 \pm .994$ | 3.723 ± 1.809 | 4.837 ± 1.267 | 4.210 ± 1.546 | 4.445 ± 1.226 | | |
| humanness | 2.167 ± 1.045 | 3.604 ± 1.395 | 2.085 ± 1.318 | 3.065 ± 1.558 | $1.820 \pm .999$ | 3.217 ± 1.535 | | |
| pleasantness | 4.000 ± 1.137 | $4.483 \pm .797$ | 4.138 ± 1.357 | 4.103 ± 1.320 | 3.830 ± 1.429 | 4.123 ± 1.122 | | |
| unpleasantness | 2.922 ± 1.083 | $3.132 \pm .623$ | 2.759 ± 1.584 | 3.666 ± 1.531 | 3.660 ± 1.443 | 3.977 ± 1.221 | | |
| approachability | $5.667 \pm .922$ | $5.527 \pm .838$ | 5.300 ± 1.471 | 4.788 ± 1.495 | 4.375 ± 1.254 | 4.530 ± 1.589 | | |

Table 9: Descriptive statistics $(M \pm SD)$ for ratings of the personalized and non-personalized robot in terms of their uniqueness, humanness, pleasantness, unpleasantness, and approachability.

| | | Study 3 | | | Study 4 | | Study 5 | | | | |
|---|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|--|--|
| | none | hat | wig | none | hat | wig | none | hat | wig | | |
| 1 | 4.292 ± 1.164 | 4.354 ± 1.068 | 4.583 ± 1.164 | 4.619 ± 1.525 | 4.754 ± 1.506 | 5.138 ± 1.489 | 4.355 ± 1.292 | 4.370 ± 1.300 | 4.610 ± 1.294 | | |
| 2 | 3.600 ± 1.480 | 3.658 ± 1.495 | 3.500 ± 1.593 | 3.060 ± 1.680 | 3.090 ± 1.730 | 3.046 ± 1.803 | 3.190 ± 1.579 | 3.300 ± 1.608 | 3.160 ± 1.697 | | |
| 3 | 4.550 ± 1.036 | 4.658 ± 1.006 | 4.067 ± 1.150 | 4.310 ± 1.502 | 4.185 ± 1.513 | 3.813 ± 1.605 | 4.235 ± 1.194 | 4.250 ± 1.228 | 3.885 ± 1.273 | | |
| 4 | $2.953 \pm .739$ | $2.942 \pm .755$ | 3.692 ± 1.135 | 3.351 ± 1.663 | 3.464 ± 1.715 | 4.184 ± 1.795 | 3.742 ± 1.335 | 3.795 ± 1.315 | 4.393 ± 1.305 | | |
| 5 | $5.583 \pm .875$ | $5.621 \pm .828$ | 5.283 ± 1.251 | 4.935 ± 1.597 | 4.881 ± 1.626 | 4.548 ± 1.785 | 4.555 ± 1.289 | 4.475 ± 1.301 | 4.095 ± 1.458 | | |

Table 10: Descriptive statistics $(M \pm SD)$ for ratings of robots with no headwear, outfitted with a cap, and outfitted with a wig in terms of uniqueness (Row 1), humanness (Row 2), pleasantness (Row 3), unpleasantness (Row 4), and approachability (Row 5).

| | | Study 3 $(df = 29)$ | | | | | Study 4 $(df = 129)$ | | | | | Study 5 ($df = 99$) | | | | |
|-----------------|-------|----------------------------|------|--------|------|-------|-----------------------------|--------|--------|------|-------|------------------------------|------|--------|------|--|
| | M_d | SE | t | p | d | M_d | SE | t | p | d | M_d | SE | t | p | d | |
| uniqueness | 063 | .082 | 760 | > .999 | 139 | 135 | .071 | -1.889 | .180 | 166 | 015 | .072 | 207 | > .999 | 021 | |
| pleasantness | 108 | .139 | 778 | > .999 | 142 | .125 | .068 | 1.840 | .201 | .161 | 015 | .088 | 170 | > .999 | 017 | |
| unpleasantness | .011 | .133 | .084 | > .999 | .015 | 113 | .085 | -1.323 | .561 | 116 | 053 | .086 | 620 | > .999 | 062 | |
| approachability | 037 | .098 | 382 | > .999 | 070 | .054 | .071 | .761 | > .999 | .067 | .080 | .087 | .918 | > .999 | .092 | |

Table 11: Post hoc comparisons (mean difference, standard error, Students' t-statistic, p-value, and Cohen's d) of participants' ratings of the personalized robots with no headwear vs. robots outfitted with hats for each significant main effect of headwear.