



# (Counter-)stereotypical Gendering of Robots in Care: Impact on Needs Satisfaction and Gender Role Concepts in Men and Women Users

Laura Moradbakhti<sup>1</sup> · Martina Mara<sup>1</sup> · Ginevra Castellano<sup>2</sup> · Katie Winkle<sup>2</sup>

Accepted: 19 July 2023 / Published online: 22 August 2023  
© The Author(s) 2023

## Abstract

We present a video-based online study (N=222) examining the impacts of gendering an in-home, socially assistive robot designed to aid with rehabilitative engagement. Specifically, we consider the potential impact on users' basic psychological need (BPN) fulfillment alongside measures of the robot's effectiveness as well as the potential impact on human caregiver gender preferences and propensity to gender stereotype more broadly. Our results suggest that the male-gendering of care robots might be particularly beneficial for men, potentially leading to greater BPN fulfillment than female-gendering. Whilst women also showed some similar gender-matching preference (i.e. preferring the female-gendered robot) this effect was less pronounced. Men who saw the male-gendered robot were also significantly more likely to indicate they would prefer a man, or had no gender preference, when asked about human caregiver preferences. Overall, and in line with (some) previous work, we find no evidence of universal positive impact from robot task-gender typicality matching. Together, our results further strengthen existing calls to challenge the default female-gendering of assistive agents seen to date, suggesting that male-gendering might simultaneously boost positive impact for men users whilst challenging stereotypes regarding who can/should do care work.

**Keywords** Social robots · Basic psychological needs · Rehabilitation · (Gender-)stereotypes

## 1 Introduction

Socially assistive robots (SARs) are those which provide assistive functionality through their social interactions with the user (in contrast or addition to e.g. physical assistance) [1]. Previous work has demonstrated their potential for use as a motivational aid for increasing engagement with therapeutic and/or rehabilitative exercise programs [2–4], driven by the fact that lack of engagement with such exercises is a

well-documented issue (e.g. [5–7]). Other works have considered implications of different robot design factors in this context, typically examining how different design choices (often drawn from psychology literature on persuasion and motivation and/or studies of domain expert practitioners) impact on user behaviour and perceptions of the robot (e.g. [8–10]).

In this work, we instead consider perception of such a robot through the lens of the theory of basic psychological needs (BPNT) [11, 12]. We identify basic psychological needs (BPN) as being particularly salient for evaluation of SAR designs because BPN fulfillment has been linked with positive effects on user engagement with a technological device [13, 14], in addition to overall wellbeing and task motivation (including motivation to exercise [15]).

Specifically in the context of rehabilitation, motivation is highly relevant as many patients struggle to stay motivated and report to feel “bored” or “following a routine” instead of being motivated to engage in the rehabilitation activities [16]. Several technologies have been designed or introduced for rehabilitation support to create higher user engagement, such as sensor based virtual reality for in-home

✉ Laura Moradbakhti  
laura.moradbakhti@jku.at

Martina Mara  
martina.mara@jku.at

Ginevra Castellano  
ginevra.castellano@it.uu.se

Katie Winkle  
katie.winkle@it.uu.se

<sup>1</sup> LIT Robopsychology Lab, Johannes Kepler University Linz, Altenberger Straße 69, 4040 Linz, Austria

<sup>2</sup> Department of Information Technology, Uppsala University, New Ångström, Lägerhyddsvägen 1, 75105 Uppsala, Sweden

rehabilitation exercises for stroke patients [17], game-based rehabilitation for elderly patients [16], or socially assistive robots for self-directed rehabilitation exercise programmes [18]. To contribute to the design and development of such technologies, we believe that the BPNT—as part of the self-determination theory [19], which is a well-established and leading theory of human motivation—can strongly contribute in explaining user's motivation to engage with technology, while simultaneously supporting their long-term well-being. As such, using impact on BPN to guide SAR design would seem like a sensible route towards delivering robots that are simultaneously acceptable and effective.

Based on previous work suggesting that gender of a digital, assistive agent can influence users' BPN satisfaction [20], we examine whether gendering of the social robot Pepper has an impact on perceptions of its potential to impact BPN satisfaction, intention to use (ITU) and acceptability. Further, building on current works questioning the potential for gendering of artificial agents to propagate and/or challenge gender stereotyping [21, 22], we look to examine whether robot gender-task stereotypicality seems to influence respondents' immediate application of such stereotypes when assessing human gender-task suitability.

We present an online, video-based experiment designed to support this investigation. For maximum real-world applicability we utilise carefully designed stimuli reflecting expert-informed design guidelines for use of a socially assistive robot to support therapeutic exercise engagement [23], and specifically recruit respondents who identify as having a long-term health condition and/or disability. The results suggest that being shown a male-gendered (counter-stereotypical) socially assistive care robot influences particularly men respondents to chose a non-stereotype conform human caregiver for themselves, thus a male or any gender caregiver, instead of a female caregiver. Moreover, BPN satisfaction positively correlated with the ascription of masculinity to the robot for men respondents. We conclude that assistive robots should not be female-gendered as a default, instead they can be designed to challenge existing stereotypes and at the same time, as our results show, have a positive impact on the user.

## 1.1 Research Questions

With this work we broadly address five research questions. Firstly, RQ1 A and B are concerned with exploring if/how binary, male versus female gendering of an in-home care robot influences how that robot is perceived by men and women.

**RQ1:** *How does gendering of a SAR for in-home healthcare influence respondents'...*

**A:** *... perception of its potential to impact on their Basic Psychological Needs (BPN)?*

**B:** *... intention to use (ITU) them?*

In particular, we look to see whether previous findings concerning gendering of a digital banking assistant replicate within our in-home care robot context.

RQ2 A and B are then concerned with the potential for gendering of our in-home care robot to influence respondents' gender preferences when thinking about human caregivers, and, more broadly, their propensity to engage gender stereotypes:

**RQ2:** *How does observation of a gendered SAR for in-home healthcare influence respondents'...*

**A:** *... immediate gender preferences for a human caregiver?*

**B:** *... immediate application of gender stereotypes when asked to assess stereotypically male/female traits and occupations?*

Finally, we look to (re-)examine the extent to which anthropomorphism might be necessary/desirable in the context of designing socially assistive robots more broadly, given that one way to avoid the risks of robot gendering would be to design less human-like robots:

**RQ3:** *How do anthropomorphic attributions affect respondents' intention to use the SAR for in-home healthcare?*

## 2 Related Work

### 2.1 (Gendered) Social Robots for Healthcare

Very few previous works specifically consider the impacts of robot gendering within a healthcare context. An exception is the work of Bryant et al. [24], who specifically investigated how (mis)matching robot gendering to (pre-validated) occupational gender role associations impacted perceptions of robot competency across a number of occupations including *home health aid*, *nurse* and *therapist*. They found no impact of robot gendering on perceptions of occupational competency or trust—i.e. they found no differences in perceptions of a male versus female-gendered Pepper robot as a home health aid, nurse, or therapist even though those occupations were associated much more with women than with men when considering human workers. The authors had hypothesized that a social robot matching gender and occupational role stereotypes would obtain higher ratings of perceived trust, based on previous results suggesting that this increased robot acceptance [25]. Work by Eyssel et al. [26] also found evidence for robot gender-role matching influencing perceptions of robots, with a short versus long haired version of the otherwise same robot being perceived as more agentic (stereotypical masculine trait) versus more communal (stereotypical feminine traits) and better suited to stereotypically male versus female tasks respectively.

Work examining robot gendering more broadly has yielded mixed results, potentially also indicating that the effects of robot gendering might emerge from interactions between robot and user gender [27–30] but also that the pertinence of gender (or lack thereof) might vary across different interaction contexts [31, 32]. Our work contributes increased understanding of the implications of robot gendering in the context of in-home healthcare, specifically with respect to perceived potential for impacting on users' Basic psychological needs (BPN) and acceptability via intention to use (ITU). We identify BPN as a particularly pertinent construct for considering SARs for in-home health support given links between BPN and task engagement, well-being and ITU demonstrated in human psychology literature, discussed in the following subsection.

## 2.2 Basic Psychological Needs: Task Engagement, Well-Being and Intention to Use

As part of their Self-Determination Theory, Deci and Ryan (SDT [19, 33]) developed six 'mini-theories' (to use their term), one of which is the theory of basic psychological Needs (BPN; [11, 12]). According to the theory, humans have three universal, psychological needs: autonomy—our desire to have control over a situation; competence—our desire to experience mastery; and relatedness—our desire to care for others and be cared for in return. If a task or activity leads to satisfaction of all three BPN, it creates autonomous motivation (volitional, self-directed motivation) and greater well-being. If one or more of these needs are not fulfilled, we lack autonomous motivation to engage in a task and our well-being can be negatively affected.

Especially in the context of physical activity, the BPN have been described as influential drivers of motivation and well-being [15, 34, 35]. In physical rehabilitation, and more specifically in the context of home-based rehabilitation, BPN fulfillment was found to be a strong predictor for motivation to engage in rehabilitation exercises [36, 37]. A study investigating patients' exercise behaviour following cardiac rehabilitation has shown that patients with higher BPN satisfaction were more motivated to exercise, subsequently leading to more independent exercise behaviour at 3 and 6 weeks post-completion of their cardiac rehabilitation program. Moreover, a recent study on physical rehabilitation [38] confirmed that BPN satisfaction had an impact on patients' (adolescents with a disability or long-term illness) meaningful engagement in physical activity, and that their need satisfaction shaped their experience of physical education (PE) participation. The authors suggest that patients' BPN satisfaction should be supported by health care professionals while patients are attending rehabilitation.

Due to the nature of physical rehabilitation, patients often need additional help at home, mostly supported by an infor-

mal caregiver. The caregiver would provide support in doing the exercises [39] and potentially nudge the patient to engage in the activities that are part of the rehabilitation process [40]. In recent years, technologies and AI systems were developed to relieve informal caregivers of their supporting task and activities (e.g. [41–43]). Nonetheless, until now, there is neither research focusing on specific design factors of robot caregivers, nor linking human–computer interaction in physical rehabilitation with BPN satisfaction. There have however been works demonstrating a link between BPN satisfaction and technology engagement. As the studies have shown, BPN satisfaction led to positive perception of technological devices [44], increased users' intention to use [20]) and engagement with technologies [13]. These works indicate that satisfaction of BPN would play an important role in simultaneously increasing patient intention to use a robot for post-rehabilitation support, and further increasing patient motivation to engage in the tasks necessary for their rehabilitation, ultimately resulting in greater overall patient wellbeing.

One of the aforementioned works specifically demonstrated that an AI Assistant's agency (high/low) and gendered design features (female/male voice) seemingly influenced participants' perceived BPN satisfaction and subsequently their ITU a digital voice assistant in a daily banking context [20]. The female-sounding, high-agency finance coach generated the lowest autonomy satisfaction and ITU scores from men, whilst the female, low-agency finance coach received the highest. Whilst the results for women participants were not significant, the authors identified a tendency in the opposite direction, suggesting that men and women have different preferences with regard to these design factors. More specifically, the results suggest that men had a preference (higher autonomy satisfaction and ITU) for a stereotype-conforming representation of a female assistant (see also e.g. [45]) with a passive interaction style, whereas women preferred a non-conforming condition (high-agency female) with high-agency as a characteristic that would typically be associated with men (c.f. [46, 47]). In the current work, we want to examine whether Moradbakhti et al.'s results [20] might (fail to) replicate in a healthcare context, given that women are typically more associated with assistive and/or domestic health and care work [24]. We also want to assess whether exposing respondents to a counter-stereotypical representation of a male robot caregiver could influence their immediate gender preferences for human caregivers, and/or reduce their propensity to stereotype when considering differences in traits and professions for men and women more broadly, building on recent interest and discourse within the HRI literature.

### 2.3 Robots and Gender Stereotyping

Recent works have called attention to the risks of robot gendering (whether done explicitly by designers or not) with respect to propagating harmful gender stereotypes. A UNESCO report on the digital skills divide calls attention to (worldwide) problematic issues of (female) gendering of digital assistants, noting that “voice assistants built for Chinese and other Asian markets are, like the assistants built by companies headquartered in North America, usually projected as women and also interact with users in ways that can perpetuate harmful gender stereotypes [21, p. 104]. The report specifically discusses risk of increasing associations between “woman” and “assistant” to the point of (real) women being (further) penalized for not being assistant-like [21, p. 108]. Yolande Strengers and Jenny Kennedy further call attention to the problematic nature of smart technologies (including social robots) as *smart wives*: technologies designed to take on archetypal, domestic wife-like such as caregiving and emotional labour and subtly characterized as the ‘*nostalgic, sometimes porn-inspired wifely figure...we are trying to move on from in most contemporary societies*’ [48]. Alesich and Rigby have also drawn attention to the possibility for gendered humanoid robots to affect our perceptions of human gender and associated cultural norms [49].

The UNESCO report identifies “genderless chatbots” as exemplifying one way to “sidestep difficult issues related to gender” [21, p. 122] and recommends exploring the feasibility of developing a “machine gender for voice assistants that is neither male nor female” [21, p. 130]. However, previous works indicate users generally show a high tendency to ascribe (binary) human gender to artificial social actors, including those that are non-anthropomorphic with minimal gender cues [50]. The Pepper robot, a common platform for HRI research, might be said to be somewhat gender ambiguous in its base embodiment design. A recent study asking participants to rate masculinity and femininity of robots based only on their appearance suggested Pepper is most often perceived as gender neutral [51]. This work also demonstrated the ways in which robot form alone can influence robot gendering, as, for Italian participants, body manipulators (e.g., legs, torso) were found to be associated with masculinity whilst surface look features (e.g., eyelashes, apparel) were found to be associated with femininity for Italian participants. A thorough evaluation of visual qualities of SARs and their effects on user perception and preferences revealed that both body structure (hourglass shaped robots were perceived as more feminine and V-shaped robots as masculine) and colour (generally male ascription to robots but white colour was more associated with the female gender) had effects on users’ gendering of SARs [52]. Nonetheless, previous works on robot gendering have generally taken the position that this gender ambigu-

ity supports simple manipulation of Pepper’s gendering via e.g. voice and name manipulation [24, 53–55]. We take the same approach here, motivated in part by Perugia et al.’s recent finding that (Italian) participants seeing a picture of the Pepper robot generally identified it as being (specifically) gender neutral [51], although it must be noted that others have found evidence indicating that Pepper’s design actually communicates a range of genders [56]. This latter finding from Seaborn and Frank suggests future work on robot gendering ought go further than the typical gendering manipulation check employed in previous works (as well as our own) to carefully examine e.g. whether researchers’ gender manipulations (mis)match with any *initial* robot gendering on the part of participants, as this has potential to induce cognitive dissonance in participants (who might have been expecting to ‘hear’ one particular gender when the robot spoke, but were surprised to instead hear something different) and influence results. We discuss this further in Sects. 5.5 and 6.

Perugia et al.’s study also indicated that, when participants did ascribe gender to robots, they overwhelmingly choose to do so in binary gender terms, even when researchers provide the opportunity to indicate androgyny via separate metric rating scales on masculinity and femininity [51]. It is not clear to what extent this reflects designers’ intentions versus observer biases, but it does indicate a difficulty and/or lack of established practice for designing non-binary robots. Future work is needed to understand this, and to explore what non-binary robots could ‘look like’, if we are to fully unlock the potential of queering robots as a mechanism for challenging stereotypes [48].

One proposed alternative to the gender ambiguous approach is to *intentionally* design gendered but norm-breaking robots that, at a minimum, avoid propagating gender stereotypes, but potentially might even aid in reducing such stereotypes in users. Taking this position, [22] demonstrated that a female-presenting university outreach robot which demonstrated gender norm-breaking behaviour in its response to inappropriate behaviour was perceived as more credible by girls (with no negative impact on the boys), and potentially reduced immediate gender bias in some observers. To our knowledge, no such work has showcased an equivalent, norm-breaking male robot. We consider use of a male-presenting robot for our in-home healthcare support scenario as norm-breaking and counter-stereotypical based on the gendered nature of care work [48], in-home health tasks generally being associated with women [24] and current gendering design trends evidenced in commercially available digital assistants [21, 45].

Whilst recent research has examined how to reduce stereotypes towards women in traditionally male jobs [57, 58] there is a research gap on inequalities and stereotypes towards men working in professions that were traditionally executed by women (e.g. care taking; [59]). Moreover, a study has shown



that exposing participants to male, family-oriented role models, had a positive impact on women's outlook on their future work-life balance [60]. These results indicate that a simple exposure to counter-stereotypical role models can positively impact participants' own expectations.

Our work continues this discussion by examining not only how robot gender-task stereotypicality matching influences observers' perceptions of the robot, but also how (mis)matching of such might influence observers' immediate application of stereotypes when thinking about humans' suitability for stereotypically gendered occupations, and/or gender-stereotypicality of particular personality traits/behaviours. Given the current knowledge gap regarding how to successfully (and ethically) design a non-binary robot, we limit ourselves here to binary robot gendering and gender-stereotypicality mismatching as our contribution towards the queering of robot gender portrayal. We hope that future work, undertaken in collaboration with the queer community, will work to go beyond these binary limitations.

### 3 Materials and Methods

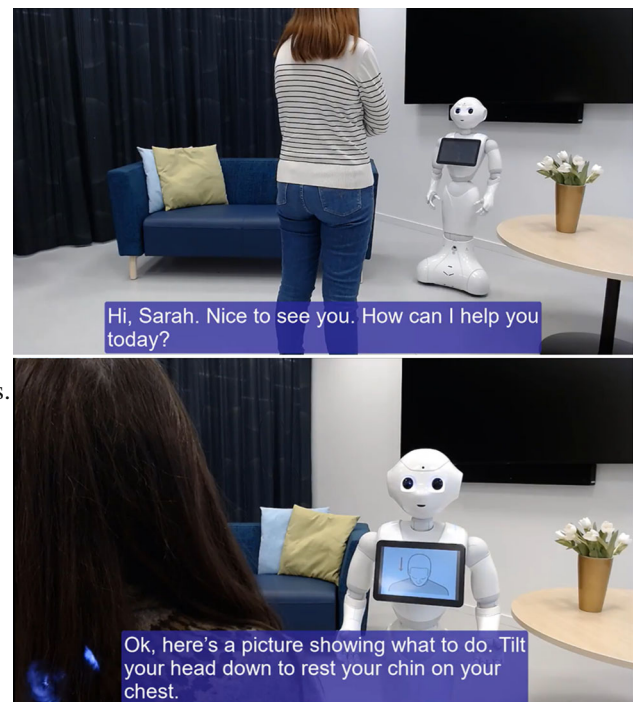
We designed a between-subjects video-based online experiment to investigate the impact of gendering an in-home socially assistive robot.

#### 3.1 Experimental Scenario and Video Production

##### 3.1.1 Experimental Use Case Scenario

Our use case scenario depicts Softbank's Pepper robot<sup>1</sup> deployed within a domestic environment. The videos showcase two users of the robot: the primary user, a care recipient (*Lisa*) and a secondary user, Lisa's informal caregiver (*Sarah*). The video narrative implies that Sarah and the robot work together to deliver Lisa's care, as Sarah asks the robot to guide Lisa through her exercises while she goes out to do other chores. This narrative is based on previous work with therapists regarding the potential role SARs might take in supporting rehabilitative therapies [23, 61]. Lisa and Sarah are fictitious users portrayed by two women on the research team for the purposes of creating these video clips.

The Pepper robot was chosen based on its gender-neutral baseline embodiment [51], which hence lends itself to gendering via voice, as has been done in previous HRI work on the implications of robot gendering [24, 53]. Further, Pepper is generally a popular platform within HRI research, in part because its CE marking<sup>2</sup> makes it a real-world viable option



**Fig. 1** Scene setup for the video stimuli showcasing the Pepper robot interacting first with an informal carer (top) and then with the primary user (bottom)

for use as a socially assistive robot. This means that design decisions regarding if/how and whether to gender the robot or not across different use cases are already happening<sup>3</sup> and we hope work like ours can minimise the risk of such gendering being done ad-hoc without thought to the potential ethical implications.

In the study preface, ahead of stimuli presentation, it was stated that Pepper is 'a robot that can be used at home to support people who need to exercise e.g. to help manage a long term health condition or disability, or to aid recovery from an injury. For now, this is a prototype and we would like to get your feedback. We will ask you questions about your preferences and perception of the robot.'

##### 3.1.2 Video Scripts and Scene Set-Up

The video clips first showed Pepper interacting with the informal carer (Sarah, Fig. 1, top) before working with the primary user/care recipient (Lisa) (Fig. 1, bottom). Pepper and the informal carer also refer, at different times, to Lisa's therapist (e.g. 'You can remind Lisa that the therapist is coming tomorrow') but this therapist never appears directly in any of the videos.

<sup>1</sup> <https://www.softbankrobotics.com/emea/en/pepper>.

<sup>2</sup> <https://www.aldebaran.com/sites/default/files/inline-files/declaration-of-conformity-pepper-1.8.pdf>.

<sup>3</sup> see e.g. "Pepper Robots Dress Up" on the Wall Street Journal: <https://www.wsj.com/video/pepper-robots-dress-up/23725372-6E6C-4C31-9646-C491A5D2D916.html>.

These references to the therapist and the initial carer-robot interaction were included to make it clear that the robot's role is to assist, rather than replace, the primary user's formal and informal human carers. The robot's role with respect to the primary user, formal and informal caregivers, as well as the robot's suggested functionality and dialogue script were all influenced by expert-informed design guidelines for robots in therapy and previous work on socially persuasive robots for exercise motivation [9, 23, 61].

Scene set-up was designed to resemble previous work using video stimuli to investigate perception of in-home care robots [62], utilising some of the same exemplar exercises (neck tilts and stretches taken from publicly available Arthritis Research UK<sup>4</sup> advice materials). Camera placement was such that the users' faces were never seen (thus removing any potential for facial expressions to inform respondents' reflections). Further, only one video was used to create all eventual stimuli by overlaying the robot's gender manipulated speech such that all user behaviour and audio was identical across videos. This removes any potential confounds arising from variations in user or robot behaviour across videos.

Both actors were women on the research team (referred to with fictitious female names) and the depicted user's therapist was also referred to as *she*. Whilst respondents were explicitly asked to imagine themselves in the user's position when watching the video, we actively chose to consistently present all female interactants (rather than e.g. a male user for men respondents) to avoid variations in assessment of the robot being affected by who it is interacting *with*. Previous HRI work has demonstrated complex interactions between interactant, robot and observer gender [53]. Accounting for all possible such gender effects would require versions of the video clip in which we systematically varied primary user, informal and formal caregiver gender as well as the robot gender, which we identify as out of scope for this study. The male<sup>5</sup> and female<sup>6</sup> robot gender conditions can be viewed online. Both videos are 3 min and 33 s long.

### 3.1.3 Robot Gendering Manipulation

Robot "gender" was first manipulated ahead of the video stimulus in the study introductory text where respondents were briefed about the '*[male/female] robot Pepper...[he/she] can support you...*' and then within the video stimuli via the robot's voice. The robot's voice was based on Amazon Web Service's (AWS) synthetic voice software "Polly" which provides a variety of different text-to-speech voices. We entered the robot's text passages in the online platform

ttsmp3.com (which offers text-to-speech downloads powered by AWS Polly) and selected the voice "Joanna" as the female and "Joey" as the male robot voice. As a manipulation check, respondents were asked how feminine and masculine they rated the robot to be on a 5-point Likert scale (note this was presented as two separate questions rather than as one semantic differential to avoid enforcing binary gender ascription).

### 3.1.4 Robot Agency Cues

Moradbakhti's previous work investigating gender and agency in perceptions of a digital banking assistant demonstrated complex interactions between agent gender and agency, indicating complex interplay between stereotypes, social norms and gender roles [63]. In particular, men were particularly unimpressed by the high agency, female version of the assistant. Given that, in this work, we are interested in gendering effects first and foremost, we specifically utilise low agency robot cues to minimise robot agency further exaggerating potential differences in perceptions of our (gendered) robot, even though one might expect (depending upon cultural context) high agency cues to be more appropriate for a healthcare assistant.

## 3.2 Study Procedure

Prior to taking part in the study, all respondents were instructed to either use headphones or keep their computer/laptop audio on high volume for the duration of the study. First, respondents read an introduction, confirmed their consent and filled out demographic information. Once these initial steps were completed, a short instruction appeared, which was followed by the video. One of the two robot gender conditions (female/male-gendered robot) was randomly assigned to each participant. After the video, respondents were asked to rate whether they understood the robot, and about their prior experience with robots, Pepper in particular. Subsequently, respondents answered questions as per the experimental measures presented in Sect. 3.3. Finally, respondents were shown a debrief page before being redirected to Prolific's platform to process their financial compensation. The study took 14 min and 34 s on average.

## 3.3 Experimental Measures

In the following subsections we detail all experimental measures implemented in our survey, in the same order in which they were presented to respondents after they watched the video stimulus. Where using Likert scales in answer options, we utilised 5-point scales as is standard in BPN literature [64–66].

<sup>4</sup> <https://www.versusarthritis.org/>.

<sup>5</sup> <https://youtu.be/Ed806gR97Uk>.

<sup>6</sup> [https://youtu.be/7XG3Gvwdq\\_Q](https://youtu.be/7XG3Gvwdq_Q).

**Table 1** Manipulation check results spearman correlation

Measure	1	2	3
Video condition female			
Video condition male	− 1.00 **		
Robot femininity	.647**	− .647**	
Robot masculinity	− .630**	.630**	− .442**

\* $p < .05$ ; \*\* $p < .01$  (two-tailed)

### 3.3.1 Human-Likeness/Robot-Likeness

Respondents' perception of the robot's human-/robot-likeness was assessed via the following two questions: (1) "On a scale from 1–5, how human-like would you rate the robot?" and (2) "On a scale from 1–5, how robot-like would you rate the robot?".

### 3.3.2 Robot Gendering Manipulation Check

Respondents' perception of the robot's gender was assessed via the following two questions: (1) "On a scale from 1–5, how feminine would you rate the robot?" and (2) "On a scale from 1–5, how masculine would you rate the robot?". Both questions were rated on a five-point Likert scale.

A strong significant correlation between the intended robot gender perception and actual perceived robot gender for both video conditions demonstrates that respondents identified the gendering of the robot in line with our expectations (see Table 1).

### 3.3.3 Perceived Benefit of the Robot

Respondents were asked to indicate their overall perception of the robot's benefit: "Do you think you would benefit from using a robot like this for daily exercises?". respondents answered this question on a five-point Likert scale from 1 (not at all) to 5 (very much).

### 3.3.4 Robot Goodwill

We originally intended to measure robot credibility, a proxy for robot persuasiveness as per previous works [9, 22, 62] but ultimately only included the *goodwill* subscale based on significant overlap between the other key subscales (*expertise*, *trustworthiness*) and our other experimental measures. We can report a good model fit based on CFA: RMSEA = 0.050; SRMR = 0.019; CFI = 1.00; TLI = 0.999. We can also report good reliability with a Cronbach's  $\alpha$  of .93.

### 3.3.5 Basic Psychological Need Satisfaction

The Basic Psychological Need Satisfaction Scale for Technology Use (BPN-TU) was created through thorough scale development and validation [67]). The items are based on existing Basic Psychological Needs Scales (e.g., [64, 68, 69]) but adapted to the context of technology interaction. The items were translated into English for the current study, having originally been developed in German. Each need was assessed with three items. Relatedness was split into Relatedness to Others and Relatedness to the Robot, assessed with three items each. The Relatedness Need is split in the BPN-TU scale, as respondents can perceive the robot as both, a mediating tool to interact with other people (Relatedness to Others), or directly as a source of social interaction (Relatedness to the Robot). The full list of items can be found in "Appendix A".

The items were rated on a five-point Likert scale from 1 (not at all) to 5 (very). To measure the validity of the scale for the current data set, the model fit was evaluated with a Confirmatory Factor Analysis using the  $\chi^2$  goodness-of-fit statistic and a combination of the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI), and the Root Mean Square Error of Approximation (RMSEA). In line with evaluation standards [70, 71], the model has a good fit: RMSEA=0.049; SRMR=0.042; CFI=0.998; TLI=0.998. Autonomy had a Cronbach's  $\alpha$  score of .86, Competence .91, Relatedness to Others .78 and Relatedness to Technology .88.

### 3.3.6 Intention to Use

We utilised two ITU items based on the ITU items from the Technology Acceptance Model (TAM3, [72]), one of the key constructs in human–computer interaction research to measure technology acceptance of users: (1) "I could imagine to use the robot in the future." and (2) "I would like to inform myself about products that are similar to this robot." These items were rated on a five-point Likert scale from 1 (not at all) to 5 (very much). We can also report good reliability with a Cronbach's  $\alpha$  of .85.

### 3.3.7 Propensity to Stereotype

To assess propensity to stereotype professions, respondents were asked: "The next question is about people in different jobs. In general, how typical do you think are the following professions for persons of different genders?" and to measure respondents' propensity to stereotype traits, they were asked: "How typical are the following traits for men and/or women?". The full list of items was added below.

Propensity to stereotype was then measured by deducting the number of stereotype non-conforming choices (for professions and traits) respondents made from the number

of stereotype-conforming choices they made. A stereotype-conform choice would be e.g. that the trait “aggressive” is rated more male than female or that the profession of a carer is typical for women. As another example, if respondents rated the profession of a carer as either “more typical for men” or “typical for all genders”, this would be counted as a stereotype non-conforming choice. Instead, if they rated this profession as “more typical for women” this would be counted as a stereotype-conforming choice. The full list of professions and traits can be found in the “Appendix B”.

The selection of stereotypical items for the current study was based on previous research on stereotypical traits (e.g. [73]) and jobs (e.g. [74]). The items were assigned as follows:

1. Female-stereotypical traits: nurturing, kind, sensitive, helpful, respectful
2. Female-stereotypical jobs: carer, secretary, teacher
3. Male-stereotypical traits: aggressive, independent, controlling, assertive, confident
4. Male stereotypical jobs: truck driver, lawyer, doctor

### 3.3.8 Caregiver Choice

Respondents were asked the following question: “*If I could choose, I would choose.... as my primary caregiver.*” with the following answer options: “a woman”; “a man”; “any gender”.

For analysis of this measure, answers to this question were coded as “a woman” (stereotype-conform) if they chose “a woman” versus “other” (stereotype non-conform) if respondents chose the “a man” or “any gender” options.

## 3.4 Respondents

A G\*Power (version 3.1.9.6) analysis ( $d=0.5$ , power = 0.95) was run to define the sample size. According to the analysis, the sample size needed for the study is 210 respondents. Based on our previous experiences with Prolific participant disqualification rates, we used participant screening tools to recruit 240 respondents, representing approximately even numbers of men and women (specifically, given our focus on binary robot gender (mis)matching) which may include trans men and trans women but inherently results in the exclusion of non-binary persons as potential respondents. This is something we hope to remedy in future works more specifically concerned with non-binary perspectives regarding gendered robot design. We further used Prolific screening tools to specifically recruit respondents who identify as having a long-term health condition and/or disability. We direct readers to Prolific’s website to see full detail on the language pertaining these screening tools and how they

**Table 2** Respondents across experimental conditions

	Male robot	Female robot
Women	63	52
Men	47	60

are administered to platform workers.<sup>7</sup> Our initial participant pool were majority UK (148) or EU (54) nationals otherwise from the US (17), Australia (6), South Africa (4), Zimbabwe (2), Canada (1) and New Zealand (1) with two respondents of unknown nationality. Respondents who successfully completed one or more of the two attention checks were compensated 2.00 via the Prolific platform (one participant was excluded on the basis of failing this). 17 respondents had to be removed from the sample as the data file indicated that they had not watched the complete video vignette. The final sample consisted of 222 respondents (107 males, 115 females,  $M(\text{age})=44.52$  years,  $SD(\text{age})=15.10$ ,  $\text{Range}(\text{age})=18\text{--}87$ ). The final distribution of respondents across conditions (whose data was utilised in data analysis) is given in Table 2. Respondents viewed an information sheet before starting their participation, they were informed of the possibility to withdraw from the study at any point, and debriefing information was provided once the study was completed. We also followed the institutions’ data collection and storage policies in line e.g., with anonymity, data retrieval, data deletion, and GDPR compliance.

## 4 Results

We generate results from our survey data as follows. We first present correlation analyses to examine whether there is any relationship between the variables of interest. If there is no significant correlation, we do not proceed with further analyses. These correlation analyses also allow for identification of unexpected findings which might be worthy of further exploratory analyses.

Based on results of the correlation analyses (presented in Sect. 4.1) we undertake additional analyses of our experimental measures according to Table 3. Table 4 includes an overview of means and standard deviations for all variables. See Table 6 as a summary of the key results for each research question.

For analysis of our data, we used the statistics software SPSS (version 27).

### 4.1 Zero-Order Correlations Between the Dependent Variables

Initial Spearman’s rank-order correlations indicate that all BPN measures, goodwill, human-likeness and ascription of

<sup>7</sup> <https://www.prolific.co/>.



**Table 3** Analyses of experimental measures

Measure	Analysis
BPN: competence	Two-way ANOVA
ITU	Two-way ANOVA
Caregiver choice	Chi-square test
Propensity to stereotype	Two-way ANOVA

masculinity to the robot significantly and positively correlated with ITU. This suggests that these variables have a positive impact on respondents' ITU the robots presented in the video vignettes.

Robot-likeness had a significant negative correlation with ITU, indicating that respondents' ITU is negatively affected by a more robot-like perception of the robot. In addition to this, robot-likeness also had a significant negative correlation with the BPN satisfaction of relatedness to the robot and goodwill. These findings demonstrate that perceived robot-likeness of the robot has negative effects on respondents' perception of the robot as e.g. caring, sensitive, concerned and also on the satisfaction of respondents' need for relatedness with the robot. Instead, higher ratings of human-likeness positively correlated with all basic needs and goodwill.

A significant, negative correlation was identified between caregiver choice and propensity to stereotype, suggesting that respondents who preferred a male caregiver, or indicated no gender preference, had overall lower propensity to stereotype.

## 4.2 Robot-Participant Gender Differences

As can be seen in the correlations Table 5, our results reveal a significant correlation between the respondents' competence satisfaction and their ascription of masculinity to the robot ( $rs(220) = .191$ ,  $p = .004$ ). No such correlation was found for relatedness or autonomy satisfaction. To further explore this result we split the correlations by participant gender (see "Appendix C"). Doing so, the results revealed that this effect

only appeared for men ( $rs(105) = .337$ ,  $p < .001$ ) and not for women ( $rs(113) = .035$ ,  $p = .713$ ). These results indicate that men had higher levels of competence satisfaction if they perceived the robot in the video vignette as male.

To better understand this relationship, we conducted a two-way ANOVA to analyze the effect of participant gender and robot video on competence satisfaction. The results revealed that there is a significant interaction effect of participant gender and robot video on competence satisfaction  $F(1,218) = 3.93$ ,  $p = .049$ ,  $\eta^2 = .018$ . The main effect for gender was also statistically significant  $F(1,218) = 4.36$ ,  $p = .038$ ,  $\eta^2 = .020$ ; with men having significantly lower ratings of competence satisfaction ( $M = 3.40$ ,  $SD = 1.02$ ) in comparison to women ( $M = 3.70$ ,  $SD = 1.01$ ). There was no significant main effect for the robot video condition  $F(1,218) = 1.25$ ,  $p = .265$ ,  $\eta^2 = .006$ . Since each variable only has two groups we cannot report post hoc tests but looking at the means, we can see that women had higher competence satisfaction ratings ( $M = 3.76$ ,  $SD = .95$ ) in comparison to men ( $M = 3.21$ ,  $SD = .96$ ) after seeing the female-gendered robot video. For the male-gendered robot video competence satisfaction ratings between women ( $M = 3.65$ ,  $SD = 1.05$ ) and men ( $M = 3.63$ ,  $SD = 1.04$ ) were very close. Overall, these findings suggest that female robot gendering had a positive impact on competence satisfaction for women, but a negative impact on competence satisfaction for men.

In addition to this, the correlations split by gender reveal that for men, all BPN and ITU significantly and positively correlate with the ascription of masculinity to the robot (see "Appendix C"). This would suggest that men who took part in the current study have an overall preference for a gender-matched robot.

Further, we conducted a two-way ANOVA to assess the effect of participant gender and robot video on ITU. There was no significant interaction effect  $F(1,218) = 1.15$ ,  $p = .285$ ,  $\eta^2 = .005$  and no significant main effect for participant gender ( $F(1,218) = 2.20$ ,  $p = .140$ ,  $\eta^2 = .010$ ) or robot video ( $F(1,218) = .94$ ,  $p = .334$ ,  $\eta^2 = .004$ ). Even though the anal-

**Table 4** Means and standard deviations of all variables

Measure	Range	Mean	Standard deviation
ITU	1–5	3.60	1.18
BPN autonomy index	1–5	3.66	0.99
BPN competence index	1–5	3.55	1.02
BPN relatedness others index	1–5	2.93	0.97
BPN relatedness bot index	1–5	3.26	1.13
Goodwill index	1–5	3.23	1.09
Prop. stereotype index	–8 to 8	–1.66	3.50
Human-like index	1–5	2.73	0.97
Robot-like index	1–5	3.89	0.96
Robot femininity index	1–5	2.24	1.19
Robot masculinity index	1–5	2.76	1.34

**Table 5** Spearman correlations between variables

Measure	1	2	3	4	5	6	7	8	9	10
ITU index										
Autonomy index	.57**									
Competence index	.73**	.68**								
Relatedness others index	.59**	.51**	.66**							
Relatedness bot index	.70**	.57**	.70**	.65**						
Goodwill index	.55**	.47**	.62**	.53**	.68**					
Prop. stereotype index	−.08	−.14*	−.11	−.10	−.10	−.04				
Human-like index	.46**	.41**	.50**	.49**	.57**	.55**	−.02			
Robot-like index	−.14*	−.08	−.20**	−.21**	−.31**	−.24**	−.00	−.48**		
Robot femininity index	.04	.01	.02	.11	.07	.07	.08	.05	−.05	
Robot masculinity index	.16*	.16*	.19**	.14*	.17*	.15*	−.09	.17*	−.00	−.44**

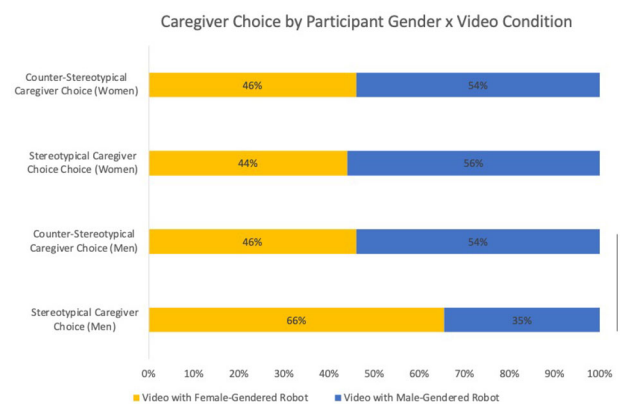
\*  $p < .05$ ; \*\*  $p < .01$  (two-tailed)

ysis was not significant, the descriptive statistics show that men's ITU ratings were considerably lower after watching the female-gendered video ( $M = 3.33$ ,  $SD = 1.23$ ) in comparison to women's ITU after seeing the female-gendered robot video ( $M = 3.73$ ,  $SD = 1.15$ ). This again supports the notion that men have a preference for the gender-matched robot, as indicated by the findings from the correlation table. The ITU ratings for women ( $M = 3.71$ ,  $SD = 1.17$ ) and men ( $M = 3.65$ ,  $SD = 1.14$ ) were more similar after watching the male-gendered robot video.

An additional exploratory two-way ANOVA analysed the effect of participant gender and robot video on respondents' perceived benefit of the robot. There was no significant interaction effect ( $F(1,218) = 1.64$ ,  $p = .682$ ,  $\eta^2 = .001$ ). However, a significant main effect for participant gender ( $F(1,218) = 11.13$ ,  $p < .001$ ,  $\eta^2 = .049$ ) suggests that overall, i.e. regardless of robot gendering, women perceived the robot as more useful, as women had significantly higher ratings for the perceived benefit from the robot ( $M = 3.60$ ,  $SD = 1.26$ ) in comparison to men ( $M = 3.01$ ,  $SD = 1.26$ ). There was no significant main effect for the robot video condition ( $F(1,218) = 1.64$ ,  $p = .202$ ,  $\eta^2 = .007$ ) but a closer look at the descriptive results shows that the male-gendered robot video yielded slightly higher ratings for men respondents ( $M = 3.17$ ,  $SD = 0.18$ ) and marginally higher for women respondents ( $M = 3.67$ ,  $SD = 0.16$ ) in comparison to the female-gendered robot video (men respondents:  $M = 2.88$ ,  $SD = 0.16$ ; women respondents:  $M = 3.52$ ,  $SD = 0.17$ ).

### 4.3 Choice of Caregiver Gender

The caregiver choice was coded as “female” (stereotype-conform) versus “other” (stereotype non-conform) which represented the questionnaire answer options “male” or “any gender”. We ran a Chi-Square Test to measure differences in respondents' caregiver choice after seeing the male versus



**Fig. 2** The figure represents respondents' care giver choice out of the two options (1) ‘woman’ (stereotypical) or (2) ‘other’ (counter-stereotypical). The bars identify the percentage of either men or women respondents who chose these options and the colouring demonstrates the two experimental robot conditions: female-gendered robot video (yellow) versus male-gendered robot video (blue). It can be seen that the majority of men respondents who said they would prefer a woman carer were in the female-gendered robot video condition. Note. \*indicates the significant differences between the caregiver choices for men respondents

female-gendered robot video. Our analysis reveals that men who saw the male-gendered robot video chose an “other” caregiver option significantly more often  $\chi^2 (1, N = 107) = 4.04$ ,  $p = .044$  (Fig. 2). The result for women was not significant (Fig. 2). These findings suggest, that men's (hypothetical) choice of a caregiver for real-life was influenced by the video vignette from the study.

### 4.4 Propensity to Stereotype

A two-way ANOVA was conducted to measure the effect of participant gender and robot video condition on respondents' propensity to stereotype (range from  $-8$  to  $8$ ,  $M = -1.66$ ,  $SD = 3.50$ ). The interaction effect  $F(1,218) = .04$ ,  $p = .849$ ,  $\eta^2 = .000$ , and main effect for participant gender  $F(1,218)$

$= 1.15$ ,  $p = .286$ ,  $\eta^2 = .005$ , were not significant. The main effect for the robot video condition just failed to reach significance  $F(1,218) = 3.59$ ,  $p = .060$ ,  $\eta^2 = .016$ . The descriptive statistics show that respondents' propensity to stereotype was overall lower after watching the male robot video ( $M = -1.24$ ,  $SD = 3.66$ ) compared to the female robot video ( $M = -2.08$ ,  $SD = 3.28$ ).

Since we only explicitly manipulated job role gender (mis)matching and not traits in the current study, we also specifically compared respondents rating of typically female versus typically male jobs. A two-way ANOVA, analysing the effect of participant gender and robot video condition on respondents' propensity to stereotype 'typically male' jobs, showed no significant interaction effect for participant gender and robot video  $F(1,218) = .72$ ,  $p = .398$ ,  $\eta^2 = .003$ . There was also no significant main effect for participant gender  $F(1,218) = .33$ ,  $p = .567$ ,  $\eta^2 = .002$ . However, the results reveal a significant main effect for robot video  $F(1,218) = 6.18$ ,  $p = .014$ ,  $\eta^2 = .028$ . The descriptive statistics show that women's propensity to stereotype 'typically male' jobs was considerably lower after watching the male-gendered robot video ( $M = 0.94$ ,  $SD = 0.40$ ) in comparison to the female-gendered robot video ( $M = 1.19$ ,  $SD = 0.56$ ). The results for men show a similar tendency (male robot:  $M = 0.96$ ,  $SD = 0.55$ ; female robot:  $M = 1.08$ ,  $SD = 0.72$ ).

In addition to this, autonomy satisfaction correlated negatively with propensity to stereotype for men only. This suggests that male respondents who have lower propensity to stereotype had their autonomy need more fulfilled with regard to the robot as a carer in comparison to men respondents with higher propensity to stereotype.

Table 6 includes an overview of the results we have conducted for each of the research questions.

## 5 Discussion

### 5.1 Gender Differences: Which Robot for Whom?

Our goal with this work was to investigate what participant-robot gender (mis)matching preferences might emerge for a socially assistive robot. Previous literature reported mixed results on such participant-robot gender matching, as well as in how gendered users do or do not seem to apply human gender stereotypes when evaluating social robots. Thus, we designed an experiment in which men and women were shown either a male or female-gendered Pepper robot in a social care context, before being asked to complete a range of measures, broadly pertaining to that robot's acceptability, and potential to support their Basic Psychological Needs (BPN) should they use it.

For our participant pool (majority UK and EU nationals) female robot gendering had a negative impact on men's com-

petence satisfaction, when they imagined working with the robot. Further, specifically for men, ascription of masculinity to the robot positively correlated not only with competence satisfaction but also with relatedness satisfaction scores (both to other people and to the robot), autonomy satisfaction scores and ITU scores. No equivalent correlations were seen for women who saw the female robot, suggesting a stronger preference for robot-participant gender matching in men than in women. For men at least, gender matching preferences seemingly over-shadowed any positive associations arising from gender-task typicality.

We identify two possible explanations for this particular impact of gender matching on men. Firstly is the possibility that this reflects a general bias of men perceiving men to be more competent than women [75]. Whilst we did not directly assess perceived robot competence in this study, results from a previous study revealed that respondents' own competence satisfaction when imagining using a chatbot links to perceived competence of that chatbot [20]. Our findings that female robot gendering had a negative impact on men's competence satisfaction could therefore indicate men perceived the male robot as being more competent.

This might be particularly triggered by the role of the robot in our scenario, as it was taking the lead on guiding the exercises and (somewhat, even though we utilised 'low agency' cues according to [20]) telling the patient what to do, perhaps thus being perceived as taking quite an authoritative role. According to previous studies [73, 76–78] men struggle with women taking the lead. Therefore, men in the male-gendered robot condition might have been more satisfied with by the idea of "working" with the robot, whereas those in the female-gendered robot condition might have disliked the idea of the robot "telling them what to do".

Alternatively, it could be that men in the male-gendered robot condition were particularly positively impressed with the male-gendered robot because it was norm-breaking. Previous studies have shown that men are more likely to stereotype than women (e.g. [79, 80]) which could be why the male-gendered, norm-breaking robot video had a stronger positive impact on men, who began the study with stronger stereotypical biases in comparison to women. This is supported by the fact that men who saw the male-gendered robot were more likely then to indicate preferences for a male or any gender carer, when compared to those who saw the female-gendered robot, although it did not impact their propensity to stereotype more broadly, something we discuss more in the following subsection.

Other gender differences between our men and women respondents might also yield clues as to their needs, and the ways in which (gendered) SARs might be leveraged (or not) in tackling these. For example, regardless of robot gendering, women perceived the robot to be more useful in managing their health condition than men. Further, men's propensity

**Table 6** Results for research questions

Research question	Result summary
RQ1 A: Robot gendering and BPN	A two-way ANOVA, to analyze the effect of participant gender and robot video on competence satisfaction, revealed a significant interaction effect ( $F(1,218)=3.93$ , $p=.049$ , $\eta^2=.018$ ). While the main effect for participant gender was also significant ( $F(1,218)=4.36$ , $p=.038$ , $\eta^2=.020$ ) with men having significantly lower ratings of competence satisfaction ( $M=3.40$ , $SD=1.02$ ) in comparison to women ( $M=3.70$ , $SD=1.01$ ), there was no significant main effect for the robot video. This demonstrates that women ( $M=3.76$ , $SD=.95$ ) had significantly higher competence satisfaction ratings in comparison to men ( $M=3.21$ , $SD=.96$ ) mainly after seeing the female-gendered robot video
RQ1 B: Robot gendering and ITU	For men specifically, ITU correlated with ascription of masculinity to the robot ( $r_s(105)=.316$ , $p<.001$ ). Between robot video conditions, men gave the female-gendered robot video lower ITU ratings ( $M=3.33$ , $SD=1.23$ ) than women ( $M=3.73$ , $SD=1.15$ ) but both interaction and main effect were not significant
RQ2 A: Human caregiver gender preferences	Men who saw the male-gendered robot video more often indicated a preference for the 'other' caregiver option ('other' being the stereotype non-conform choice: either a male or any gender caregiver) than those who saw the female-gendered robot video $\chi^2(1, N=107)=4.04$ , $p=.044$ . The result for women was not significant
RQ2 B: Application of stereotypes	Results of a two-way ANOVA on the effect of participant gender and robot video on respondents' propensity to stereotype (range from $-8$ to $8$ , $M=-1.66$ , $SD=3.50$ ), did not result in a significant interaction effect and also no significant main effect for participant gender. The main effect for robot video just failed to reach significance $F(1,218)=3.59$ , $p=.060$ , $\eta^2=.016$ . However, the descriptive statistics show that respondents' propensity to stereotype was lower after watching the male robot video ( $M=-1.24$ , $SD=3.66$ ) compared to the female robot video ( $M=-2.08$ , $SD=3.28$ ). Moreover, a two-way ANOVA analysing the propensity to stereotype 'typically male' jobs revealed a significant main effect for the robot video condition ( $F(1, 218)=6.18$ , $p=.014$ , $\eta^2=.028$ ): it was considerably lower in women who saw the male-gendered robot video ( $M=0.94$ , $SD=0.40$ ) than those who saw the female-gendered robot video ( $M=1.19$ , $SD=.056$ ). The results for men show a similar tendency (male-gendered robot video: $M=0.96$ , $SD=0.55$ ; female-gendered robot video: $M=1.08$ , $SD=0.72$ ). Neither the interaction effect, nor the main effect for participant gender were significant
RQ3: Human-likeness and ITU	Human-likeness significantly correlated with ITU ( $r_s(220)=.456$ , $p<.001$ )

to stereotype negatively correlated with their autonomy satisfaction, perhaps painting a (stereotypical) picture of the socially conservative man as being most resistant to/feeling most threatened by accepting assistive help. Such nuances must be considered by robot designers if we are to create/deploy SARs in ways that are not just acceptable but also generate positive real-world impact on such users lives.

Again it must be noted that perceptions of gender, gender roles and gender stereotyping vary significantly across cultures. For example, ideas about men being more competent than women, (wo)men struggling with women taking the lead and men engaging in caregiving are grounded in western and/or patriarchal social structures of domination. Our discussions and speculations regarding social phenomena that may underlie our results is therefore limited to such settings, reflecting our own cultural background and experience.

## 5.2 Challenging Gender Stereotypes with Robots in Care

Throughout this work, we posited the male-gendered care robot as being norm-breaking, based on typical gender-role associations linking women with in-home caregiving and nursing roles [81, 82]. Our idea was that subjecting respondents to this gender norm-breaking robot might then reduce

their propensity to gender-stereotype humans when thinking about caregivers or other gender-stereotypical role and trait associations. We found partial support for this. Women who watched the male-gendered care robot showed lower propensity to stereotype men's professions compared to those who saw the female-gendered care robot. Further, men who saw the male-gendered robot were then more likely to report preference for a male or any gender caregiver, compared to those who saw the female-gendered robot.

Perhaps more interestingly however, is that our results clearly indicate a suggestion that men would seemingly be much more satisfied with a male-gendered care robot than a female-gendered one, even when that robot has been designed to be low-agency (arguably passive and submissive) reflecting current (problematic) design trends around female agents that supposedly represent user preferences [21]. It could be that, on an individual level, this preference reflects (and maybe reinforces) norms regarding gender and competence, i.e. that men are more competent than women. However, this contrasts from Moradbakhti et al.'s previous work, in which men did show a particular preference for low-agency, female digital assistants (over high-agency female and high/low-agency male equivalents) in a banking context [20]. Either way, at a broader societal level this suggests we can deploy male-gendered care robots to male users in a way



which simultaneously maximises positive impact on those users, avoids harmful propagation of gender norms regarding women being subservient and possibly reduces stereotypical thinking with regards to who can do care work/what tasks particularly suit men and women. In short, we think this provides good evidence, once again, that we can challenge stereotypes whilst also increasing ‘effectiveness’ (acceptability, impact) of human–robot interaction, rather than perceiving there to be a disconnect between delivering ‘what is ethical’ versus ‘what people want’.

### 5.3 The Risks and Opportunities of Anthropomorphism

Increasing robot-likeness could be one way to avoid ethical issues with regard to robot gendering, and previous work in robot ethics has called for a minimization of anthropomorphism [83]. However, the current results clearly indicate that the perception of human-likeness had a positive impact on respondents’ ITU, satisfaction of all BPN and goodwill perception, while robot-likeness negatively correlated with ITU, goodwill perception and relatedness satisfaction to the robot. This suggests that an increase in robot-likeness leads to lower acceptability and more negative perception of a SAR; replicating similar findings in previous work considering the ethical risks versus impact benefits of anthropomorphism in SARs [62]. Ideally, perception of and interaction with social robots would be guided by sociomorphing (perception of actual non-human social capacities in direct interaction with social robots, see [84]) instead of anthropomorphism. Future studies should therefore continue to explore ways in which we design somewhat human-like but ethically considerate SAR, e.g. to allow for sociomorphing instead of anthropomorphism, and avoid the uncanny valley [85, 86]. Creating norm-breaking SAR, as presented in the current study via the video vignettes, is one way to challenge existing stereotypes and leave a positive impact on respondents’ perception of stereotypical gender-roles.

### 5.4 On BPN and ITU for Evaluating Socially Assistive Robots

In line with previous findings [13, 44], our results confirmed that BPN had a strong, positive correlation with respondents’ ITU the robot. Especially in the physical rehabilitation sector, where motivation and task engagement play a crucial role for patients’ recovery, BPN offer an important measure for technology acceptance and usage. This is particularly relevant, as patients often “struggle” to stay motivated to continue their physical rehabilitation [16] and recent studies and programs have been focused on ways to increase patients’ motivation [17]. In addition to this, a previous study on BPN and physical activity has shown [87] that BPN satisfaction

can positively affect patients’ adherence to long-term physical activity. Moreover, according to Deci and Ryan [11], BPN satisfaction does not only lead to greater task motivation, it also leads to overall well-being. Future studies could focus on long-term effects of SAR rehabilitation on patients’ motivation and wellbeing.

Our study is one of the first to draw this line between BPN satisfaction, the foundation of a well-established theory on human motivation, and its influence for the design of SAR. Assessing users’ need satisfaction and, more importantly, understanding which design factors of a SAR can increase their BPN satisfaction, should be a part of future design processes. Moreover, as our results confirmed, the BPN-TU scale we applied to measure BPN satisfaction provides a valid construct to measure users’ need satisfaction, when interacting with SAR. Therefore, we recommend future HRI studies to include BPN satisfaction as a measure to ensure not only higher acceptance of SARs but also respondents’ motivation to engage with SARs.

### 5.5 Limitations

One limitation of the current study is that the male and the female voice might differ in intonation, pitch etc., resulting in more differences between the voices than just gender. The limited number of voices available (combined with the complex interplay of factors that contribute to voice gendering) makes it difficult to identify two equivalent-apart-from-gender voices that are similar in all other ways. Future studies might compare multiple male/female voices to increase confidence that these effects are solely gender related. In addition, given Pepper’s somewhat gender ambiguous base embodiment, some respondents could have experienced a mismatch between the gender they initially ascribed to Pepper before our manipulations became clear (particularly on hearing the gendered voice manipulations). Such dissonance would likely influence their mental model of the robot’s gender/identity construction, and, relatedly, their responses to our experimental measures. Given that the majority of our respondents are UK/EU nationals, overlapping (somewhat) culturally with Perugia et al.’s (majority Italian) participant pool who generally perceived Pepper as being gender neutral [51], and our own respondents relatively strong gender ascriptions to Pepper, we assume any such dissonance to be minimal and/or evenly distributed across our experimental condition. However, this potential for dissonance, and its implications for HRI represents an interesting avenue for further work well motivated by the findings of Seaborn and Frank [56]. Given that gender is socially constructed and diverse across cultures, with there being no universal perception of gendered characteristics, our results should not be considered universally applicable. Such differences might also exist within our population pool

and exasperate e.g. the above mentioned dissonance issue. Further exploration of (individual respondents') strength of gender ascription compared to other pertinent experimental measures could offer one way to start examining this idea.

On user gender, our video clips featured only woman as the robot's users. This may have resulted in women respondents being more able to imagine themselves in these positions when completing the experimental measures. However, we expect interaction effects between robot, primary user, informal and former carer gender [53], thus trying to match participant-actor gender for all of these roles would significantly increase our number of conditions and respondents required, in a way that was infeasible for this study. The choice of using all-women instead of men robot users was based on the fact that women are more often in the caregiving role than men [24], and since we wanted to introduce and compare effects of a stereotypical versus non-stereotypical care-taker—as represented by the social robot—we chose the stereotypical option for the human caregiver in the video clip to avoid confounding effects. We prefer to recognise instead the limitation that men respondents might have found it slightly harder to imagine themselves in the actors' place than women respondents, but we do not feel this is a strong barrier to their participation and inclusion in the study.

Another key limitation is the video-based, hypothetical nature of our study. The COVID situation prevented us from safely and ethically conducting an in-person version of this experiment at the time this work was carried out. Moreover, conducting an "in-home" study would add an additional layer of risk for patients' safety. Conducting an online experiment with video clips that demonstrate this in-home interaction offered respondents the possibility to imagine themselves in the scenario. We believe that this online study is a good alternative to gain first results on this topic and we highly encourage follow-up in-person studies. Since the current online study was able to demonstrate relevant findings, it might be expected that an in-person interaction and particularly longitudinal, multi-interaction studies yield stronger effects with regard to respondents' stereotypical perception of male/female jobs/traits. Studies undertaken 'in the wild', involving real-world deployment, would also yield more accurate intention to use and need satisfaction results.

Notably, the current study only assessed BPN satisfaction of primary robot users (those expected to receive care from the robot). Both formal and informal caregivers would need to interact with the robot in a real-life setting, with the potential for this to impact on their BPN. Future studies on design factors of SAR and BPN satisfaction should therefore consider and investigate BPN satisfaction of all three user groups.

Finally, we want to highlight again the limited binary conceptualization of gender in our work, which does not reflect the true, diverse human gender spectrum and necessi-

tates the exclusion of non-binary respondents. Future work might consider, more broadly and not broken down by participant gender, the extent to which (counter)stereotypical interventions impact people depending on the strength of their preconceptions about gender norms (which could be women, men, or people identifying as non-binary). Designing for non-binary and gender-fluid robots might also offer opportunities for getting away from traditional stereotype propagation and/or work towards further normalizing of these gender identities, work on which represents a great opportunity to better engage with the non-binary participant population who have thus far often been excluded from gender-related research in HRI.

## 6 Conclusion

This article presented a video-based online experiment examining the impact of robot gendering on perceptions of an in-home, socially assistive robot designed to better engage patients in rehabilitative/condition managing exercises. We specifically examined respondents' perceptions regarding such a robot's impact on their BPN - a measure we identified as highly pertinent for socially assistive robots given links between BPN, wellbeing and task motivation.

Within the context of a majority UK/EU national participant pool, our results identified significant differences in the way robot gendering does or does not impact men and women. User-robot gender matching seemed to be more important for men than for women; over-shadowing any potential impact of robot-task gender typicality. Based on previous results, we suggested this could be because men generally perceive men to be more competent than women, and/or may have their own competence threatened when feeling under the authority of a female figure. Regardless, we have posited that this represents another reason to challenge the status quo of default female gendering in assistive systems. Our results indicated men might prefer male-gendered care robots, and such robots might challenge broader social norms regarding who does care work, thus simultaneously delivering better experience at the individual user level whilst avoiding the propagation of gender norms at the societal level. Broadly, our results also add to previous evidence indicating a positive influence of anthropomorphism on SAR effectiveness and acceptability, indicating some potential difficulty in 'avoiding the gendering issue' by moving towards more robot-like designs more broadly. Future studies should therefore continue to explore ways in which we design anthropomorphic but ethically considerate SARs.

It is important to recognise that perceptions and constructions of gender, and expectations, stereotypes and biases surrounding gender can vary significantly between and within cultures. Future work on robot gendering, when looking to

manipulate gender perceptions of a fixed robot platform, should carefully consider such the potential that this generating might induce dissonance within their respondents.

**Acknowledgements** This work was partly supported by the Wallenberg AI, Autonomous Systems and Software Program-Humanities and Society (WASP-HS) funded by the Marianne and Marcus Wallenberg Foundation and the Marcus and Amalia Wallenberg Foundation.

**Funding** Open access funding provided by Johannes Kepler University Linz.

**Data Availability** The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

**Conflict of interest** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adap-

tation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix A Full Basic Psychological Need Scale

See Fig. 3.

**Fig. 3** Full basic psychological need fulfillment scale

Imagine working with the robot every day. How would you feel?

	Not at all	Not really	Neutral	Somewhat	Very Much
When I use the robot, I feel like I can perform actions in the way I want to.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When I use the robot, I can act independently.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When I use the robot, I feel like I am in control.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Imagine working with the robot every day. How would you feel?

	Not at all	Not really	Neutral	Somewhat	Very Much
When I use the robot, I feel empowered in my own abilities.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When I use the robot, I feel confident that I can reach my goals.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When I use the robot, I feel competent.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Imagine working with the robot every day. How would you feel?

	Not at all	Not really	Neutral	Somewhat	Very Much
When I use the robot, I feel like my social circle reacts positively to my use of the robot.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When I use the robot, I feel like I look good in front of my social circle.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When I use the robot, I feel less alone.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Imagine working with the robot every day. How would you feel?

	Not at all	Not really	Neutral	Somewhat	Very Much
When I use the robot, I feel like the interaction goes both ways.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have a friendly feeling towards the robot.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I can imagine building a bond with the robot.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Appendix B Propensity to Stereotype Questions

See Fig. 4.

**Fig. 4** Propensity to stereotype questions

The next question is about people in different jobs. In general, how typical do you think are the following professions for persons of different genders?

	More typical for women	More typical for men	Typical for all genders
carer (childcare, elderly care etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
secretary	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
truck driver	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
lawyer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
teacher	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
doctor	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

How typical are the following traits for men and/or women?

	More typical for women	More typical for men	Typical for all genders
nurturing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
kind	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
sensitive	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
helpful	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
respectful	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
aggressive	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
independent	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
controlling	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
assertive	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
confident	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



## Appendix C Zero-Order Spearman Correlation Tables Split by Gender

See Figs. 5 and 6.

See Tables 7 and 8.

**Table 7** Spearman correlations between variables for men respondents

Measure	M	SD	1	2	3	4	5	6	7	8	9	10
ITU index	3.47	1.20										
Autonomy index	3.65	.94	.52**									
Competence index	3.40	1.02	.69**	.64**								
Relatedness others index	2.91	.93	.60**	.42**	.65**							
Relatedness bot index	3.13	1.17	.66**	.53**	.68**	.65**						
Goodwill index	3.11	1.13	.53**	.36**	.55**	.46**	.67**					
Prop. stereotype index	−1.87	3.65	−.09	−.21*	−.11	−.07	−.11	.04				
Human-like index	2.64	1.01	.49**	.37**	.54**	.54**	.60**	.53**	−.00			
Robot-like index	3.96	1.07	−.07	−.06	−.14	−.17	−.37**	−.23*	−.00	−.41**		
Robot femininity index	2.24	1.17	.03	−.08	−.08	.13	.07	.10	.05	.15	−.12	
Robot masculinity index	2.60	1.27	.32**	.20*	.34**	.24*	.25**	.12	−.02	.16	.10	−.34**

\* $p < .05$ ; \*\* $p < .01$  (two-tailed)

**Table 8** Spearman correlations between variables for women respondents

Measure	M	SD	1	2	3	4	5	6	7	8	9	10
ITU index	3.72	1.16										
Autonomy index	3.66	1.03	.59**									
Competence index	3.70	1.01	.76**	.72**								
Relatedness others index	2.95	1.01	.58**	.58**	.69**							
Relatedness bot index	3.39	1.08	.72**	.61**	.73**	.66**						
Goodwill index	2.65	1.04	.54**	.55**	.68**	.61**	.69**					
Prop. stereotype index	−1.46	3.35	−.09	−.08	−.11	−.12	−.10	−.14				
Human-like index	2.81	.93	.41**	.44**	.45**	.45**	.54**	.57**	−.05			
Robot-like index	3.83	.85	−.19*	−.09	−.23*	−.25**	−.24**	−.24*	.02	−.56**		
Robot femininity index	2.23	1.20	.06	.08	.11	.10	.09	.05	.11	−.05	.00	
Robot masculinity index	2.91	1.38	−.01	.11	.04	.04*	.07	.16	.15	.15	−.06	−.54**

\* $p < .05$ ; \*\* $p < .01$  (two-tailed)

**Fig. 5** Video script 1

Scene 1 (informal caregiver with robot)	
<b>Robot</b>	Hi Sarah. Nice to see you. How can I help you today?
<b>Informal Caregiver (Sarah):</b>	I need to go food shopping so can you do Lisa's exercises with her? She seemed a bit demotivated but maybe you can remind her that the therapist is coming tomorrow.
<b>Robot:</b>	Yes, I got the exercises and will ask Lisa if she wants to do them when she is ready. Would you like me to send out a message to you once we have completed the exercises for today?
<b>Informal Caregiver (Sarah):</b>	Great. Yes, please let me know when you are done.
<b>Robot:</b>	Sure, is there anything else you would like me to do?
<b>Informal Caregiver (Sarah):</b>	Yes, please make sure to enter the data in the training plan so I can track her progress.
<b>Robot:</b>	No problem.
<b>Informal Caregiver (Sarah):</b>	Great, see you later.
<b>Robot:</b>	Thank you, see you later Sarah!

**Fig. 6** Video script 2

Scene 2 (patient with robot):	
<b>Robot:</b>	Hi Lisa, how are you feeling today?
<b>Patient (Lisa):</b>	I am ok, just a bit stiff and tired, I did not sleep very well.
<b>Robot:</b>	I am sorry to hear this. Would you like to start with today's exercises? The therapist is visiting tomorrow so we should try to make sure we get through everything she has sent for us.
<b>Patient (Lisa):</b>	Ok, we can start, I just don't know how much I can do today.
<b>Robot:</b>	Do not worry, I will keep you motivated! Which exercise should we start with? We have to do some neck tilts and neck stretches.
<b>Patient (Lisa):</b>	Yes, let's do the neck tilts.
<b>Robot:</b>	Ok, here is a picture showing what to do. Tilt your head down to rest your chin on your chest. Gently tense your neck muscles and hold for 5 seconds.
Scene Break	
<b>Robot:</b>	Great job! You managed all five repetitions. Would you like to take a quick break?
<b>Patient (Lisa):</b>	Ok, sounds good.
Scene Break	
<b>Robot:</b>	Ok, shall we move on?
<b>Patient (Lisa):</b>	Actually, I don't really like doing any more. Maybe we could leave it there?
<b>Robot:</b>	It is ok to be tired, as long as the exercises aren't causing any pain. Do you think you could give the neck stretches a go? You did so well with the neck tilts.
<b>Patient (Lisa):</b>	Ok, I'll give it a try.
<b>Robot:</b>	Great! Here is the picture. Keeping the rest of your body straight ...
Scene Break	
<b>Robot:</b>	Well done, you made it. That's all of the exercises for today. You did really well! How are you feeling?
<b>Patient (Lisa):</b>	Actually, I feel a bit better now. I think the stretching helped with my stiffness.
<b>Robot:</b>	I am glad to hear that! I would like to inform Sarah and the therapist that you have completed the exercises and share your progress with them. Is this ok for you?
<b>Patient (Lisa):</b>	Ok, that's fine. Thank you. I will go and grab something to eat now.
<b>Robot:</b>	You are welcome. Goodbye for now!

## References

1. Feil-Seifer D, Mataric MJ (2005) Defining socially assistive robotics. In: 9th International conference on rehabilitation robotics, 2005. ICORR 2005. IEEE, pp 465–468
2. Lara JS, Casas J, Aguirre A, Munera M, Rincon-Roncancio M, Irfan B, Senft E, Belpaeme T, Cifuentes CA (2017) Human–robot sensor interface for cardiac rehabilitation. In: 2017 International conference on rehabilitation robotics (ICORR), pp 1013–1018
3. Fasola J, Mataric MJ (2010) Robot exercise instructor: a socially assistive robot system to monitor and encourage physical exercise for the elderly. In: RO-MAN, 2010 IEEE. IEEE, pp 416–421
4. Fitter NT, Mohan M, Kuchenbecker KJ, Johnson MJ (2020) Exercising with Baxter: preliminary support for assistive social-physical human–robot interaction. J Neuroeng Rehabil 17(1):19. <https://doi.org/10.1186/s12984-020-0642-5>
5. O'Shea SD, Taylor NF, Paratz JD (2007) but watch out for the weather: factors affecting adherence to progressive resistance exercise for persons with COPD. J Cardiopulm Rehabil Prev 27(3):166–174 (quiz 175–176)

6. Forkan R, Pumper B, Smyth N, Wirkkala H, Ciol MA, Shumway-Cook A (2006) Exercise adherence following physical therapy intervention in older adults with impaired balance. *Phys Ther* 86(3):401–410
7. Visser M, Brychta RJ, Chen KY, Koster A (2014) Self-reported adherence to the physical activity recommendation and determinants of misperception in older adults. *J Aging Phys Act* 22(2):226–234. <https://doi.org/10.1123/japa.2012-0219>
8. Sussenbach L, Riether N, Schneider S, Berger I, Kummert F, Lutkebohle I, Pitsch K (2014) A robot as fitness companion: towards an interactive action-based motivation model. In: The 23rd IEEE international symposium on robot and human interactive communication, pp 286–293
9. Winkle K, Lemaignan S, Caleb-Solly P, Leonards U, Turton A, Bremner P (2019) Effective persuasion strategies for socially assistive robots. In: 2019 14th ACM/IEEE international conference on human–robot interaction (HRI), pp 277–285
10. Rea DJ, Schneider S, Kanda T (2000) "Is this all you can do? harder!": The effects of (im)polite robot encouragement on exercise effort. In: Proceedings of the 2021 ACM/IEEE international conference on human–robot interaction, ser. HRI '21. Association for computing machinery, pp 225–233. [Online]. <http://doi.org/10.1145/3434073.3444660>
11. Deci EL, Ryan RM (2000) The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. *Psychol Inq* 11(4):227–268
12. Ryan RM, Deci EL (2000) The darker and brighter sides of human existence: basic psychological needs as a unifying concept. *Psychol Inq* 11(4):319–338
13. De Vreede T, Raghavan M, De Vreede G-J (2021) Design foundations for ai assisted decision making: a self determination theory approach
14. Jiménez-Barreto J, Rubio N, Molinillo S (2021) Find a flight for me, Oscar! Motivational customer experiences with chatbots. *Int J Contemp Hosp Manag* 33:3860–3882
15. Balaguer I, González L, Fabra P, Castillo I, Mercé J, Duda JL (2012) Coaches' interpersonal style, basic psychological needs and the well-and ill-being of young soccer players: a longitudinal analysis. *J Sports Sci* 30(15):1619–1629
16. Randriambelonoro M, Perrin C, Blocquet A, Kozak D, Fernandez JT, Marfaing T, Bolomey E, Benhissen Z, Frangos E, Geissbuhler A et al (2020) Hospital-to-home transition for older patients: using serious games to improve the motivation for rehabilitation-a qualitative study. *J Popul Ageing* 13(2):187–205
17. Wittmann F, Held JP, Lamercy O, Starkey ML, Curt A, Höver R, Gassert R, Luft AR, Gonzenbach RR (2016) Self-directed arm therapy at home after stroke with a sensor-based virtual reality training system. *J Neuroeng Rehabil* 13(1):1–10
18. Winkle K, Turton A, Caleb-Solly P, Bremner P (2018) Patient engagement with rehabilitative therapy programmes: therapist strategies and affecting influences. *Proceedings of SRR. Clin Rehabil* 32(10):1414–1415. <https://doi.org/10.1177/0269215518784346>
19. Deci EL, Ryan RM (1985) Self-determination theory, vol 25, p 2019
20. Moradbakhti L, Schreibelmayer S, Mara M (2022) Do men have no need for feminist AI? Agentic and gendered voice assistants in the light of basic psychological needs and technology acceptance. *Front Psychol* 13:2404
21. West M, Kraut R, El Chew H (2019) I'd blush if i could: closing gender divides in digital skills through education. Available: <https://unesdoc.unesco.org/ark:/48223/pf0000367416.page=1>
22. Winkle K, Melsiön GI, McMillan D, Leite I (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, ser. HRI '21. ACM
23. Winkle K, Caleb-Solly P, Turton A, Bremner P (2018) Social robots for engagement in rehabilitative therapies: design implications from a study with therapists. In: Proceedings of the 2018 ACM/IEEE international conference on human–robot interaction, ser. HRI '18. ACM, pp 289–297. Available: <http://doi.acm.org/10.1145/3171221.3171273>
24. Bryant D, Borenstein J, Howard A (2020) Why should we gender? The effect of robot gendering and occupational stereotypes on human trust and perceived competency. In: Proceedings of the 2020 ACM/IEEE international conference on human–robot interaction, ser. HRI '20. Association for Computing Machinery, pp 13–21. Available <https://doi.org/10.1145/3319502.3374778>
25. Tay B, Jung Y, Park T (2014) When stereotypes meet robots: the double-edge sword of robot gender and personality in human–robot interaction. *Comput Hum Behav* 38:75–84
26. Eyssel F, Hegel F (2020) s)he's got the look: gender stereotyping of robots. *Handb Gestalt Digit Vernetzter Arbeitswelten* 42(9):2213–2230. <https://doi.org/10.1111/j.1559-1816.2012.00937.x>
27. Ghazali AS, Ham J, Barakova E, Markopoulos P (2019) "Assessing the effect of persuasive robots interactive social cues on users' psychological reactance, liking, trusting beliefs and compliance. *Adv Robot* 33(7–8):325–337. <https://doi.org/10.1080/01691864.2019.1589570>
28. Nomura T (2017) Robots and gender. *Gender Genome* 1(1):18–25. <https://doi.org/10.1089/gg.2016.29002.nom>
29. Crowell CR, Villanoy M, Scheutzz M, Schermerhornz P (2009) Gendered voice and robot entities: perceptions and reactions of male and female subjects. In: 2009 IEEE/RSJ international conference on intelligent robots and systems, pp 3735–3741. ISSN: 2153-0866
30. Paetzel M, Peters C, Nyström I, Castellano G (2016) Congruency matters—how ambiguous gender cues increase a robot's uncanniness. In: Agah A, Cabibihan J-J, Howard AM, Salichs MA, He H (eds) Social robotics. Lecture notes in computer science. Springer, Berlin, pp 402–412
31. Siegel M, Breazeal C, Norton MI (2009) Persuasive robotics: the influence of robot gender on human behavior. In: 2009 IEEE/RSJ international conference on intelligent robots and systems, pp 2563–2568
32. Forgas-Coll S, Huertas-Garcia R, Andriella A, Alenyâ G (2022) The effects of gender and personality of robot assistants on customers' acceptance of their service. *Serv Bus* 16(2):359–89. <https://doi.org/10.1007/s11628-022-00492-x>
33. Deci EL, Ryan RM (2008) Self-determination theory: a macrotheory of human motivation, development, and health. *Can Psychol Psychol Can* 49(3):182
34. Gunnell KE, Crocker PR, Wilson PM, Mack DE, Zumbo BD (2013) Psychological need satisfaction and thwarting: a test of basic psychological needs theory in physical activity contexts. *Psychol Sport Exer* 14(5):599–607
35. Li C, Wang CJ, Kee YH et al (2013) Burnout and its relations with basic psychological needs and motivation among athletes: a systematic review and meta-analysis. *Psychol Sport Exerc* 14(5):692–700
36. Farholm A, Halvari H, Niemiec CP, Williams GC, Deci EL (2017) Changes in return to work among patients in vocational rehabilitation: a self-determination theory perspective. *Disabil Rehabil* 39(20):2039–2046
37. Russell KL, Bray SR (2009) Self-determined motivation predicts independent, home-based exercise following cardiac rehabilitation. *Rehabil Psychol* 54(2):150
38. Bentzen M, Malmquist LK (2021) Differences in participation across physical activity contexts between adolescents with and



- without disability over three years: a self-determination theory perspective. *Disabil Rehabil* 44:1–9
39. Vloothuis J, Depla M, Hertogh C, Kwakkel G, van Wegen E (2020) Experiences of patients with stroke and their caregivers with caregiver-mediated exercises during the care4stroke trial. *Disabil Rehabil* 42(5):698–704
  40. Lou S, Carstensen K, Møldrup M, Shahla S, Zakharia E, Nielsen CP (2017) Early supported discharge following mild stroke: a qualitative study of patients' and their partners' experiences of rehabilitation at home. *Scand J Car Sci* 31(2):302–311
  41. Joerin A, Rauws M, Ackerman ML (2019) Psychological artificial intelligence service, tess: delivering on-demand support to patients and their caregivers: technical report. *Cureus* 11(1):1–5
  42. Meske C, Amojó I, Poncette A-S, Balzer F (2019) The potential role of digital nudging in the digital transformation of the healthcare industry. In: *International conference on human–computer interaction*. Springer, pp 323–336
  43. Zhan Y, Haddadi H (2019) Activity prediction for improving well-being of both the elderly and caregivers. In: *Adjunct proceedings of the 2019 ACM international joint conference on pervasive and ubiquitous computing and proceedings of the 2019 ACM international symposium on wearable computers*, pp 1214–1217
  44. Hassenzahl M, Diefenbach S, Göritz A (2010) Needs, affect, and interactive products-facets of user experience. *Interact Comput* 22(5):353–362
  45. Curry AC, Robertson J, Rieser V (2020) Conversational assistants and gender stereotypes: public perceptions and desiderata for voice personas. In: *Proceedings of the second workshop on gender bias in natural language processing*, pp 72–78
  46. Brems C, Johnson ME (1990) Reexamination of the BEM sex-role inventory: the interpersonal BSRI. *J Personal Assess* 55(3–4):484–498
  47. Guo Y, Yin X, Liu D, Xu SX (2020) She is not just a computer: gender role of ai chatbots in debt collection
  48. Strengers Y, Kennedy J (2023) *The smart wife: why Airi, Alexa, and other smart home devices need a feminist reboot*. MIT Press, Cambridge
  49. Alesich S, Rigby M (2017) Gendered robots: implications for our humanoid future. *IEEE Technol Soc Mag* 36(2):50–59
  50. Nass C, Moon Y, Green N (1997) Are machines gender neutral? gender-stereotypic responses to computers with voices. *J Appl Soc Psychol* 27(10):864–876. <https://doi.org/10.1111/j.1559-1816.1997.tb00275.x>
  51. Perugia G, Guidi S, Bicchì M, Parlangeli O (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*, ser. HRI '22. IEEE Press, pp 110–119
  52. Liberman-Pincu E, Parmet Y, Oron-Gilad T (2023) Judging a socially assistive robot by its cover: the effect of body structure, outline, and color on users' perception. *ACM Trans Hum Robot Interact* 12(2):1–26
  53. Jackson RB, Williams T, Smith N (2020) Exploring the role of gender in perceptions of robotic noncompliance. In: *Proceedings of the 2020 ACM/IEEE international conference on human–robot interaction*, ser. HRI '20. Association for Computing Machinery, pp 559–567. [Online]. Available <https://doi.org/10.1145/3319502.3374831>
  54. Rogers K, Bryant D, Howard A (2020) Robot gendering: influences on trust, occupational competency, and preference of robot over human. In: *Extended abstracts of the CHI conference on human factors in computing systems*, pp 1–7
  55. Cai Z (2020) Gendered robot voices influencing trust-towards a robot recommendation system. *USCCS* 2020, p 1
  56. Seaborn K, Frank A (2022) What pronouns for pepper? A critical review of gender/ing in research. In: *CHI conference on human factors in computing systems*, pp 1–15
  57. Butkus R, Serchen J, Moyer DV, Bornstein SS, Hingle ST, Health and Public Policy Committee of the American College of Physicians (2018) Achieving gender equity in physician compensation and career advancement: a position paper of the American college of physicians. *Ann Intern Med* 168(10):721–723
  58. Olsson M, Martiny SE (2018) Does exposure to counter stereo typical role models influence girls' and women's gender stereotypes and career choices? a review of social psychological research. *Front Psychol* 9:2264
  59. Croft A, Schmader T, Block K (2015) An underexamined inequality: cultural and psychological barriers to men's engagement with communal roles. *Personal Soc Psychol Rev* 19(4):343–370
  60. Block K (2012) Communal male role models: how they influence identification with domestic roles and anticipation of future involvement with the family. *University of British Columbia's Undergraduate Journal of Psychology*, vol 1
  61. Winkle K, Caleb-Solly P, Turton A, Bremner P (2020) Mutual shaping in the design of socially assistive robots: a case study on social robots for therapy. *Int J Soc Robot*. <https://doi.org/10.1007/s12369-019-00536-9>
  62. Winkle K, Caleb-Solly P, Leonards U, Turton A, Bremner P (2021) Assessing and addressing ethical risk from anthropomorphism and deception in socially assistive robots. In: *2021 16th ACM/IEEE international conference on human–robot interaction (HRI)*
  63. Moradbakhti L, Schreibelmayer S, Mara M (2022) Do men have no need for feminist artificial intelligence? agentic and gendered voice assistants in the light of basic psychological needs. *Front Psychol* 13:855091. <https://doi.org/10.3389/fpsyg.2022.855091>
  64. Chen B, Vansteenkiste M, Beyers W, Boone L, Deci EL, der Kaap-Deeder V, Duriez B, Lens W, Matos L, Mouratidis A et al (2015) Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motiv Emot* 39(2):216–236
  65. Cordeiro P, Paixão P, Lens W, Lacante M, Luyckx K (2016) The Portuguese validation of the basic psychological need satisfaction and frustration scale: concurrent and longitudinal relations to well-being and ill-being. *Psychol Belg* 56(3):193
  66. Costa S, Ingoglia S, Inguglia C, Liga F, Lo Coco A, Larcan R (2018) Psychometric evaluation of the basic psychological need satisfaction and frustration scale (BPNSFS) in Italy. *Meas Eval Couns Dev* 51(3):193–206
  67. Moradbakhti L, Leichtmann B, Mara M (2022) Development and validation of a basic psychological needs scale for technology use is available as a preprint at: <https://osf.io/4eabq>
  68. Heissel A, Pietrek A, Flunger B, Fydrich T, Rapp MA, Heinzl S, Vansteenkiste M (2019) The validation of the German basic psychological need satisfaction and frustration scale in the context of mental health. *Eur J Health Psychol*. <https://doi.org/10.1027/2512-8442/a000017>
  69. Sheldon KM, Hilpert JC (2012) The balanced measure of psychological needs (BMPN) scale: an alternative domain general measure of need satisfaction. *Motiv Emot* 36(4):439–451
  70. Brown TA (2015) *Confirmatory factor analysis for applied research*. Guilford Publications, New York
  71. Mair P (2018) *Modern psychometrics with R*. Springer, Berlin
  72. Venkatesh V, Bala H (2008) Technology acceptance model 3 and a research agenda on interventions. *Decis Sci* 39(2):273–315
  73. Eagly AH, Karau SJ (2002) Role congruity theory of prejudice toward female leaders. *Psychol Rev* 109(3):573
  74. Bhaskaran J, Bhallamudi I (2019) Good secretaries, bad truck drivers? Occupational gender stereotypes in sentiment analysis. *arXiv preprint arXiv:1906.10256*

75. Prime JL, Carter NM, Welbourne TM (2009) Women take care, men take charge: managers' stereotypic perceptions of women and men leaders. *Psychol Manag J* 12(1):25–49
76. Carli LL (2001) Gender and social influence. *J Soc Issues* 57(4):725–741
77. Norman L (2010) Bearing the burden of doubt: female coaches' experiences of gender relations. *Res Q Exerc Sport* 81(4):506–517
78. Walker NA, Bopp T (2011) The underrepresentation of women in the male-dominated sport workplace: Perspectives of female coaches. *J Workplace Rights* 15(1):47–64
79. Koch AJ, D'Mello SD, Sackett PR (2015) A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *J Appl Psychol* 100(1):128
80. Carli LL, Alawa L, Lee Y, Zhao B, Kim E (2016) Stereotypes about gender and science: women ≠ scientists. *Psychol Women Q* 40(2):244–260
81. Anthony AS (2004) Gender bias and discrimination in nursing education: Can we change it? *Nurse Educ* 29(3):121–125
82. Rudman LA, Phelan JE (2010) The effect of priming gender roles on women's implicit gender beliefs and career aspirations. *Soc Psychol*. <https://doi.org/10.1027/1864-9335/a000027>
83. BSI. BS 8611:2016 robots and robotic devices: guide to the ethical design and application of robots and robotic systems
84. Seibt J, Vestergaard C, Damholdt MF (2020) Sociomorphing, not anthropomorphizing: towards a typology of experienced sociality. In: *Culturally sustainable social robotics—proceedings of robophilosophy*, pp 51–67
85. Mara M, Appel M, Gnambs T (2022) Human-like robots and the uncanny valley: a meta-analysis of user responses based on the godspeed scales. *Z Psychol* 230(1):33
86. Appel M, Weber S, Krause S, Mara M (2016) On the eeriness of service robots with emotional capabilities. In: *2016 11th ACM/IEEE international conference on human-robot interaction (HRI)*. IEEE, pp 411–412
87. Springer JB, Lamborn SD, Pollard DM (2013) Maintaining physical activity over time: the importance of basic psychological need satisfaction in developing the physically active self. *Am J Health Promot* 27(5):284–293

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.