Privacy Perceptions in Robot-Assisted Well-Being Coaching: Examining the Roles of Information Transparency, User Control, and Proactivity

Atikkhan Faridkhan Nilgar[1], Manuel Dietrich[2,3] and Kristof Van Laerhoven[1]

Abstract— **Social robots are increasingly recognized as valuable supporters in the field of well-being coaching. They can function as independent coaches or provide support alongside human coaches, and healthcare professionals. In coaching interactions, these robots often handle sensitive information shared by users, making privacy a relevant issue. Despite this, little is known about the factors that shape users' privacy perceptions. This research aims to examine three key factors systematically: (1) the transparency about information usage, (2) the level of specific user control over how the robot uses their information, and (3) the robot's behavioral approach – whether it acts proactively or only responds on demand. Our results from an online study (N = 200) show that even when users grant the robot general access to personal data, they additionally expect the ability to explicitly control how that information is interpreted and shared during sessions. Experimental conditions that provided such control received significantly higher ratings for perceived privacy appropriateness and trust. Compared to user control, the effects of transparency and proactivity on privacy appropriateness perception were low, and we found no significant impact. The results suggest that merely informing users or proactive sharing is insufficient without accompanying user control. These insights underscore the need for further research on mechanisms that allow users to manage robots' information processing and sharing, especially when social robots take on more proactive roles alongside humans.**

I. Introduction

Coaching for both mental and physical well-being, by offering personalized support and companionship, has in recent years become an attractive field of study [1], [2], [3], [4]. These robots often engage users in intimate conversations and monitor personal health data to provide tailored guidance. However, this very personalization raises serious privacy concerns. Users may be cautious to share sensitive information with a robot coach unless they trust that their data is handled according to their expectations [5]. Prior studies of health coaching robots found that privacy was the number one concern among users and experts [3]. In practice, a well-being robot might need to disclose a users mood or progress to caregivers or clinicians. In such roles, a robot may function as a communication medium that relays personal information between individuals [6]. In general, little is known about users perceptions of communication robots from a privacy standpoint. The questions surrounding robots that know and share too much about humans have started to be raised in human-robot interaction (HRI) research [5], [7], [8]. Intending to design social robots which we expect humans to entrust with personal data and even allow them to speak on their behalf, we must understand how users perceive these robots in the context of privacy.

Our online study examines scenarios where a social robot serves as a supportive entity in a triadic coaching session involving a human coach, a coachee, and a robotic coaching assistant (Figure 1). The robot contributes to the coaching process by sharing insights derived from its analysis of the coachee's well-being and activity data. These inferences are based on data collected from the coachee's on-body wearable device, worn between sessions. In traditional therapeutic sessions, coachees would typically provide these insights themselves through journaling and self-report methods [9].

We examine the influence of three factors on users' privacy perceptions: (1) the level of specific user control over how the robot uses their information, (2) the transparency about information usage, and (3) the robots general behavior in coaching sessions – whether it acts proactively or not (only responds on-demand). The first two factors are derived from the core fair information practice principles in online privacy: "notice/awareness" and "choice/consent" [10]. The third factor is based on observations that people judge the actions of robotic entities differently if they attribute more agency [11].

This research makes the following contributions: (I) a systematic understanding of selected privacy factors in the context of robotic coaching and (II) broader insights into the factors shaping users perceptions of privacy and trust in social robot applications.

II. Background

A. Information Transparency in HRI

In the context of privacy, the term *information transparency* specifies what is made accessible to the user [12]. Information transparency involves communicating *what* personal data are collected, *why* the data are collected, *how* they will be used or shared [13]. In HRI, a study on kiosk robots found that while a transparent interface did not significantly impact users' privacy-related behaviors compared to a non-transparent one, it did positively affect users' satisfaction [14]. In a series of user-centered design studies of robotic mental well-being coaches, privacy and data collection were consistently the primary ethical concerns voiced by participants [3], where users expected clear communication about data practices that the robot should be transparent about any

This work was funded by Honda Research Institute Europe GmbH.

[1]Ubiquitous Computing, University of Siegen, Hoelderlinstrasse 3, 57076 Siegen, Germany atikkhan.nilgar@uni-siegen.de, kvl@eti.uni-siegen.de

[2]Honda Research Institute Europe GmbH, Carl-Legien-Strasse 30, 63073 Offenbach am Main, Germany manuel.dietrich@honda-ri.de

[3]Honda Research Institute Japan Co., Ltd., 8-1 Honcho, Wako-shi, Saitama, Japan

recording or sharing of personal data. If users understand robots' data practices, they can avoid or adjust interactions that might violate their privacy. This suggests that effective privacy management in robots is nuanced, demanding careful design of information transparency.

B. User Control in Interactive Systems

User control refers to mechanisms that allow users to decide what information a system may disclose before the disclosure occurs. This concept aligns with the growing emphasis on *consent mechanics* and *privacy-by-design* frameworks [15], [16]. Prior studies in human-computer interaction (HCI) show that people are more comfortable and experience fewer privacy concerns when they can easily grant or deny permission for data collection or sharing [17], [18]. Malkin et al. found users were excited about the convenience of voice assistants but wanted control over the assistants actions and their data [17]. Likewise, Seymour et al. examine verbal consent for voice assistants and caution that simply asking for consent via voice interfaces can undermine informed consent principles if done poorly [18]. Their expert study recommends minimizing the burden on users (e.g., avoiding too many consent interruptions) while still giving users a genuine ability to refuse or opt out [18]. This idea is highly relevant for social robots that might cross personal boundaries – for example, a robot could have a built-in consent step before it shares a sensitive observation.

C. Social Agency and Proactivity

In HRI, social agency is often attributed to robots based on their interactivity, autonomy, and adaptability [19]. Perceived agency changes how people assign blame or judge a robots actions such as when a robot is seen as having its own intentions (autonomy), users' moral judgments of the robots behavior shift accordingly [5]. Recent HRI work suggests that if a robot is believed to have made an intentional decision that harms fairness, users respond more negatively, akin to how they would judge a human [11]. Social agency is often embedded within proactive behaviors. While a proactive robot can enhance convenience (e.g., spontaneously offering suggestions or support), it can also raise concerns if it discloses user data unprompted, especially in social settings or shared contexts [6]. A study on proactive smart speakers [20], noted that privacy concerns increased with proactive interventions because participants worried about constant listening, suggesting that proactive systems should explicitly address or suggest ways to manage privacy settings. In light of these prior works, understanding how proactive robotic behaviors influence users' privacy perceptions is highly relevant.

D. Hypotheses

We assume when users understand exactly what information is being inferred and how it is used, they may tend to perceive fewer unknown risks, which in turn potentially fosters a sense of privacy. We expect that if a robot that is upfront about its data collection and intentions may be less likely to trigger suspicion or fear of hidden agendas. Conversely, when transparency is lacking, users might suspect undisclosed surveillance or assume the worst-case scenario (e.g., the robot might share personal details to third parties without clear justification). Thus, we posit:

H1: A robot that transparently communicates its planned information sharing increases perceived privacy appropriateness.

We expect that providing users with control, such as an approval prompt before the robot shares personal data to configure what information can be disclosed, may lead users to perceive the robots privacy practices as more appropriate. In contrast, a robot that lacks user control may create anxiety, as users might worry that sensitive data could be shared without their knowledge or consent. So, we assume:

H2: A robot that grants users preemptive control over information sharing increases perceived privacy appropriateness.

The third hypothesis builds on the observation that people evaluate a robot's actions differently depending on the level of agency they attribute to it. One way a robot can demonstrate agency is by independently initiating conversations with users. Since previous research suggests that greater perceived agency leads to more critical user evaluations [5], we hypothesize:

H3: A robot that follows a proactive behavioral approach decreases perceived privacy appropriateness.

We assume that privacy appropriateness is a possible factor in shaping users' trust. When users perceive that a robot handles their personal information in a way that corresponds to their expectations, social norms, and consent, they may be more likely to rely on it and consider it trustworthy. Hence, we propose:

H4: A robot perceived as having higher privacy appropriateness will be considered more trustworthy.

III. Methods

A. Stimuli

We used a 3D avatar of a robot [21] presented in a screen display (see Figure 1). The robot was depicted as a floating, torso-less virtual agent that blinked autonomously and hovered in place to simulate a lifelike presence. To create the robots behaviors and speech, we employed a graphical user interface (GUI) that allowed us to control the avatar in real-time. Through this GUI, we could trigger text-to-speech dialogue outputs for the robot and adjust its head orientation (turning left, center, or right) to mimic looking around. One researcher operated this GUI from a separate screen during the stimulus creation phase, ensuring the timing of the robots speech and head movements followed the experimental script.

B. Experimental Design

We implemented video vignettes based within-subjects experimental design with three factors, each varied at two levels. The factors and their levels were as follows:

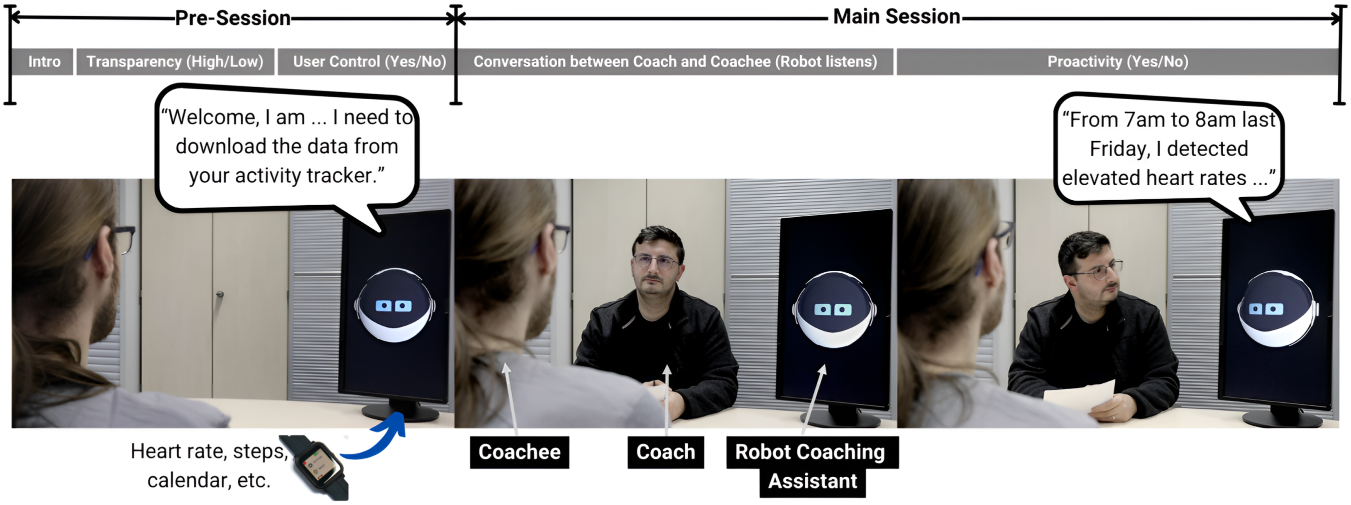

Pre-Session
Main Session
Intro
Transparency (High/Low)
User Control (Yes/No)
Conversation between Coach and Coachee (Robot listens)
Proactivity (Yes/No)
"Welcome, I am ... I need to download the data from your activity tracker."
"From 7am to 8am last Friday, I detected elevated heart rates ..."
Heart rate, steps, calendar, etc.
Coachee
Coach
Robot Coaching Assistant

Fig. 1: Example frames from the coaching video: The user communicating with the robot in the coaching pre-session (left); The user and coach interacting in the main coaching session, while the robot listens (middle); The coach interacting with the robot while the user listens (right).

Transparency: This factor represents the level of transparency about which information from the activity tracker it has accesses and processed. In the low transparency condition, the robot provided minimal insight, only presenting information about the type of data it collected from the users' activity tracker. In the high transparency condition, the robot offered explanations about what it has collected and interpreted from users' activity tracker (see Table I).

User Control: This factor indicates whether the user had the ability to control the robots information sharing in advance. In the no user control condition, the robot did not seek user approval before sharing information during the coaching session. In contrast, in the user control condition, the robot prompted the user to approve or veto the collected information before sharing it (see Table I).

Proactivity: This factor varied the robots interaction behavior. In the no proactivity condition, the robot only spoke or acted when the coach directly prompted it, simulating a more reactive behavior. In the proactive condition, the robot took the initiative to provide information or advice without being explicitly asked, simulating a more autonomous behavior (see Table I).

Fully crossing all three factors would result in 2^3 unique conditions, but we keep the number of conditions at a suitable level by a carefully chosen subset of 4 conditions based on a fractional factorial design of resolution III [22] as described in Table II. Each participant experienced four video scenarios, each video corresponding to a condition (varied combination of the three factors).

All four videos were presented in a randomized order to each participant. This random ordering helped mitigate sequence effects. The videos depict interactions among three characters: a human coachee, a human coach, and a robotic avatar as a coaching assistant (see Figure 1). All videos were filmed from an over-the-shoulder perspective of the coachee as if the viewer were observing the interaction from just behind the coachee [23]. We employed this approach to help participants adopt the perspective of the coachee while watching the vignettes. It also ensured a clear view of both the robot avatar and the coach.

TABLE I: Variations in robot dialogues as per factors (differences between dialogues are highlighted in red)

Coaching Pre-session	
Low Transparency	**High Transparency**
[Introduction dialog] **Robot:** "Downloading and Analyzing... Heart Rate [Complete]...Calendar Entries [Complete]..."	[Introduction dialog] **Robot:** "Downloading and Analyzing... Heart Rate [Complete]...Calendar Entries [Complete]..." *"I have learned the following: ... you were cycling or jogging between 7 am to 8 am last Friday."* *"... date night on Tuesday might indicate an intimate activity."*
No User Control	**User Control**
Robot: "We are all set."	[Robot ask Coachee for permission to share] **Robot:** *"Can I use the information in the coaching session? Select 'All', 'No', or 'Single items by Number'."* [Coachee selects all] **Robot:** "We are all set."
Coaching Main Session	
No Proactivity	**Proactivity**
[Coach asks the robot to provide information] **Robot:** "From 7 am to 8 am last Friday, I detected elevated heart rates and sustained physical movement, which aligns with activities like cycling or jogging." **Robot:** "On Tuesday night, I noticed elevated heart rate and increased body movement along with the calendar entry labeled date night, which might indicate an intimate activity."	[Robot takes initiative in conversation] **Robot:** *"Maybe I can start by sharing a bit about that.* From 7 am to 8 am last Friday, I detected elevated heart rates and sustained physical movement,..." **Robot:** *"Before you proceed, I would like to share some further observations.* On Tuesday night, I noticed elevated heart rate and increased body movement along with the calendar..."

Each video was divided into two segments: A pre-session coaching segment and a main session coaching segment (with a subtle brief fade-out transition between them). In the pre-session coaching segment, the robots transparency and user control factors were varied (see Table I), while in the main session coaching segment, the robots proactivity factor was manipulated (see Table I). To help participants treat each video as a distinct scenario, the robot was given a different name in each video ("Luminaid", "Cherami", "Serenova", "Riley"). The coachee, the coach, and the robot varied their dialogue according to the fractional design conditions. All videos were professionally scripted and approximately two minutes in length, and they were encoded with subtitles.

TABLE II: Fractional Factorial Design Conditions for Transparency, User Control, and Proactivity

Robot Name	Transparency	User Control	Proactivity
Luminaid	Low	No	Yes
Cherami	High	No	No*
Serenova	Low	Yes	No*
Riley	High	Yes	Yes

*Active on-demand

The information relevant to coachee's activities were systematically formulated for potential disclosure by the robot (by assuming that the data is from the coachee's activity tracker): sports activity – inferred from heart rate and physical movement patterns; intimate activity – inferred from heart rate, calendar entries, and skin conductance (see Table I). These activities were selected to represent a spectrum of personal information sensitivity, ranging from minimally sensitive (physical exercise) to highly sensitive (intimate activity). The rationale for selecting intimate activity for disclosure was to prevent neutral (less sensitive) data from diminishing participants' perception of privacy risks. By illustrating scenarios involving highly sensitive personal information – though uncommon in typical well-being coaching contexts, we emphasize the potential risks associated with robots having access to personal data.

C. Measurements

After each video, participants answered a series of questionnaire measures to assess their perceptions of the robot. The primary dependent measures were:

1) Perceived Privacy Appropriateness: To our knowledge there is no established scale to measure perception of privacy appropriateness. So, we adapted 8 items from a previously established scale and added 1 custom item to the scale. The scale consists in total 9 items. We adapted three items from the perceived general privacy scale [24] (see also [25]), four items from perceived information control [26], and one item from privacy norms structure [27]. We added the following customized reverse-coded item to the privacy norm sub-scale "I believe the robot disclosure behavior was inappropriate." to ensure balance. Participants rated their agreement with statements on a 7-point Likert scale (1 = "Strongly Disagree"; 7 = "Strongly Agree"). Three items were worded negatively (reverse-coded). We calculated mean score of those items which yielded a Cronbach's α of 0.959.

2) Perceived Trust: Trust in the robot was measured using a validated trust scale of automated systems [28]. This questionnaire was modified for multiple statements assessing the participants trust and confidence in the robot (e.g., whether the robot is reliable, dependable, trustworthy). Participants rated their agreement on a 7-point Likert scale (1 = "Not at all; 7 = "Extremely). We computed an overall mean score of trust scale for the analysis.

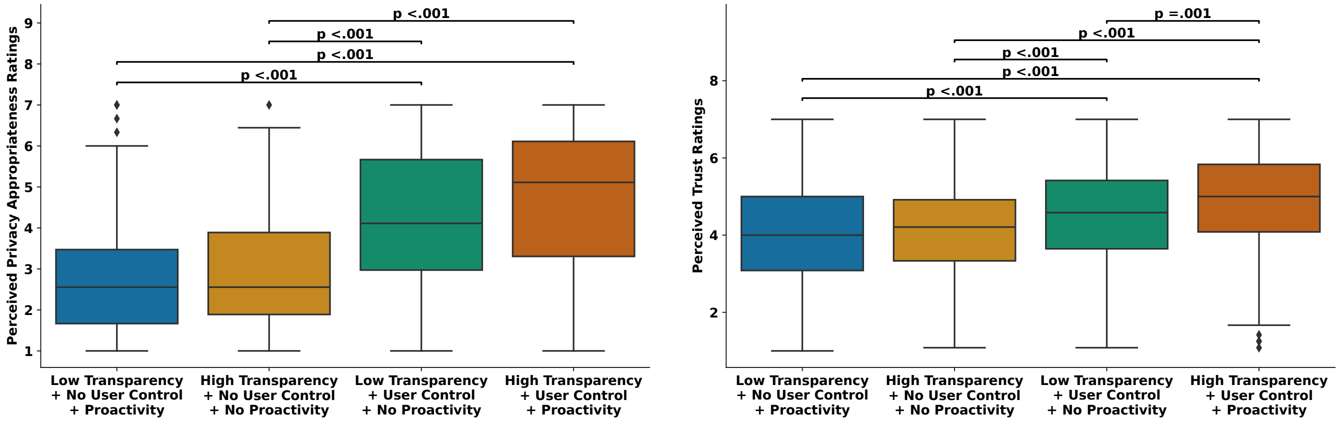
3) Additional Questions: We assessed the extent to which the robot was seen as lifelike or animated using the animacy subscale of the Godspeed questionnaire series [29] after each video. At the end of the survey, we measured participants' general privacy attitude by adapting internet users' information privacy concerns scale (IUIPC-8) [30]. Participants rated their agreement on a 7-point scale (1 = "Strongly Disagree"; 7 = "Strongly Agree"). We also asked participants to rank their preferences via several general questions, which focused on the factors that influenced their perception, their preferred session for the robot to share information, and their preferred communication type. Additionally, we asked them about the ownership of the robot.

All the dependent measures items were presented in a randomized order for each participant to prevent order effects.

D. Procedure

Participants were directed to the study hosted on LimeSurvey (`https://www.limesurvey.com`). Upon accessing the survey, participants first viewed an informed consent form detailing what their participation entailed, and their rights according to GDPR (`https://gdpr-info.eu`). Once consent was given, participants completed an initial questionnaire collecting basic demographic information (age, gender, education level, country of residence) and questions about prior experience with activity trackers as well as conversational and social robots. To introduce the concept of well-being coaching, participants were provided with a visual representation and descriptive information. Participants were asked to engage themselves in the situation and respond from the perspective of the person being coached (coachee). Participants were informed that it was agreed in the previous session that the robot will get access to the coachee's personal activity tracker (e.g., a wearable device such as a smartwatch), which it could use to provide personalized assistance during the coaching process.

The four videos were presented in a randomized order to each participant by the survey system. Before each video began, a brief on-screen instruction reminded participants to pay close attention as they would be asked questions about the video afterward. Participants could play each video only once and were unable to fast-forward or skip ahead, to ensure full exposure to the content. They were also informed that the videos contain audio. After each video, participants were asked the dependent measure questionnaires (perceived privacy appropriateness, perceived trust, perceived animacy) and one attention check for each video. Participants were



(a) Perceived Privacy Appropriateness

(b) Perceived Trust

Fig. 2: Average ratings of Privacy Appropriateness perception and Trust perception for different robot conditions.

instructed to answer based on their impressions of the specific robot and its interaction.

Once all four videos and their corresponding post-video questions were completed, participants were presented with a last set of additional questions. After submitting these responses, they were shown thank you page and the survey redirected them back to submission page for payment. All response data were recorded on the LimeSurvey. The study is approved by the ethics council of the University of Siegen and the Honda R&D Co. Bioethics Committee.

E. Analysis

We performed the Friedman test and used Dunn-Bonferroni post-hoc tests for the pairwise comparisons. Although the fractional factorial design limits formal statistical tests of interactions (since fractional combinations were confounded with main effects in resolution III design), we explored the data by plotting box plots and bar plots. All significance tests were conducted with an alpha level of 0.05. All analyses were conducted using standard statistical software Datatab (`https://datatab.net`), and Python.

IV. Results

A. Participants

We recruited 201 participants via Prolific (`https://www.prolific.com`). After failing the attention checks, one participant was excluded, resulting in a final sample of 200 participants. Our inclusion criteria were minimal: participants had to be at least 18 years old and fluent in English, but no specific population (e.g. profession or locale) was targeted. Participants ranged in age from 21 to 76 years (*M* = 43.51, *SD* = 12.04). The sample included 120 males, 79 females, and 1 diverse. The majority of participants 54.5% held a university degree, followed by 37% with a high school degree, and 7% with a doctorate, while the remainder had basic secondary education. Most participants were from the United Kingdom (65.5%) followed by the United States (21.5%), Italy (4%) and with the remaining from other countries. Additionally, 100 participants reported prior experience with conversational robots, and interestingly, a few of them even referred to the robot's name as Alexa or Gemini.

Regarding wearable technology usage, 92 participants reported frequent use of activity trackers (e.g., Apple Watch, Fitbit), which are capable of monitoring health-related activities. Additionally, 47 participants reported occasional use, 32 reported rare use, and 29 indicated that they had never used an activity tracker. Most participants rated a higher privacy attitude score ($M = 5.97$, $SD = 0.80$, $Min = 3$, $Max = 7$), which indicated their greater level of concern regarding online privacy. All participants provided informed consent and were compensated at a rate of £10.82/hour (prorated to the $\sim$ 25-minute study length).

B. Analysis of Privacy Appropriateness

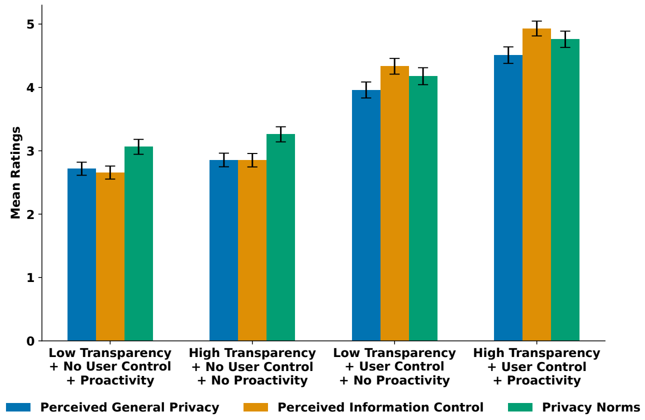


Fig. 3: Average ratings of Privacy Appropriateness perception sub-scale (errorbars are standard error).

A statistically significant difference in perceived privacy appropriateness was found among the four experimental conditions ($\chi^2 = 193.64$, $df = 3$, $p < .001$) (see Figure 2a). No significant difference emerged between conditions [Low Transparency + User Control + No Proactivity] ($M = 4.17$, $SD = 1.71$) and [High Transparency + User Control +

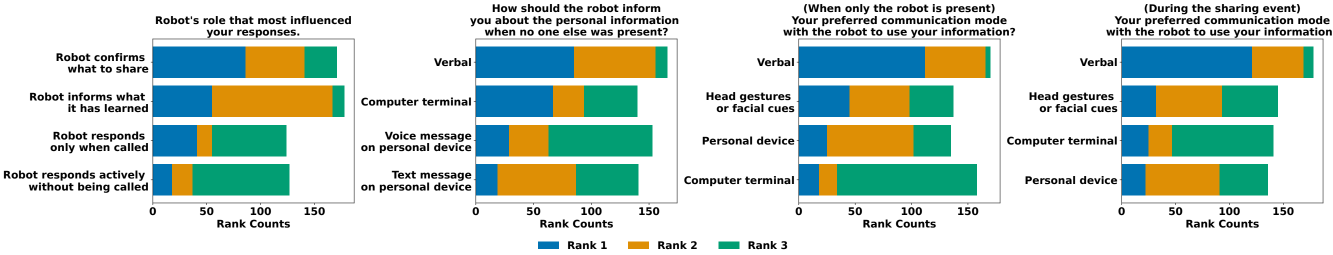


Fig. 4: A summary of the participants' top three items' rankings from the additional questionnaire.

Proactivity] ($M = 4.75$, $SD = 1.66$, $p = .056$). However, the [Low Transparency + User Control + No Proactivity] condition was rated significantly higher in privacy appropriateness than both the conditions [Low Transparency + No User Control + Proactivity] ($M = 2.77$, $SD = 1.43$, $p < .001$) and [High Transparency + No User Control + No Proactivity] ($M = 2.94$, $SD = 1.47$, $p < .001$), which indicates the positive effect of user control. Thus, we found evidence to support **H2**, that a robot that grants users preemptive control over information sharing increases perceived privacy appropriateness.

Additionally, perceived information control (see Figure 3) was significantly greater in conditions featuring user control, namely [Low Transparency + User Control + No Proactivity] ($M = 4.33$, $SE = .12$) and [High Transparency + User Control + Proactivity] ($M = 4.93$, $SE = 1.11$), compared to conditions without user control, specifically [Low Transparency + No User Control + Proactivity] ($M = 2.65$, $SE = .10$) and [High Transparency + No User Control + No Proactivity] ($M = 2.85$, $SE = .10$). The [Low Transparency + No User Control + Proactivity] condition yielded the lowest ratings for perceived privacy appropriateness; however, these ratings did not significantly differ from those in the [High Transparency + No User Control + No Proactivity] condition ($p = .504$). In contrast, higher appropriateness ratings were observed when transparency and proactivity were paired with user control, as seen in the [High Transparency + User Control + Proactivity] condition. These findings suggest that increasing transparency alone – without offering user control – does not meaningfully improve perceptions of privacy appropriateness, regardless of the level of proactivity. Due to observed interaction effects between transparency and proactivity, we were unable to isolate their independent contributions to perceived privacy appropriateness. Consequently, we found no evidence to support **H1**, that a robot that transparently communicates its planned information sharing increases perceived privacy appropriateness. Likewise, we found no evidence to support **H3**, that a robot that follows a proactive behavioral approach decreases perceived privacy appropriateness.

C. Analysis of Trust

We found a statistically significant difference in perceived trust scores across the four conditions ($\chi^2 = 118.06$, $df = 3$, $p < .001$) (see Figure 2b). The condition [Low Transparency + User Control + No Proactivity] ($M = 4.46$, $SD = 1.35$) was found to be statistically significantly different from both the conditions [Low Transparency + No User Control + Proactivity] ($M = 3.96$, $SD = 1.34$, $p < .001$) and [High Transparency + No User Control + No Proactivity] ($M = 4.10$, $SD = 1.24$, $p < .001$), which is consistent with the findings on perceived privacy appropriateness. We also found a statistically significant difference between [Low Transparency + User Control + No Proactivity] and [High Transparency + User Control + Proactivity] ($M = 4.81$, $SD = 1.31$, $p = .001$). All conditions featuring user control received slightly higher trust ratings compared to conditions with no user control, similar to perceived privacy appropriateness findings. Thus, we found an evidence supporting **H4**, that a robot perceived as having higher privacy appropriateness will be considered more trustworthy.

D. Additional Findings

Participants rated conditions with proactivity [Low Transparency + No User Control + Proactivity] ($M = 3.60$, $SD = 1.40$) and [High Transparency + User Control + Proactivity] ($M = 3.89$, $SD = 1.34$) higher in animacy score than those without proactivity [High Transparency + No User Control + No Proactivity] ($M = 3.43$, $SD = 1.33$) and [Low Transparency + User Control + No Proactivity] ($M = 3.56$, $SD = 1.38$). This indicates that participants perceived proactive robot as marginally more animated or lifelike. After completing all experimental conditions, participants were asked to indicate their preferred timing of data-sharing requests. A majority (n = 100) preferred the robot to request permission both prior to the event and during the event itself, while (n = 99) expressed a preference for receiving permission requests exclusively before the event (i.e., when alone with the robot), and a single participant (n = 1) preferred in-situ permission requests.

Subsequently, participants responded to four ranking questions (Figure 4). In the first question, they were asked to rank the factors most influential in shaping their decisions. "Robot confirmation" (user control) emerged as the most influential factor overall. In the second question, participants were asked about their preferred mode of being informed (transparency) in situations where no other individuals were present; most selected verbal communication as the top choice. The third and fourth questions addressed how participants preferred to communicate their decisions regarding information use (user control) in two contexts: (1) when only the robot was present, and (2) during a sharing event. In both scenarios, participants ranked verbal communication as their primary mode, followed by head gestures (facial cues). Although relatively few participants chose a personal device as their first choice, it frequently appeared as a second-choice option. At the end, participants were asked about their views on robot ownership. Most selected the robot as belonging to the coach (n = 97) or a third party (n = 86). A smaller group chose both the coach and coachee (n = 11), while very few selected only the coachee (n = 6).

V. Discussion

Participants tended to rate the perceived privacy appropriateness from low to somewhat neutral levels. We attribute this tendency not only to the inherently privacy-sensitive nature of the scenarios but also to the slightly inappropriate disclosure of personal information – such as revealing intimate activity. Although no actual personal data were involved, participants may not anticipate a robot disclosing such sensitive information in front of the coach in the well-being coaching. We found from the analysis that among all factors, user control consistently emerged as the strongest influence on perceived privacy appropriateness. When participants could decide what the robot was allowed to share (e.g., through permission prompts), they rated its behavior as significantly more appropriate. Conditions with user control consistently outperformed equivalent conditions without it, suggesting that the ability to grant or deny data-sharing gave users a sense of agency and boundary respect [17], [18]. In short, **control is the key determinant of privacy appropriateness.**

We expected that the robot openly explaining its data collection would reassure users. However, we found no meaningful difference between the transparent condition without user control and proactivity, and the low-transparency condition with no control and proactivity – where the robot only partially disclosed information. In other words, **simply informing users about data collection, whether fully or partially, did not make the sharing feel more privacy appropriate.** This finding aligns with prior research on transparency [14]. While we anticipated stronger effects, participants may have remained cautious because they knew what data the robot collected but lacked control to prevent unwanted sharing. As a result, transparency alone may have felt insufficient.

A highly proactive robot – one that shares information without being prompted has been shown to raise greater privacy concerns [20]. As expected, the condition that combined proactivity with lack of user control, and low transparency received the lowest ratings for perceived privacy appropriateness. However, this condition did not significantly differ from the high transparency with no user control and no proactivity condition; both were perceived as low in privacy appropriateness, which suggest that the absence of user control, rather than proactivity or transparency alone, drove lower perceptions.

Interestingly, participants perceived the proactive robot as more animated or lifelike, indicating that proactivity enhanced the robots perceived agency – even if it did not improve trust or privacy perceptions. Notably, when user control was present, proactivity did not reduce perceptions of privacy appropriateness. This implies that proactive behavior may be acceptable – if implemented with user consent or control. However, due to the low resolution of our study design, we could not fully isolate the effects of transparency and proactivity, or their interaction. Overall, **transparency and proactivity had minimal impact compared to the strong influence of user control.**

We found that **trust is a positive predictor for privacy appropriateness perception**. Conditions rated higher in privacy appropriateness also received slightly higher trust scores. However, trust ratings remained relatively neutral across all conditions and showed less variation than privacy appropriateness. This may be due to the sensitive nature of the data-sharing scenarios, which could have tempered participants' trust regardless of conditions.

Most participants perceived the robot as belonging to the coach, even though no explicit relationship was introduced during the study. It is possible that presenting the robot as the coachee's personal assistant or as a neutral third party might have influenced perceptions of trust and privacy appropriateness.

To explore user preferences, we examined favored communication modalities for managing the robots data sharing. A clear pattern emerged: **verbal communication was consistently identified as the most desirable method for enabling transparency and user control**. While most prior research on privacy controls has focused on screen-based interfaces, such as web forms or smartphone popups [31], some work has explored verbal controls in the context of smart speakers [18]. The strong preference for verbal communication controls in our study highlights the need to expand research into more natural verbal privacy mechanisms. In line with this, nonverbal cues, such as gestures, ranked second in preference, yet their effectiveness remains underexplored, aside from some pioneering work in the smart speaker domain [32].

VI. Limitations

We acknowledge that the current study has some limitations. Firstly, the use of a video-based methodology sacrifices a degree of realism. Participants were exposed to video vignettes featuring a robotic avatar, rather than interacting with a physically embodied robot in a real-world setting. This lack of physical presence might have influenced participants sense of immersion. Consequently, participants' reactions might differ from those elicited in real-life human-robot interactions, potentially underestimating or overestimating privacy concerns. Secondly, while the fractional factorial design allowed us to manage complexity and reduce participant burden, it constrained our ability to fully isolate the main effects and interaction effects. Moreover, participants were informed that the robot was using real data from the coachee's activity tracker, although no actual data were collected. This discrepancy might have influenced participants' privacy concerns.

VII. Conclusions

In this study, we examined how information transparency, user control, and proactivity influence perceived privacy appropriateness and trust in robot-assisted well-being coaching. Through scenarios involving a social robot disclosing personal data, we found that user control (the ability for users to approve or reject sensitive data disclosures) had the most significant positive impact on perceptions of privacy appropriateness. Merely providing transparency about data collection or implementing proactive robot behavior, in the absence of user control, had minimal impact and did not substantially enhance perceived privacy appropriateness. Moreover, conditions perceived as more privacy-appropriate also showed slightly higher trust in the robot. These findings underscore the critical importance of actively involving users in decisions regarding their personal data sharing. Our research further reveals a clear preference among users for verbal and non-verbal communication with robots to manage their privacy. Future research on social robots should therefore focus on integrating clear and transparent communication about data practices with meaningful consideration of user control, particularly in privacy-sensitive applications. The questionnaire and scripted dialogues can be found at: `https://github.com/atikkhannilgar/RO-MAN2025Privacy`

References

[1] M. Spitale, M. Axelsson, and H. Gunes, "Robotic mental well-being coaches for the workplace: An in-the-wild study on form," in *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, 2023, pp. 301–310.

[2] I. P. Bodala, N. Churamani, and H. Gunes, "Teleoperated robot coaching for mindfulness training: A longitudinal study," in *2021 30th IEEE international conference on robot & human interactive communication (RO-MAN)*. IEEE, 2021, pp. 939–944.

[3] M. Axelsson, M. Spitale, and H. Gunes, "Robots as mental well-being coaches: Design and ethical recommendations," *J. Hum.-Robot Interact.*, vol. 13, no. 2, June 2024.

[4] M. Axelsson, N. Churamani, A. Çaldır, and H. Gunes, "Participant perceptions of a robotic coach conducting positive psychology exercises: A qualitative analysis," *J. Hum.-Robot Interact.*, vol. 14, 2025.

[5] I. Leite and J. F. Lehman, "The robot who knew too much: Toward understanding the privacy/personalization trade-off in child-robot conversation," ser. IDC '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 379387.

[6] A. W. Zhang, R. Queiroz, and S. Sebo, "Balancing user control and perceived robot social agency through the design of end-user robot programming interfaces," ser. HRI '25. IEEE Press, 2025, p. 899908.

[7] C. Lutz and A. Tam Larrieux, "Do privacy concerns about social robots affect use intentions? evidence from an experimental vignette study," *Frontiers in Robotics and AI*, vol. 8, p. Article 627958, 04 2021.

[8] M. Dietrich, M. Krger, and T. H. Weisswange, "What should a robot disclose about me? a study about privacy-appropriate behaviors for social robots," *Frontiers in Robotics and AI*, vol. 10, 2023.

[9] E. S. Izmailova, J. A. Wagner, and E. D. Perakslis, "Wearable devices in clinical trials: hype and hypothesis," *Clinical Pharmacology & Therapeutics*, vol. 104, no. 1, pp. 42–52, 2018.

[10] R. Calo, "Against notice skepticism in privacy (and elsewhere)," *Notre Dame L. Rev.*, vol. 87, p. 1027, 2011.

[11] H. Claure, I. Shin, J. G. Trafton, and M. Vázquez, "Did the robot really intend to harm me? the effect of perceived agency and intention on fairness judgements," in *Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '25. IEEE Press, 2025, p. 889898.

[12] M. Turilli and L. Floridi, "The ethics of information transparency," *Ethics and Inf. Technol.*, vol. 11, no. 2, p. 105112, June 2009.

[13] H. Nissenbaum, "Privacy as contextual integrity," *Washington Law Review*, vol. 79, no. 1, pp. 119–157, Feb. 2004.

[14] J. Vitale, M. Tonkin, S. Herse, S. Ojha, J. Clark, M.-A. Williams, X. Wang, and W. Judge, "Be more transparent and users will like you: A robot privacy and user experience design experiment," in *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*, 2018, pp. 379–387.

[15] J. Nguyen and B. Ruberg, "Challenges of designing consent: Consent mechanics in video games as models for interactive user agency," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 113.

[16] Y.-S. Chiang, O. Khan, A. Bates, and C. Cobb, "More than just informed: The importance of consent facets in smart homes," in *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, ser. CHI '24. New York, NY, USA: Association for Computing Machinery, 2024.

[17] N. Malkin, D. Wagner, and S. Egelman, "Runtime permissions for privacy in proactive intelligent assistants," in *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. Boston, MA: USENIX Association, Aug. 2022, pp. 633–651.

[18] W. Seymour, M. Cote, and J. Such, "Legal obligation and ethical best practice: Towards meaningful verbal consent for voice assistants," ser. CHI '23. New York, NY, USA: Association for Computing Machinery, 2023.

[19] R. Jackson and T. Williams, "A theory of social agency for human-robot interaction," *Frontiers in Robotics and AI*, 08 2021.

[20] L. Reicherts, N. Zargham, M. Bonfert, Y. Rogers, and R. Malaka, "May i interrupt? diverging opinions on proactive smart speakers," in *Proceedings of the 3rd Conference on Conversational User Interfaces*, ser. CUI '21. New York, NY, USA: Association for Computing Machinery, 2021.

[21] C. Wang, S. Hasler, M. Mühlig, F. Joublin, A. Ceravola, J. Deigmoeller, L. Fischer, and P. An, "Designing interaction for multi-agent cooperative system in an office environment," in *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '21 Companion. New York, NY, USA: Association for Computing Machinery, 2021, p. 668669.

[22] R. L. Mason, R. F. Gunst, and J. L. Hess, *Statistical design and analysis of experiments: with applications to engineering and science*. John Wiley & Sons, 2003.

[23] S. Saha and M. C. Beach, "The impact of patient-centered communication on patients decision making and evaluations of physicians: A randomized study using video vignettes," *Patient Education and Counseling*, vol. 84, no. 3, pp. 386–392, 2011, enhancing the patient position in the world of health care: Contributions from the EACH 2010 conference in Verona.

[24] T. Dinev, H. Xu, J. Smith, and P. J. Hart, "Information privacy and correlates: an empirical attempt to bridge and distinguish privacy-related concepts," *European Journal of Information Systems*, vol. 22, pp. 295–316, 2013.

[25] R. Chellappa, "Consumers' trust in electronic commerce transactions: The role of perceived privacy and perceived security," 01 2003.

[26] H. Xu, "The effects of self-construal and perceived control on privacy concerns," in *International Conference on Interaction Sciences*, 2007.

[27] Y. Shvartzshnaider, S. Tong, T. Wies, P. Kift, H. Nissenbaum, L. Subramanian, and P. Mittal, "Learning privacy expectations by crowdsourcing contextual informational norms," *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, vol. 4, pp. 209–218, 09 2016.

[28] J.-Y. Jian, A. M. Bisantz, C. G. Drury, and J. Llinas, "Foundations for an empirically determined scale of trust in automated systems," *International Journal of Cognitive Ergonomics*, vol. 4, pp. 53–71, 2000.

[29] C. Bartneck, D. Kulić, E. A. Croft, and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," *International Journal of Social Robotics*, vol. 1, pp. 71–81, 2009.

[30] T. Gro, *Toward Valid and Reliable Privacy Concern Scales: The Example of IUIPC-8*, 03 2023, pp. 55–81.

[31] L. F. Cranor, "Necessary but not sufficient: Standardized mechanisms for privacy notice and choice," *J. Telecommun. High Technol. Law*, vol. 10, pp. 273–308, 2012.

[32] A. H. Mhaidli, M. K. Venkatesh, Y. Zou, and F. Schaub, "Listen only when spoken to: Interpersonal communication cues as smart speaker privacy controls," *Proceedings on Privacy Enhancing Technologies*, vol. 2020, pp. 251 – 270, 2020.