

Telepresence Social Robotics towards Co-Presence: A Review

Luis Almeida ^{1,2,*} , Paulo Menezes ²  and Jorge Dias ^{2,3}¹ Ci2—Smart Cities Research Center, Polytechnic Institute of Tomar, 2300-313 Tomar, Portugal² Institute of Systems and Robotics, University of Coimbra, 3030-290 Coimbra, Portugal; paulomenezes@isr.uc.pt (P.M.); jorge@deec.uc.pt (J.D.)³ Khalifa University Center for Autonomous Robotic Systems (KUCARS), Khalifa University of Science and Technology (KU), Abu Dhabi 127788, United Arab Emirates

* Correspondence: laa@ipt.pt

Abstract: Telepresence robots are becoming popular in social interactions involving health care, elderly assistance, guidance, or office meetings. There are two types of human psychological experiences to consider in robot-mediated interactions: (1) telepresence, in which a user develops a sense of being present near the remote interlocutor, and (2) co-presence, in which a user perceives the other person as being present locally with him or her. This work presents a literature review on developments supporting robotic social interactions, contributing to improving the sense of presence and co-presence via robot mediation. This survey aims to define social presence, co-presence, identify autonomous “user-adaptive systems” for social robots, and propose a taxonomy for “co-presence” mechanisms. It presents an overview of social robotics systems, applications areas, and technical methods and provides directions for telepresence and co-presence robot design given the actual and future challenges. Finally, we suggest evaluation guidelines for these systems, having as reference face-to-face interaction.

Keywords: social robotics; co-presence; copresence; social presence; telepresence; cognitive robotics; telerobotics



Citation: Almeida, L.; Menezes, P.; Dias, J. Telepresence Social Robotics towards Co-Presence: A Review. *Appl. Sci.* **2022**, *12*, 5557. <https://doi.org/10.3390/app12115557>

Academic Editors: Alessandro Gasparetto, Manuel Armada and Keun-Chang Kwak

Received: 10 April 2022

Accepted: 19 May 2022

Published: 30 May 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Telepresence robots are becoming popular in the context of social interactions. Typically, these systems enable people to look at a distant place via teleoperating a robot and interacting with another person at a remote location using the built-in communication devices. Some relevant applications include health care, elderly assistance, autism therapy, guidance, and office meetings [1–7].

This literature review aims to gather knowledge to help roboticists design improved user- and environment-adaptive systems and technical methods that contribute to enhancing the sense of presence or co-presence via social robot mediation. Reviews have addressed user-adaptive systems [2,8] and environment-adaptive systems [9] for social robotics (in which the robot is generally an autonomous agent serving the bystander user). However, we further explore telepresence social robotics, with an emphasis placed on the relationship between the robot’s operator and the bystander user.

Within social telepresence robots interactions, two types of human psychological experiences can be considered (see Figure 1). The first one involves the *remote user*, in which he or she should sense being in the *local environment* (i.e., telepresence) [10,11], and the second type involves the *local user*, in which ultimately he or she should sense that the remote user is with him or her in the local environment (i.e., *co-presence*) [12,13]. This research will focus on this last type of interaction, or how to enhance the sense of *co-presence* via robot mediation. To clarify the role of each agent in the interaction, the following terminology is adopted:

1. Mobile robotic telepresence (MRP) system: remotely controllable mobile platform with video conferencing equipment that allows remote users to navigate within a local

environment and socially interact with other persons. These systems can incorporate semi-autonomous functionalities to mitigate operation loads such as navigation aids, points to follow, and obstacle avoidance.

2. Robotic telepresence (RP) system: remotely controllable or semi-autonomous robotic device with video conferencing capabilities that enable social interaction with people in the local environment without locomotion means. Remote users can explicitly control parts of the robot (e.g., the head’s panning, swinging, tilting, eye gazing, and facial expressions, as well as arm or hand gestures) or enable some semi-autonomous behaviors (e.g., blinking, face tracking, eye saccade, and breathing).
3. Remote user: user that steers the robot from a distant location or simply connects to the robot through a computer interface.
4. Local user: user that shares the physical environment with the robot (bystander).
5. Local environment: environment shared by the local user and robot.

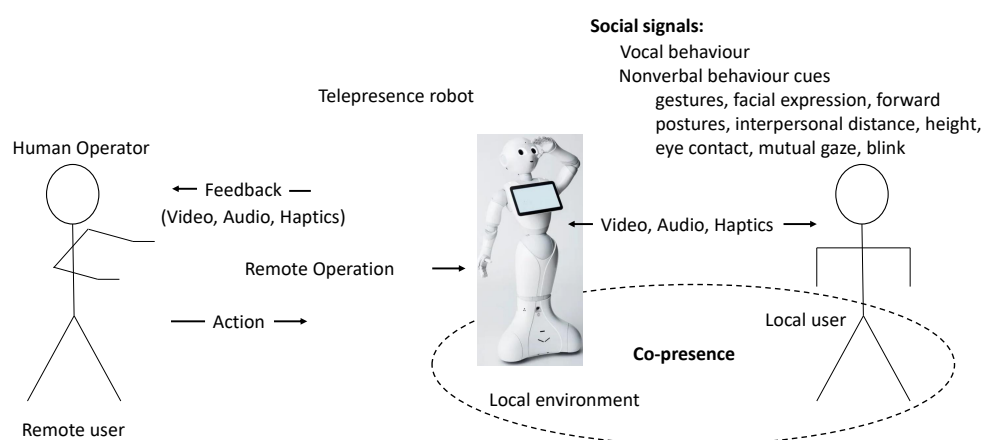


Figure 1. Interaction scenario with telepresence and co-presence.

Presence is often defined as the sense of *being there* in a mediated environment [14,15]. Additionally, Sheridan [16] differentiates presence (virtual) from telepresence (experiential). Presence describes the experience of being present within a virtual world, while telepresence refers to the sense of being in a mediated remote real environment. Co-presence has been used to refer to the *sense of being together* with others in a mediated (either in remote real or virtual) environment [13,17–19].

Marvin Minsky introduced the telepresence concept in the teleoperation context to describe the phenomenon in which a human operator feels physically present at a remote location through interaction with the human’s sensing systems [11] (i.e., “through actions of the user and the corresponding perceptual feedback provided by the teleoperation technology”) [10].

Paulos and Canny [20] developed one of the first telepresence robots and referred to it as a personal roving presence (PRoP) device. The goal was to “provide a physical mobile proxy, controllable over the Internet to provide tele-embodiment”. The system consisted of a simple controllable mobile platform with a video conference set-up (microphone, speaker, and a video camera with 16x zoom and a 30-cm screen on the top of a plastic pole). Additionally, the robot enabled simple gesturing through a two-DoF pointer. They introduced the concept of tele-embodiment in the robotics context to describe the sensation of embodiment of a human in a real distant location [21]. Tele-embodiment was defined as telepresence with a personified perceptible body [22]. However, they did not address key conditions such as body ownership [23] or agency [24]. Li [25] surveyed and compared 33 experimental works involving people’s interactions with virtual agents, telepresence robots, and co-present robots, concluding that robots are more persuasive and positively perceived when they are physically present in the user’s environment.

Short [12] introduced the concept of *social presence*, defining it as the degree of salience of the participants involved in an interaction and their interpersonal relationship. He mentioned that *social presence* relied on two concepts: *intimacy* and *immediacy*. *Intimacy* senses the degree of connectedness between the interactants, and *immediacy* refers to the psychological sense of togetherness between the communicators. Taking face to face (FtF) as the reference, both concepts are determined by a set of verbal and nonverbal cues such as vocal cues, gestures, facial expressions, and physical appearance. The capability to deliver such cues differs from communication means, so Short considered social presence as the *quality of the medium itself*. Later, Biocca [13] referred to *social presence* as the effect on one person's behavior caused by the presence of another or caused by knowing that he or she could be observed. *Co-presence*, defined as the "psychological connection to and with another person" [17,26], has been explored in several works [27–30].

Cognitive robotics aims to provide robots with intelligent behavior through a processing architecture that involves perception, long- and short-term memory, learning, and reasoning. These approaches try to deal with people's behavior unpredictability and with real-world complexity. Cognitive technologies are a form of hyper-automation that may combine areas such as symbolic representation, automation, prediction, user-adaptive systems, computer vision (CV), machine learning (ML), deep learning (DL), or artificial intelligence (AI) [2,9,31]. Nevertheless, the use of AI methodologies to emulate or interpret human subjective experiences, such as emotions, should be inspired by neurophysiologic-psychological foundations [32].

An inner issue related to teleoperated telepresence robots is the time delay issue (mainly due to the communication channel and less due to the hardware performance). This can affect synchronicity (rate of message exchange between operator and bystander), compromising the social presence [33] (e.g., degradations in audio and video streams, control streams, and haptic feedback). Problems regarding latency, bandwidth limitations, and channel corruptions should be mitigated, and while early solutions involved user interface design and control theory-based models (e.g., supervisory control or passivity-based teleoperation), the approaches evolved to predictive displays and control. Advanced solutions for time delay issues are using time series prediction methods to predict the time delay, a robot's movements, and user intentions. These new adaptive-based control methods make use of nonlinear statistical models and neural network (NN) or machine learning (ML) techniques (e.g., recurrent neural networks, sequence to sequence, long short-term memory, or generative adversarial networks) [31].

The method for this literature review and article selection consisted of retrieving and collecting review studies on social presence, co-presence, and the principles and heuristics of human–robot interaction (HRI), with emphasis on teleoperated telepresence robots. Searches were performed on bibliographic scientific databases, such as ACM's digital library, Google Scholar, MIT Press Direct, Elsevier's ScienceDirect, IEEE's Xplore, PubMed, Scopus, and Springer. Queries included general keywords such as social robots, social robots survey, co-presence taxonomy, copresence or co-presence robots, telepresence robots, adaptive systems, and more specifically, the compositions of these keywords. The selection of papers for in-depth reading was determined by the number of citations, being a recent publication, being a journal (e.g., IEEE transactions, Elsevier's, or Springer), being a book, including user evaluations studies, or being an article in a reputable conference in the field (e.g., ACM HRI conferences, IEEE Robotics and Automation Society, ICRA, or the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)). Citations in these papers directed new readings and paper selections. Figure 2a depicts the article's citation distribution per main topic, Figure 2b is from the *Co-Presence Taxonomy/Predictors* topic, Figure 2c is from the *From Telepresence to Co-Presence Design* topic, and Figure 2d shows the article citation distribution per year.

This survey presents an overview of social robotics systems and focuses on how to enhance the sense of co-presence via robot mediation. It reviews the literature to define social presence and co-presence, identifies predictors, and proposes a taxonomy for "co-presence" and "user-adaptive systems" mechanisms. It provides technical methods to support robotic

social interactions. The structure of this article is composed of four parts. Section 2 identifies potential predictors for social presence, suggesting a taxonomy for “co-presence”. Section 3 presents several robotic telepresence systems currently available in the market or used in research. It also reviews autonomous user-adaptive systems for social robots, aiming for a taxonomy, and additionally provides design guidelines for mechanisms that enhance the sense of co-presence in communications through a teleoperated telepresence robot. It includes guidelines for the evaluation of these systems, having as reference the face-to-face interaction. Finally, Section 4 presents the conclusions and future work.

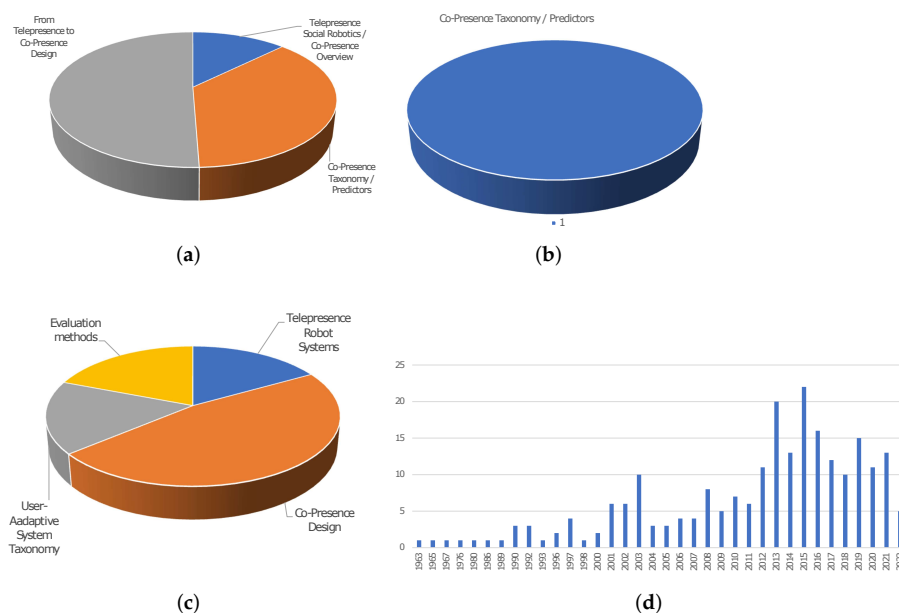


Figure 2. (a) Article’s citation distribution per main topic, (b) per *Co-Presence Taxonomy / Predictors* topic, (c) per *From Telepresence to Co-Presence Design* topic, and (d) the article’s citation distribution per year.

2. Co-Presence Taxonomy

Social presence has been defined as the *sense of being together with another*, which includes primitive reactions to social cues and automatic creation of simulations or mental models of “*other minds*” [13]. Short et al. [12] defined *social presence* as “the degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationship”.

Co-presence is a different concept, introduced by Goffman [26] to describe the active state in which a person perceives his or her interlocutor, and the interlocutor also perceives him or her. *Copresence* refers to a “psychological connection to and with another person”, in which “interactants feel they were able to perceive their interaction partner and that their interaction partner actively perceived them” [17]. With co-presence being a subjective concept, it involves different dimensions and interpretations depending on the social science discipline and application area (e.g., sociology or psychology) [18,29,30].

Social presence appears in the literature as being related to the *quality of the communication’s medium* [12] and the user’s perception of the medium. Therefore, preliminary studies have focused on the effect of modality on social presence. They identified potential predictors of social presence by analyzing the technology’s capability to reproduce social cues (e.g., visual representation, audio, and haptic feedback). The findings were biased by the considered concept definitions. Some predictors contribute directly to presence, co-presence, or social presence, while others affect them indirectly by acting on a person’s involvement and immersion. Therefore, it is important to distinguish the “immersion” concept and the “presence” concept [15,34].

Immersion, also known as *sensorimotor immersion*, refers to the extent and fidelity of physical stimulation affecting the human sensory systems and the system’s responsiveness

to the motor inputs. The immersive level depends on the number and range of sensors and the motor channels connected to a remote agent in a real environment (e.g., a robot) or to a mediated virtual environment. Immersion is determined by the naturalness and coherence between actions (head, body, and gesture movements) and the expected sensory feedback [35–38].

Presence is the psychological product of technological immersion, defined as *the perceptual illusion of non-mediation* [39] or simply referred to as the sense of *being there* in a mediated virtual environment [15,34]. Sheridan [16] differentiates presence (virtual) from telepresence (experiential), in which *presence* describes the experience of being present within a virtual world while *telepresence* refers to the sense of being in a mediated remote real environment [19,40].

Co-presence has been used to refer to the *sense of being together* with others in a mediated environment, either remote real or virtual [18,41]. As described in the definitions, the use of concepts such as co-presence and social presence should not be confused as they are assessed differently [17].

In the context of social robotics, there are agents with autonomous and semi-autonomous behaviors that are seen by the local person as the “other”. Additionally, some agents simply mediate the communications between two persons (the remote and local users). In the former case, the sense of co-presence is assessed between an artificial system and a person, while in the second scenario, co-presence involves two humans. Typically, in robotic telepresence, the representation of the remote real person is shaped by the technology that mediates communication. This affects the perception of thoughts and emotions when compared with actual face-to-face (FtF) interaction. Such representation of remote humans may be supported through text, images, video, 3D avatars, 3D reconstruction, virtual human agents, computers, and robots. Zhao [18], Cummings [33], and Oh, Bailenson, and Welch [29] reviewed the concepts of social presence and co-presence, and their studies suggest a classification for co-presence predictors. This paper adopts some of these literature predictors, framing them in the context of telepresence social robotics.

To unveil a list of technological predictors of social presence, the authors of [33] performed a literature review of empirical studies and grouped them according to similar manipulations. They performed a bottom-up analysis process and identified the following predictors (Table 1):

Table 1. Technological predictors.

Technological predictors of social presence [33]	Behavioral realism. Anthropomorphism. Perceived agency of interactant Level of embodiment Synchronicity Inclusion of imagery Inclusion of imagery (dynamic) Inclusion of voice Inclusion of haptic feedback Others
--	---

Initial studies were centered on immersive qualities, but the recent literature also began to address contextual and individual factors, given the subjectivity of the social presence concept [29]. Nevertheless, studies on technological predictors dominate the literature, enlarging the *immersive qualities* class.

The categorization of predictors that affect social presence or co-presence, based on related works, point to (1) immersive qualities, (2) contextual and social properties, and (3) individual traits (see Table 2).

Table 2. Categorization of predictors.

	Immersive qualities	Modality, visual representation, interactivity, haptic feedback, audio quality, depth cues, video and display.
Co-presence factors	Contextual and social properties	Personality / traits of virtual human, agency, physical proximity, task type, social cues, identity cues.
	Individual traits	Demographic variables, psychological traits.

2.1. Immersive Qualities

2.1.1. Modality

The first studies on social presence analyzed the effect of the modality on the levels of presence achieved, given that the immersion degree varies. These studies on general modality identified technological features with an impact on the social presence (e.g., visual representation, interactivity, depth cues, audio quality, and display). However, medium communication comprises multiple features, and it is a challenge to discriminate the contribution of each affordance. In [42], the *media richness theory* refers to varying the technological qualities of the medium affording distinct levels of social presence. General modality was also identified in [43] as a predictor of telepresence while analyzing the influence of immersion. Initial studies analyzed the influence of modality on social presence by comparing (1) Face-to-face (FtF) real interactions with computer-mediated communication (CMC), (2) text-based CMC with mediums supporting visual and audio modalities, and (3) immersive virtual environments with non-immersive virtual environments.

Face-to-face (FtF) interaction is considered the ground truth for social presence [44], and several works compare face-to-face (FtF) interaction and CMC, evaluating the capability of these mediated communications to elicit social presence. In general, these studies reveal that the sense of social presence is higher in an FtF interaction when compared with CMC conversation. Cortese et al. [45] designed a task in which participants had to discuss a news article for 20 min, either with FtF interaction or through computer-mediated communication (CMC) (chat). Communication apprehension was one of the psychological factors to be assessed (i.e., “the level of anxiety or fear associated with either real or anticipated communication with another person or persons”). They found that the CMC participants experienced a lower level of social presence. Researchers assessing the sociability of a partner and the level of co-location again found higher social presence levels for FtF interaction.

In studies involving decision-making scenarios [44] and in online learning achievement [46], the results privileged FtF interaction. One study [47] involving a series of online seminars for 2 months (the same teacher teaching the same contents online and via in-person, FtF interaction) reported no differences in the levels of social presence between both forms of interactions. One justification might be related to the fact that students had enough time to adapt their communication skills to an online learning platform, the evolution of e-learning technologies, and the fact that the students felt more comfortable not moving to a classroom for 2 months. However, this study did not address the characterization of the subjects’ ages in their concluding remarks, which could reveal a tendency.

Video and audio modalities guarantee higher degrees of social presence when compared with text-based CMC. However, this difference is not so clear when comparing video-audio modality against audio-only modality. Studies have shown that the introduction of video modality increases the social presence feeling if participants are required to perform visual tasks [48,49]. In studies that compared video-audio vs. audio modalities when involving tasks that do not require visual feedback, such as interview tasks or decision-making tasks, the researchers did not report a significant difference in the social presence [50].

These studies suggest that increasing the quality of an immersive component, such as a video feature, may not be proportional to the social presence felt. There seems to exist a threshold from which further enhancements of a given modality may not produce an additional contribution.

Table 3 summarizes the relevant aforementioned and next predictor's references, their significant conclusions, and insights on statistics comparisons.

2.1.2. Visual Representation

In communication, the visual representation of interactants is a feature with an impact on social presence. Research has explored to what extent a representation form of the partner can contribute to the sense of social presence. Typically, studies manipulate (1) inclusion or no inclusion of visual representation and (2) the level of realism of the visual representation. The authors of [51] defined *realism* as the extent to which a digital human representation behaves and appears like a real human. The overall concept is referred to as being based on three components: photographic, anthropomorphic, and behavior's communicative realism.

The photographic component assesses the human-like visual appearance in a representation. Most studies report that the existence of a visual representation of the partner leads to higher social presence levels. In [52], participants who spoke with their partners through an avatar while shopping in a virtual mall felt a higher social presence in comparison with those who talked without seeing any partner representation. In [53], in an online support-seeker activity, users reported a higher sense of social presence when a profile picture of the counselor was present, as opposed to not having a picture. The users also demonstrated a higher willingness to answer questions when the pictures were available.

Anthropomorphism contributes to communicative realism because physical attributes such as the mouth, eyes, arms, and legs are involved in speech to generate facial expressions, gestures, and movements. It addresses the level of interpretation of what is not human or personal in terms of human or personal characteristics. Apart from behavioral realism, this manipulation focuses on the degree to which interactants are presented as human-like on the visual and auditory plane. For instance, users would interact via video as opposed to users who would interact via motion capture-controlled cartoons or by means of anthropomorphic agents or avatars vs. animated forms or emoji.

Communicative realism addresses the degree to which a digital representation of the partner presents physical and social human-like behavior (e.g., breathing, natural blinking, and posture changes). Behavior realism studies manipulate the presence or absence of nonverbal behavior (e.g., animation) or the degree to which the nonverbal behavior of a virtual human resembles a real human (e.g., with or without eye gazing). The effect of communicative (behavioral) realism is more evident when the behavior of an agent or avatar reflects the awareness of the partner's presence (e.g., nodding at the right time, mutual gazing, or blushing). Von der Pütten et al. [54] found that nodding the head of a computer-controlled agent during an interaction contributed to a higher degree of social presence in opposition to no nodding. In another study [55], the participants of an interaction with a virtual agent reported a higher level of social presence when they saw the agent blushing, a consequence of some mistake during a presentation. Study 1 in [50] found that the participants felt higher levels of social presence when the partner's representation (e.g., avatar) was able to maintain a mutual eye gaze in opposition to the absence of eye gazing. However, Study 2 in [50] realized that maintaining a mutual eye gaze for too excessive a time for video and avatars decreased social presence (i.e., an unnatural behavior). Bent's studies carefully tracked participants' nonverbal behavior using the head's orientation and position sensors, eye gaze trackers, a breath-monitoring chest belt, and data glove-based finger movement trackers. In the avatar's mediation condition, the tracked data were used to animate the avatars in real time (head and body movements, eye movements, and hand and finger movements). Their findings showed a similar activity in terms of visual attention and nonverbal activity either in video or in avatar conditions, contributing both positively and quite similarly to eliciting social presence. This suggests

that avatars can be used as a tool to assess social presence, with the advantage that they enable behavior cue segmentation. Another interesting fact is that users tend to direct their heads to their partner's image but their gazes towards the workspace. A justification for this behavior, even knowing that it is a computer representation of the partners, might be related to the human's unconscious social etiquette, that being to keep the face directed to the interaction partner.

Studies show that behavioral realism tends to contribute consistently and positively to social presence [28,30], while photographic and anthropomorphic realism presents varying effects (positive [56], neutral [50,57], or even negative contributions [17]). The justification for these discrepancies might be related to several facts reported in the literature: (1) photographic realism is not the main contributor to social presence (i.e., the appearance of the visual representation has a secondary role in comparison with cues from social behavior) [58], (2) manipulations of small features of the visual representation may not be reflected in social presence questionnaires, and (3) the degrees of behavioral realism differ from study to study, making a quantitative comparison with photographic realism vary [59].

In [60,61], the researchers evaluated the effect of visual and behavioral realism on the perceived quality of communication using avatars. They found a positive effect on social presence when there was consistency between both realism components [60]; that is, although visual representation does not represent a major contribution, the participants felt a greater social presence when the avatar, demonstrating more realistic behaviors (e.g., inferred vs. random eye gaze) which were complemented by a higher level of photographic realism (avatar with a human-like face instead of a dummy face). Bailenson [61] pointed out the consistency between photographic and behavioral realism as a positive social presence predictor.

In [62], the effect of the 3D avatar type (character-like vs. realistically reconstructed) on users' trust and co-presence in a mixed reality-based collaborative teleconference was explored. Visual representation based on realistically reconstructed avatars has been shown to elicit the user's sense of co-presence.

In [30], virtual humans that demonstrated higher responsiveness to events (behavioral realism) contributed positively to the co-presence in mixed reality environments. Experimental conditions involved remote collaboration, in-person collaboration, and communication interactions via mixed reality, augmented reality, virtual reality, video chat, text apps, and virtual assistants.

2.1.3. Interactivity

The definitions of social telepresence, in the context of robot-mediated interaction, rely on the capability to put forward the robot operator's presence to the local person (bystander). Additionally, the extent to which this person is aware that he or she is talking to or interacting with a human being has an impact on social presence [15] and co-presence [63]. Studies on this subject try to understand the effect of the interactivity of the agent on social presence. Such interactivity may refer to a computer agent, a person's avatar, or a telepresence robot, but the focus of this analysis is on the use of a telepresence robot for conversation mediation. Thus, the level of social presence depends on the fidelity of the medium to support the interactivity that characterizes persons' conversations. It includes visual and audio cues, nonvisual sensing (i.e., directional sound and haptics like force feedback and touch) and environmental interactivity (e.g., response rate to user input, reciprocity of the interaction capability between the remote user and local user, and clarity of causal relationships between remote user actions and local user reactions).

2.1.4. Haptic Feedback

To improve the sense of reality, it is important to provide some type of physical feedback to the operator or bystander. Useful contributions include providing tactile cues to let the user recognize the surface texture and materials or support kinesthetic feedback

to help the user experience the weight of a virtual object. These kinesthetic and tactile sensations enable haptic perception.

Haptic feedback is a challenge. However, it may improve the degree of presence considerably. Considerable progress has been made in the field of visual and auditory displays, but haptic feedback is in its early stages, gaining much attention nowadays [64–66]. Touch contact plays an important role in human interactions. From an early age, babies explore their surroundings with their hands and feel physical contact with parents holding them, and at older ages, handshakes, kisses, and embraces trigger emotions and strengthen relationships. Nevertheless, the physical contact of a robot with a person raises safety issues, and that may justify why haptic feedback is not so prevalent.

Touching a robot's part (e.g., hand or body) or sensing that component pulling our hand, operated by a remote party, can improve co-presence. Remote hand-shaking has been explored [67], and examples include the Nao robot hand-shaking the bystander while the robot's operator uses a low-cost haptic device (WiiMote) to feel it [68]. In [69], a robot hand was attached under a videoconferencing terminal's display, and their evaluation demonstrated that mutual touch enhances the feeling of being close. However, the partner's action should not appear in the video. Gregory Welch et al. [70] developed a tactile telepresence system prototype that enables a remote visitor to convey touching patterns on the forehead of an isolated patient through a tablet touch video interface. Regarding human-robot first encounters [5] and greetings, the authors of [71] used Kendon's model [72] to develop an interaction that included six phases (initiation of approach, distance salutation, head dip, approach, final approach, and close salutation). Human tracked gestures were the inputs for a decision module (based on the hidden Markov model and behavior tree [73]) that initiated a specific phase at the right moment.

Teleexistence surrogate anthropomorphic robot (TELESAR) VI can mimic the user's movements and gestures from a mechanically unconstrained full-body master cockpit and provide haptic feedback to the operator [74]. The 10 fingers of the teleoperated robot are equipped with vibration, force, and temperature sensors that can realistically deliver these components of haptic information. Operators can shake the hand of another person through the robot and feel it.

2.1.5. Depth Cues (Stereoscopy and Motion Parallax)

Considering an interaction between two persons through a teleoperated robot, the depth cues become more important for the remote user, since the local user is with the robot, and has natural depth cues. On the other side, if the remote user can perceive the local user in a 3D space, it improves the scene's realism and the co-presence. The use of 3D displays or head-mounted displays (HMDs) by the remote user are common approaches to delivery depth cues. However, this requires 3D sensors in the robot's side (e.g., stereo cameras and RGBD sensors). Additionally, with the inclusion of an autostereoscopic or 3D display in the robot to present the remote user to the local user, it is possible to enhance the closeness [75–77].

2.1.6. Audio Quality

As mentioned earlier, audio modalities guarantee higher levels of social presence. The audio channel should provide bidirectional communication between the remote and local users to exchange messages. Recognition of a person's voice plays an important role in person identification, contributing to the sense of co-presence [78]. Voice transmission is expected to be fluid without cuts or delays. Telepresence robots quite often make use of an array of microphones to acquire spatial sound, enabling the remote user to identify the direction of the sound source [79] or simply detect the movements of the local user.

2.1.7. Video and Display

The sense of being telepresent is also determined by the fidelity and capability of the medium to present the remote environment, including the visualization of persons (face expressions, gestures, postural behaviors, etc.). To this end, there are mediation technology

requirements that include visual display parameters (e.g., latency, frame rate, field of view (FOV), point of view (egocentric vs exocentric), image resolution, color quality, and image clarity) and environment presentation consistency across displays [80]. Display type comparisons reveal a positive effect on co-presence using immersive 3D displays in nonverbal interactions [81]. For example, the Willow Garage Texai robot rely on the principle “reciprocity of vision (if I see you, you must see me)”, while Excite robot designers defend that “The visitor’s [user’s] environment should be immersive so that the user would have a first-person experience of the destination [remote environment] including full sensory stimulation focusing on immersive vision, audio, and haptics” [3].

2.2. Contextual and Social Properties

Early studies on predictors of social presence focused on immersive qualities; however, the research began to address contextual and individual properties. Given the subjectivity associated to social presence experience, and aside from the physical distance and the medium’s technological qualities, analyses started to consider a psychological distance between the interactants [29,82,83]. These include factors such as Personality/Traits of Virtual Human, Agency [17], Physical Proximity [57,59], Task Type [48,84,85], Social Cues [27,48,86], or Identity Cues [27].

In [30], a contextual responsiveness predictor is explored that assesses the capability of a virtual human (VH) to detect and respond to events and cues that happen in the shared space of the VH and the user (e.g., to a broom that falls in the user’s physical environment or that falls into the virtual VH space). They showed that when the VH detects and directs gazing at the event or orients itself in that direction, the user presents higher levels of co-presence. Studies suggest that users’ perception of the physical space affects their co-presence in mixed reality. Ignoring events in the background, such as objects moving or a person walking [87], or the inability to shift attention to an external event does not contribute to co-presence. In [62], the robot that plays a game with the user uses a “cheat” function to trick the user, which affects the user’s trust, contributing positively to co-presence.

In [27], users reported a higher level of social presence when communicating simultaneously with several remote interlocutors through a telepresence robot than with a single remote person. In [27], a second study showed that users felt the presence of the remote interlocutor more when the telepresence robot had a low identity than a higher identity (e.g., robot’s head LCD with or without a face drawing).

2.3. Individual Traits

Gender and age: social studies showed that female subjects tend to experience higher degrees of social presence when compared with males [57,88], but age is not a relevant factor [89].

Attractiveness: in [59], a human’s avatar that looks more attractive in a virtual mirror raised the person’s level of self-confidence in the next encounters with other person’s avatars and eventually in the real world (distances between avatars are reduced (proxemics)). Such findings provide traits for telepresence and co-presence robot design.

Height: in [59], a human’s avatar that looked taller than its interlocutor tended to make that person more persuasive in new interactions with others.

Psychological traits: a person with a higher immersive tendency showed higher degrees of social presence [89]. Additionally, people more prone to human social interactions reported higher levels of social presence in experiments involving social robots [90]. In [45], persons low in communication apprehension (CA) experienced higher levels of social presence than those high in CA. Less sociable people tended to show lower scores on social presence assessments.

Table 3 summarizes the relevant aforementioned and next predictor’s references, their significant conclusions, and insights on statistics comparisons (e.g., N = number of subjects, μ = mean, σ = standard deviation; subscripts refer to the condition, where superscript $+$ =

significant condition, df = degree of freedom, F = ANOVA statistic F , p = p -value, η^2 = eta squared, χ^2 = chi square, β = standardized path coefficient, and r = correlation coefficient).

Table 3. Co-presence studies.

Predictor Category	Predictor	Evaluation Process	Study	Quantitative Comparison (Statistics)
Immersion	Modality	FtF vs. CMC (Net-Metting teleconference)	[44]	$N = 70$ (38 pairs), $\mu_{FtF} = 34.6$, $\mu_{CMC} = 32.1$, $F = 40.2$, $df = 1$, $p = 0.00$
Immersion	Modality	FtF vs. CMC (chat)	[45]	$N = 152$, $\chi^2(8, N = 152) = 6.267$, $p = 0.617$; ($\beta = -0.948$, $p < 0.001$)
Immersion	Modality	FtF vs. CMC	[46]	$N = 257$, $\mu_{FtF} = 3.63$, $\sigma_{FtF} = 0.62$; $\mu_{CMC} = 3.48$, $\sigma_{CMC} = 0.57$; $t(255) = 2.077$, $p = 0.0039$
Immersion	Modality	FtF vs. CMC (on-line teaching and learning)	[47]	$N = 50$, $\mu_{FtF} = 38.9$, $\sigma_{FtF} = 1.2$; $\mu_{CMC} = 36.91$, $\sigma_{CMC} = 1.36$; $F(1, 48) = 1.194$, $p = 0.28$
Immersion	Modality	Audio vs. Audio + Video	[48]	$N = 34$ (17 pairs), male: $\mu_{audio} \approx 53.5$, $\mu_{audio+video} \approx 71.75$; $F(1, 18) = 9.9$, $p = 0.04$
Immersion	Modality	Text vs. Audio vs. Audio + Video vs. Audio + Avatar	[50]	$N = 150$, Factor scores: $\mu_{text} = -0.48$, $\mu_{audio} = 0.26$, $\mu_{audio+video} = 0.22$, $\mu_{audio+LFavatar} = 0.09$; $\mu_{audio+HFavatar} = 0.10$; $F(4, 137) = 2.59$, $p = 0.04$, $\eta_p^2 = 0.09$
Immersion	Visual Representation	Photographic Realism (Low- vs. High-Fidelity Avatar)	[50]	$N = 150$, Factor scores: $\mu_{audio+LFavatar} = 0.09$; $\mu_{audio+HFavatar} = 0.10$; $F(4, 137) = 2.59$, $p = 0.04$, $\eta_p^2 = 0.09$
Immersion	Visual Representation	Photographic Realism	[52]	$N = 80$, embodiment index: $voice = 1.68$, $voice + avatar = 5.2$, $df = 4$, $p < 0.01$, $\mu_{embodiment} = 3.41$, $\sigma_{embodiment} = 1.94$; $\mu_{copresence} = 5.27$, $\sigma_{copresence} = 1.44$;
Immersion	Visual Representation	Photographic Realism	[57]	$N = 50$, $\mu_{flat_shaded_face} \approx \mu_{photographic_texture_face}$
Immersion	Visual Representation	Anthropomorphic	[17]	$N = 134$, copresence index: $low_{anthropomorphic_image}^+$, $more_{anthropomorphic_image}$, no_{image} , $R = 0.18$, $F = 4.23$, $p = 0.04$
Immersion	Visual Representation	Anthropomorphic, Behavioral Realism	[28]	Definitions and it uses, digital representations
Immersion	Visual representation	Behavioral realism (mutual gaze)	[57]	$N = 50$, women's social presence score: $\mu_{no_mutual_gaze} = -13.25$, $\sigma_{no_mutual_gaze} = 18.58$; $\mu_{high_mutual_gaze} = 2.5$, $\sigma_{high_mutual_gaze} = 15.55$; 5 conditions ($r = 0.30$, $p < 0.03$)
Immersion	Visual Representation	Consistency between Visual and Behavioral Realism	[60]	$N = 48$, $low_realism$: $\mu_{random_gaze} = 1.2$, $\sigma_{random_gaze} = 0.2$; $\mu_{inferred_gaze} = 0.7$, $\sigma_{inferred_gaze} = 0.2$; $high_realism$: $\mu_{random_gaze} = 0.3$, $\sigma_{random_gaze} = 0.1$; $\mu_{inferred_gaze} = 1.1$, $\sigma_{inferred_gaze} = 0.3$;
Immersion	Visual Representation	Consistency between Visual and Behavioral Realism	[61]	$N = 146$, copresence: behavioral realism ⁺ , $F(3, 133) = 2.72$, $p < 0.05$, $\eta^2 = 0.06$; visual representation ⁺ $F(6, 133) = 2.18$, $p < 0.05$, $\eta^2 = 0.09$
Immersion	Visual Representation	Avatar Behavioral Realism to Events	[30]	$N = 65$, copresence: $\mu_{responsive}^+ = 4.31$, $\sigma_{responsive} = 0.11$; $\mu_{nonresponsive} = 3.96$, $\sigma_{nonresponsive} = 0.12$; σ ; $F(1, 63) = 5.06$, $p = 0.02$

Table 3. Cont.

Predictor Category	Predictor	Evaluation Process	Study	Quantitative Comparison (Statistics)
Immersion	Visual Representation	HMD vs. Desktop	[80]	$N = 21$, presence Q5: $\mu_{HMD} = 5.28$, $\sigma_{HMD} = 1.58$; $\mu_{Desktop} = 3.42$, $\sigma_{Desktop} = 1.77$; $F(1, 20) = 26.54$, $p < 0.0001$
Immersion	Visual Representation	HMD vs. Desktop	[91]	$N = 26$, presence Q5: $\mu_{HMD} = 5.88$, $\sigma_{HMD} = 0.52$; $\mu_{Desktop} = 2.48$, $\sigma_{Desktop} = 1.75$; $F(4, 95) = 32.19$, $p < 0.0001$
Immersion	Visual Representation	2D vs. 3D vs. Verbal vs. Nonverbal	[81]	$N = 40$, copresence: $3D_{nonverbal}^+$, $3D_{verbal}$, $t(16.35) = 7.48$, $p < 0.05$; $2D_{nonverbal}$, $2D_{verbal}^+$, $t(17.967) = -8.05$, $p < 0.05$
Immersion	Interactivity	Whole Body Interaction	[92]	$N = 13$, embodiment: <i>immersive+body intention-based robot control</i> ⁺ , $F(3, 44) = 19.11$, $p < 0.0001$; $N = 24$, Embodiment score: Haptic feedback ⁺ = 49.8, $F(1, 23) = 29.67$, $p < 0.0001$; Realism score: Haptic feedback = 33.0, $F(1, 23) = 22.97$, $p < 0.0001$;
Immersion	Haptic Feedback	Present vs. Absent	[93]	$N = 144$, copresence: $\mu_{stereo}^+ = 3.85$, $\sigma_{stereo} = 1.34$; $\mu_{mono} = 3.25$, $\sigma_{mono} = 1.48$; $F(1, 140) = 6.97$, $p < 0.01$, $\eta_p^2 = 0.05$
Immersion	Depth Cues	Stereoscopy (stereo vs. mono)	[94]	$N = 82$, presence: <i>binaural</i> ⁺ , <i>mono</i> ⁻ , $t(4, N = 79) = 10.7$, $p = 0.031$
Immersion	Audio Quality	Binaural vs. Stereophonic vs. Monophonic	[78]	active perception (visuo-auditory, vestibular emulation, Bayesian models), $f = 6-10$ Hz.
Immersion	Audio Quality	Attention, Binaural	[95]	3D capture, maintain face-directed gaze through robot positioning, $f = 2.12$ Hz
Immersion	Display	Face-to-Face Point of View	[75]	$N = 144$, copresence: $\mu_{humansize_display}^+ = 3.94$, $\sigma_{humansize_display} = 1.46$; $\mu_{smallsize_display} = 3.17$, $\sigma_{smallsize_display} = 1.30$; $F(1, 140) = 11.41$, $p < 0.001$, $\eta_p^2 = 0.08$
Immersion	Display	Three 55-inch Screens vs. One 55-inch Screen	[94]	3D capture, 3D display, eye/head tracking, frame rates: 34, 48, 74 Hz
Immersion	Display	Autostereoscopic Telepresence	[96]	$N = 144$, computer voice with a personality (extrovert/introvert) similar to human interlocutor, $F(1, 67) = 11.13$, $p < 0.001$, $\eta_p^2 = 0.14$; <i>voice</i> ⁺ _{extrovert} , <i>voice</i> _{introvert} , $F(1, 71) = 17.91$, $p < 0.001$, $\eta_p^2 = 0.20$
Context	Personality or Traits of Virtual Human	Personality manifested by voice and match between content	[97]	$N = 134$, copresence index: $\frac{agency_{human_human_interaction}}{agency_{human_computer_interaction}}$, $R \approx 0.03$, $F = 0.15$, $p = 0.7$;
Context	Agency	Avatar vs. Agent	[17]	Illusory self-identification
Context	Agency	Conscious experience of being someone	[24]	$N = 90$, $\frac{agency_{human_human_interaction}^+}{agency_{human_computer_interaction}}$, $F(1, 90) = 10.870$, $p = 0.001$, $\eta^2 = 0.112$
Context	Agency	Avatar vs. Agent	[82]	$N = 134$, male social presence: $std_{path_{estlocation_accessibility_cues}^+} = 0.21$, $std_{path_{est}^+}_{richer_medium} = 0.06$
Context	Physical Proximity	Close vs. Distant (spatial proximity)	[98]	

Table 3. Cont.

Predictor Category	Predictor	Evaluation Process	Study	Quantitative Comparison (Statistics)
Context	Task Type	Caregiver: Human vs. Robot	[99]	$N = 60$, social presence: $\mu_{robot_as_caregiver}^+ = 5.56$, $\sigma_{robot_as_caregiver} = 1.04$, $\mu_{human_as_caregiver} = 4.20$, $\sigma_{human_as_caregiver} = 0.83$;
Context	Social Cues	Online Buddy: Present vs. Absent	[85]	
Context	Identity Cues	Telepresence Robot, Identity Cues: High vs. Low	[27]	
Individual	Demographic Variables	Gender: Female vs Male	[48]	
Individual	Psychological Traits	Communication Apprehension	[45]	
Individual	Psychological Traits	Belonging Feeling	[99]	

3. From Telepresence to Co-Presence Design

Presently, the market [100] offers full solutions for mobile robotic telepresence (MRP) systems [5–7,9] (see Table 4), and the research presents telepresence robot solutions such as the ones listed in Table 5. There are also unmovable robotic telepresence (RP) systems, which are listed in Table 6. These robotic telepresence systems are depicted in Figures 3 and 4.

Table 4. Mobile robotic telepresence (MRP) systems: full market solutions.

References	Robotic Telepresence Systems	Application Area	Expression or Manipulation	Navigation Features	Cost
[101]	Giraff	Eldery	Head tilt (screen display/camera)	No	USD 11,900.00
[102]	Double 2, 3	Office, education, hospital	Motorized height	Accelerometer and gyroscope for balance, kickstands when static	USD 2749.00
[103]	PadBot 2	Office, education, hospital	Tilt head (screen)	Obstacle detection, collision avoidance, anti-falling system	USD 1297.00
[103]	PadBot U1—v2	Office, education, hospital	Tilt head (screen)	Collision-prevention sensors. Edge detection and anti-falling sensors.	USD 797.00
[103]	PadBot T1	Office, home	No	Collision prevention, cliff sensor	USD 185.00
[104]	Beam Pro	Corporate, manufacturing, medical	No	Crash avoidance, assisted driving	USD 14,945.00
[105]	Ava 500	Healthcare, office	No	2D or 3D imaging, sonars and lasers for autonomous navigation, omnidirectional navigation, scheduling capabilities, cliff sensor	USD 32,000.00
[106]	Ohmni Super-Cam	Office, home	No	Includes downward-facing camera for full visibility	USD 2195.00
[107]	VGo	Office, education	Tiltable head	Crash avoidance, notification of obstacles locations, cliff sensor	USD 3995.00
[108]	TeleMe 2	Office, education	Laser pointer option	Crash avoidance: infrared sensors detect obstacles and will automatically reduce robot’s speed	USD 3995.00
[109]	RP-Vita	Healthcare, FDA clearance	Active patient monitoring	Obstacle avoidance, omnidirectional and autonomous navigation	USD 80,000.00
[110]	Teleporter	Office, factory, hospitals	Laser pointer, secondary webcam	Crash avoidance: infrared, 3D, or sonar sensors	USD 14,995.00

Table 5. Mobile robotic telepresence (MRP) systems: research-oriented solutions.

References	Robotic Telepresence Systems	Application Area	Expression or Manipulation	Navigation Features	Cost
[22]	PRoP	Research	Laserpointer, 2 DOF hand and arm	-	-
[111]	FURo-i	Home	No	Bumpers	USD 1800.00
[112]	MeBot	Research	3 DOF neck for screen and 3 DOF arms	Collision prevention, cliff sensor	-
[113]	Origibot	Research	Tiltable head for screen, 1 DOF arm (180°), 2 DOF gripper	No	Low cost
[114]	Nao	Research	Humanoid, 25 DOF, tiltable head, arms, legs, 4 directional microphones and speakers, 2 cameras	No	-
[114]	Pepper	Research	DOF (head: 2, shoulder: 2, elbow: 2, wrist: 1, hands (5 fingers): 1, waist: 2, knees: 1, base: 3 wheels), 2D and 3D cameras and sonars	Autonomous navigation, bumpers	USD 30,000.00
[115,116]	GrowMeUp	Eldery research	Robot expression, directional microphones and speakers, 2D and 3D cameras, sonars, touchscreen	Obstacle avoidance, autonomous navigation, service, expression, behaviors	-

Table 6. Unmovable robotic telepresence (RP) systems.

References	Robotic Telepresence System	Application Area	Expression or Manipulation	Cost
[117]	Kubi	Office, education	Pan 300°, tilt 900°(screen)	USD 675.00
[118]	TableTop TeleMe	Office, education	Pan 360°, tilt head (screen)	USD 3995.00
[119]	SelfieBot	Office, education	Pan 180°, tilt 180° head (screen)	USD 195.00
[120]	Meeting Owl Pro	Office	Static 360° camera (1080p)	USD 999
[121]	Robovie mR2	Research	Expression, arms, gestures, eye blinking (cameras), 18 joints (3 in each eye), 18 servo motors	-

3.1. Co-Presence Design

Mechanisms that contribute to enhancing co-presence in telepresence robots should consider the robot-side systems and the remote user (robot's operator) side solutions. Robot-side interfaces support interactions between the robot and the local user (bystander) and between the robot and the remote user (operator). Human-robot interfaces can be classified into sight, hearing, touch, and body-sensing technologies. Technological advances include robust robot sensory (vision, face and expression recognition, object recognition, activity identification, pressure, touch, temperature, speech understanding, sound localization, etc.), acting (mobility, proxemics, gestures, gazing, facial expressions, speech synthesis, etc.), reasoning (localization, planning, context awareness, grasping, etc.), and appearance (familiar, unfamiliar, human-like, and mechanical) [3,122,123].



Figure 3. Mobile robotic telepresence (MRP) systems: (a) PRoP, (b) Giraff, (c) Double 2, 3, (d) PadBot 2, (e) PadBot 3, (f) PadBot T1, (g) Beam Pro, (h) Ava 500, (i) Ohmni SuperCam, (j) VGo, (k) TeleMe, (l) RP-Vita, (m) Teleporter, (n) FURo-i, (o) MeBot, (p) Oritbot2, (q) Nao, (r) Pepper, and (s) GrowMeUp.



Figure 4. Unmovable robotic telepresence (RP) systems: (a) Kubi, (b) TableTop TeleMe, (c) SelfieBot, (d) Meeting Owl Pro, and (e) Robovie mR2.

3.1.1. Sensing

Robotic sensing technologies are becoming more efficient, lighter, and cheaper. Early human–robot interfaces used to integrate few sensors and relied mainly on video and audio data and low-resolution proximity sensors (e.g., sonars). Current robots can be equipped with 3D or 2D cameras (e.g., low-cost RGB-D cameras), pressure sensors, touch sensors, directional sound sensors (arrays), high-precision proximity sensors (e.g., range laser finder (lidar)), and robot pose and position sensors (e.g., gyroscopes, accelerometers, and GPS). The fusion of these sensors combined with high-accuracy robots, person localization algorithms (e.g., simultaneous localization and mapping (SLAM) or Open Pose), and deep learning approaches, have improved robot operations in an environment, enhancing HRI between operators and bystanders. Valuable information can be extract due to advances in sensor technologies and software, such as sound locations [95,124,125], speech segregation [126,127] and recognition [128,129], attention [79], gesture recognition [130,131], human action analysis [132,133], human intentions [134,135], object recognition [136], and scene understanding [137,138].

3.1.2. Action Capabilities

Advances in robot software and hardware, lighter and stronger materials, component miniaturization, and lighter and more powerful batteries have broadened robots' capabilities. Robot mobility has improved significantly, enabling robust navigation in an unstructured environment and in rough terrain [139,140], and they can climb stairs, walk fast, and run, such as the Boston Dynamics ATLAS robot [141] or Honda ASIMO Robot [142]. Advances in humanoid mobility and equilibrium are remarkable, including compliant interactions and variable speed [143]. Having arms, hands, and fingers with more degrees of freedom (DOFs) enabled new types of interactions such as high-fidelity gestures, grasping objects smoothly [144], or even open doors and pass through them [141,145]. Whole-body expressive movements [146], facial features to support expression synthesis [147], and speech synthesis technologies are enabling better HRIs.

3.1.3. Reasoning

Robots are designed to perform several tasks, but task execution is not always perfect (e.g., motion constraints, impaired sensory, control, and communication delays). Thus, advancements in software reasoning processes have been developed to supervise tasks, aiming not for perfect execution but optimum performance. Namely, notable advances have been made in localization and mapping [148,149] and in grasping [144]. In [150], the authors explored approaches for a telepresence robot to detect and position itself with a group of people for social interactions (maintaining an egocentric perspective). The inclusion of these autonomous algorithms can help operators and bystanders in their interactions with the robot, simplifying the control, reducing the effort, and improving the intuitiveness.

3.1.4. Appearance

The acceptability of the robots enrolled in a human assistive task also depends on their appearance. Designers have created robots with human-like appearances [151]. The Geminoid robot has an incredibly realistic head and facial features [152]. This approach enables more effective communication through facial expressions and natural gestures. Additionally, given the human-like robot morphology, it is simpler to map the human gestures and movements in the robot. The search for realism, however, suggests some warnings regarding Mori's "uncanny valley" [153,154]; that is, if a robot or agent is an imperfect replica of a human being, people may feel defrauded in their expectations regarding the affinity as a pair, triggering strange, familiar feelings of unease and revulsion.

3.1.5. Managing Robot Autonomy in Telepresence Systems

Advances in robot autonomy do not eliminate the role of human operators. Human skills remain crucial in an unstructured environment or when dealing with unpredicted events. The integration of autonomous mechanisms aims for process simplification, and it changes the nature of human–robot interaction (HRI). However, there are cases where the complexity increases [155] (e.g., 2019 Boeing 737 Max autopilot problems with deadly consequences). The availability of automated behaviors for telepresence or in humanoid robots may lead people to use them indiscriminately, diverting attention from the interaction essentials. Nevertheless, autonomous mechanisms aim to reduce users' mental workload, performing increasingly complex tasks and now being part of our daily lives (e.g., self-driving cars, autonomous vacuum cleaners, and chatbots). The literature refers to methods to integrate autonomy mechanisms in telerobotics [156], and they can be classified into direct control, supervisory control, shared control, traded control, collaborative control, and cooperative control [157,158].

Direct control: The robot has no autonomy. An operator controls all the robot's functions manually. Mirroring is a type of direct control in which the robot replicates the human's movements and expressions.

Supervisory control: The robot is programmed intermittently according to the continuous information received from the robot. The human and the robot integrate a closed control loop focused on task performance [159].

Shared control: The human operator controls the robot continuously. However, those commands may be strictly followed by the robot (similar to direct control) or be modified by the robot's system to improve performance or run safely.

Collaborative control: The operator and the robot work together as peers to determine the robot's behavior. In Fong's work [160], there is an explicit semantic dialogue between humans and robots to mediate the sharing of control.

Traded control: The human operator starts a behavior or task that is autonomously performed by the robot. At any time, the operator can stop that behavior or task and start a new one.

Cooperative control: The behavior of a single robot results from the controlled cooperation of several operators using any of the aforementioned methodologies.

In the shared control method, the operator provides continuous commands to the robot, aiming for high-level behavior from the robot. However, the robot may change those inputs to reach the perceived system goals [161]. The method assumes that the operator knows how to direct the robot's high-level behaviors but may not be sufficiently skilled to express the right commands due to a lack of situation awareness, embodiment, telepresence, or lack of robot motor accuracy and sensor information. Typically, the shared control method includes "safeguard" mechanisms, in which the operators' command actions are overwritten if they violate the robot's safety rules, such as collision with a wall or person or losing balance [162]. The software of HPR-1s or HRP-5P humanoid robots was developed to discard commands that could make the robot lose balance, limiting joint angles [163–165]. In Almeida et al. [92], given the robot's height, the wheel's initial acceleration provided by the operator had to be supervised by the robot to avoid its falling down. In Crandall and Goodrich's works, the robot's desired trajectory was provided through a joystick as the intended general direction and not as low-level position commands [166].

In the traded control method, a task or subtask is performed autonomously by the robot, but it is initiated by the operator and may be stopped at any time. The method is useful for simultaneously controlling multiple robot's appendages, such as in teleoperation of humanoids [167–169]. The Geminoid HI-1 robot [169] relies on a traded control known as state-based control, in which the operator selects the state from a library of states. It includes five conscious behaviors, namely right looking, left looking, listening, speaking, and being idle. For each state, the robot assumes autonomous behaviors (i.e., motion files), avoiding an explicit operator's control of 50 robot actuators. The integration of multiple semi-autonomous mechanisms is essential while controlling the eyes, head, torso, arms, hands, and fingers simultaneously in a humanoid robot. Quite often, operators need to control low-level robot behaviors and additionally focus their attention on high-level tasks, such as (1) robot navigation, (2) obstacle avoidance, (3) triggering robot's unconscious and conscious behaviors [169], (4) object and scene understanding [137,170,171], (5) mission planning, or (6) people's interaction. Osawa et al. [172] evaluated the automation of involuntary and voluntary movements using a teleoperated telepresence robot (robovie-mR2). The implemented behavior generation architecture (bi-layered architecture [173]) enabled the combination of autonomous movements and manual movements controlled by a remote operator. The results showed that bystander users evaluated both the involuntary and voluntary movements positively but also revealed that from the remote operator's point of view, the automation of voluntary movements should require additional care (agency issue conflicts).

3.1.6. Time Delay Mitigation

The dynamic nature of the communication medium has an impact on the complexity of teleoperated systems. Time delay, jitter, distance, bandwidth constraints, packet loss, or blackout in internet-based solutions can delay or distort interactions. This can affect the synchronicity (rate of message exchange between operator and bystander), compro-

mismanaging the social presence [33] (e.g., degradations in audio and video streams, control streams, or haptic feedback). Traditional methods to mitigate time delay in telerobotics involved user interface design and control theory-based models (e.g., supervisory control or passivity-based teleoperation) and evolved into predictive displays and control [159]. Recent solutions for time delay issues use *time series prediction* methods to predict the time delay, robot movements, and user intentions (e.g., user's gaze prediction [174]). These new adaptive-based control methods make use of nonlinear statistical models and neural network (NN) or machine learning (ML) techniques.

Ferrell and Sheridan [175] determined that a time delay affects human operators' performance while teleoperating manipulators. They realized that the person within the control loop of teleoperated systems under time delays used to adopt a *move-and-wait* strategy to accomplish certain tasks. To address this problem, they proposed supervisory control [176], in which the robot is preprogrammed or programmed online to perform certain subtasks autonomously. By transmitting only high-level commands, there is a data communication reduction, and task time completion improves. Meanwhile, several extensions of supervisory control were developed, including specific languages to chain tasks or predictive displays (i.e., visualization of a phantom robot model that predicts the motion of the real robot) [177,178].

Control-based approaches for time delay mitigation in teleoperation systems can be clustered into two classes [31,179]: (1) *predictive control-based methods* (e.g., a discrete linear quadratic Gaussian (LQG) controller for teleoperation acting on the sampling rate or output feedback control of multiple-input and multiple-output (MIMO) systems) and (2) *passivity-based methods* that model the master–slave operator systems and unsure stability and performance under time delay variability (e.g., a two-port network, hybrid matrix, impedance matrix, constant time delay, scattering approach, wave variable, scaling, and geometric scattering).

Time series prediction approaches for time delay mitigation in teleoperation systems try to compensate for the time delay, observing past intrinsic patterns to predict the future values [31,156]. They integrate trends, seasonality, and white noise and can be clustered into two types: *statistical methods* and *neural network (NN) or machine learning (ML) methods*:

(1) *Statistical methods* (e.g., moving average (MA), linear auto-regression (AR), auto-regression + moving average (ARMA), and auto-regression + moving average + nonlinear component (ARIMA) [180]);

(2) *NN or AI methods* (e.g., recurrent neural networks (RNNs) [181,182], long short-term memory networks (LSTMs) [183,184], sequence to sequence (Seq2Seq) [185,186], and generative adversarial networks (GANs) [187,188]).

Statistical methods have the advantage of not requiring training with data and are simpler to implement. Although time series prediction traditionally relied on statistical approaches, it has difficulties in modeling the entire set of nonstationary signals. Nevertheless, methods like ARIMA can cope with nonstationary signals. Statistical methods are not appropriate for modeling complex tasks, being more suitable for short-term predictions. Neural networks, on the other hand, have an advantage over statistical approaches in that they enable data description without explicit knowledge of its distribution and can model more complex time series data based on past observations. Neural networks are more prone to adapt their behaviors as the input data increases [31].

3.2. User-Adaptive Systems Taxonomy

Social robots aim to assist people, enable telesurveillance of elderly people, guide people on tours, promote physical and mental exercise, keep company, or entertain [1,5,6]. In short, they contribute to the user's well-being, adapting to people, to the environment, and ultimately to the context. Case studies include interaction of a service robot for 1 week in an elderly care center [116]. Several types of user-adaptive mechanisms are described in the robotics literature [2,8,9,31,189,190].

Typically, a framework for a user-adaptive system comprises two components (see Figure 5): the *interface* layer that is used for the exchange of information between the user and the system (It integrates sensors for the system to perceive the user and actuators

to provide stimuli.) and the *decision-making module* which, based on perceived information, makes algorithmic decisions and generates response actions to be synthesized by the interface.

Robot systems with autonomous and semi-autonomous behaviors can be classified with the following taxonomy [8]:

Autonomous and semi-autonomous behaviors supported by	<ul style="list-style-type: none"> - Adaptive systems with no user model - Systems based on static user models - Systems based on dynamic user models
---	--

1. Adaptive systems with no user model: systems with reactive behavior regarding the user’s immediate feedback and with no cache of the user’s information (see Figure 6);
2. Adaptive systems based on static user models: systems that rely on pre-loaded knowledge retrieved from the relevant attributes of the user and used to adjust the system’s behavior (see Figure 7);
3. Adaptive systems based on dynamic user models: similar to the previous example, these systems explicitly maintain user models. They are task-oriented models, updated with users’ information during their interactions (see Figure 8).

User-adaptive systems require information about the user which is typically stored in the form of a user model [9,191]. As reported in an early survey [192], a new field of research emerged concerning with acquisition, organization, and representation of the system’s user.

Adaptive systems without user modeling can implicitly map the characteristics of a generic user in the architecture of the decision-making module (Figure 6). Nevertheless, it is a reactive adaptation that shapes the system’s behavior directly based on the user’s feedback. The user’s behavior changes are monitored and trigger an immediate switch to a new system’s operational state, while no storage or user model update is performed. Table 7 summarizes several works that adopt this type of architecture.

Adaptive systems based on static user models assume that the person’s profile does not evolve during the interaction. These static models can be built during an initial phase of the interaction (Figure 7), similar to the calibration process, or the user’s profile can be pre-supplied using external questionnaires. These types of systems are not able to dynamically learn the characteristics of the user. Examples of related works are listed in Table 8.

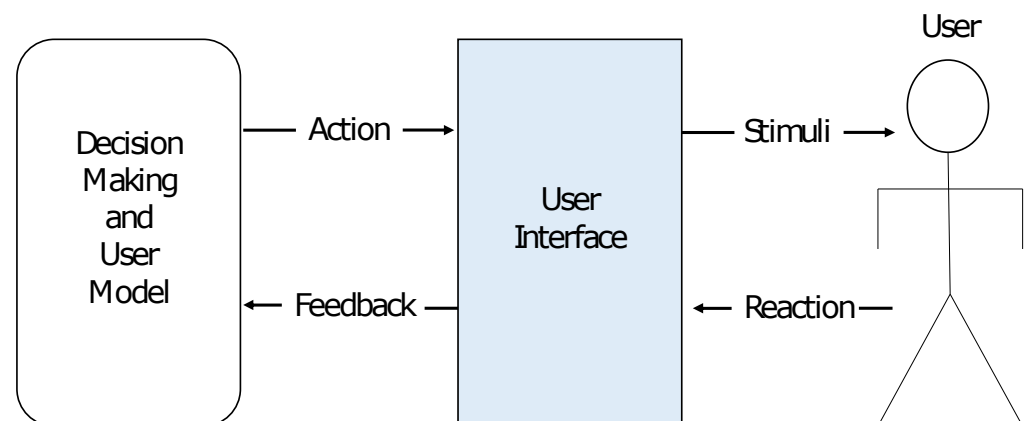


Figure 5. Overview of a generic user-adaptive system, which includes a user interface layer and a decision-making module.

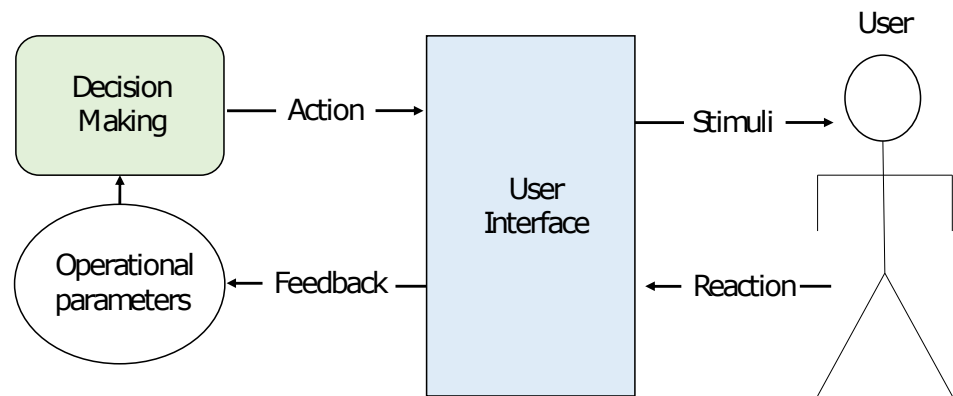


Figure 6. General schematic of a user-adaptive system without the user’s model. The system’s behaviors are direct reactions to the user’s feedback, and decisions are made without the user’s previous knowledge.

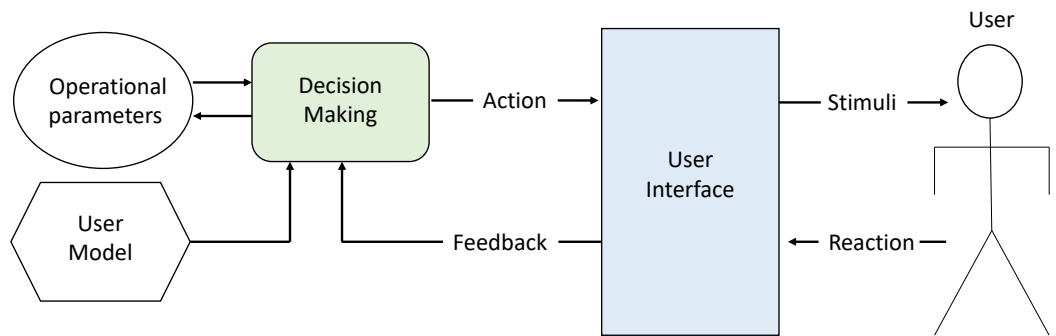


Figure 7. General schematic of a user-adaptive system based on static user models.

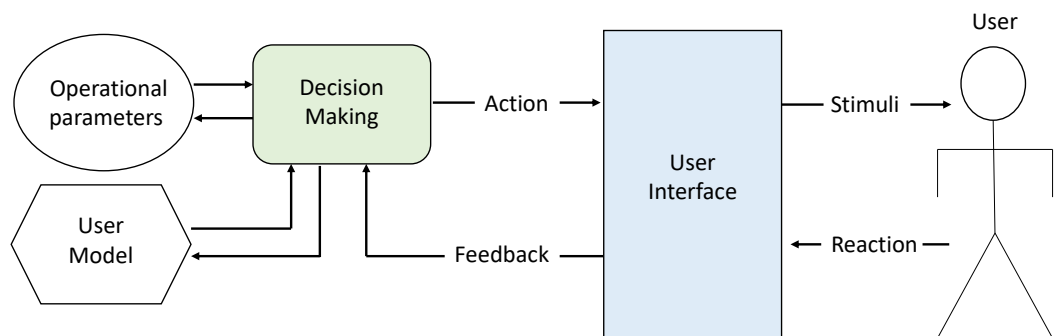


Figure 8. General schematic of a user-adaptive system based on dynamic user models. The user’s feedback reactions are used to continuously update robot knowledge and consequently tune the system’s behavior.

Adaptive systems based on dynamic user models perceive, learn, and update the knowledge regarding the context model and the user model. The stored user model is updated during the interaction based on the user’s reactions. This category of systems is considered the best performing user-adaptive solution, although its implementation is more complex [191,193,194]. Table 9 compiles several references for systems based on dynamic user models.

Additionally, one of the described categories, such as adaptive systems based on dynamic user models, can coexist in a telepresence teleoperated robot [1,3,172,173], thus adding adaptiveness functionalities either for the robot’s operator (remote user) or for the local user that is with the robot. The general architecture of a teleoperation system with user adaptiveness is depicted in Figure 9.

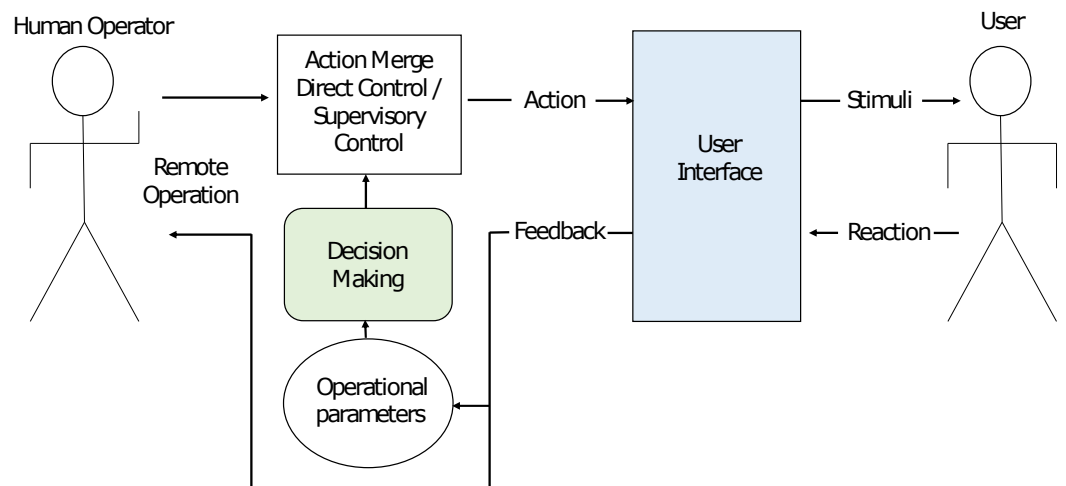


Figure 9. A schematic of the general architecture of a teleoperation system that includes an adaptive system.

The decision-making modules of the listed user-adaptive systems include different frameworks, such as the Markov decision process (MDP), partially observable Markov decision process (POMDP) (α POMDP [195]), Mixed Observability Markov Decision Processes (MOMDP), fuzzy control, rule-based, hidden mode stochastic hybrid system (HMSHS), Bayes-adaptive, dynamic factor graph (DFG), active leaning, or reinforcement learning. Recent approaches for these frameworks are described in [196].

Table 7. Adaptive parameters, input modalities, framework of decision, output modalities, and social robot evaluation with no user model.

Adaptive Parameters	Study	Input Modality	Decision Making	Output Modality	Evaluation Process	Evaluation Metrics
Robot’s Navigation Goal	[197]	Brain-actuated controls	Rule-based	Robot commands	Measurements, questionnaires,	Robot path
Decisions takeleftturn	[198]	Physical controls	POMDP	Image and sound	Measurements, questionnaires	POMDP rewards, perceived control, driving performance, similarity to real world, naturalness, social appropriateness
Robot Speed	[199]	User’s pose and speed	MOMDP	Motor control	Measurements	Speed difference and distance to the user
Decisions (object to move)	[200]	Speech, gaze	Rule-based	Robot arm movement	Measurements, questionnaires	Prediction accuracy, projection accuracy, perceived awareness, response time and intentionality
Robot Speed	[201]	Odometry, Physical controls	Fuzzy control	motor controls	-	-
Decisions (warn driver or intervene)	[202]	Physical controls	Hidden mode stochastic hybrid system	Image and sound	Measurements	Time in unsafe and safe states
Decisions (room to clean)	[203]	User locations, task success	Motor control	Rule-based	-	-
Robot’s Navigation Goal	[204]	Physical controls	Rule-based	Robot commands	Measurements	Recognition accuracy
Voice Pitch	[205]	User speech	Rule-based	Robot speech	Questionnaires	Persistence and learning gain, rapport, perceived social presence

Table 7. *Cont.*

Adaptive Parameters	Pa-rameters	Study	Input Modality	Decision Making	Output Modality	Evaluation Process	Evaluation Metrics
Robot's Gestures		[206]	Vision, speech	Rule-based	Robot com-mands	Measurements, questionnaires	Information distance, perceived behavior performance, perceived gesture recognition, enjoyment, perceived social interaction
Robot Speed and Path		[207]	Physical con-trols	Rule-based	Robot com-mands	-	-
Decisions (navi-gation goal)		[208]	Physical con-trols	POMDP	Robot com-mands	Measurements	State variables, robot path, des-tination probabilities
Decisions (what objects to move, when to speak)		[209]	Speech, vi-sion, depth	Rule-based	Speech, robot com-mands	Measurements	User's speech time

Table 8. Adaptive parameters, input modalities, framework of decision, output modalities and social robot evaluation with static user model.

Adaptive Parameters	Study	Input Modality	Decision Making	Output Modality	Evaluation Process	Evaluation Metrics
Robot's Gestures and Speech	[210]	Speech	Rule-based	Gestures, speech	Questionnaires	Preference toward a type of adaptation
Decisions (placement of objects)	[211]	Crowd-sourced data	Rule-based	Robot con-trols	Measurements	F-scores
Robot Location, Inter-face Complexity, Warn-ing Levels, Font Size	[212]	-	Rule-based	-	-	-
Decisions (how to dress the user)	[213]	User's pose, speech	Rule-based	Robot com-mands	Measurements	Task completion speed
Decisions (how to dress users)	[214]	User's pose, speech	Rule-based	Robot com-mands	Measurements	Classification accuracy
Speech Output Gender, Sound Volume, Robot's Name, Robot Speed	[215]	Speech, touch	Rule-based	Robot commands, speech	Questionnaires	Acceptance, perceived usability
Sequence of dance movements	[216]	User's pose	Rule-based	Robot com-mands	Questionnaire, manual classifi-cation	Gaze position, body lan-guage, facial emotion, perceived bond, amuse-ment, satisfaction, enjoy-ment, anxiety, observed leadership, expectancy

Table 9. Adaptive parameters, input modalities, framework of decision, output modalities, and evaluation of the social robots with a dynamic user model.

Adaptive Parameters	Study	Input Modality	Decision Making	Output Modality	Evaluation Process	Evaluation Metrics
Promote Regular Physical Activity Habits	[84]	User's position, pose (exercise performance), speech	Rule-based	Navigation, robot commands, speech (avatar coach)	Measurements, questionnaires	User's exercise performance, flow
Decisions (service, navigate, turn to the person, stop, smile)	[116,195]	Person detection, speech, emotion recognition, touchscreen	α POMDP, SOA-based model	Navigation, robot commands, approach, speech, robot expression, recognition, service	Measurements, questionnaires	Usability, appearance, satisfaction
Human Robot Greetings Phase	[71]	Tracking human gestures	MDP	Kendom phase trigger (initiate approach, distance salutation, head dip, approach, final approach, and close salutation)	Measurements, observation	Sequence estimation accuracies
Colors of LEDs	[217]	Physical controls	Rule-based	LED colors	Measurements	Cumulative reward from users, error estimation
Decisions (what interactions to perform with the user)	[218]	Physical controls robot	Rule-based	Commands	Measurements, questionnaires	Child learning rate, human intervention ratio
Reading Difficulty Level	[219]	Speech, touch	Active learning	Number of words learned	Measurements, questionnaires	Images, speech
Decisions (adaptation to user's subtask selection)	[220]	Vision, speech	Rule-based	Robot commands, speech	Measurements	Number of communications required of the user
Decisions (moments to take action, including parameter adjustment and services)	[221]	Gesture sound, projected images	MDP	Questionnaires	Perceived coherence, user satisfaction, ease of use, perceived helpfulness, originality, perceived adaptivity	
Decisions (when to deploy services)	[222]	Speech, touch	Equilibrium maintenance	Speech, images, robot commands	Measurements	Opportunity relevance for the selected service
Decisions (dialogue to play)	[223]	Tactile sensors, sound, touch	Dynamic factor graph	Image, speech, robot commands	Questionnaires	User's opinion
Decisions (select learning content type)	[224]	Speech, physical controls	Rule-based	LEDs, robot commands	-	-

Table 9. Cont.

Adaptive Parameters	Study	Input Modality	Decision Making	Output Modality	Evaluation Process	Evaluation Metrics
Decisions (sounds to play)	[225]	Physical controls	Context-free stochastic grammars	Sound (music)	Measurements, questionnaires	Engagement, perceived difficulty, progression, conformity, number of user interventions, speed
Decisions (placing a shared object)	[226]	Vision, physical controls	MAMDP	Robot commands	Measurements, questionnaires	Perceived trustworthiness, ratio of users that change strategies
Decisions (positive, negative, or neutral output)	[227]	Facial expressions, RGBD, electrodermal data, touch screen	Rule-based	Images, speech, gestures	Questionnaires	Understanding, perceived enjoyment, trust
Decisions (where to guide the user)	[228]	Vision, user's attention, robot position, odometry, speech	Rule-based	Robot commands, navigation	Questionnaires	User's opinion (score)

Co-presence mechanisms: the availability of robotic autonomous mechanisms enables a robot's voluntary or involuntary behaviors that contribute to enhancing co-presence [172], such as those listed in Table 10.

Table 10. Robotic mechanisms to enhance co-presence.

Type	Voluntary	Involuntary
Eye contact	X	-
Gaze following	X	-
Gazing at the closest face	X	-
Gazing at a random face	X	-
Gazing at the closest object	X	-
Gazing at a random object	X	-
Gazing at a moving object	X	-
Looking around the gazing position	X	-
Joint attention	X	-
Sleeping	X	-
Changing LED colors	X	-
Mouth movement	X	-
Nodding in response to human speech	X	-
Waving both hands at a random human	X	-
Waving left hand at a random human	X	-
Waving right hand at a random human	X	-
Waving left hand in response to palms	X	-
Waving right hand in response to palms	X	-
Waving both hands in response to palms	X	-
Reflexive blinking with eye movement	-	X
Spontaneous blinking	-	X
Avoiding objects at close range	-	X
Eye saccade	-	X
Breathing	-	X

3.3. Evaluation Methods

To assess co-presence, telepresence systems require objective and qualitative metrics. Quantitative measures may include physiological signals (such as heart rate, skin temperature, electrodermal activity (EDA), and skin conductance responses (SCRs) [229], eye scan patterns, electroencephalography (EEG), or functional magnetic resonance imaging (fMRI)) [34,230,231], as well as other metrics that are simpler to obtain, such as accuracy, time to perform a task, and the number of errors or communication delays. However, given the human factor and the psychological components of interaction, questionnaires remain essential tools. There are methodologies for measuring the *presence*, *social presence* or *co-presence*, and *flow state* of the users using technological devices [19,232–235].

Flow is a psychological state that people describe when they are fully engaged in some events to the point of forgetting time, fatigue, and everything else but the activity itself [40,236]. Table 11 lists the available questionnaires to measure the levels of presence, co-presence, immersion, and flow.

Table 11. List of questionnaires to assess presence, flow, and game.

Psychological Phenomena	Questionnaire	Number of Questions	Ref.
Presence	Slater, Usoh and Steed (SUS)	6	[237]
Presence	Temple Presence Inventory (TPI)	42	[238]
Presence	Igroup Presence Questionnaire (IPQ)	14	[239]
Presence	Sense of Presence Inventory (ITC-SOPI)	38	[240]
Presence	Presence Questionnaire, version 3 (PQ)	29	[241]
Presence	Networked Minds Social Presence Inventory (NMSPI)	34	[41,242]
Presence	Multimodal Presence Scale (MPS)	15	[243]
Presence	Spatial Presence Experience Scale (SPES)	8	[244]
Flow	EduFlow Scale (EFS)	12	[245]
Flow	Flow Short Scale (FSS)	13	[246]
Flow	Reading Flow Short Scale	8	[247]
Game and Flow	EGameFlow (EGF)	42	[248]
Usability	Nielsen Norman Group	-	[249,250]

Usability, testing and accessibility—Jakob Nielsen, one of the most active proponents of usability processes, referred to the following elements that comprise a definition of usability [249–251]:

1. Ease of use: the use of products or tasks should be natural and easily performed by the user.
2. Simplicity of learning: tasks and product features must be intuitive and present a logical and consistent sequence to simplify learning.
3. Improved reliability: levels of satisfaction and performance are increased when the action's results correspond to the user's expectations.
4. Reduction in errors: usability can be increased if designers attribute the errors to the product or task (rather than the user), redesigning it based on the user's feedback.
5. Enhanced user satisfaction: the user's satisfaction principle must guide all of the design process, making the product or task pleasing to use or perform.

In [252], a taxonomy of usability guidelines for the design of telepresence teleoperated robots (interaction effectiveness and efficiency, information presentation, interface visual design, robot surroundings and environment awareness, robot state awareness, and cognitive factors) is proposed. The usability testing process is an effective use of materials and time [17,249,253,254] that should not be overlooked.

4. Conclusions

This work presented a survey of recent works, proposing the development of support for social robotic interactions with applications in health care, elderly assistance, guidance, or office meetings. It focused on enhancing social presence via telepresence robot mediation, in which a user should sense his or her remote interlocutor as being locally present with him or her. The research gathered knowledge to help roboticists design improved user- and environment-adaptive systems and technical methods that contribute to enhancing the sense of presence or co-presence. This literature review aimed to define social presence, identify autonomous “user-adaptive systems” for social robots, and propose a taxonomy for “co-presence” mechanisms. The referred works address robot sensing, perception, action, reasoning, appearance, automation, and cognitive approaches (e.g., statistics models and AI). Additionally, it presents an overview of social robotics systems and application areas and provides directions for telepresence and co-presence robot design, considering the actual and future challenges. Finally, some guidelines for the evaluation of these systems are left, having as reference face-to-face interactions. Based on survey findings in engineering and psychology, our future work includes the design of telepresence and co-presence robots that better emulate or interpret human subjective experiences.

Author Contributions: Conceptualization, L.A., P.M. and J.D.; methodology, L.A., P.M. and J.D.; validation, L.A. and P.M.; data curation, L.A.; writing—original draft preparation, L.A., P.M. and J.D.; writing—review and editing, L.A., P.M. and J.D.; supervision, P.M. and J.D.; project administration, P.M. and J.D.; funding acquisition, P.M. and J.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially developed in the framework of the project “ACTIVAS – Supporting Habitat for an Active, Safe and Healthy Life” (POCI-01-0247-FEDER-046101) which is supported by European Regional Development Fund (ERDF), through the Incentive System to Research and Technological development, within the Portugal2020 Competitiveness and Internationalization Operational Program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors acknowledge the support of the University of Coimbra, Institute of Systems and Robotics in Coimbra, Portugal, and the Polytechnic Institute of Tomar, Portugal.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sheridan, T.B. A review of recent research in social robotics. *Curr. Opin. Psychol.* **2020**, *36*, 7–12. [CrossRef] [PubMed]
2. Nocentini, O.; Fiorini, L.; Acerbi, G.; Sorrentino, A.; Mancioffi, G.; Cavallo, F. A Survey of Behavioral Models for Social Robots. *Robotics* **2019**, *8*, 54. [CrossRef]
3. Tsui, K.; Yanco, H. Design Challenges and Guidelines for Social Interaction Using Mobile Telepresence Robots. *Rev. Hum. Factors Ergon.* **2013**, *9*, 227–301. [CrossRef]
4. Alabdulkareem, A.; Alhakbani, N.; Al-Nafjan, A. A Systematic Review of Research on Robot-Assisted Therapy for Children with Autism. *Sensors* **2022**, *22*, 944. [CrossRef] [PubMed]
5. Avelino, J.; Garcia-Marques, L.; Ventura, R.; Bernardino, A. Break the ice: A survey on socially aware engagement for human–robot first encounters. *Int. J. Soc. Robot.* **2021**, *13*, 1851–1877. [CrossRef] [PubMed]
6. Isabet, B.; Pino, M.; Lewis, M.; Benveniste, S.; Rigaud, A.S. Social Telepresence Robots: A Narrative Review of Experiments Involving Older Adults before and during the COVID-19 Pandemic. *Int. J. Environ. Res. Public Health* **2021**, *18*, 3597. [CrossRef]
7. IEEE. Robots: Your Guide to the World of Robotics. 2022. Available online: <https://robots.ieee.org/> (accessed on 5 April 2022).
8. Martins, G.S.; Santos, L.; Dias, J. User-adaptive interaction in social robots: A survey focusing on non-physical interaction. *Int. J. Soc. Robot.* **2019**, *11*, 185–205. [CrossRef]
9. Hellou, M.; Gasteiger, N.; Lim, J.Y.; Jang, M.; Ahn, H.S. Personalization and Localization in Human-Robot Interaction: A Review of Technical Methods. *Robotics* **2021**, *10*, 120. [CrossRef]
10. IJsselsteijn, W. Presence in the past: What can we learn from media history? In *Being There: Concepts, Effects and Measurements of User Presence in Synthetic Environments. Emerging Communication: Studies in New Technologies and Practices in Communication*; IOS Press: Amsterdam, The Netherlands, 2003; pp. 17–40.
11. Minsky, M. Telepresence. *Omni* **1980**, *2*, 45–51.
12. Short, J.; Williams, E.; Christie, B. *The Social Psychology of Telecommunications*; John Wiley & Son: Hoboken, NJ, USA, 1976.
13. Biocca, F.; Harms, C.; Burgoon, J.K. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence Teleoperators Virtual Environ.* **2003**, *12*, 456–480. [CrossRef]
14. Steuer, J. Defining Virtual Reality: Dimensions Determining Telepresence. *J. Commun.* **1992**, *42*, 73–93. [CrossRef]
15. Biocca, F. The Cyborg’s Dilemma: Embodiment in Virtual Environments. In Proceedings of the Second International Conference on Cognitive Technology Humanizing the Information Age, Aizu-Wakamatsu, Japan, 25–28 August 1997; p. 12.
16. Sheridan, T.B. Musings on telepresence and virtual presence. *Presence* **1992**, *1*, 120–126. [CrossRef]
17. Nowak, K.L.; Biocca, F. The effect of the agency and anthropomorphism on users’ sense of telepresence, copresence, and social presence in virtual environments. *Presence Teleoperators Virtual Environ.* **2003**, *12*, 481–494. [CrossRef]
18. Zhao, S. Toward a taxonomy of copresence. *Presence* **2003**, *12*, 445–455. [CrossRef]
19. Lombard, M.; Biocca, F.; Freeman, J.; IJsselsteijn, W.; Schaevitz, R. *Immersed in Media: Telepresence Theory, Measurement & Technology*; Springer International Publishing: Cham, Switzerland, 2015.
20. Paulos, E.; Canny, J. PRoP: Personal roving presence. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Los Angeles, CA, USA, 18–23 April 1998; pp. 296–303.
21. Paulos, E.; Canny, J. Ubiquitous Tele-embodiment: Applications and Implications. *Int. J. Hum. Stud.* **1997**, *46*, 861–877. [CrossRef]
22. Paulos, E.; Canny, J. Social tele-embodiment: Understanding presence. *Auton. Robot.* **2001**, *11*, 87–95. [CrossRef]
23. Slater, M.; Sanchez-Vives, M.V. Transcending the self in immersive virtual reality. *Computer* **2014**, *47*, 24–30. [CrossRef]
24. Blanke, O.; Metzinger, T. Full-body illusions and minimal phenomenal selfhood. *Trends Cogn. Sci.* **2009**, *13*, 7–13. [CrossRef]

25. Li, J. The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *Int. J. Hum. Stud.* **2015**, *77*, 23–37. [[CrossRef](#)]
26. Goffman, E. *Behavior in Public Places*; A Free Press paperback; Free Press: New York, NY, USA, 1963.
27. Choi, J.J.; Kwak, S.S. Who is this?: Identity and presence in robot-mediated communication. *Cogn. Syst. Res.* **2017**, *43*, 174–189. [[CrossRef](#)]
28. Nowak, K.L.; Fox, J. Avatars and computer-mediated communication: A review of the definitions, uses, and effects of digital representations. *Rev. Commun. Res.* **2018**, *6*, 30–53. [[CrossRef](#)]
29. Oh, C.S.; Bailenson, J.N.; Welch, G.F. A systematic review of social presence: Definition, antecedents, and implications. *Front. Robot.* **2018**, *5*, 114. [[CrossRef](#)] [[PubMed](#)]
30. Pimentel, D.; Vinkers, C. Copresence with virtual humans in mixed reality: The impact of contextual responsiveness on social perceptions. *FRontiers Robot.* **2021**, *8*, 25. [[CrossRef](#)] [[PubMed](#)]
31. Farajiparvar, P.; Ying, H.; Pandya, A. A Brief Survey of Telerobotic Time Delay Mitigation. *Front. Robot.* **2020**, *7*, 198. [[CrossRef](#)]
32. Assuncao, G.; Patrao, B.; Castelo-Branco, M.; Menezes, P. An Overview of Emotion in Artificial Intelligence. *IEEE Trans. Artif. Intell.* **2022**. [[CrossRef](#)]
33. Cummings, J.J.; Wertz, B. Technological predictors of social presence: A foundation for a meta-analytic review and empirical concept explication. In Proceedings of the 10th Annual International Workshop on Presence, Prague, Czech Republic, 3–6 July 2018.
34. Bohil, C.J.; Alicea, B.; Biocca, F.A. Virtual reality in neuroscience research and therapy. *Nat. Rev. Neurosci.* **2011**, *12*, 752–762. [[CrossRef](#)]
35. Slater, M. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2009**, *364*, 3549–3557. [[CrossRef](#)]
36. Slater, M.; Sanchez-Vives, M.V. Enhancing Our Lives with Immersive Virtual Reality. *Front. Robot.* **2016**, *3*, 74. [[CrossRef](#)]
37. Slater, M.; Spanlang, B.; Sanchez-Vives, M.V.; Blanke, O. First Person Experience of Body Transfer in Virtual Reality. *PLoS ONE* **2010**, *5*, e10564. [[CrossRef](#)]
38. Blanke, O.; Slater, M.; Serino, A. Behavioral, neural, and computational principles of bodily self-consciousness. *Neuron* **2015**, *88*, 145–166. [[CrossRef](#)]
39. Lombard, M.; Ditton, T. At the Heart of It All: The Concept of Presence. *J. Comput. Commun.* **1997**, *3*, JCMC321. [[CrossRef](#)]
40. Saari, T.; Laarni, J.; Ravaja, N.; Kallinen, K.; Turpeinen, M. Virtual Ba and Presence in Facilitating Learning from Technology Mediated Organizational Information Flows. In Proceedings of the Annual International Workshop on Presence, Valencia, Spain, 13–14 October 2004; Technical University of Valencia: Valencia, Spain, 2004; pp. 133–140.
41. Biocca, F.; Harms, C. Defining and measuring social presence: Contribution to the networked minds theory and measure. *Proc. PRESENCE* **2002**, *2002*, 1–36.
42. Daft, R.L.; Lengel, R.H. Organizational information requirements, media richness and structural design. *Manag. Sci.* **1986**, *32*, 554–571. [[CrossRef](#)]
43. Cummings, J.J.; Bailenson, J.N. How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychol.* **2016**, *19*, 272–309. [[CrossRef](#)]
44. Biocca, F.; Harms, C.; Gregg, J. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. In Proceedings of the 4th Annual International Workshop on presence, Philadelphia, PA, USA, 21–23 May 2001; pp. 1–9.
45. Cortese, J.; Seo, M. The role of social presence in opinion expression during FtF and CMC discussions. *Commun. Res. Rep.* **2012**, *29*, 44–53. [[CrossRef](#)]
46. Zhan, Z.; Mei, H. Academic self-concept and social presence in face-to-face and online learning: Perceptions and effects on students' learning achievement and satisfaction across environments. *Comput. Educ.* **2013**, *69*, 131–138. [[CrossRef](#)]
47. Francescato, D.; Porcelli, R.; Mebane, M.; Cuddetta, M.; Klobas, J.; Renzi, P. Evaluation of the efficacy of collaborative learning in face-to-face and computer-supported university contexts. *Comput. Hum. Behav.* **2006**, *22*, 163–176. [[CrossRef](#)]
48. De Greef, P.; Ijsselstein, W.A. Social presence in a home tele-application. *Cyberpsych. Behav.* **2001**, *4*, 307–315. [[CrossRef](#)]
49. de Greef, H. Video communication best for female friends? In Proceedings of the ISPR 2014: 15th International Workshop on Presence (PRESENCE 2014), Vienna, Austria, 17–19 March 2014; pp. 187–193.
50. Bente, G.; Rüggenberg, S.; Krämer, N.C.; Eschenburg, F. Avatar-mediated networking: Increasing social presence and interpersonal trust in net-based collaborations. *Hum. Commun. Res.* **2008**, *34*, 287–318. [[CrossRef](#)]
51. Blascovich, J. *Social Influence within Immersive Virtual Environments*; Springer: Berlin/Heidelberg, Germany, 2002; pp. 127–145.
52. Kim, H.; Suh, K.S.; Lee, U.K. Effects of collaborative online shopping on shopping experience through social and relational perspectives. *Inf. Manag.* **2013**, *50*, 169–180. [[CrossRef](#)]
53. Feng, B.; Li, S.; Li, N. Is a profile worth a thousand words? How online support-seeker's profile features may influence the quality of received support messages. *Commun. Res.* **2016**, *43*, 253–276. [[CrossRef](#)]
54. von der Pütten, A.M.R.; Krämer, N.C.; Gratch, J.; Kang, S.H. "It doesn't matter what you are!" Explaining social effects of agents and avatars. *Comput. Hum. Behav.* **2010**, *26*, 1641–1650. [[CrossRef](#)]
55. Pan, X.; Gillies, M.; Slater, M. The impact of avatar blushing on the duration of interaction between a real and virtual person. In Proceedings of the Presence 2008: The 11th Annual International Workshop on Presence, Padova, Italy, 16–18 October 2008; pp. 100–106.

56. Kang, S.H.; Watt, J.H. The impact of avatar realism and anonymity on effective communication via mobile devices. *Comput. Hum. Behav.* **2013**, *29*, 1169–1181. [[CrossRef](#)]
57. Bailenson, J.N.; Blascovich, J.; Beall, A.C.; Loomis, J.M. Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence Teleoperators Virtual Environ.* **2001**, *10*, 583–598. [[CrossRef](#)]
58. Blascovich, J.; Loomis, J.; Beall, A.C.; Swinth, K.R.; Hoyt, C.L.; Gerson, J.N. Immersive virtual environment technology as a methodological tool for social psychology. *Psychol. Inq.* **2002**, *13*, 103–124. [[CrossRef](#)]
59. Blascovich, J.; Bailenson, J. *Infinite Reality: Avatars, Eternal Life, New Worlds, and the Dawn of the Virtual Revolution*; William Morrow & Co: New York, NY, USA, 2011.
60. Garau, M.; Slater, M.; Vinayagamoorthy, V.; Brogni, A.; Steed, A.; Sasse, M.A. The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In Proceedings of the 2003 Conference on Human Factors in Computing Systems, CHI 2003, Ft. Lauderdale, FL, USA, 5–10 April 2003; Cockton, G., Korhonen, P., Eds.; ACM: New York, NY, USA, 2003; pp. 529–536. [[CrossRef](#)]
61. Bailenson, J.N.; Swinth, K.; Hoyt, C.; Persky, S.; Dimov, A.; Blascovich, J. The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence Teleoperators Virtual Environ.* **2005**, *14*, 379–393. [[CrossRef](#)]
62. Jo, D.; Kim, K.H.; Kim, G.J. Effects of avatar and background types on users' co-presence and trust for mixed reality-based teleconference systems. In Proceedings of the 30th Conference on Computer Animation and Social Agents, Seoul, Korea, 22–24 May 2017; pp. 27–36.
63. Herath, D.C.; Jochum, E.; Vlachos, E. An experimental study of embodied interaction and human perception of social presence for interactive robots in public settings. *IEEE Trans. Cogn. Dev. Syst.* **2017**, *10*, 1096–1105. [[CrossRef](#)]
64. Perret, J.; Vander Poorten, E. Touching Virtual Reality: A Review of Haptic Gloves. In Proceedings of the ACTUATOR 2018: 16th International Conference on New Actuators, Bremen, Germany, 25–27 June 2018; pp. 1–5.
65. Fernando, C.; Furukawa, M.; Kurogi, T.; Kamuro, S.; Sato, K.; Minamizawa, K.; Tachi, S. Design of TELESAR V for transferring bodily consciousness in telexistence. In Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura, Algarve, Portugal, 7–12 October 2012; pp. 5112–5118. [[CrossRef](#)]
66. Fisch, A.; Mavroidis, C.; Melli-Huber, J.; Bar-Cohen, Y. Haptic devices for virtual reality, telepresence, and human-assistive robotics. In *Biologically Inspired Intelligent Robots*; SPIE Digital Library: Washington, DC, USA, 2003; pp. 73–101. [[CrossRef](#)]
67. Prasad, V.; Stock-Homburg, R.; Peters, J. Human-robot handshaking: A review. *arXiv* **2021**, arXiv:2102.07193.
68. Bevan, C.; Fraser, D.S. Shaking hands and cooperation in tele-present human-robot negotiation. In Proceedings of the 2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Portland, OR, USA, 2–5 March 2015; pp. 247–254.
69. Nakanishi, H.; Tanaka, K.; Wada, Y. Remote handshaking: Touch enhances video-mediated social telepresence. In Proceedings of the CHI Conference on Human Factors in Computing Systems, CHI'14, Toronto, ON, Canada, 26 April–1 May 2014; Jones, M., Palanque, P.A., Schmidt, A., Grossman, T., Eds.; ACM: New York, NY, USA, 2014; pp. 2143–2152. [[CrossRef](#)]
70. Mostofa, N.; Avendano, I.; McMahan, R.P.; Conner, N.E.; Anderson, M.; Welch, G.F. Tactile Telepresence for Isolated Patients. In Proceedings of the 2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct), Bari, Italy, 4–8 October 2021; IEEE Computer Society: Los Alamitos, CA, USA, 2021; pp. 346–351. [[CrossRef](#)]
71. Carvalho, M.; Avelino, J.; Bernardino, A.; Ventura, R.M.M.; Moreno, P. Human-Robot greeting: Tracking human greeting mental states and acting accordingly. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2021, Prague, Czech Republic, 27 September–1 October 2021; pp. 1935–1941. [[CrossRef](#)]
72. Kendon, A. *Conducting Interaction: Patterns of Behavior in Focused Encounters*; Cambridge University Press: Cambridge, UK, 1990.
73. Colledanchise, M.; Ögren, P. *Behavior Trees in Robotics and AI: An Introduction*; CRC Press: Boca Raton, FL, USA, 2018. [[CrossRef](#)]
74. Tachi, S.; Inoue, Y.; Kato, F. TELESAR VI: Telexistence Surrogate Anthropomorphic Robot VI. *Int. J. Humanoid Robot.* **2020**, *17*, 2050019:1–2050019:33. [[CrossRef](#)]
75. Almeida, L.; Menezes, P.; Seneviratne, L.; Dias, J. Incremental 3d body reconstruction framework for robotic telepresence applications. In Proceedings of the Robo 2011: The 2nd IASTED International Conference on Robotics, Pittsburgh, PA, USA, 7–9 November 2011; pp. 286–293.
76. Almeida, L.; Menezes, P.; Dias, J. Augmented reality framework for the socialization between elderly people. In *Handbook of Research on ICTs for Human-Centered Healthcare and Social Care Services*; IGI Global: Hershey, PA, USA, 2013; pp. 430–448.
77. Plüss, C.; Ranieri, N.; Bazin, J.C.; Martin, T.; Laffont, P.Y.; Popa, T.; Gross, M. An immersive bidirectional system for life-size 3d communication. In Proceedings of the CASA '16: 29th International Conference on Computer Animation and Social Agents, Geneva, Switzerland, 23–25 May 2016; pp. 89–96.
78. Dicke, C.; Aaltonen, V.; Rämö, A.; Vilermo, M. Talk to me: The influence of audio quality on the perception of social presence. In Proceedings of the HCI 2010 24, University of Abertay, Dundee, UK, 6–10 September 2010; pp. 309–318.
79. Lanillos, P.; Ferreira, J.F.; Dias, J. Designing an artificial attention system for social robots. In Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 28 September–3 October 2015; pp. 4171–4178. [[CrossRef](#)]
80. Almeida, L.; Menezes, P.; Dias, J. Interface Transparency Issues in Teleoperation. *Appl. Sci.* **2020**, *10*, 6232. [[CrossRef](#)]
81. Kim, D.; Jo, D. Effects on Co-Presence of a Virtual Human: A Comparison of Display and Interaction Types. *Electronics* **2022**, *11*, 367. [[CrossRef](#)]

82. Appel, J.; von der Pütten, A.; Krämer, N.C.; Gratch, J. Does Humanity Matter? Analyzing the Importance of Social Cues and Perceived Agency of a Computer System for the Emergence of Social Reactions during Human-Computer Interaction. *Adv. Hum.-Comp. Int.* **2012**, *2012*. [CrossRef]
83. Quintas, J.; Martins, G.S.; Santos, L.; Menezes, P.; Dias, J. Toward a Context-Aware Human-Robot Interaction Framework Based on Cognitive Development. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, *49*, 227–237. [CrossRef]
84. Menezes, P.; Rocha, R.P. Promotion of active ageing through interactive artificial agents in a smart environment. *SN Appl. Sci.* **2021**, *3*, 583. [CrossRef]
85. Kim, H.S.; Shyam Sundar, S. Can online buddies and bandwagon cues enhance user participation in online health communities? *Comput. Hum. Behav.* **2014**, *37*, 319–333. [CrossRef]
86. Li, S.; Feng, B.; Li, N.; Tan, X. How social context cues in online support-seeking influence self-disclosure in support provision. *Commun. Q.* **2015**, *63*, 586–602. [CrossRef]
87. Kim, K.; Schubert, R.; Hochreiter, J.; Bruder, G.; Welch, G. Blowing in the wind: Increasing social presence with a virtual human via environmental airflow interaction in mixed reality. *Comput. Graph.* **2019**, *83*, 23–32. [CrossRef]
88. Johnson, R.D. Gender Differences in E-Learning: Communication, Social Presence, and Learning Outcomes. *J. Organ. End User Comput.* **2011**, *23*, 79–94. [CrossRef]
89. Lim, J.; Richardson, J.C. Exploring the effects of students' social networking experience on social presence and perceptions of using SNSs for educational purposes. *Internet High. Educ.* **2016**, *29*, 31–39. [CrossRef]
90. Jin, S.A.A. Parasocial Interaction with an Avatar in Second Life: A Typology of the Self and an Empirical Test of the Mediating Role of Social Presence. *Presence Teleoperators Virtual Environ.* **2010**, *19*, 331–340. [CrossRef]
91. Garcia, J.C.; Patrao, B.; Almeida, L.; Perez, J.; Menezes, P.; Dias, J.; Sanz, P.J. A Natural Interface for Remote Operation of Underwater Robots. *IEEE Comput. Graph. Appl.* **2017**, *37*, 34–43. [CrossRef]
92. Almeida, L.; Menezes, P.; Dias, J. Improving robot teleoperation experience via immersive interfaces. In Proceedings of the 2017 4th Experiment@International Conference (exp.at'17), Faro, Algarve, Portugal, 6–8 June 2017; pp. 87–92. [CrossRef]
93. Lee, S.; Kim, G.J. Effects of haptic feedback, stereoscopy, and image resolution on performance and presence in remote navigation. *Int. J. Hum. Stud.* **2008**, *66*, 701–717. [CrossRef]
94. Ahn, D.; Seo, Y.; Kim, M.; Kwon, J.H.; Jung, Y.; Ahn, J.; Lee, D. The effects of actual human size display and stereoscopic presentation on users' sense of being together with and of psychological immersion in a virtual character. *Cyberpsychol. Behav. Soc. Netw.* **2014**, *17*, 483–487. [CrossRef]
95. Ferreira, J.; Lobo, J.; Bessiere, P.; Castelo-Branco, M.; Dias, J. A Bayesian framework for active artificial perception. *IEEE Trans. Cybern.* **2013**, *43*, 699–711. [CrossRef]
96. Maimone, A.; Bidwell, J.; Peng, K.; Fuchs, H. Enhanced personal autostereoscopic telepresence system using commodity depth cameras. *Comput. Graph.* **2012**, *36*, 791–807. [CrossRef]
97. Lee, K.M.; Nass, C. Social-psychological origins of feelings of presence: Creating social presence with machine-generated voices. *Media Psychol.* **2005**, *7*, 31–45. [CrossRef]
98. Jung, S.; Roh, S.; Yang, H.; Biocca, F. Location and modality effects in online dating: Rich modality profile and location-based information cues increase social presence, while moderating the impact of uncertainty reduction strategy. *Cyberpsychol. Behav. Soc. Netw.* **2017**, *20*, 553–560. [CrossRef]
99. Kim, K.J.; Park, E.; Sundar, S.S. Caregiving role in human-robot interaction: A study of the mediating effects of perceived benefit and social presence. *Comput. Hum. Behav.* **2013**, *29*, 1799–1806. [CrossRef]
100. telepresencerobots.com. Telepresence Robots Shop. Available online: <https://telepresencerobots.com/robots/orbis-robotics-teleporter/> (accessed on 5 April 2022).
101. Orlandini, A.; Kristoffersson, A.; Almquist, L.; Björkman, P.; Cesta, A.; Cortellessa, G.; Galindo, C.; Gonzalez-Jimenez, J.; Gustafsson, K.; Kiselev, A.; et al. ExCITE Project: A Review of Forty-Two Months of Robotic Telepresence Technology Evolution. *Presence Teleoperators Virtual Environ.* **2016**, *25*, 204–221. [CrossRef]
102. Double; Robotics. Double Robotics, Inc. Available online: <https://www.doublerobotics.com/> (accessed on 5 April 2022).
103. InbotTechnology. PADBOT, Inbot Technology, Ltd. Available online: <https://www.padbot.com/> (accessed on 5 April 2022)
104. OceanRobotics. Beam Pro, GoBe Robots, OceanRobotics, Inc. Available online: <https://gobe.blue-ocean-robotics.com/robots> (accessed on 5 April 2022).
105. Robotics, A.; iRobot. AVA 500, Ava Robotics and iRobot, Inc. Available online: <https://www.avarobotics.com/> (accessed on 5 April 2022).
106. OhmniLabs. Ohmni Telepresence Robot, OhmniLabs, Inc. Available online: <https://ohmnilabs.com/> (accessed on 5 April 2022).
107. VGo. VGo Robotic Telepresence, Vecna Technologies, Inc. Available online: <https://www.vgocom.com/> (accessed on 5 April 2022).
108. MantaroBot1. TeleMe—TelePresence Robot, Mantaro Inc. Available online: <http://www.mantarobot.com/products/teleme-2/index.htm> (accessed on 5 April 2022).
109. InTouchHealth; iRobot. RP-VITA Telepresence Robot, InTouch Health and iRobot, Inc. Available online: <https://intouchhealth.com/> (accessed on 5 April 2022).
110. Aubot. Teleporter robot, Aubot Inc. Available online: <https://aubot.com/> (accessed on 5 April 2022).
111. FutureRobot. FURo-i, Future Robot CO., Lda. Available online: <http://www.futurerobot.com/default/> (accessed on 5 April 2022).

112. Adalgeirsson, S.O.; Breazeal, C. MeBot: A robotic platform for socially embodied telepresence. In Proceedings of the 2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Osaka, Japan, 2–5 March 2010; pp. 15–22.
113. OriginRobotics. ORIGIBOT Telepresence Robot, Origin Robotics, Inc. Available online: <https://www.originrobotics.com/> (accessed on 5 April 2022).
114. SoftBankRobotics. NAO and Pepper Robots, SoftBank Robotics, Lda. Available online: <https://www.softbankrobotics.com/> (accessed on 5 April 2022).
115. Martins, G.S.; Santos, L.; Dias, J. The GrowMeUp project and the applicability of action recognition techniques. In Proceedings of the Third Workshop on Recognition and Action for Scene Understanding (REACTS), Valletta, Malta, 5 September 2015.
116. Portugal, D.; Alvito, P.; Christodoulou, E.; Samaras, G.; Dias, J. A study on the deployment of a service robot in an elderly care center. *Int. J. Soc. Robot.* **2019**, *11*, 317–341. [[CrossRef](#)]
117. Xandex. kubi Telepresence Robots, Xandex Inc. Available online: <https://www.kubiconnect.com/> (accessed on 5 April 2022).
118. MantaroBot2. TableTop TeleMe—TelePresence Robot, Mantaro Inc. Available online: http://www.mantarobot.com/products/tabletop_teleme/index.htm (accessed on 5 April 2022).
119. SELFIEBOT.CO. Selfie Bot, SELFIEBOT.CO. Available online: <https://www.selfiebot.co/> (accessed on 5 April 2022).
120. OwlLabs. Owl Pro, Owl Labs, Inc. Available online: <https://owllabs.com/> (accessed on 5 April 2022).
121. Matsumura, R.; Shiomi, M.; Nakagawa, K.; Shinozawa, K.; Miyashita, T. A desktop-sized communication robot: “robovie-mr2”. *J. Robot. Mechatron.* **2016**, *28*, 107–108. [[CrossRef](#)]
122. Goodrich, M.A.; Crandall, J.W.; Barakova, E. Teleoperation and Beyond for Assistive Humanoid Robots. *Rev. Hum. Factors Ergon.* **2013**, *9*, 175–226. [[CrossRef](#)]
123. Goodrich, M.A.; Schultz, A.C. Human–Robot Interaction: A Survey. *Found. Trends Hum. Comput. Interact.* **2008**, *1*, 203–275. [[CrossRef](#)]
124. Hornstein, J.; Lopes, M.; Santos-Victor, J.; Lacerda, F. Sound Localization for Humanoid Robots—Building Audio-Motor Maps based on the HRTF. In Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, 9–15 October 2006; pp. 1170–1176. [[CrossRef](#)]
125. Ferreira, J.F.; Pinho, C.; Dias, J. Implementation and calibration of a Bayesian binaural system for 3D localisation. In Proceedings of the 2008 IEEE International Conference on Robotics and Biomimetics, Bangkok, Thailand, 22–25 February 2009; pp. 1722–1727. [[CrossRef](#)]
126. Roman, N.; Wang, D.L.; Brown, G.J. Speech segregation based on sound localization. *J. Acoust. Soc. Am.* **2003**, *114*, 2236–2252. [[CrossRef](#)]
127. Wang, D.; Chen, J. Supervised speech separation based on deep learning: An overview. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2018**, *26*, 1702–1726. [[CrossRef](#)]
128. Anusuya, M.; Katti, S. Front end analysis of speech recognition: A review. *Int. J. Speech Technol.* **2011**, *14*, 99–145. [[CrossRef](#)]
129. Nassif, A.B.; Shahin, I.; Attili, I.; Azzeh, M.; Shaalan, K. Speech Recognition Using Deep Neural Networks: A Systematic Review. *IEEE Access* **2019**, *7*, 19143–19165. [[CrossRef](#)]
130. Stiefelhagen, R.; Fugen, C.; Gieselmann, R.; Holzapfel, H.; Nickel, K.; Waibel, A. Natural human-robot interaction using speech, head pose and gestures. In Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Sendai, Japan, 28 September–2 October 2004; Volume 3, pp. 2422–2427. [[CrossRef](#)]
131. Liu, H.; Fang, S.; Zhang, Z.; Li, D.; Lin, K.; Wang, J. MFDNet: Collaborative Poses Perception and Matrix Fisher Distribution for Head Pose Estimation. *IEEE Trans. Multimed.* **2021**. [[CrossRef](#)]
132. Ji, Y.; Yang, Y.; Shen, F.; Shen, H.T.; Li, X. A Survey of Human Action Analysis in HRI Applications. *IEEE Trans. Circuits Syst. Video Technol.* **2020**, *30*, 2114–2128. [[CrossRef](#)]
133. Beddiar, D.R.; Nini, B.; Sabokrou, M.; Hadid, A. Vision-based human activity recognition: A survey. *Multimed. Tools Appl.* **2020**, *79*, 30509–30555. [[CrossRef](#)]
134. Kelley, R.; Nicolescu, M.; Tavakkoli, A.; Nicolescu, M.; King, C.; Bebis, G. Understanding human intentions via Hidden Markov Models in autonomous mobile robots. In Proceedings of the 2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI), Amsterdam, The Netherlands, 12–15 March 2008; pp. 367–374.
135. Quintas, J.; Almeida, L.; Brito, M.; Quintela, G.; Menezes, P.; Dias, J. Context-based understanding of interaction intentions. In Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2012), Paris, France, 9–13 September 2012.
136. Roth, P.M.; Winter, M. Survey of appearance-based methods for object recognition. In *Institute for Computer Graphics and Vision*; Technical report ICGTR0108 (ICG-TR-01/08); Graz University of Technology: Graz, Austria, 2008.
137. Li, L.J.; Socher, R.; Fei-Fei, L. Towards total scene understanding: Classification, annotation and segmentation in an automatic framework. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 2036–2043. [[CrossRef](#)]
138. Ye, C.; Yang, Y.; Mao, R.; Fermüller, C.; Aloimonos, Y. What can i do around here? Deep functional scene understanding for cognitive robots. In Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 29 May–3 June 2017; pp. 4604–4611. [[CrossRef](#)]
139. Garcia Bermudez, F.L.; Julian, R.C.; Haldane, D.W.; Abbeel, P.; Fearing, R.S. Performance analysis and terrain classification for a legged robot over rough terrain. In Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura, Algarve, Portugal, 7–12 October 2012; pp. 513–519. [[CrossRef](#)]

140. Dynamics, B. Spot—The Agile Mobile Robot | Boston Dynamics. Available online: <https://www.bostondynamics.com/products/spot> (accessed on 5 April 2022).
141. Dynamics, B. ATLAS—The Most Dynamic Humanoid Robot | Boston Dynamics. Available online: <https://www.bostondynamics.com/atlas> (accessed on 5 April 2022).
142. Honda. ASIMO Robot | Honda. Available online: <https://asimo.honda.com/default.aspx> (accessed on 5 April 2022).
143. Hubicki, C.; Abate, A.; Clary, P.; Rezazadeh, S.; Jones, M.; Peekema, A.; Van Why, J.; Domres, R.; Wu, A.; Martin, W.; et al. Walking and running with passive compliance: Lessons from engineering: A live demonstration of the atrias biped. *IEEE Robot. Autom. Mag.* **2018**, *25*, 23–39. [\[CrossRef\]](#)
144. Sahbani, A.; El-Khoury, S.; Bidaud, P. An overview of 3D object grasp synthesis algorithms. *Robot. Auton. Syst.* **2012**, *60*, 326–336. [\[CrossRef\]](#)
145. Gray, S.; Chitta, S.; Kumar, V.; Likhachev, M. A single planner for a composite task of approaching, opening and navigating through non-spring and spring-loaded doors. In Proceedings of the 2013 IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, 6–10 May 2013; pp. 3839–3846.
146. Venture, G.; Kulić, D. Robot Expressive Motions: A Survey of Generation and Evaluation Methods. *J. Hum. Robot Interact.* **2019**, *8*. [\[CrossRef\]](#)
147. Abdollahi, H.; Mahoor, M.; Zandie, R.; Sewierski, J.; Qualls, S. Artificial Emotional Intelligence in Socially Assistive Robots for Older Adults: A Pilot Study. *IEEE Trans. Affect. Comput.* **2022**. [\[CrossRef\]](#)
148. Montemerlo, M.; Thrun, S. *FastSLAM: A Scalable Method for the Simultaneous Localization and Mapping Problem in Robotics*; Springer: Berlin/Heidelberg, Gernay, 2007; Volume 27.
149. Sualeh, M.; Kim, G.W. Simultaneous localization and mapping in the epoch of semantics: A survey. *Int. J. Control. Autom. Syst.* **2019**, *17*, 729–742. [\[CrossRef\]](#)
150. Pathi, S.K.; Kiselev, A.; Loutfi, A. Detecting Groups and Estimating F-Formations for Social Human-Robot Interactions. *Multimodal Technol. Interact.* **2022**, *6*, 18. [\[CrossRef\]](#)
151. Ishiguro, H. Android science: Conscious and subconscious recognition. *Connect. Sci.* **2006**, *18*, 319–332. [\[CrossRef\]](#)
152. Sakamoto, D.; Kanda, T.; Ono, T.; Ishiguro, H.; Hagita, N. Androids as a Telecommunication Medium with a Humanlike Presence. In *Geminoid Studies: Science and Technologies for Humanlike Teleoperated Androids*; Ishiguro, H., Dalla Libera, F., Eds.; Springer: Singapore, 2018; pp. 39–56. [\[CrossRef\]](#)
153. Mori, M.; MacDorman, K.F.; Kageki, N. The uncanny valley [from the field]. *IEEE Robot. Autom. Mag.* **2012**, *19*, 98–100. [\[CrossRef\]](#)
154. Becker-Asano, C.; Ogawa, K.; Nishio, S.; Ishiguro, H. Exploring the uncanny valley with Geminoid HI-1 in a real-world application. In Proceedings of the IADIS International Conference Interfaces and Human Computer Interaction, Freiburg, Germany, 26–30 July 2010; pp. 121–128.
155. Strauch, B. Ironies of automation: Still unresolved after all these years. *IEEE Trans. Hum. Syst.* **2017**, *48*, 419–433. [\[CrossRef\]](#)
156. Lichiardopol, S. A survey on teleoperation. In *Technische Universitat Eindhoven*; DCT report 2007.155; Technische Universitat Eindhoven: Eindhoven, The Netherlands, 2007; Volume 20, pp. 40–60.
157. Siciliano, B.; Khatib, O.; Kröger, T. *Springer Handbook of Robotics*; Springer: Berlin/Heidelberg, Germany, 2008. [\[CrossRef\]](#)
158. Clabaugh, C.; Matarić, M. Escaping oz: Autonomy in socially assistive robotics. *Annu. Rev. Control. Robot. Auton. Syst.* **2019**, *2*, 33–61. [\[CrossRef\]](#)
159. Sheridan, T.B. *Telerobotics, Automation, and Human Supervisory Control*; MIT Press: Cambridge, MA, USA, 1992.
160. Fong, T.; Thorpe, C.E.; Baur, C. Robot, asker of questions. *Robot. Auton. Syst.* **2003**, *42*, 235–243. [\[CrossRef\]](#)
161. Anderson, R. Autonomous, teleoperated, and shared control of robot systems. In Proceedings of the IEEE International Conference on Robotics and Automation, Minneapolis, MN, USA, 22–28 April 1996; Volume 3, pp. 2025–2032. [\[CrossRef\]](#)
162. Fong, T.W.; Thorpe, C.; Baur, C. A Safeguarded Teleoperation Controller. In Proceedings of the IEEE International Conference on Advanced Robotics (ICAR '01), Budapest, Hungary, 22–25 August 2001.
163. Sian, N.; Yokoi, K.; Kajita, S.; Kanehiro, F.; Tanie, K. Whole body teleoperation of a humanoid robot—development of a simple master device using joysticks. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Lausanne, Switzerland, 30 September–4 October 2002; Volume 3, pp. 2569–2574. [\[CrossRef\]](#)
164. Kaneko, K.; Kaminaga, H.; Sakaguchi, T.; Kajita, S.; Morisawa, M.; Kumagai, I.; Kanehiro, F. Humanoid Robot HRP-5P: An Electrically Actuated Humanoid Robot with High-Power and Wide-Range Joints. *IEEE Robot. Autom. Lett.* **2019**, *4*, 1431–1438. [\[CrossRef\]](#)
165. Stanton, C.; Bogdanovych, A.; Ratanasena, E. Teleoperation of a humanoid robot using full-body motion capture, example movements, and machine learning. In Proceedings of the Australasian Conference on Robotics and Automation, Wellington, New Zealand, 3–5 December 2012; Volume 8, p. 51.
166. Crandall, J.W.; Goodrich, M.A. Characterizing efficiency of human robot interaction: A case study of shared-control teleoperation. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Lausanne, Switzerland, 30 September–4 October 2002; Volume 2, pp. 1290–1295.
167. Lu, Y.; Huang, Q.; Li, M.; Jiang, X.; Keerio, M. A friendly and human-based teleoperation system for humanoid robot using joystick. In Proceedings of the 2008 7th World Congress on Intelligent Control and Automation, Chongqing, China, 25–27 June 2008; pp. 2283–2288.
168. Harutyunyan, V.; Manohar, V.; Gezehei, I.; Crandall, J.W. Cognitive telepresence in human-robot interactions. *J. Hum. Interact.* **2013**, *1*, 158–182. [\[CrossRef\]](#)

169. Sakamoto, D.; Kanda, T.; Ono, T.; Ishiguro, H.; Hagita, N. Android as a telecommunication medium with a human-like presence. In Proceedings of the 2007 2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI), Arlington, VA, USA, 10–12 March 2007; pp. 193–200.
170. Mi, J.; Tang, S.; Deng, Z.; Goerner, M.; Zhang, J. Object affordance based multimodal fusion for natural human-robot interaction. *Cogn. Syst. Res.* **2019**, *54*, 128–137. [[CrossRef](#)]
171. Garg, S.; Sünderhauf, N.; Dayoub, F.; Morrison, D.; Cosgun, A.; Carneiro, G.; Wu, Q.; Chin, T.J.; Reid, I.; Gould, S.; et al. Semantics for Robotic Mapping, Perception and Interaction: A Survey. *Found. Trends Robot.* **2020**, *8*, 1–224. [[CrossRef](#)]
172. Osawa, M.; Okuoka, K.; Takimoto, Y.; Imai, M. Is Automation Appropriate? Semi-autonomous Telepresence Architecture Focusing on Voluntary and Involuntary Movements. *Int. J. Soc. Robot.* **2020**, *12*, 1119–1134. [[CrossRef](#)]
173. Takimoto, Y.; Hasegawa, K.; Sono, T.; Imai, M. A simple bi-layered architecture to enhance the liveness of a robot. In Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 24–28 September 2017; pp. 2786–2792.
174. Arita, R.; Suzuki, S. Maneuvering Assistance of Teleoperation Robot Based on Identification of Gaze Movement. In Proceedings of the 2019 IEEE 17th International Conference on Industrial Informatics (INDIN), Helsinki, Finland, 22–25 July 2019; Volume 1, pp. 565–570. [[CrossRef](#)]
175. Ferrell, W.R. Remote manipulation with transmission delay. *IEEE Trans. Hum. Factors Electron.* **1965**, *HFE-6*, 24–32. [[CrossRef](#)]
176. Ferrell, W.R.; Sheridan, T.B. Supervisory control of remote manipulation. *IEEE Spectr.* **1967**, *4*, 81–88. [[CrossRef](#)]
177. Bejczy, A.K.; Kim, W.S. Predictive displays and shared compliance control for time-delayed telemanipulation. In Proceedings of the IEEE International Workshop on Intelligent Robots and Systems, towards a New Frontier of Applications, Ibaraki, Japan, 3–6 July 1990; pp. 407–412.
178. Bejczy, A.K.; Kim, W.S.; Venema, S.C. The phantom robot: Predictive displays for teleoperation with time delay. In Proceedings of the IEEE International Conference on Robotics and Automation, Cincinnati, OH, USA, 13–18 May 1990; pp. 546–551.
179. Uddin, R.; Ryu, J. Predictive control approaches for bilateral teleoperation. *Annu. Rev. Control.* **2016**, *42*, 82–99. [[CrossRef](#)]
180. Lorek, K.S.; Willinger, G.L. A Multivariate Time-Series Prediction Model For Cash-Flow Data. *Account. Rev.* **2016**, *71*, 81–102.
181. Mikolov, T.; Karafiát, M.; Burget, L.; Cernocký, J.; Khudanpur, S. Recurrent neural network based language model. In Proceedings of the INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association, Makuhari, Chiba, Japan, 26–30 September 2010; Kobayashi, T., Hirose, K., Nakamura, S., Eds.; ISCA: Tokyo, Japan, 2010; pp. 1045–1048.
182. Su, H.; Hu, Y.; Karimi, H.R.; Knoll, A.; Ferrigno, G.; De Momi, E. Improved recurrent neural network-based manipulator control with remote center of motion constraints: Experimental results. *Neural Netw.* **2020**, *131*, 291–299. [[CrossRef](#)] [[PubMed](#)]
183. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
184. Lipton, Z.C.; Berkowitz, J.; Elkan, C. A critical review of recurrent neural networks for sequence learning. *arXiv* **2015**, arXiv:1506.00019
185. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to Sequence Learning with Neural Networks. In Proceedings of the Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, Montreal, QC, Canada, 8–13 December 2014; pp. 3104–3112.
186. Mariet, Z.; Kuznetsov, V. Foundations of sequence-to-sequence modeling for time series. In Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics, Naha, Okinawa, Japan, 16–18 April 2019; Volume 89, pp. 408–417.
187. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., Weinberger, K., Eds.; Curran Associates, Inc.: New York, NY, USA, 2014; Volume 27.
188. Zhang, K.; Zhong, G.; Dong, J.; Wang, S.; Wang, Y. Stock market prediction based on generative adversarial network. *Procedia Comput. Sci.* **2019**, *147*, 400–406. [[CrossRef](#)]
189. Akalin, N.; Loutfi, A. Reinforcement Learning Approaches in Social Robotics. *Sensors* **2021**, *21*, 1292. [[CrossRef](#)]
190. Hemminahaus, J.; Kopp, S. Towards adaptive social behavior generation for assistive robots using reinforcement learning. In Proceedings of the 2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Vienna, Austria, 6–9 March 2017; pp. 332–340.
191. Norcio, A.F.; Stanley, J. Adaptive human-computer interfaces: A literature survey and perspective. *IEEE Trans. Syst. Man Cybern.* **1989**, *19*, 399–408. [[CrossRef](#)]
192. McTear, M.F. User modelling for adaptive computer systems: A survey of recent developments. *Artif. Intell. Rev.* **1993**, *7*, 157–184. [[CrossRef](#)]
193. Albrecht, S.V.; Stone, P. Autonomous agents modelling other agents: A comprehensive survey and open problems. *Artif. Intell.* **2018**, *258*, 66–95. [[CrossRef](#)]
194. Rossi, S.; Ferland, F.; Tapus, A. User profiling and behavioral adaptation for HRI: A survey. *Pattern Recognit. Lett.* **2017**, *99*, 3–12. [[CrossRef](#)]
195. Martins, G.S.; Al Tair, H.; Santos, L.; Dias, J. α POMDP: POMDP-based user-adaptive decision-making for social robots. *Pattern Recognit. Lett.* **2019**, *118*, 94–103. [[CrossRef](#)]
196. Xiang, X.; Foo, S. Recent Advances in Deep Reinforcement Learning Applications for Solving Partially Observable Markov Decision Processes (POMDP) Problems: Part 1—Fundamentals and Applications in Games, Robotics and Natural Language Processing. *Mach. Learn. Knowl. Extr.* **2021**, *3*, 554–581. [[CrossRef](#)]

197. Lopes, A.C.; Pires, G.; Nunes, U. Assisted navigation for a brain-actuated intelligent wheelchair. *Robot. Auton. Syst.* **2013**, *61*, 245–258. [[CrossRef](#)]
198. Broz, F.; Nourbakhsh, I.; Simmons, R. Planning for human–robot interaction in socially situated tasks. *Int. J. Soc. Robot.* **2013**, *5*, 193–214. [[CrossRef](#)]
199. Fiore, M.; Khambhaita, H.; Milliez, G.; Alami, R. An adaptive and proactive human-aware robot guide. In Proceedings of the Social Robotics—7th International Conference, ICSR 2015, Paris, France, 26–30 October 2015; Springer: Cham, Switzerland, 2015; pp. 194–203.
200. Huang, C.M.; Mutlu, B. Anticipatory robot control for efficient human-robot collaboration. In Proceedings of the 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, New Zealand, 7–10 March 2016; pp. 83–90.
201. Chiang, H.H.; Chen, Y.L.; Lin, C.T. Human-robot interactive assistance of a robotic walking support system in a home environment. In Proceedings of the 2013 IEEE International Symposium on Consumer Electronics (ISCE), Hsinchu, Taiwan, 3–6 June 2013; pp. 263–264.
202. Lam, C.; Yang, A.Y.; Driggs-Campbell, K.; Bajcsy, R.; Sastry, S.S. Improving human-in-the-loop decision making in multi-mode driver assistance systems using hidden mode stochastic hybrid systems. In Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 28 September–2 October 2015; pp. 5776–5783.
203. Kim, H.G.; Yang, J.Y.; Kwon, D.S. Experience based domestic environment and user adaptive cleaning algorithm of a robot cleaner. In Proceedings of the 2014 11th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), Kuala Lumpur, Malaysia, 12–15 November 2014; pp. 176–178.
204. Matsubara, T.; Miro, J.V.; Tanaka, D.; Poon, J.; Sugimoto, K. Sequential intention estimation of a mobility aid user for intelligent navigational assistance. In Proceedings of the 2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Kobe, Japan, 31 August–4 September 2015; pp. 444–449.
205. Madureira, A.; Cunha, B.; Pereira, J.P.; Gomes, S.; Pereira, I.; Santos, J.M.; Abraham, A. Using personas for supporting user modeling on scheduling systems. In Proceedings of the IEEE 2014 14th International Conference on Hybrid Intelligent Systems, Kuwait, Kuwait, 14–16 December 2014; pp. 279–284.
206. Shen, Q.; Dautenhahn, K.; Saunders, J.; Kose, H. Can real-time, adaptive human–robot motor coordination improve humans' overall perception of a robot? *IEEE Trans. Auton. Ment. Dev.* **2015**, *7*, 52–64. [[CrossRef](#)]
207. Moustiris, G.P.; Geravand, M.; Tzafestas, C.; Peer, A. User-adaptive shared control in a mobility assistance robot based on human-centered intention reading and decision making scheme. In Proceedings of the IEEE International Conference on Robotics and Automation Workshop: Human-Robot Interfaces for Enhanced Physical Interactions, Stockholm, Sweden, 16 May 2016.
208. Schadenberg, B.R.; Neerinx, M.A.; Cnossen, F.; Looije, R. Personalising game difficulty to keep children motivated to play with a social robot: A Bayesian approach. *Cogn. Syst. Res.* **2017**, *43*, 222–231. [[CrossRef](#)]
209. Smith, J.S.; Chao, C.; Thomaz, A.L. Real-time changes to social dynamics in human-robot turn-taking. In Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 28 September–2 October 2015; pp. 3024–3029.
210. Aly, A.; Tapus, A. A model for synthesizing a combined verbal and nonverbal behavior based on personality traits in human-robot interaction. In Proceedings of the 2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Tokyo, Japan, 3–6 March 2013; pp. 325–332.
211. Abdo, N.; Stachniss, C.; Spinello, L.; Burgard, W. Robot, organize my shelves! Tidying up objects by predicting user preferences. In Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA), Seattle, WA, USA, 26–30 May 2015; pp. 1557–1564.
212. Duque, I.; Dautenhahn, K.; Koay, K.L.; Christianson, B. A different approach of using Personas in human-robot interaction: Integrating Personas as computational models to modify robot companions' behaviour. In Proceedings of the 2013 IEEE RO-MAN, Gyeongju, Korea, 26–29 August 2013; pp. 424–429.
213. Klee, S.D.; Ferreira, B.Q.; Silva, R.; Costeira, J.P.; Melo, F.S.; Veloso, M. Personalized assistance for dressing users. In Proceedings of the 7th International Conference on Social Robotics, ICSR 2015, Paris, France, 26–30 October 2015; Springer: Cham, Switzerland, 2015; pp. 359–369.
214. Gao, Y.; Chang, H.J.; Demiris, Y. User modelling for personalised dressing assistance by humanoid robots. In Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 28 September–2 October 2015; pp. 1840–1845.
215. Fischinger, D.; Einramhof, P.; Papoutsakis, K.; Wohlkinger, W.; Mayer, P.; Panek, P.; Hofmann, S.; Koertner, T.; Weiss, A.; Argyros, A.; et al. Hobbit, a care robot supporting independent living at home: First prototype and lessons learned. *Robot. Auton. Syst.* **2016**, *75*, 60–78. [[CrossRef](#)]
216. Ros, R.; Baroni, I.; Demiris, Y. Adaptive human–robot interaction in sensorimotor task instruction: From human to robot dance tutors. *Robot. Auton. Syst.* **2014**, *62*, 707–720. [[CrossRef](#)]
217. Baraka, K.; Veloso, M. Adaptive interaction of persistent robots to user temporal preferences. In Proceedings of the 7th International Conference on Social Robotics, Paris, France, 26–30 October 2015; Springer: Cham, Switzerland, 2015; pp. 61–71.
218. Senft, E.; Baxter, P.; Kennedy, J.; Belpaeme, T. Sparc: Supervised progressively autonomous robot competencies. In Proceedings of the 7th International Conference on Social Robotics, Paris, France, 26–30 October 2015; Springer: Cham, Switzerland, 2015; pp. 603–612.

219. Gordon, G.; Breazeal, C. Bayesian active learning-based robot tutor for children's word-reading skills. In Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, TX, USA, 25–30 January 2015; pp. 1343–1349.
220. Devin, S.; Alami, R. An implemented theory of mind to improve human-robot shared plans execution. In Proceedings of the 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, New Zealand, 7–10 March 2016; pp. 319–326.
221. Karami, A.B.; Sehaba, K.; Encelle, B. Adaptive artificial companions learning from users' feedback. *Adapt. Behav.* **2016**, *24*, 69–86. [[CrossRef](#)]
222. Grosinger, J.; Pecora, F.; Saffiotti, A. Making Robots Proactive through Equilibrium Maintenance. In Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI), New York, NY, USA, 9–15 July 2016; pp. 3375–3381.
223. Müller, S.; Sprenger, S.; Gross, H.M. Online adaptation of dialog strategies based on probabilistic planning. In Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication, Edinburgh, UK, 25–29 August 2014; pp. 692–697.
224. Lim, G.H.; Hong, S.W.; Lee, I.; Suh, I.H.; Beetz, M. Robot recommender system using affection-based episode ontology for personalization. In Proceedings of the 2013 IEEE RO-MAN, Gyeongju, Korea, 26–29 August 2013; pp. 155–160.
225. Sarabia, M.; Lee, K.; Demiris, Y. Towards a synchronised Grammars framework for adaptive musical human-robot collaboration. In Proceedings of the 2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Kobe, Japan, 31 August–4 September 2015; pp. 715–721.
226. Nikolaidis, S.; Kuznetsov, A.; Hsu, D.; Srinivasa, S. Formalizing human-robot mutual adaptation: A bounded memory model. In Proceedings of the 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, New Zealand, 7–10 March 2016; pp. 75–82.
227. Aylett, R.; Kappas, A.; Castellano, G.; Bull, S.; Barendregt, W.; Paiva, A.; Hall, L. I know how that feels—An empathic robot tutor. In Proceedings of the eChallenges e-2015 Conference, Vilnius, Lithuania, 25–27 November 2015; pp. 1–9.
228. Sekmen, A.; Challa, P. Assessment of adaptive human-robot interactions. *Knowl. Based Syst.* **2013**, *42*, 49–59. [[CrossRef](#)]
229. Braithwaite, J.J.; Watson, D.G.; Jones, R.; Rowe, M. A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. *Psychophysiology* **2013**, *49*, 1017–1034.
230. Slater, M.; Brogni, A.; Steed, A. Physiological responses to breaks in presence: A pilot study. In Proceedings of the Presence 2003: The 6th Annual International Workshop on Presence, Aalborg, Denmark, 6–8 October 2003; Volume 157.
231. Meehan, M.; Insko, B.; Whitton, M.; Brooks, F., Jr. Physiological Measures of Presence in Stressful Virtual Environments. *Acm Trans. Graph.* **2002**, *21*, 645–652. [[CrossRef](#)]
232. Pianzola, F. Presence, flow, and narrative absorption questionnaires: A scoping review. *Open Res. Eur.* **2021**, *1*, 11. [[CrossRef](#)]
233. Bulu, S.T. Place presence, social presence, co-presence, and satisfaction in virtual worlds. *Comput. Educ.* **2012**, *58*, 154–161. [[CrossRef](#)]
234. Rhee, T.; Thompson, S.; Medeiros, D.; dos Anjos, R.; Chalmers, A. Augmented Virtual Teleportation for High-Fidelity Telecollaboration. *IEEE Trans. Vis. Comput. Graph.* **2020**, *26*, 1923–1933. [[CrossRef](#)] [[PubMed](#)]
235. Han, J.; Conti, D. The Use of UTAUT and Post Acceptance Models to Investigate the Attitude towards a Telepresence Robot in an Educational Setting. *Robotics* **2020**, *9*, 34. [[CrossRef](#)]
236. Redaelli, C.; Riva, G. Flow for Presence Questionnaire. In *Digital Factory for Human-Oriented Production Systems*; Springer: London, UK, 2011; pp. 3–22.
237. Usoh, M.; Catena, E.; Arman, S.; Slater, M. Using Presence Questionnaires in Reality. *Presence Teleoperators Virtual Environ.* **2000**, *9*, 497–503. [[CrossRef](#)]
238. Lombard, M.; Ditton, T.B.; Crane, D.; Davis, B.; Gil-Egui, G.; Horvath, K.; Rossmann, J.; Park, S. Measuring presence: A literature-based approach to the development of a standardized paper-and-pencil instrument. In Proceedings of the Third International Workshop on Presence, Delft, The Netherlands, 27–28 March 2000; Volume 240, pp. 2–4.
239. Schubert, T. The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. *Z. Für Medien.* **2003**, *15*, 69–71. [[CrossRef](#)]
240. Lessiter, J.; Freeman, J.; Keogh, E.; Davidoff, J. A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence* **2001**, *10*, 282–297. [[CrossRef](#)]
241. Witmer, B.G.; Jerome, C.J.; Singer, M.J. The Factor Structure of the Presence Questionnaire. *Presence Teleoperators Virtual Environ.* **2005**, *14*, 298–312. [[CrossRef](#)]
242. Harms, C.; Biocca, F. Internal Consistency and Reliability of the Networked Minds Social Presence Measure. In Proceedings of the Seventh Annual International Workshop: Presence 2004, Universidad Politecnica de Valencia, Valencia, Spain, 12–15 October 2004.
243. Makransky, G.; Lilleholt, L.; Aaby, A. Development and Validation of the Multimodal Presence Scale for Virtual Reality Environments: A Confirmatory Factor Analysis and Item Response Theory Approach. *Comput. Hum. Behav.* **2017**, *72*, 276–285. [[CrossRef](#)]
244. Hartmann, T.; Wirth, W.; Schramm, H.; Klimmt, C.; Vorderer, P.; Gysbers, A.; Böcking, S.; Ravaja, N.; Laarni, J.; Saari, T.; et al. The Spatial Presence Experience Scale (SPES). *J. Media Psychol. Theor. Methods Appl.* **2015**, *1*, 1–15. [[CrossRef](#)]
245. Heutte, J.; Fenouillet, F.; Boniwell, I.; Martin-Krumm, C.; Csikszentmihalyi, M. EduFlow: Proposal for a New Measure of Flow in Education. Previous Paper. 2014. Available online: <http://jean.heutte.free.fr/spip.php?article201> (accessed on 5 April 2022).
246. Engeser, S.; Rheinberg, F. Flow, performance and moderators of challenge-skill balance. *Motiv. Emot.* **2008**, *32*, 158–172. [[CrossRef](#)]

247. Thissen, B.; Menninghaus, W.; Schlotz, W. Measuring Optimal Reading Experiences: The Reading Flow Short Scale. *Front. Psychol.* **2018**, *9*, 2542. [[CrossRef](#)] [[PubMed](#)]
248. Fu, F.L.; Su, R.C.; Yu, S.C. EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Comput. Educ.* **2009**, *52*, 101–112. [[CrossRef](#)]
249. Nielsen, J. Usability 101: Introduction to Usability. Nielsen Norman Group. Available online: <http://www.nngroup.com/articles/usability-101-introduction-to-usability> (accessed on 7 October 2014).
250. Nielsen, J.; Budiu, R. *Mobile Usability*; Pearson Education: Upper Saddle River, NJ, USA, 2012.
251. Nielsen, J. Usability metrics: Tracking interface improvements. *IEEE Softw.* **1996**, *13*, 1–2. [[CrossRef](#)]
252. Adamides, G.; Christou, G.; Katsanos, C.; Xenos, M.; Hadzilacos, T. Usability guidelines for the design of robot teleoperation: A taxonomy. *IEEE Trans. Hum. Syst.* **2014**, *45*, 256–262. [[CrossRef](#)]
253. Kristoffersson, A.; Severinson Eklundh, K.; Loutfi, A. Measuring the Quality of Interaction in Mobile Robotic Telepresence: A Pilot Perspective. *Int. J. Soc. Robot.* **2013**, *5*, 89–101. [[CrossRef](#)]
254. Kurosu, M.; Hashizume, A. ERM-AT Applied to Social Aspects of Everyday Life. In Proceedings of the Human-Computer Interaction. Theory, Methods and Tools, 23rd HCI International Conference, HCII 2021, Virtual Event, 24–29 July 2021; Kurosu, M., Ed.; Springer International Publishing: Cham, Switzerland, 2021; pp. 280–290.