Somebody That I Used to Know: The Risks of Personalizing Robots for Dementia Care

Alyssa Kubota, Maryam Pourebadi, Sharon Banh, Soyon Kim, and Laurel D. Riek
Department of Computer Science and Engineering
University of California San Diego
{akubota, pourebadi, s1banh, sok020, lriek}@ucsd.edu

Abstract Robots have great potential to support people with dementia (PwD) and their caregivers. They can provide support for daily living tasks, conduct household chores, provide companionship, and deliver cognitive stimulation and training. Personalizing these robots to an individual's abilities and preferences can help enhance the quality of support they provide, increase their usability and acceptability, and alleviate caregiver burden. However, personalization can also introduce many risks, including risks to the safety and autonomy of PwD, the potential to exacerbate social isolation, and risks of being taken advantage of due to dark patterns in robot design. In this article, we weigh the risks and benefits by drawing on empirical data garnered from the existing ecosystem of robots used for dementia caregiving. We also explore ethical considerations for developing personalized cognitively assistive robots for PwD, including how a robot can practice beneficence to PwD, where responsibility falls when harm to a PwD occurs because of a robot, and how a robot can acquire informed consent from a PwD. We propose key technical and policy concepts to help robot designers, lawmakers, and others to develop personalized robots that protect users from unintended consequences, particularly for people with cognitive impairments.

Keywords— Personalized Robots, Cognitively Assistive Robots, Robot-Delivered Health Interventions, Robot Ethics, Healthcare Robotics, Human-Robot Interaction

Kubota, A., Pourebadi, M., Banh, S., Kim, S., and Riek, L. D. "Somebody That I Used to Know: The Risks of Personalizing Robots for Dementia Care". In *We Robot 2021*.

1 Introduction

Imagine the following scenario: A person with dementia (PwD) lives at home with a full-time caregiver. That caregiver is overburdened, stressed, and is an older adult with their own health problems. This scenario is experienced by over 16 million informal dementia caregivers in the US, and this number will only continue to grow, representing a global health emergency [5].

Over the past 20 years, many researchers have explored the use of assistive robots that can aid both PwD and their caregivers with a range of daily living tasks, including to provide companionship, deliver cognitive stimulation, and support household chores (see Figure 1) [24, 47, 50, 75, 89, 92, 138, 141]. For example, PARO (see Figure 1.3) is a robotic seal which has been shown to provide companionship to PwD and reduce stress, anxiety, and pain among PwD and their caregivers [44, 93, 138]. In addition, researchers have used socially assistive robots to administer cognitive stimulation and training (i.e., a cognitively assistive robot (CAR)) with the goal of slowing the progression of dementia and/or reducing the severity of its impact [8, 24, 75, 127, 132]. For instance, JESSIE (see Figure 1.1) teaches people metacognitive strategies to reduce the impact of cognitive impairment on their daily lives [75].

For these assistive robots, a key concept discussed in the health technology community is personalization, which reflects how well a system can adapt to a person longitudinally. Personalization offers many benefits, such as improving adherence to cognitive interventions, increasing engagement with intervention content, and enabling goal-oriented health management [76, 107, 128].

Personalizing assistive robots for PwD is especially important due to the progressive and unpredictable nature of dementia, as one's individual needs and preferences will evolve over time [50, 78].

Figure 1: Three exemplar robots used to support PwD and their caregivers. From left to right: 1) **JESSIE** is a tabletop robot developed by our lab, and is used to support people with mild cognitive impairment and early-stage dementia. It provides personalized, adaptive cognitive training to help teach users metacognitive strategies to minimize the impact of cognitive impairment on their daily lives [75]. 2) **Spoonbot** is a tabletop robot radio developed by our lab, and is used to support people with late-stage dementia who have trouble eating. It leverages embodied cueing, mimicry, and music to encourage eating [50]. 3) **PARO** is a robotic seal and is used to stimulate social interaction and support therapy for PwD, including robot-assisted therapy and sensory therapy [110, 138].

For example, for people with early stage dementia, a robot can interact verbally with someone (e.g., provide medication reminders) [79], whereas for those with late stage dementia verbal prompts will not work, and physical and nonverbal aural cues are more appropriate [50, 58]. Spoonbot (see Figure 1.2) is an example of a robot our team built for people with late-stage dementia - it uses mimicry cues and a person's favorite music to assist with eating. Therefore, it is important to adapt to a PwD's physical and cognitive abilities, personal preferences, care setting, and other life circumstances [78].

Roboticists have made great strides in developing personalized systems for PwD and their caregivers. However, a large body of this work has been either primarily technology-focused or health-outcomes focused, yet there's a growing need for further investigation into the potential negative consequences assistive robots could have on this population. For example, researchers have raised concerns about some robots used in dementia caregiving, such as PARO use being associated with irritability, hallucinations, and disinhibition among people with severe dementia, or overstimulation of PwD in group settings [69, 131]. PwD are a vulnerable population who are already at high risk of manipulation and abuse [88], so one must think critically about how these robots could cause harm, and possible means for mitigation.

We are currently at an inflection point, where it is becoming relatively easy and inexpensive to develop and deploy CARs to deliver personalized interventions to PwD, and many companies are vying to capitalize on this trend. However, it is important to carefully consider the ramifications: What are the potential consequences of introducing underdeveloped personalized CARs to care for PwD? Furthermore, what are some unintended consequences of a highly personalized CAR for PwD?

In this article, we draw upon empirical data from our own work, as well as from the literature, to explore these questions. We contextualize concerns regarding inaccurate personalization of CARs for PwD, and the potential unintended consequences of personalizing robot behavior accurately. We also propose key technical and policy concepts to enable robot designers, law-makers, and others to develop CARs that protect users from unintended consequences, particularly those designed for people with cognitive impairments. We hope that our work will inspire roboticists to consider the potential risks and benefits of robot personalization, and support future ethically-focused robot design.

2 Background

2.1 Dementia and Caregiving

Dementia is an irreversible, neurodegenerative syndrome characterized by cognitive, functional, and/or behavioral decline. Worldwide, approximately 50 million people live with dementia, most of whom are older adults. It is most commonly caused by a neurodegenerative disease such as Alzheimer's disease, vascular disease, or Lewy body disease. Dementia is progressive, and symptoms can range across the spectrum from early stage (e.g. difficulty with attention or problem solving) to late stage (e.g. loss of communication abilities). The decline in cognitive abilities such as reasoning and short-term memory can lead to hazardous behaviors such as wandering or driving without supervision, and leave PwD particularly susceptible to domestic or financial abuse, neglect, and exploitation [78].

As PwD lose their ability to live independently, they become reliant on caregivers to complete activities of daily living (ADLs) (e.g. personal hygiene, eating) and instrumental ADLs (IADLs) (e.g. managing medication, scheduling medical appointments) [78, 94]. As a result, family members (e.g. spouses, children) must often assume the role of informal caregivers and shoulder the responsibility of caring for a PwD [78]. However, caregiving can place heavy emotional, mental, physical and financial strain on an individual, placing them at higher risk of additional comorbidities such as cardiovascular diseases and depression [19]. Furthermore, this transition to a caregiving role can place strain on the relationship between PwD and their caregivers, which can cause feelings of guilt, anxiety, and depression in both a PwD and their caregivers [42].

Person-centered care is one predominant care approach that can help maintain the relationship between a PwD and their caregivers by encouraging caregivers to recognize the personhood and individuality of a PwD. As a strengths-based approach, person-centered care recognizes an individual's goals, abilities, and preferences, such as by understanding their culture or building on their strengths and current abilities, rather than trying to replace the abilities they have lost, to promote their well-being [36, 38, 78].

One important aspect of person-centered care is supporting the autonomy of PwD. Being active in daily decision making can help a PwD preserve their dignity and identity, which can help them lead more full and rewarding lives [38, 78]. These decisions may be major, such as deciding which health interventions to receive, or relatively minor, such as choosing what food to eat. It is generally agreed that PwD should be able to make and act on their own decisions whenever possible, and caregivers should structure interactions to support the autonomy of PwD.

However, as dementia progresses, PwD often begin to lose their capacity to make or communicate decisions. Thus, they may not know or be able to reason about what is best for their health. Caregivers may be forced to choose between supporting a PwD's autonomy vs. ensuring their health or safety. For example, if a PwD refuses to maintain basic personal hygiene (e.g. bathing), should a caregiver respect their desire to not do so and risk causing harm (e.g. a urinary tract infection, social ostracism), or should they override the PwD's wishes and force them to complete these activities [119]? Neither scenario is ideal for satisfying both the PwD's autonomy and health, so there is much debate surrounding whether to prioritize respecting a person's autonomy or abiding by the principle of non-maleficence (i.e. preventing harm). The caregiver's decision may change based on culture, situation, and personal preference, but a PwD's independence and privacy are often considered secondary to harm prevention [38, 52].

2.2 Assistive Robots for People with Dementia

There are a wide range of technologies that can support and enhance the care of PwD, as well as extend their independence. These technologies can be categorized into devices used "by" PwD such as for prompts and reminders, devices used "with" PwD such as for communication with caregivers, or devices used "on" PwD such as for location or activity monitoring (see Table 1) [45, 52]. Commercially available assistive technologies used "by" PwD include memory aids (e.g. electronic medication reminders), while those used "with" PwD may include communication aids (e.g. telephones, telepresence robots) or collaborative devices (e.g. electronic games, reminiscence software) [45, 122]. Caregivers may also use products "on" a PwD for safety and security, such as wearable devices to detect if a PwD has fallen or to track their sleep patterns [84].

Many types of robots exist to cognitively and socially assist PwD (see Figure 1). These *robot-delivered health interventions* have many benefits, including the potential to expand access to healthcare by extending it into a person's home, reducing treatment time and cost, and prolonging a PwD's independence [47]. Robots can provide support on multiple dimensions, including socially and cognitively.

Robots can also provide social support and provide non-pharmaceutical therapeutic interventions for PwD. Often, zoomorphic robots such as PARO or AIBO act as companions for PwD, or therapists may use them to augment interventions such as animal-assisted therapy or multi-sensory behavior therapy [61, 90]. These robots can help reduce negative feelings such as stress and anxiety among PwD and caregivers, and even improve their mood [44, 67, 93, 96].

Researchers have also explored the use of CARs

Table 1: Assistive technologies for PwD can be categorized into a) devices used "by" PwD, b) devices used "with" PwD, and c) devices used "on" PwD.

Type of Device	Description	Example
Devices used "by" PwD	Devices that PwD can use independently to support the completion of everyday activities.	Electronic medication reminders, reminder alarms, memory aids
Devices used "with" PwD	Devices that are used to connect with PwD, such as for communication with caregivers or collaborative devices.	Communication aids: telephones, telepresence robots Collaborative devices: electronic games, reminiscence software
Devices used "on" PwD	Devices that are used to monitor PwD for safety and security, and provide alerts if necessary.	Fall detectors, Activity monitors, GPS locators, Geofencing products

to support cognitive training and cognitive stimulation among PwD which can help slow the progression of the disease [76, 132]. These robots can remind PwD of appointments, medication, and dietary requirements to reduce reliance on their memory [91]. In addition, they may assist PwD with cognitive training games to support their memory, or accompany human clinicians with memory training programs [98, 126]. Researchers are also exploring the use of robots to teach PwD metacognitive strategies which help strengthen memory, planning, and executive functioning in order to help them manage their impairment in their daily life [75]. Often, the cognitive support that an assistive robot provides can extend or supplement a health intervention.

However, the majority of these technologies have been designed and developed without thorough consideration or understanding of the needs and perspectives of PwD, particularly those in the later stages of their disease. In addition, many commercial technologies center themselves around PwD's cognitive limitations, rather than their strengths which can lead to them being stigmatized and disempowered [50, 66].

Fortunately, researchers are increasingly adopting more inclusive approaches when designing robots for PwD through a *critical dementia* lens [50, 80]. Critical dementia encapsulates person-centered dementia care, and focuses on understanding and supporting the strengths and personhood of PwD in technology design. It explores how embodiment, context, and emotional and sensorial experience impact how PwD interact with the world around them [80]. It draws from approaches including participatory design, which aims to involve all stakeholders (e.g. PwD, caregivers, clinicians) throughout the design/development process in order to ensure that the end product is usable and valuable to them [83, 114]. It also includes user-centered design, which prioritizes the interests and needs of users and entails gathering iterative feedback at each stage of the development process [1, 31]. These approaches enable technology creators to move away from a deficit model of aging, which focuses on a person's potential disabilities and loss of ability, and instead incorporate a social model of aging, which better captures the preferences and contexts of users [81]. This framing can help promote the dignity and personhood of PwD when designing assistive robots.

While it is vital to include the perspectives of PwD and caregivers in the development of robots, their respective values may not always align, particularly in relation to a PwD's autonomy. For example, caregivers may use cameras to surveil a PwD to ensure their safety (e.g. see if they are wandering or have fallen) which can infringe upon the PwD's privacy. Caregivers may also imagine using robots to encourage PwD to do something that they do not want to do (e.g. eat, bathe), or prevent them from doing something unsafe that they want to do (e.g. go out alone, eat unhealthy food) which can limit their autonomy [89, 95, 119]. Thus, it is important to also consider these design tensions and how they might impact the autonomy of a PwD and the PwD-caregiver relationship. We further explore this in Section 3.2.

2.3 Personalization of CARs that Deliver Health Interventions

To ensure assistive robots are usable and acceptable for individuals living with dementia, it is critical that the robots are personalized. Personalization is tailoring a health intervention or system to suit an individual's factors such as their preferences, abilities, and goals, and it reflects how well a system can adapt an intervention to a person longitudinally. Personalization is essential in this space because there is no singular experience shared by everyone living with or caring for someone with dementia. Personalization can maximize the utility and efficacy of interventions for each PwD's individual situation by enabling assistive robots to address the heterogeneity of PwD including cultural and personal backgrounds, living situations (e.g. at home, in a long-term care facility), how dementia progresses in different people, and individual preferences. For instance, an assistive robot can be personalized to a user physically (e.g. adjusting movement speed, proxemics), cognitively (e.g. adjusting the difficulty level of cognitive training tasks), and socially (e.g. referring to a person by name) [76, 109]. In this work, we primarily focus on the personalization of CARs that deliver health interventions [132].

Personalized CARs offer many benefits, including improving adherence to and adoption of an intervention, as well as adoption of and engagement with the technology. For example, cognitive stimulation is most effective and enjoyable if it meets a person where they are in terms of their cognitive abilities [39, 40]. Tapus et al. [127] demonstrated that adjusting the difficulty of a robot-delivered cognitive stimulation game to a PwD's performance improved their overall task performance, engagement with the intervention, and enjoyment during the task. In contrast, if a health intervention is not personalized, it can provoke frustration and depression in a person with cognitive impairments and their caregivers [6, 118].

Research also demonstrates that adapting robot behavior to an individual can help improve its adoption and engagement with users. For example, a robot can adapt its behavior in real time to a user to maintain engagement, such as changing its tone of voice to draw their attention back if they are distracted during an interaction. This can help maintain engagement with users for longer periods of time, and thus improve retention of material (such as in a therapeutic intervention) [124]. A robot may also adapt its behavior to be more acceptable to a user, such as adjusting its communication style to be more passive or assertive depending on a user's cultural background or which they respond better to [109].

There are also health applications for which personalization to PwD is necessary. For instance, reminiscence therapy encourages PwD to recall memories from their past. So, robots that provide this intervention must have some knowledge of a user's history in order to ask relevant questions and guide the therapy. The MARIO robot is one such example, which could store and retrieve user-specific knowledge provided by family members and caregivers in order to facilitate reminiscence therapy [4].

Robots that can autonomously personalize their behavior are particularly important in this space to fulfill the needs of PwD. The ability to adapt with little to no input from users is especially important when users have low technology literacy and do not have the time or resources to learn how to use the system, as is often the case for clinicians, informal caregivers, and PwD [51, 75].

2.4 Key Technical Concepts for Personalization

From a technical perspective, developers often use machine learning (ML) algorithms to enable robots to autonomously personalize their behavior to users. This can be decomposed into two main phases: preference learning and behavior adaptation. In this section, we will provide a brief overview of these topics and why existing computational approaches may not be appropriate for use with PwD.

One technique for personalizing robots to an individual is to learn and understand what that person's likes and dislikes are, i.e., learning their preferences. **Preference learning** aims to predict what a person will prefer based on their known preferences, often inferred from previous behavior [41]. For instance, in the context of assistive robots for PwD, a robot might take note of songs that elicited a positive response in order to play new music for a PwD. Common computational approaches to preference learning include classification algorithms such as *k*-nearest neighbors and decision trees [41]. While most work in preference learning is limited to only ranking a specific set of items, more recent work aims to infer a user's underlying preferences so that learned preferences can be generalized across contexts [142, 143, 144]. For a more detailed discussion of current research in preference learning, please see [41, 140].

Once a system has an understanding of a person's preferences, it can adapt its behavior to suit those preferences. **Behavior adaptation** refers to how a robot adjusts its behavior, usually in response to external stimuli. This adaptation can occur over short periods of time (e.g. making a noise to draw attention to itself if a user is distracted), or longer periods of time (e.g. adopting an encouraging personality if a user responds better to that during a therapy). Reinforcement learning (RL) approaches are among the most common for autonomous behavior adaptation, though researchers are also exploring other methods such as neural networks and Gaussian processes [76]. Inverse RL is another approach which enables systems to learn from human experts; for instance, a therapeutic robot might observe how a human therapist interacts with a user in order to learn how it should behave. For a more thorough overview of current research in robot behavior adaptation, please see [76, 109].

While existing approaches to preference learning and behavior adaptation have proven effective for applications such as post-stroke rehabilitation and teaching social skills to children [10, 105], most may not be appropriate for longitudinally personalizing robots to PwD for a few reasons.

First, the majority of approaches assume people's preferences and abilities stay constant. However, learning the preferences of PwD can be difficult due to the complex and fluctuating nature of dementia. For example, many PwD experience "sundowning," or increased confusion, anxiety, and agitation later in the day [71], as well as fluctuating levels of lucidity, which may make it challenging for robots to learn what types behaviors to communicate when. Furthermore, one cannot assume that a person will have the same preferences over time or in different contexts. This concern is especially true for PwD whose cognitive abilities may change dramatically as their condition progresses, so a personalized assistive robot must be able to keep up with these changes. For instance, the roles of an assistive robot may transition from delivering a one-on-one cognitive intervention, to observing and sharing information with a caregiver as a PwD needs more support from caregivers as their condition progresses.

Next, many approaches suffer from the "cold start" problem, where a system must begin interacting with a user with no prior knowledge about them [112]. In order to learn more about a user, many approaches rely on exploration of possible actions. However, depending on the possible actions and the context, acting without knowing the preferences or abilities of a PwD may have the potential to harm a PwD. For instance, a robot may need to know a person's level of dementia, tolerance for sensory input, or emotional state to avoid overstimulating them or causing distress during an interaction.

Finally, there are existing computational approaches that learn from human experts (e.g. inverse RL). These approaches require stakeholders to commit much time and effort to use them effectively. However, in the context of assistive robots for PwD, these experts are often caregivers and clinicians who are already overburdened, and may lack the time and technical literacy to communicate their expertise to a robot. Thus, these robots may not interact with a PwD to their full potential or in the way that these experts intend.

3 Risks to personalizing robots for people with dementia

While personalizing robots offers many benefits, there are also risks associated with doing so, even for people without dementia. Inherently, personalization requires the collection of personal information, including health-related information, which raises privacy concerns. Furthermore, these data are often collected longitudinally, fused with other data, and then used by other machine learning systems to infer and predict behavioral patterns of individuals. This not only raises the risks of bias and proxy discrimination [101], but also violates users' ability to provide informed consent as they are unwitting recipients of these opaque systems [68].

Personalizing technology to users has led to a rise in concerns such as privacy violations, over attachment to the technology, "echo chambers" (i.e. only conveying content that reinforces a user's existing beliefs), and manipulation of users [55, 60, 102]. For instance, social media platforms have become adept at presenting users personalized content in order to maintain engagement with the platform. They can even use a person's personal information to show them targeted advertisements in order to maximize advertisement revenue, sometimes at the cost of a user's well-being, through the dissemination of inaccurate information or falsely advertised products [12]. In addition, researchers have identified content personalization as a mechanism that has amplified extreme behavior among radicalist groups, including violent extremists, by enabling large-scale personal expression and collective action with little moderation [11, 129, 139].

The physical embodiment of robots introduces additional concerns that are not present in virtual systems. Research shows that a robot's physical embodiment affords it many advantages that can increase engagement and trustworthiness in social interactions (e.g. richer communication channels, physical presence) [29]. However, these characteristics can be problematic if used in a careless or manipulative manner. For instance, perceived trustworthiness of a robot based on its appearance can significantly influence a user's intention to purchase the device [120] or cause them to find a robot more authoritative [54]. Thus, a robot could exploit a user's trust and manipulate them to behave in ways they might not otherwise (e.g. share personal information, purchase products). Researchers predict that robots will be used to surveil users and market products to them, but with access to far richer and more intimate data than can be gathered by a web-based system, resulting in more persuasive advertisement [32].

Personalizing robots could further exacerbate these risks by leading to the development of "Spybot" robots which gather personal information and can lead to more effective deception by "Scambot" robots or manipulation by "Nudgebot" robots [55]. For instance, a CAR might be perceived as more trustworthy by a PwD if it resembles a family member or clinician, which could inadvertently deceive users and give the robot more authority [89].

Lying and deception are widely discussed concerns in both the dementia caregiving and personalized technology communities [13, 33, 55, 87, 96]. In human-robot interaction, deception can occur if a robot leads someone to believe something that is not true. This deception may happen intentionally or unintentionally, and there are many ways it might occur when interacting with PwD, including Turing Deceptions and misconceptions of a robot's capabilities.

Turing Deceptions: PwD may experience Turing Deceptions when interacting with a robot, i.e. believe they are interacting with a human when in fact they are interacting with a robot, and assume the robots have their own motives, goals, beliefs, and feelings [97, 105]. For instance, if a robot visually or aurally resembles a trusted caregiver or clinician, PwD may be more willing to cooperate with the robot, share personal information with it, or otherwise act in a way that they would not otherwise [89, 104]. While this may be desirable in some cases (e.g. using a caregiver's face or voice to encourage PwD to return to bed, take medication, or eat [89, 95]), this is also a form of deception.

Misconceptions of a robot's capabilities. Another major source of unintentional deception is a disconnect between the actual capabilities of a robot and the capabilities that users think it is capable of. This disconnect is problematic because over- or underestimation of a robot's capabilities can impede users from making informed decisions regarding how robots are involved in their care [136]. This can also affect the level of trust that a user has in a robot, which may lead users to overtrust it and attribute it too much authority during an interaction, or undertrust it and not follow the guidance it provides. Either scenario may result in negative health outcomes, or misuse of the robot [7, 145].

Vandemeulebroucke et al. [136] suggest that increased use and familiarity of robots earlier in life can moderate such deception. However, the memory challenges of PwD may prohibit them from becoming familiar with a robot in this way once the condition has progressed. To proactively help circumvent this, an increasing number of older adults integrate assistive technology into their lives in preparation for the potential development of future memory challenges [30]. In addition, dementia community health workers and family caregivers suggest incorporating features that PwD may already be familiar with (e.g. touch screens, common objects) into robot design in order to increase their usability and acceptability among PwD [50, 89]. Furthermore, Moharana et al. [89]
found that PwD have greater trust in robots that resemble people they are already comfortable and familiar with. Integrating familiar features into robot design can help convey its capabilities to a PwD and build trust between a robot and PwD.

Personalizing CARs to PwD amplifies these concerns and introduces many others. In this paper, we identify and discuss four major risks of personalizing CARs to PwD (see Figure 2), which include: 1) Safety risks that arise from inaccurate personalization, 2) (Human) autonomy infringement risks, 3) Social isolation risks, and 4) Risks of being taken advantage of due to dark patterns in robot design. To support discussion of these risks, we introduce three exemplar robots representative of those currently in use in dementia care, as shown in Figure 1.

3.1 Risk 1: Inaccurate personalization can lead to safety risks

While carrying benefits such as autonomy, ML approaches have the potential to cause physical or mental harm when they are not adequately personalized to PwD, such as not understanding the context of care, perpetuating bias, or simply being inaccurate. One of the most significant risks of inaccurately personalized CARs is providing a PwD with inadequate care or care that is misaligned with their stage of dementia. A robot may not necessarily understand the complexities of care, so there is a rising concern that automating these decisions without human supervision will cause them to be ineffective or harmful. For example, the automation of a medical diagnosis system may cause physical harm without supervision of a human medical professional [14], demonstrating the challenges of health automation technology even before introducing the additional complexity levels of physical embodiment, home settings, or cognitive impairments. Also, the robot may provide self-care instructions that do not account for a PwD's comorbidities, which may contribute to harm.

As a PwD's condition progresses, they may require different levels of support to use or understand a CAR effectively. However, if a robot fails to adequately understand or adapt to a PwD's con-

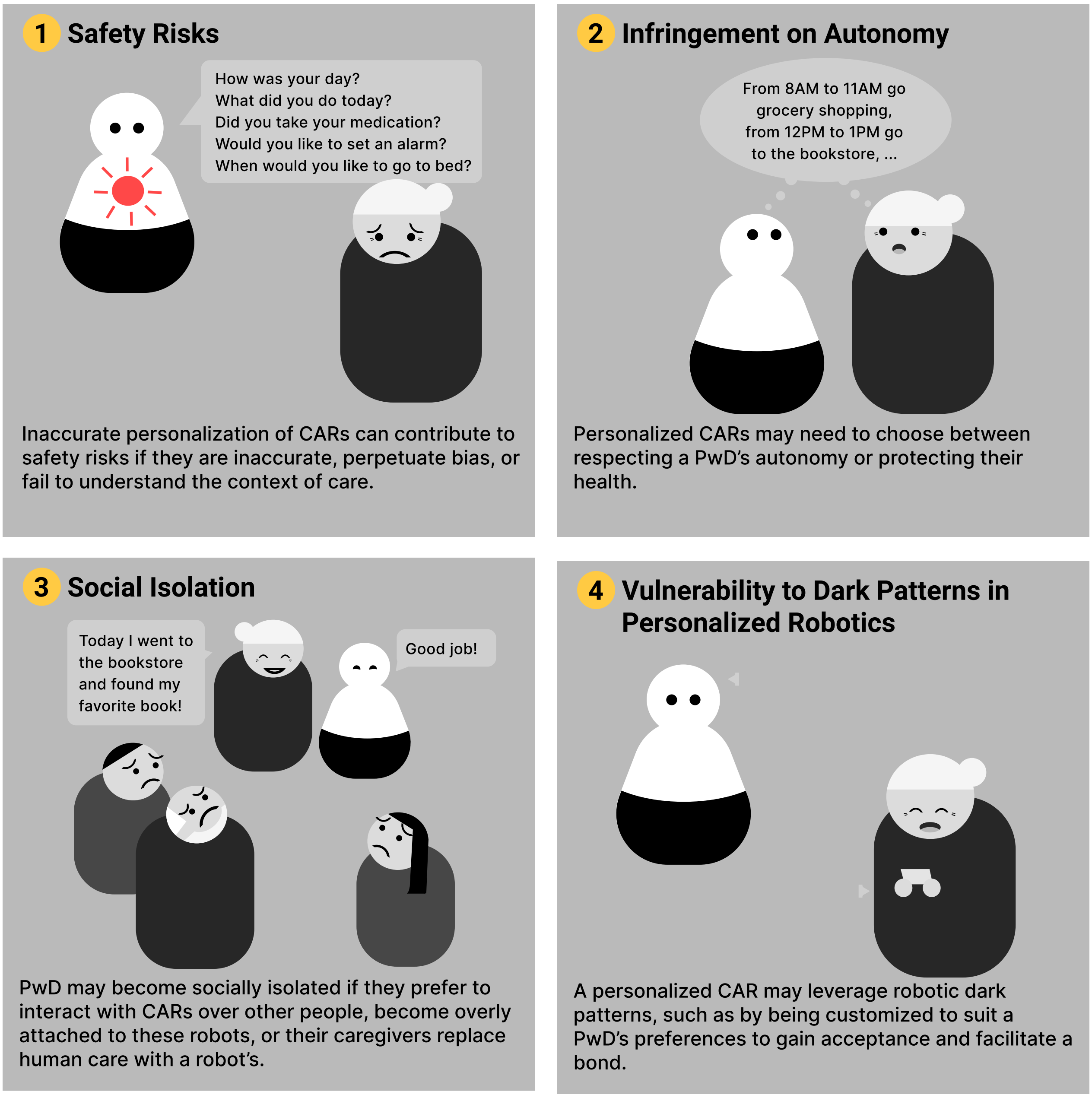

Figure 2: Personalizing CARs to PwD introduces many ethical concerns, including 1) Safety risks that arise from inaccurate personalization, 2) (Human) autonomy infringement risks, 3) Social isolation risks, and 4) Risks of being taken advantage of due to dark patterns in robot design.

ditions, this can limit their ability to fully utilize an assistive robot. This failure may be considered an error of omission (if the robot did not adapt its behavior at all) or commission (if the robot adapted its behavior incorrectly) [108]. In either case, these errors may have negative effects including reducing the usability of the robot, which can reduce a PwD's use of the robot, lower a PwD's self-confidence in their cognitive abilities, and possibly lead to depression and anxiety [6, 43]. PwD may also be unable to communicate what they want or need to the robot if it fails to account for their physical or cognitive considerations, which can leave PwD feeling as though they have lost their autonomy and dignity [43]. Although this problem could be avoided through caregiver supervision, caregivers are often already overburdened by existing caregiving responsibilities, and having to monitor an adaptive robot would simply add further cognitive load.

ML algorithms for therapeutic interventions could demonstrate biases that unintentionally exclude or harm PwD. In the case of robots for PwD, most existing work frames dementia and aging as a series of losses, rather than acknowledging the full life and identity of the PwD. This narrow understanding of PwD may perpetuate existing biases about PwD, limit an algorithm's performance, and ultimately place the mental and physical health of PwD at risk [125]. Thus, it is vital to develop concrete guidelines for assessing the potential mental and physical impact of inaccurate personalization on PwD and ways to avoid harm.

While researchers will ideally test ML algorithms extensively before using them in real world applications, there are still limitations to what these algorithms can achieve, such as for uncommon scenarios (i.e. "edge cases") or populations not reflected in test data (e.g. PwD). Developers may be tempted to naively apply an algorithm to the context of personalizing robots to PwD. However, if that algorithm was not tested or validated with this population and possible scenarios were not considered in the design while personalizing the algorithm, it may lead to inaccurate physical and mental health assessments of PwD (e.g. depression, detection of pain in non-verbal individuals) which can cause serious harm to their mental and physical health. For instance, pre-trained models of facial analysis technology such as Facial Alignment Network (FAN) achieve relatively high accuracy for older adults without dementia, but the accuracy drops significantly for PwD [125].

Caregivers may rely on such algorithms to automatically identify and alert them of agitation and aggression in PwD so they can reliably intervene in a timely and appropriate manner. However, if the algorithm was not developed with or trained on data from PwD, a system may not alert a caregiver of agitation or aggression until the harm has already occurred (e.g. causing distress, emotional withdrawal, physical harm) [72]. On the other hand, a system may give the caregiver false alarms, which may cause them unnecessary stress and cause them to become desensitized to alarms, so they are unprepared to react in the case of a true agitated or aggressive episode.

3.2 Risk 2: Infringement on the autonomy of PwD

As discussed in Section 2.2, respecting the autonomy of PwD empowers them to construct their lives based on their values and personality. This entails supporting their freedom, independence, and privacy. Although PwD may not be able to execute all their decisions (i.e., *agent autonomy*), they often can express their interests (i.e., *choice autonomy*), which caregivers or assistive robots can consider in order to make choices that support their values [119]. The choices a person makes reflect their unique identity, personality, and lifestyle (i.e., *actual autonomy*). It is important for family members, caregivers, and technology developers to support PwD's choice autonomy and respect their actual autonomy longitudinally [119]. Studies suggest using personalized assistive robots can promote the autonomy of PwD and support person-centered care [65].

However, just as caregivers may be forced to choose between respecting the autonomy or protecting the health and safety of a PwD, personalized assistive robots may also be forced to make this decision. This places personalized robots for PwD in a peculiar position where their actions (or lack thereof) may depend on how autonomy and control are distributed between themselves and PwD.

Consider a scenario encountered in our prior work, told to us by a dementia caregiver [89]. A PwD, who also has diabetes and cancer, was feeling ill, and only wanted to eat popsicles, to the point where she wanted to have more than ten each day. Even though the popsicles brought her joy, consuming too many popsicles upset her stomach and detrimentally affected her blood sugar. Caregiver participants in our study suggested an assistive robot could "be the bad guy," denying the frequent popsicle requests, so that the caregiver did not have to. On the one hand, offloading emotional labor from an overwhelmed caregiver may be beneficent; however, the scenario raises questions regarding the autonomy of the PwD. A fully autonomous robot may, indeed, be configured to limit the consumption of sweets and keep the PwD on a strict diet to promote their wellbeing. However, situations like this can also create conflicts between the autonomy of the PwD and the beneficence of the caregiver [119].

One approach to sharing autonomy is to always give users ultimate control of assistive robots. In the context of supporting older adults, Sharkey and Sharkey suggest this will have positive effects on a user's sense of autonomy and can reduce the risks of infringing on their privacy [116]. However, PwD may have impaired judgment, so respecting their autonomy may be at the cost of their own health. Furthermore, personalized robots that must wait for approval from a PwD before acting may be limited in their ability to protect users, such as if the robot recognizes a dangerous situation but cannot autonomously take steps to prevent it (e.g. if a PwD tries to reach a tall cupboard by climbing on a precarious chair [116]). Thus, it is infeasible and potentially harmful to give PwD full control over assistive robots.

On the other hand, assistive robots may act in opposition to or without PwD feedback. Sharkey and Sharkey [116] suggested that assistive robots might help PwD as "autonomous supervisors" to help protect the safety of PwD. This can happen by designing robots such that they autonomously take steps to prevent the dangerous situation, or restraining PwD from performing a potentially dangerous action. In the case of the PwD who loves popsicles, a personalized robot might recognize her dietary restrictions and choose to protect the PwD's health by limiting her popsicle consumption, even if doing so defies her wishes. As PwD often have impaired judgment abilities, a robot may be unable to obtain accurate (or any) feedback when trying to make decisions, which further raises the risk of autonomy infringement [56, 85]. However, the ethical problem here is that restraining a PwD to prevent potential harm could be "a slippery slope towards authoritarian robotics" [116].

Another alternative to sharing the autonomy between assistive robots and PwD suggests developing robot technology with the aim of having robots provide care to the older adults while considering their autonomy. This means that instead of overriding the PwD, the system should allow them to make decisions about their daily activities, and warn them to stop a potentially dangerous activity, if needed [63]. Including PwD in decision making can decrease infantilization and improve PwD's independence [63]. So rather than just outright refusing to give the PwD a popsicle, the robot might try to explain why it cannot give her a popsicle right now or distract her from the topic altogether [50]. While a downside to this is that the PwD may not pay attention to the robot, understand the suggestions the robot is making, or simply not trust the capabilities of the robot to make reliable suggestions, a personalized CAR may be able to more effectively understand and coordinate with a PwD to reach a satisfactory outcome.

How a personalized robot should behave in these situations is still an open question, as each of these approaches requires considering the tradeoffs between a PwD's autonomy and safety. While a personalized assistive robot will likely be forced to decide which tradeoffs to make, the ideal solution will be much more nuanced than simply adapting to a user's preferences. However, as with human caregivers, a robot will likely be expected to prioritize a PwD's health and safety over their autonomy. Personalized robots may be particularly adept at distracting or redirecting a PwD from their potentially unsafe desire (e.g. to eat a popsicle), such as by knowing what alternatives they might like or how to change the topic of conversation. This can preserve the health of PwD, but it effectively restricts their autonomy. Thus, roboticists need to consider developing personalized assistive robots that suit a PwD's individual personality to support their actual autonomy as well as their needs and choices to support their choice autonomy. However, when designing CARs, roboticists should be mindful of the fact that the cognitive limitations that restrict a PwD's agent autonomy may also limit how they can interact with these robots.

3.3 Risk 3: Social Isolation

Social isolation has significant effects on the physical and mental health of PwD. Anxiety, boredom, depression, and lack of meaningful activities are prevalent among PwD living in assisted living facilities [3, 18, 53]. Given that having strong social connections has been shown to protect against various adverse health outcomes, including depression [111], it is important that the social support needs of PwD are considered when designing assistive technology.

In an effort to encourage social connectedness for PwD, caregivers use technology such as robots to connect PwD to family members [92] and to provide companionship to PwD [69]. To better address the social support needs of a PwD and encourage social interactions, many researchers are exploring how to personalize these systems and tailor them to an individual's interests and capabilities. However, while personalized robots are intended to be more effective, using them in the context of dementia care may pose more risks than benefits, including preferring to interact with robots over other people, over-attachment to these robots, and the supplanting of human interaction by a robot.

A personalized assistive robot that aims to combat social isolation among PwD may have the unintended consequence of over-attachment. For instance, it might emulate the "perfect companion" by learning the likes and dislikes of a PwD. A PwD might find the companionship of such a robot to be preferable to another person's, so they might choose the company of these robots over other people. This problem becomes even more pronounced as people with cognitive impairments may believe they are interacting with another person when they are actually interacting with a robot (i.e., a Turing deception) [104]. While some people believe that robots can help mitigate feelings of isolation and help improve social connectedness (e.g. serving as a social facilitator, establishing virtual visits with family and friends) [116], many scholars question whether the relationship between a PwD and a robot can be considered meaningful or moral [96, 116, 136].

If PwD prefer the companionship of a robot to that of another person, there is also the risk of PwD becoming overly attached to the robots. A personalized CAR could understand how and when a PwD would be receptive to social cues such as touch and eye contact in order to establish and maintain a bond with the PwD [133]. Over attachment can lead to distress and loss of therapeutic benefits when robots are taken away, and further exacerbate social isolation [136].

As CARs become more adept at providing care and more personalized to suit a PwD's individual needs, they can help relieve some care responsibilities of human caregivers [136]. However, some researchers are concerned that robots that are adept at providing care to PwD could lead to reduced interaction between a PwD and caregivers [136]. Furthermore, a robot that is highly personalized to a PwD may be able to provide comprehensive care to a PwD, potentially replacing human caregivers entirely [116, 136]. Caregivers may also trust a personalized robot to be more proficient at providing care than a robot that is not personalized, leading them to leave the PwD under the care of a robot for longer periods of time, further reducing human interaction and exacerbating the potential for social isolation.

In addition to the aforementioned risks of highly personalized care robots, non-personalized (or poorly personalized) robots may also lead to social isolation, such as by causing confusion or lowering the confidence of PwD. For instance, a CAR that fails to appropriately adapt to a PwD's capabilities could cause them to lose confidence in their communicative or cognitive abilities. This can lead to anxiety or depression, and cause them to withdraw from their friends and family [6].

3.4 Risk 4: Vulnerability to Dark Patterns in Personalized Robotics

Dementia gradually diminishes an individual's communication abilities and judgment, making it more difficult for them to avoid, prevent, and report deception. While some researchers believe that introducing assistive robots for dementia care can reduce the abuse many PwD experience [96], it is not a stretch to imagine a scenario where a robot could take advantage of a PwD, particularly if these robots follow a model similar to existing adaptive technologies (e.g. maximizing engagement, prioritizing advertising revenue over user well-being) [55].

In the field of user interface design, dark patterns are user experience (UX) and user interface (UI) interactions designed to mislead or trick users to make them do something they do not want to do. In existing technologies such as online social media, designers have been known to leverage dark patterns, or use their knowledge of human behavior and the desires of end users, to implement deceptive functionality that is not in the user's best interest [17, 48]. For instance, on social media platforms, dark patterns may be used to increase engagement with the platform, increase ad revenue, or get users to share personal information. While these behaviors are beneficial for the platform, they can be detrimental to users, as over-engagement with these media can lead to addiction, social isolation, anxiety, and depression [28, 74].

In the context of CARs that personalize their interactions based on the data collected from a PwD, there may be dark patterns that designers could use to take advantage of PwD. Thus, as robots become more sophisticated and autonomous, it is important to research how personalized robots that collect personal information from users may be designed to leverage or exploit this data to facilitate deceptive interactions with PwD.

Dark patterns in robotics is a largely unexplored area. Lacey et al. [77] discuss how cuteness of robots can be a deceptive tactic that roboticists use to gather information from users. For example, Blue Frog's Buddy is an emotional robot whose marketing website states: "How not to resist to his cuteness and not want to adopt him?". Prior research has found that "cute" technology is "lovable" and fosters an affectionate relationship [46, 77]. Among PwD, a personalized CAR may be customized to suit a user's preferences (e.g. have a "cute" appearance or friendly personality) to gain acceptance and facilitate a bond.

However, a PwD may therefore more readily share sensitive information with a personalized robot, unwittingly give them access to private accounts, or be manipulated into purchasing other products from the robot's developer. Additionally, because a personalized robot could have information on the wants and needs of a PwD, and PwD and caregivers often have low technology familiarity, developers may have the power to intentionally make turning off the robot or disengaging from the robot difficult. It is important that dark patterns in the context of personalized CARs among PwD are further explored to avoid negative consequences for PwD and to hold technology creators accountable.

4 Additional Ethical Considerations

In addition to the risks discussed in Section 3, there are some additional ethical considerations when developing personalized CARs for PwD. These include: a) how a robot can practice beneficence to PwD, b) where responsibility falls should harm to a PwD occur because of a robot, and c) how a robot can acquire informed consent from a PwD. While these considerations are not necessarily unique to personalized CARs for PwD, it is important that roboticists keep them in mind in order to understand how these robots may impact users in real world environments. Thus, we explore each of these considerations in this section.

4.1 How can a robot practice beneficence toward people with dementia?

Human caregivers are often very intentional with their language and actions in order to set a PwD up for success and practice beneficence to PwD (e.g. minimizing confusion or agitation). For instance, they will purposefully phrase their sentences to be short and simple, and ask closed questions such as those that can be answered with a "yes" or "no" response. In addition, caregivers will use non-verbal cues such as gestures or visual aids to help communicate with PwD, particularly as a PwD's verbal communication abilities deteriorate with the progression of the disease. This can help improve comprehension of PwD, increase their ability to respond successfully, and reduce the chances of causing frustration or confusion [27, 50].

Even so, frustrating or confusing interactions are largely inevitable, especially for people with advanced dementia. These individuals may have difficulty processing abstract language or not even recognize they have dementia (i.e., anosognosia), which can lead to reduced confidence or perceiving themselves as being "faulty". In addition, as the disease progresses and prospective memory becomes weaker, technologies that were previously helpful (e.g., reminder technologies) may become less effective and can cause tensions and frustration between a PwD and caregivers [50]. Therefore, it is not a stretch to imagine that even the most accurately personalized CARs are likely to inadvertently cause confusing or frustrating feelings in interactions with PwD.

It is not uncommon for human caregivers to deceive PwD, often coming from a place of compassion with the goal of minimizing disorientation or distress that might come along with correcting a PwD's perception of the world [33]. In fact, human caregivers may deceive PwD to help improve their sense of self-agency and autonomy [15, 21, 23, 136, 115]. For instance, in our prior work [50], a professional dementia caregiver told us about a PwD who used to be an accountant. The professional caregiver allowed her to think that she was the current accountant of their dementia caregiving organization, thereby respecting and acknowledging her domain expertise. On the other hand, many researchers argue that deceiving people in such a way is a "moral failure" because this alters the PwD's perception of reality and may lead them to believe a different reality than those around them [33, 113, 121, 136].

However, these experiences beg the question of whether it is appropriate for assistive robots to actively deceive PwD, as a human caregiver might. Thus, while these robots could leverage the knowledge, background, and expertise of a PwD in order to respect their autonomy, whether or to what extent they (or human caregivers) should deceive a PwD is still an open question. Some scholars argue that this deception is benign and permissible as long as it is in the best interest of a PwD [49, 146]. In addition, regardless of whether a PwD can tell if a robot's empathic response is real or not, PwD may still experience real feelings of comfort and companionship [26] which many argue is acceptable [22, 117]. On the other hand, some people express discomfort with the idea that PwD might perceive and engage with robots (even non-personalized ones) as living agents [16].

It is generally agreed that assistive robots should, whenever possible, practice nonmaleficence and not bring harm to people [37, 73, 137]. Indeed, under the beneficence principle, assistive robots should actively behave in a person's best interest. This might entail telling white lies to promote a PwD's dignity, provide comfort, or avoid confusion or distress [115]. But while it is possible that a personalized CAR may be more effective at practicing beneficence, they cannot necessarily avoid frustrating or confusing interactions with PwD. It can be difficult to avoid frustrating or confusing interactions even for human caregivers, so we propose that these robots can practice beneficence by: a) taking appropriate precautions to mitigate frustrating or confusing interactions before they occur, b) taking steps to alleviate feelings of frustration or confusion should they occur (even if that means notifying a human caregiver), and c) striving to ensure that these feelings are no worse than that which the PwD might experience with a human caregiver.

4.2 Responsibility for harm

There has been much debate around who or what should be held responsible if a machine causes harm to a person, particularly in healthcare contexts [70, 135]. Traditionally, care providers are required to assume responsibility for the outcome of a medical intervention [134]. Historically, if a machine causes harm to a person, there is a clear entity at fault. For instance, an operator controlling a machine may be blamed if they make a mistake, or a manufacturer may take responsibility for defective hardware.

However, when considering personalized CARs for PwD, there are many gray areas that arise due to impaired reasoning abilities and the "responsibility gap" (i.e., the inability to trace responsibility to any particular entity due to the unpredictable nature of an autonomous robot's future behavior) [86]. With the introduction of personalized robots into dementia caregiving, there needs to be a sense of moral, legal, and/or fiscal responsibility in order to ensure that PwD and other users of personalized assistive robots are safe.

There is much discussion about who should be held responsible for the actions of an autonomous robot. Some researchers suggest that the autonomous systems themselves should be held responsible [59]. However, others argue that machines cannot understand the consequences of their actions and thus hold the concept of responsibility meaningless [82]. Others argue that the manufacturers (e.g. researchers, developers, designers) should take responsibility for the robots they created, since they have a professional responsibility to follow proper ethical and professional design protocols before getting into the hands of users [82, 135].

Alternatively, others adopt antiquated views of user blaming [62], suggesting users are responsible, as they supervise and manage the robots, and they are the ones that the system is learning from [9, 135]. Elish [35] coined the term "moral crumple zone" to describe such scenarios, in which responsibility for harm caused by an autonomous agent may be misattributed to a human who in fact had little control over the agent's actions.

The responsibility gap can make it difficult to attribute responsibility for harm caused by autonomous systems [86]. Inherently, personalized assistive robots must learn to behave in ways that were not explicitly defined by a human programmer. This can lead to unpredictable robot behavior, so nobody, from the programming team to the end user, can be seen as clearly responsible for a robot's behavior. As a personalized robot learns from more people and must base its decisions on potentially conflicting information (e.g. if a PwD enjoys baking, but a caregiver does not want them to use an oven, and a clinician suggests avoiding desserts), this can add an additional layer of complexity and uncertainty when trying to attribute responsibility for harm, should it occur.

So, how does one determine who should be held responsible in the case that a personalized robot harms a PwD? It is imperative that the field of autonomous systems protect users from misattributed responsibility and avoid moral crumple zones [35]. Instead, there is an increasing emphasis on "responsible robotics" which places the responsibility on the researchers and developers [135]. This requires that an organization determines ethical issues that arise from use of the robot, as well as to assign people to resolve those issues [135]. Furthermore, some researchers suggest that determining how to regulate the responsible use of these robots will require more thorough exploration and testing across populations and cultures with multidisciplinary studies and collaborations [37, 136].

4.3 Acquiring consent from people with dementia

Informed consent is a person's adequate comprehension and subsequent voluntary choice to participate in some event, such as a medical intervention [25]. It is important for both human and artificial agents providing healthcare services to obtain consent (or assent) from PwD to protect a person's wellbeing and agency. However, in the context of dementia, the problem of acquiring informed consent is difficult because it is challenging to determine whether their condition affected their capacity of giving informed consent [63], and their capacity to provide consent may change as their dementia progresses. Difficulty obtaining consent can be especially problematic for personalized health interventions such as assistive robots personalized to PwD, as data collection and processing are essential for a robot to learn a person's preferences.

To help address this challenge in the medical and research spaces, organizations have developed various recommendations for acquiring informed consent from PwD. In general, there are three ways to acquire informed consent from or on behalf of PwD: (i) direct consent from a person with acceptable level of competence and cognitive capacity, (ii) proactive consent through advanced directives (i.e. externalizations of PwD's wishes, decisions, and choices about future actions), or (iii) through proxy decision making (e.g. assent from a third party) [63]. Ienca et al. [63] suggest that the combination of the three may better protect a PwD's autonomy. In the context of personalized robots for PwD, third parties (e.g. children, spouses) can help identify the aspects of an assistive robot that they would like to adjust [63].

However, there are no standard protocols for obtaining consent from PwD across institutions, sectors, or countries, including in the context of personalized robots for PwD [57, 99]. Even the question of who is responsible for providing informed consent (the PwD, caregivers, researchers, or another party) has no clear answer [43]. Researchers and developers across numerous communities (e.g. dementia caregiving, robotics, gerontechnology) have proposed recommendations for obtaining consent from PwD and have called for regulatory frameworks to standardize this process [63, 130].

In the case of personalized CARs for PwD, researchers recommend using an iterative model known as "ongoing consent" [64]. For instance, a robot learning to personalize its behavior to fit an individual's personality and goals should obtain consent at multiple intervals during an intervention. It should be able to answer questions or provide additional information in a clear and transparent manner (e.g. employing visual aids). A robot should also communicate with PwD using well-designed communication modalities suitable to the person's stage of dementia (e.g. non-verbal embodied cueing, mimicry, and music) to better convey meaning, improve PwD self-agency, and reduce caregiver burden [50]. The specific points at which a robot might provide this information and ask for consent might vary depending on the context, but it is generally agreed that PwD may withdraw consent at any time, whether verbally or by expressing signs of distress [64].

In addition, as a personalized CAR further learns from a PwD's individual's choices and decisions, it may be able to help clinicians with more in-depth competency assessments, by being able to provide insights into longitudinally observed behaviors. Caregivers and clinicians would then have a better understanding to reprogram the robot (or remove it) as needed [75].

5 Key Policy Concepts

There are some key policy concepts that robot designers, law-makers, and others should keep in mind to develop safe and ethically-informed approaches for longitudinal robot-delivered health interventions, particularly those designed for people with cognitive impairments. These include a) Community care approaches to design, b) Justice and accessibility, c) Educating caregivers and clinicians, and d) Promoting the agency of PwD (see Table 2). In this section, we provide a brief overview of each of these concepts and how they relate to personalized CARs for PwD.

5.1 Community care approaches to design

In order to ensure CARs will accurately address and personalize their behavior to the needs of PwD, robot developers should adopt community-centered care approaches to design and closely involve key stakeholders such as PwD, their caregivers, and clinicians throughout the development process. In particular, adopting user-centered design approaches and offering "whole person care" (i.e. care that aims to improve a person's situation as a whole by addressing their social and/or behavioral needs in addition to their physical health) is essential to recognizing a user as a person and addressing their well-being as individuals beyond simply someone living with dementia [78].

Our earlier work suggests several design guidelines to contextualize new roles and behaviors for assistive robots within the PwD's family caregiving paradigm, including: a) relieving a caregiver's emotional burden by communicating facts and information PwD may not want to hear or make PwD do things they may not want to do, b) redirect PwD to more positive interactions during emotionally difficult times, and c) accentuating positive shared moments [89]. Furthermore, our recent work on community-centered design for PwD suggests that using non-verbal, embodied action prompts as a health intervention for caregiving technology can help convey meaning, improve a PwD's sense of self, and reduce the burden of caregivers [50].

As researchers continue this avenue of exploration, it will be important to consider the needs and goals of the community in addition to those of individual end users, which may require closer collaborations between ethicists, engineers, and other stakeholders [50, 65]. This will help empower PwD and their caregivers by supporting their independence and promoting their agency, as well as mitigate the risks of social isolation, objectification, and deception that personalization might cause [136].

Both Dixon et al. [31] and Guan et al. [50] suggest several methods for engaging in these research practices, such as conducting interviews, community design workshops, and family meetings. Robotics researchers have used tools including low fidelity design probes, sketches, and foam blocks to help stakeholders communicate their ideas and envision interactions with a robot during these

Table 2: Key policy concepts to help guide the creation of safe and ethically-informed robot-delivered health interventions, and help protect PwD from unintended consequences.

Policy Concept	Description	Example
Community-centered care approaches to design	In order to ensure robots can accurately address and adapt to the needs of PwD, robot developers must closely involve key stakeholders including PwD, caregivers, and clinicians throughout the development process.	User-centered design approaches and offering ``whole person care".
Justice and accessibility	Roboticists should support and encourage accessibility of care robots in order to ensure that they are affordable and usable by PwD and their caregivers.	Curb the cost of production, use affordable materials, and leverage open-source solutions. Partner with health systems to understand local community needs and barriers with regard to technology adoption.
Educating caregivers and clinicians	It is important for caregivers and clinicians to be educated on the potential risks of using a personalized care robot in order to mitigate its potential for harm.	Provide caregivers the knowledge, resources, and skills to be able to use robots to best support PwD through in-person education sessions and resources they can refer back to later.
Promoting the agency of PwD	A robot's morphology and behaviors should support the autonomy of PwD when possible in order to support their dignity and individuality.	Practice user-centered design, make systems intuitive and easy to control, and establish systems through which PwD (and/or caregivers) can express their preferences.

design sessions [50, 89]. In addition, employing these methods remotely is particularly important because PwD and their caregivers are primarily older adults, and thus at a higher risk for severe illness and death from COVID-19. Furthermore, remote studies provide the opportunity of having PwDs in their normal home environment rather than controlled environments during the study [31], which can help provide better contextualization to researchers.

5.2 Justice and accessibility

There is a growing movement among the robotics and caregiving communities to support fair distribution and universal access to technologies for care [50, 63]. Nonetheless, robots that can adapt their behavior to be personalized to PwD may be more expensive to develop and produce than their non-personalized counterparts due to more complex hardware or software. However, due to the limited low-cost and open-source technologies currently available or in development, the adoption of personalized care technologies is likely to be limited by socio-economic factors, or even exacerbate a growing socio-economic divide [65]. This concern underlies the fact that many PwD may live in poverty, may not have access to broadband internet, and caregivers often have low technology literacy [50]. Thus, it is crucial that those developing and deploying personalized assistive robots for dementia caregiving consider the unique needs of this population to prioritize access.

There are many steps roboticists can take to support accessibility of these robots, which may traditionally be prohibitively expensive for PwD and caregivers to adopt into their homes. These include curbing the cost of production, using affordable materials, and utilizing and developing open-source solutions [63]. These steps can help reduce the cost of a product for end users, or potentially make it possible for them to create their own (e.g. 3D printing hardware and downloading software). Decommodification of assistive technology for PwD is an alternative solution to lowering cost for users and improving the accessibility of these products (e.g. offering robots through a rental service or long-term care insurance system) [20].

In addition, roboticists across industry and academia can partner with health systems to learn more about the cost-related barriers to technology adoption and sustainment unique to the populations they serve. Such partnerships can help roboticists create robots that are more likely to be purchased by healthcare systems, rather than patients. Most healthcare systems are incentivized by payers to improve health outcomes (e.g., reducing unplanned hospital visits among home health patients [100]), which in turn incentivizes them to adopt new interventions in support of those goals.

Increasing accessibility to technology also entails ensuring it is intuitive and usable by the intended end users: PwD and their caregivers. However, this population tends to be older adults with low technology literacy. One approach to improving usability for this population is to integrate familiar features into the design of a device, such as touch screens or verbal communication. Developers can also design technologies that are based off of or extend the functionality of existing items in a PwD's home, such as a smart photo frame. Leveraging aspects of objects and technologies that PwD may already be familiar with can improve the acceptability and usability of new technologies for PwD and caregivers.

5.3 Educating caregivers and clinicians

As the number and quality of robots to support PwD increase, so too will the number of caregivers who will adopt personalized assistive robots into their caregiving routine. While these robots will ideally help alleviate their caregiving responsibilities and enable them to have more productive interactions with PwD, it is essential that caregivers and clinicians understand how to use these robots, as well as the potential risks associated with using them. This will help ensure that stakeholders have realistic expectations of the robot's capabilities and expected impacts, as well as help them understand how to regulate responsible use of these robots. For example, it is important that caregivers are aware of the possibility that the personalized behavior of these robots can lead PwD to form stronger attachments with them, so caregivers can recognize signs of over-attachment and know what steps to take to prevent escalation to social isolation.

To help facilitate this education, there are several approaches the robotics and caregiving communities can take, including face-to-face content delivery and providing easily accessible information that stakeholders can refer back to. Research shows that the majority of education about dementia caregiving in general is delivered in face-to-face interactions [103], so this is a natural way to teach caregivers about personalized CARs as well. In fact, in our conversations with clinicians who work with people with cognitive impairments and their caregivers, they recommended having an individual in-person session with stakeholders to teach them about technology before they use it. This enables roboticists to immediately answer any questions a caregiver might have, show them demonstrations of the robot in a controlled environment, and help lower technical barriers to use. Berridge et al. [13] similarly recommends including PwD in the installation and onboarding processes of new technologies.

In addition to an in-person education session, our conversations with clinicians revealed that it is beneficial to provide important information in a form that stakeholders can easily refer back to later. For instance, developers might give stakeholders a manual that covers the main points that they should know, or print a QR code on the robot itself that links to a digital version of the information. Regardless of the form, the information should be written in common language, preferably accompanied by icons or images to improve its accessibility, as caregivers often have low technology literacy [51]. Thus, they can easily find and refer back to the information if they have questions.

In addition to the many opportunities for personalized CARs to support PwD, robots can also be a powerful tool that support training and education for caregivers and other stakeholders. Neither formal nor informal caregivers receive adequate training and support to provide effective care for PwD [103, 106]. However, studies have shown that having this knowledge can help improve both the quality of care they can provide to PwD [2], as well as health outcomes for caregivers themselves [2, 123]. Thus, it is crucial that caregivers are provided the knowledge, resources, and skills to be able to use personalized CARs to best support PwD, while also maintaining their own health and well-being [78].

5.4 Promoting the agency of PwD

As discussed in Section 3.2, it is extremely important to encourage the autonomy of PwD when designing personalized assistive robots in order to support their dignity and individuality. There are multiple steps developers can take to help promote a PwD's agency, including practicing user-centered design to make systems intuitive and easy to control and establishing systems through which PwD (or caregivers, in their place) can express their preferences.

Developing robots that are intuitive for PwD will help ensure that PwD can easily communicate their needs to the robot. This can improve their ability to modify the robot's behavior, thus promoting their agency [34]. As discussed in Section 5.2, intuitive interaction can be achieved by leveraging familiar features such as voice commands or touch screens, and using or alluding to common objects such as a radio as shown in Figure 1.2 [50]. Applying a critical dementia lens to the design of personalized robots (and care technology in general) is essential to ensuring that robots can best support stakeholder needs and interests while also preserving their agency and personhood.

In addition, developers can also promote the agency of PwD by establishing systems through which they can express their preferences. For example, PwD may provide advanced directives (i.e.

specifying their desires before the onset of dementia) or consent by proxies (i.e. delegating decisions to a trusted individual such as a family member) [63]. Moreover, developers can design robotic systems that require PwD to be active participants in decision making. For instance, the robot can offer a variety of stimulating activities for a PwD and prompt the PwD to choose. While these are not foolproof methods to understanding the wishes of a PwD, these approaches may be the closest that an assistive robot can get to understanding the desires of a PwD when they are not necessarily in a state of mind to communicate or fully reason about a decision.

6 Discussion

Personalized robots have the potential to vastly improve whole person care for PwD, but it is also important to minimize the risks they might pose. The risks raised in this work are but a few potential challenges that accompany these technologies, demonstrating the need for continued and critical exploration into the potential consequences of personalizing CARs, particularly for PwD and other cognitive impairments.

Weighing the benefits and risks of behavior adaptation in this domain can help guide robot developers, policy makers, and other stakeholders as they help shape a world where robots can assist with care in homes, hospitals, and other community care settings. Moving forward, it will be essential for these stakeholders to acknowledge and address the potential risks of these technologies when developing technology, policy, and other advancements in this space. Promoting this culture of ethical awareness will be more likely to produce safe and ethically-informed personalized technologies which mitigate their risks while augmenting their benefits. We hope that our work will inspire roboticists to consider the potential risks and benefits of robot personalization, and support future ethically-focused robot design.

References

[1] C. Abras, D. Maloney-Krichmar, J. Preece, et al. User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications*, 37(4):445–456, 2004.

[2] E. Aksoydan, A. Aytar, A. Blazeviciene, R. L. van Bruchem-Visser, A. Vaskelyte, F. Mattace-Raso, S. Acar, A. Altintas, E. Akgun-Citak, S. Attepe-Ozden, et al. Is training for informal caregivers and their older persons helpful? a systematic review. *Archives of gerontology and geriatrics*, 83:66–74, 2019.

[3] O. Almeida, G. Hankey, B. Yeap, J. Golledge, and L. Flicker. Depression as a modifiable factor to decrease the risk of dementia. *Translational psychiatry*, 7(5):e1117–e1117, 2017.

[4] L. Asprino, A. Gangemi, A. G. Nuzzolese, V. Presutti, D. R. Recupero, and A. Russo. Ontology-based knowledge management for comprehensive geriatric assessment and reminiscence therapy on social robots. In *Data Science for Healthcare*, pages 173–193. Springer, 2019.

[5] A. Association. 2019 alzheimer's disease facts and figures. *Alzheimer's & dementia*, 15(3):321–387, 2019.

[6] A. Bahar-Fuchs, L. Clare, and B. Woods. Cognitive training and cognitive rehabilitation for mild to moderate alzheimer's disease and vascular dementia. *Cochrane database of systematic reviews*, (6), 2013.

[7] A. L. Baker, E. K. Phillips, D. Ullman, and J. R. Keebler. Toward an understanding of trust repair in human-robot interaction: Current research and future directions. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 8(4):1–30, 2018.

[8] E. Barrett, M. Burke, S. Whelan, A. Santorelli, B. L. Oliveira, F. Cavallo, R.-M. Dröes, L. Hopper, A. Fawcett-Henesy, F. J. Meiland, et al. Evaluation of a companion robot for individuals with dementia: quantitative findings of the mario project in an irish residential care setting. *Journal of gerontological nursing*, 45(7):36–45, 2019.

[9] S. Beck. The problem of ascribing legal responsibility in the case of robotics. *AI & society*, 31(4):473–481, 2016.

[10] T. Belpaeme, J. Kennedy, A. Ramachandran, B. Scassellati, and F. Tanaka. Social robots for education: A review. *Science robotics*, 3(21), 2018.

[11] W. L. Bennett and A. Segerberg. The logic of connective action: Digital media and the personalization of contentious politics. *Information, communication & society*, 15(5):739–768, 2012.

[12] P. Bernal. Facebook: Why facebook makes the fake news problem inevitable. *N. Ir. Legal Q.*, 69:513, 2018.

[13] C. Berridge, G. Demiris, and J. Kaye. Domain experts on dementia-care technologies: Mitigating risk in design and implementation. *Science and Engineering Ethics*, 27(1):1–24, 2021.

[14] E. Bird, J. Fox-Skelly, N. Jenner, R. Larbey, E. Weitkamp, and A. Winfield. The ethics of artificial intelligence: Issues and initiatives. *European Parliamentary Research Service, Technical Report PE*, 634, 2020.

[15] J. Borenstein and Y. Pearson. Robot caregivers: harbingers of expanded freedom for all? *Ethics and Information Technology*, 12(3):277–288, 2010.

[16] H. L. Bradwell, R. Winnington, S. Thill, and R. B. Jones. Ethical perceptions towards real-world use of companion robots with older people and people with dementia: survey opinions among younger adults. *BMC geriatrics*, 20(1):1–10, 2020.

[17] H. Brignull, M. Miquel, J. Rosenberg, and J. Offer. Dark patterns-user interfaces designed to trick people, 2015.

[18] A. L. Byers and K. Yaffe. Depression and risk of developing dementia. *Nature Reviews Neurology*, 7(6):323–331, 2011.

[19] S.-T. Cheng. Dementia caregiver burden: a research update and critical analysis. *Current psychiatry reports*, 19(9):1–8, 2017.

[20] Y.-h. Chou, S.-y. B. Wang, and Y.-t. Lin. Long-term care and technological innovation: the application and policy development of care robots in taiwan. *Journal of Asian Public Policy*, 12(1):104–123, 2019.

[21] M. Coeckelbergh. Health care, capabilities, and ai assistive technologies. *Ethical theory and moral practice*, 13(2):181–190, 2010.

[22] M. Coeckelbergh. Are emotional robots deceptive? *IEEE Transactions on Affective Computing*, 3(4):388–393, 2011.

[23] M. Coeckelbergh. Care robots and the future of ict-mediated elderly care: a response to doom scenarios. *AI & society*, 31(4):455–462, 2016.

[24] D. Cruz-Sandoval, A. Morales-Tellez, E. B. Sandoval, and J. Favela. A social robot as therapy facilitator in interventions to deal with dementia-related behavioral symptoms. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, pages 161–169, 2020.

[25] F. K. Dankar, M. Gergely, and S. K. Dankar. Informed consent in biomedical research. *Computational and structural biotechnology journal*, 17:463–474, 2019.

[26] K. Darling. 'who's johnny?'anthropomorphic framing in human-robot interaction, integration, and policy. *Anthropomorphic Framing in Human-Robot Interaction, Integration, and Policy (March 23, 2015). ROBOT ETHICS*, 2, 2015.

[27] K. De Vries. Communicating with older people with dementia. *Nursing older people*, 25(4), 2013.

[28] J. Deleuze, J. Long, T.-Q. Liu, P. Maurage, and J. Billieux. Passion or addiction? correlates of healthy versus problematic use of videogames in a sample of french-speaking regular players. *Addictive Behaviors*, 82:114–121, 2018.

[29] E. Deng, B. Mutlu, and M. Mataric. Embodiment in socially interactive robots. *arXiv preprint arXiv:1912.00312*, 2019.

[30] E. Dixon, A. M. Piper, and A. Lazar. "taking care of myself as long as i can": How people with dementia configure self-management systems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2021.

[31] E. Dixon, A. Shetty, S. Pimento, and A. Lazar. Lessons learned from remote user-centered design with people with dementia. In *Dementia Lab Conference*, pages 73–82. Springer, 2021.

[32] J. Donath. The robot dog fetches for whom? In *A networked self and human augmentics, artificial intelligence, sentience*, pages 10–24. Routledge, 2018.

[33] R. Dresser. A tangled web: Deception in everyday dementia care. *Journal of Law, Medicine & Ethics*, 49(2):257–262, 2021.

[34] Y. Du, S. Tiomkin, E. Kiciman, D. Polani, P. Abbeel, and A. Dragan. Ave: Assistance via empowerment. *Advances in Neural Information Processing Systems*, 33, 2020.

[35] M. C. Elish. Moral crumple zones: Cautionary tales in human-robot interaction (pre-print). *Engaging Science, Technology, and Society (pre-print)*, 2019.

[36] S. Fazio, D. Pace, J. Flinner, and B. Kallmyer. The fundamentals of person-centered care for individuals with dementia. *The Gerontologist*, 58(suppl_1):S10–S19, 2018.

[37] D. Feil-Seifer and M. J. Matarić. Socially assistive robotics. *IEEE Robotics & Automation Magazine*, 18(1):24–31, 2011.

[38] D. Fetherstonhaugh, L. Tarzia, and R. Nay. Being central to decision making means i am still here!: The essence of decision making for people with dementia. *Journal of aging studies*, 27(2):143–150, 2013.

[39] J. M. Fleming, D. Shum, J. Strong, and S. Lightbody. Prospective memory rehabilitation for adults with traumatic brain injury: A compensatory training programme. *Brain Injury*, 19(1):1–10, 2005.

[40] E. Frank Lopresti, A. Mihailidis, and N. Kirsch. Assistive technology for cognitive rehabilitation: State of the art. *Neuropsychological rehabilitation*, 14(1-2):5–39, 2004.

[41] J. Fürnkranz and E. Hüllermeier. Preference learning and ranking by pairwise comparison. In *Preference learning*, pages 65–82. Springer, 2010.

[42] L. Garand, M. Amanda Dew, L. R. Eazor, S. T. DeKosky, and C. F. Reynolds III. Caregiving burden and psychiatric morbidity in spouses of persons with mild cognitive impairment. *International journal of geriatric psychiatry*, 20(6):512–522, 2005.

[43] J. Gerłowska, M. Furtak-Niczyporuk, and K. Rejdak. Robotic assistance for people with dementia: a viable option for the future? *Expert Review of Medical Devices*, 17(6):507–518, 2020.

[44] N. Geva, F. Uzefovsky, and S. Levy-Tzedek. Touching the social robot paro reduces pain perception and salivary oxytocin levels. *Scientific reports*, 10(1):1–15, 2020.

[45] G. Gibson, L. Newton, G. Pritchard, T. Finch, K. Brittain, and L. Robinson. The provision of assistive technology products and services for people with dementia in the united kingdom. *Dementia*, 15(4):681–701, 2016.

[46] J. Gn. The technology of the cute body. *Eidos. A Journal for Philosophy of Culture*, 2(4 (6)), 2018.

[47] S. Góngora Alonso, S. Hamrioui, I. de la Torre Díez, E. Motta Cruz, M. López-Coronado, and M. Franco. Social robots for people with aging and dementia: a systematic review of literature. *Telemedicine and e-Health*, 25(7):533–540, 2019.

[48] S. Greenberg, S. Boring, J. Vermeulen, and J. Dostal. Dark patterns in proxemic interactions: a critical perspective. In *Proceedings of the 2014 conference on Designing interactive systems*, pages 523–532, 2014.

[49] F. S. Grodzinsky, K. W. Miller, and M. J. Wolf. Developing automated deceptions and the impact on trust. *Philosophy & Technology*, 28(1):91–105, 2015.

[50] C. Guan, A. Bouzida, R. M. Oncy-Avila, S. Moharana, and L. D. Riek. Taking an (embodied) cue from community health: Designing dementia caregiver support technology to advance health equity. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2021.

[51] E. Guisado-Fernández, G. Giunti, L. M. Mackey, C. Blake, and B. M. Caulfield. Factors influencing the adoption of smart health technologies for people with dementia and their informal caregivers: scoping review and design framework. *JMIR aging*, 2(1):e12192, 2019.

[52] A. Hall, C. B. Wilson, E. Stanmore, and C. Todd. Moving beyond 'safety' versus 'autonomy': a qualitative exploration of the ethics of using monitoring technologies in long-term dementia care. *BMC geriatrics*, 19(1):1–13, 2019.

[53] M. Handley, F. Bunn, and C. Goodman. Supporting general hospital staff to provide dementia sensitive care: A realist evaluation. *International journal of nursing studies*, 96:61–71, 2019.

[54] K. S. Haring, A. Mosley, S. Pruznick, J. Fleming, K. Satterfield, E. J. de Visser, C. C. Tossell, and G. Funke. Robot authority in human-machine teams: effects of human-like appearance on compliance. In *International Conference on Human-Computer Interaction*, pages 63–78. Springer, 2019.

[55] W. Hartzog. Unfair and deceptive robots. *Md. L. Rev.*, 74:785, 2014.

[56] K. B. Hirschman, S. X. Xie, C. Feudtner, and J. H. Karlawish. How does an alzheimer's disease patient's role in medical decision making change over time? *Journal of geriatric psychiatry and neurology*, 17(2):55–60, 2004.

[57] J. Hirt, N. Ballhausen, A. Hering, M. Kliegel, T. Beer, and G. Meyer. Social robot interventions for people with dementia: A systematic review on effects and quality of reporting. *Journal of Alzheimer's Disease*, 79(2):773, 2021.

[58] J. Hoey, P. Poupart, A. von Bertoldi, T. Craig, C. Boutilier, and A. Mihailidis. Automated handwashing assistance for persons with dementia using video and a partially observable markov decision process. *Computer Vision and Image Understanding*, 114(5):503–519, 2010.

[59] Y. Hu. Robot criminals. *U. Mich. JL Reform*, 52:487, 2018.

[60] A. Huber, A. Weiss, and M. Rauhala. The ethical risk of attachment how to identify, investigate and predict potential ethical risks in the development of social companion robots. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 367–374. IEEE, 2016.

[61] L. Hung, C. Liu, E. Woldum, A. Au-Yeung, A. Berndt, C. Wallsworth, N. Horne, M. Gregorio, J. Mann, and H. Chaudhury. The benefits of and barriers to using a social robot paro in care settings: a scoping review. *BMC geriatrics*, 19(1):1–10, 2019.

[62] W. Hyman. Blaming the user for safety system failures: Back where we started. *Biomedical Safety & Standards*, 34(22):174–175, 2004.

[63] M. Ienca, F. Jotterand, C. Vică, and B. Elger. Social and assistive robotics in dementia care: ethical recommendations for research and practice. *International Journal of Social Robotics*, 8(4):565–573, 2016.

[64] M. Ienca and E. F. Villaronga. Privacy and security issues in assistive technologies for dementia. *Intelligent Assistive Technologies for Dementia: Clinical, Ethical, Social, and Regulatory Implications*, page 221, 2019.

[65] M. Ienca, T. Wangmo, F. Jotterand, R. W. Kressig, and B. Elger. Ethical design of intelligent assistive technologies for dementia: a descriptive review. *Science and engineering ethics*, 24(4):1035–1055, 2018.

[66] W. IJsselsteijn, A. Tummers-Heemels, and R. Brankaert. Warm technology: A novel perspective on design for and with people living with dementia. In *HCI and Design in the Context of Dementia*, pages 33–47. Springer, 2020.

[67] K. Inoue, K. Wada, and R. Uehara. How effective is robot therapy?: Paro and people with dementia. In *5th European Conference of the International Federation for Medical and Biological Engineering*, pages 784–787. Springer, 2011.

[68] M. L. Jones, E. Kaufman, and E. Edenberg. Ai and the ethics of automating consent. *IEEE Security & Privacy*, 16(3):64–72, 2018.

[69] M. M. Jung, L. van der Leij, and S. M. Kelders. An exploration of the benefits of an animallike robot companion with more advanced touch interaction capabilities for dementia care. *Frontiers in ICT*, 4:16, 2017.

[70] I. R. Kerr, J. Millar, and N. Corriveau. Robots and artificial intelligence in health care. *Ian Kerr & Jason Millar & Noel Corriveau," Robots and Artificial Intelligence in Health Care" in Joanna Erdman, Vanessa Gruben, Erin Nelson, eds, Canadian Health Law and Policy, 5th ed (Toronto: LexisNexis Canada, 2017)*, 257, 2017.

[71] N. Khachiyants, D. Trinkle, S. J. Son, and K. Y. Kim. Sundown syndrome in persons with dementia: an update. *Psychiatry investigation*, 8(4):275, 2011.

[72] S. S. Khan, B. Ye, B. Taati, and A. Mihailidis. Detecting agitation and aggression in people with dementia using sensors—a systematic review. *Alzheimer's & Dementia*, 14(6):824–832, 2018.

[73] T. Körtner. Ethical challenges in the use of social service robots for elderly people. *Zeitschrift für Gerontologie und Geriatrie*, 49(4):303–307, 2016.

[74] E. Krossbakken, S. Pallesen, R. A. Mentzoni, D. L. King, H. Molde, T. R. Finserås, and T. Torsheim. A cross-lagged study of developmental trajectories of video game engagement, addiction, and mental health. *Frontiers in psychology*, 9:2239, 2018.

[75] A. Kubota, E. I. Peterson, V. Rajendren, H. Kress-Gazit, and L. D. Riek. Jessie: Synthesizing social robot behaviors for personalized

neurorehabilitation and beyond. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, pages 121–130, 2020.

[76] A. Kubota and L. D. Riek. Methods for robot behavior adaptation for cognitive neurorehabilitation. *Annual Review of Control, Robotics, and Autonomous Systems*, 5, 2021.

[77] C. Lacey and C. Caudwell. Cuteness as a 'dark pattern'in home robots. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 374–381. IEEE, 2019.

[78] E. B. Larson and C. Stroud. Meeting the challenge of caring for persons living with dementia and their care partners and caregivers: A way forward. 2021.

[79] M. Law, C. Sutherland, H. S. Ahn, B. A. MacDonald, K. Peri, D. L. Johanson, D.-S. Vajsakovic, N. Kerse, and E. Broadbent. Developing assistive robots for people with mild cognitive impairment and mild dementia: a qualitative study with older adults and experts in aged care. *BMJ open*, 9(9):e031937, 2019.

[80] A. Lazar, C. Edasis, and A. M. Piper. A critical lens on dementia and design in hci. In *CHI*, pages 2175–2188, 2017.

[81] H. R. Lee and L. D. Riek. Reframing assistive robots to promote successful aging. *ACM Transactions on Human-Robot Interaction (THRI)*, 7(1):1–23, 2018.

[82] G. Lima, N. Grgić-Hlača, and M. Cha. Human perceptions on moral responsibility of ai: A case study in ai-assisted bail decision-making. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2021.

[83] S. Lindsay, K. Brittain, D. Jackson, C. Ladha, K. Ladha, and P. Olivier. Empathy, participatory design and people with dementia. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 521–530, 2012.

[84] K. Lorenz, P. P. Freddolino, A. Comas-Herrera, M. Knapp, and J. Damant. Technology-based tools and services for people with dementia and carers: Mapping technology onto the dementia care pathway. *Dementia*, 18(2):725–741, 2019.

[85] V. W. Lui, L. C. Lam, D. N. Luk, L. H. Wong, C. W. Tam, H. F. Chiu, and P. S. Appelbaum. Capacity to make treatment decisions in chinese older persons with very mild dementia and mild alzheimer disease. *The American Journal of Geriatric Psychiatry*, 17(5):428–436, 2009.

[86] A. Matthias. The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and information technology*, 6(3):175–183, 2004.

[87] A. Matthias. Robot lies in health care: When is deception morally permissible? *Kennedy Institute of Ethics Journal*, 25(2):169–162, 2015.

[88] B. McCausland, L. Knight, L. Page, and K. Trevillion. A systematic review of the prevalence and odds of domestic abuse victimization among people with dementia. *International Review of Psychiatry*, 28(5):475–484, 2016.

[89] S. Moharana, A. E. Panduro, H. R. Lee, and L. D. Riek. Robots for joy, robots for sorrow: community based robot design for dementia caregivers. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 458–467. IEEE, 2019.

[90] E. Mordoch, A. Osterreicher, L. Guse, K. Roger, and G. Thompson. Use of social commitment robots in the care of elderly people with dementia: A literature review. *Maturitas*, 74(1):14–20, 2013.

[91] W. Moyle. The promise of technology in the future of dementia care. *Nature Reviews Neurology*, 15(6):353–359, 2019.

[92] W. Moyle, C. Jones, M. Cooke, S. O'Dwyer, B. Sung, and S. Drummond. Connecting the person with dementia and family: a feasibility study of a telepresence robot. *BMC geriatrics*, 14(1):1–11, 2014.

[93] W. Moyle, C. J. Jones, J. E. Murfield, L. Thalib, E. R. Beattie, D. K. Shum, S. T. O'Dwyer, M. C. Mervin, and B. M. Draper. Use of a robotic seal as a therapeutic tool to improve dementia symptoms: a cluster-randomized controlled trial. *Journal of the American Medical Directors Association*, 18(9):766–773, 2017.

[94] W. H. Organization et al. Global action plan on the public health response to dementia 2017–2025. 2017.

[95] R. Orpwood, T. Adlam, N. Evans, J. Chadd, and D. Self. Evaluation of an assisted-living smart home for someone with dementia. *Journal of Assistive Technologies*, 2008.

[96] F. O'Brolcháin. Robots and people with dementia: Unintended consequences and moral hazard. *Nursing ethics*, 26(4):962–972, 2019.

[97] M. Pino, M. Boulay, F. Jouen, and A. S. Rigaud. "are we ready for robots that care for us?" attitudes and opinions of older adults toward socially assistive robots. *Frontiers in aging neuroscience*, 7:141, 2015.

[98] O. Pino, G. Palestra, R. Trevino, and B. De Carolis. The humanoid robot nao as trainer in a memory program for elderly people with mild cognitive impairment. *International Journal of Social Robotics*, 12(1):21–33, 2020.

[99] E. Portacolone, J. Halpern, J. Luxenberg, K. L. Harrison, and K. E. Covinsky. Ethical issues raised by the introduction of artificial companions to older adults with cognitive impairment: A call for interdisciplinary collaborations. *Journal of Alzheimer's Disease*, 76(2):445–455, 2020.

[100] A. Pozniak, M. Turenne, E. Lammers, P. Mukhopadhyay, V. Slanchev, J. Doherty, L. Green, C. Schur, N. Byrd, C. Cogan, Z. Dietrich, Z. Ding, J. Dreyfus, K. Hanslits, S. Isaac, N. Ji, Y. Jin, R. Mandell, K. Milkovich, K. Repeck, J. Schrager, B. Simmons, A. Szymanski, J. Xing, and E. Young. Evaluation of the home health value-based purchasing (hhvbp) model fourth annual report. *Arbor Research Collaborative for Health and L&M Policy Research*, pages 1–163, 2021.

[101] A. E. Prince and D. Schwarcz. Proxy discrimination in the age of artificial intelligence and big data. *Iowa L. Rev.*, 105:1257, 2019.

[102] S. Reig, M. Luria, E. Forberger, I. Won, A. Steinfeld, J. Forlizzi, and J. Zimmerman. Social robots in service contexts: Exploring the rewards and risks of personalization and reembodiment. In *Designing Interactive Systems Conference 2021*, pages 1390–1402, 2021.

[103] N. V. Resciniti, W. Tang, M. Tabassum, J. L. Pearson, S. M. Spencer, M. C. Lohman, D. K. Ehlers, D. Al-Hasan, M. C. Miller, A. Teixeira, et al. Knowledge evaluation instruments for dementia caregiver education programs: A scoping review. *Geriatrics & gerontology international*, 20(5):397–413, 2020.

[104] L. Riek and D. Howard. A code of ethics for the human-robot interaction profession. *Proceedings of we robot*, 2014.

[105] L. D. Riek. Healthcare robotics. *Communications of the ACM*, 60(11):68–78, 2017.

[106] A. Robinson, C. Eccleston, M. Annear, K.-E. Elliott, S. Andrews, C. Stirling, M. Ashby, C. Donohue, S. Banks, C. Toye, et al. Who knows, who cares? dementia knowledge among nurses, care workers, and family members of people living with dementia. *Journal of Palliative Care*, 30(3):158–165, 2014.

[107] H. Robinson, B. A. MacDonald, N. Kerse, and E. Broadbent. Suitability of healthcare robots for a dementia unit and suggested improvements. *Journal of the American Medical Directors Association*, 14(1):34–40, 2013.

[108] T. L. Rodziewicz, B. Houseman, and J. E. Hipskind. Medical error prevention. 2018.

[109] S. Rossi, F. Ferland, and A. Tapus. User profiling and behavioral adaptation for hri: A survey. *Pattern Recognition Letters*, 99:3–12, 2017.

[110] S. Šabanović, C. C. Bennett, W.-L. Chang, and L. Huber. Paro robot affects diverse interaction modalities in group sensory therapy for older adults with dementia. In *2013 IEEE 13th international conference on rehabilitation robotics (ICORR)*, pages 1–6. IEEE, 2013.

[111] Z. I. Santini, A. Koyanagi, S. Tyrovolas, C. Mason, and J. M. Haro. The association between social relationships and depression: a systematic review. *Journal of affective disorders*, 175:53–65, 2015.

[112] A. I. Schein, A. Popescul, L. H. Ungar, and D. M. Pennock. Methods and metrics for cold-start recommendations. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 253–260, 2002.

[113] M. Schermer. Nothing but the truth? on truth and deception in dementia care. *Bioethics*, 21(1):13–22, 2007.

[114] D. Schuler and A. Namioka. *Participatory design: Principles and practices*. CRC Press, 1993.

[115] A. T. Seaman and A. M. Stone. Little white lies: Interrogating the (un) acceptability of deception in the context of dementia. *Qualitative Health Research*, 27(1):60–73, 2017.

[116] A. Sharkey and N. Sharkey. Granny and the robots: ethical issues in robot care for the elderly. *Ethics and information technology*, 14(1):27–40, 2012.

[117] A. Sharkey and N. Sharkey. We need to talk about deception in social robotics! *Ethics and Information Technology*, pages 1–8, 2020.

[118] G. W. Small, P. V. Rabins, P. P. Barry, N. S. Buckholtz, S. T. DeKosky, S. H. Ferris, S. I. Finkel, L. P. Gwyther, Z. S. Khachaturian, B. D. Lebowitz, et al. Diagnosis and treatment of alzheimer disease and related disorders: consensus statement of the american association for geriatric psychiatry, the alzheimer's association, and the american geriatrics society. *Jama*, 278(16):1363–1371, 1997.

[119] K. L. Smebye, M. Kirkevold, and K. Engedal. Ethical dilemmas concerning autonomy when persons with dementia wish to live at home: a qualitative, hermeneutic study. *BMC health services research*, 16(1):1–12, 2015.

[120] Y. Song and Y. Luximon. The face of trust: The effect of robot face ratio on consumer preference. *Computers in Human Behavior*, 116:106620, 2021.

[121] R. Sparrow and L. Sparrow. In the hands of machines? the future of aged care. *Minds and Machines*, 16(2):141–161, 2006.

[122] V. Sriram, C. Jenkinson, and M. Peters. Informal carers' experience of assistive technology use in dementia care at home: a systematic review. *BMC geriatrics*, 19(1):1–25, 2019.

[123] C. A. Surr, C. Gates, D. Irving, J. Oyebode, S. J. Smith, S. Parveen, M. Drury, and A. Dennison. Effective dementia education and training for the health and social care workforce: a systematic review of the literature. *Review of educational research*, 87(5):966–1002, 2017.

[124] D. Szafir and B. Mutlu. Pay attention! designing adaptive agents that monitor and improve user engagement. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 11–20, 2012.

[125] B. Taati, S. Zhao, A. B. Ashraf, A. Asgarian, M. E. Browne, K. M. Prkachin, A. Mihailidis, and T. Hadjistavropoulos. Algorithmic bias in clinical populations—evaluating and improving facial analysis technology in older adults with dementia. *IEEE Access*, 7:25527–25534, 2019.

[126] A. Tapus, C. Tapus, and M. Mataric. The role of physical embodiment of a therapist robot for individuals with cognitive impairments. In *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pages 103–107. IEEE, 2009.

[127] A. Tapus, C. Tapus, and M. J. Mataric. The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia. In *2009 IEEE international conference on rehabilitation robotics*, pages 924–929. IEEE, 2009.

[128] I. Testad, A. Corbett, D. Aarsland, K. O. Lexow, J. Fossey, B. Woods, and C. Ballard. The value of personalized psychosocial interventions to address behavioral and psychological symptoms in people with dementia living in care home settings: a systematic review. 2014.

[129] K. Thomas, D. Akhawe, M. Bailey, D. Boneh, E. Bursztein, S. Consolvo, N. Dell, Z. Durumeric, P. G. Kelley, D. Kumar, et al. Sok: Hate, harassment, and the changing landscape of online abuse. 2021.

[130] A. Thorogood, A. Mäki-Petäjä-Leinonen, H. Brodaty, G. Dalpé, C. Gastmans, S. Gauthier, D. Gove, R. Harding, B. M. Knoppers, M. Rossor, et al. Consent recommendations for research and international data sharing involving persons with dementia. *Alzheimer's & Dementia*, 14(10):1334–1343, 2018.

[131] M. Valentí Soler, L. Agüera-Ortiz, J. Olazarán Rodríguez, C. Mendoza Rebolledo, A. Pérez Muñoz, I. Rodríguez Pérez, E. Osa Ruiz, A. Barrios Sánchez, V. Herrero Cano, L. Carrasco Chillón, et al. Social robots in advanced dementia. *Frontiers in aging neuroscience*, 7:133, 2015.

[132] R. Van Patten, A. V. Keller, J. E. Maye, D. V. Jeste, C. Depp, L. D. Riek, and E. W. Twamley. Home-based cognitively assistive robots: maximizing cognitive functioning and maintaining independence in older adults without dementia. *Clinical Interventions in Aging*, 15:1129, 2020.

[133] A. Van Wynsberghe. Designing robots for care: Care centered value-sensitive design. *Science and engineering ethics*, 19(2):407–433, 2013.

[134] A. van Wynsberghe. To delegate or not to delegate: Care robots, moral agency and moral responsibility. In *50th Anniversary AISB Convention*, 2014.

[135] A. van Wynsberghe. Responsible robotics and responsibility attribution. *Robotics, AI, and Humanity*, page 239, 2021.

[136] T. Vandemeulebroucke, B. D. de Casterlé, and C. Gastmans. The use of care robots in aged care: A systematic review of argument-based ethics literature. *Archives of gerontology and geriatrics*, 74:15–25, 2018.

[137] D. Vanderelst and J. Willems. Can we agree on what robots should be allowed to do? an exercise in rule selection for ethical care robots. *International Journal of Social Robotics*, pages 1–10, 2019.

[138] K. Wada, T. Shibata, T. Musha, and S. Kimura. Robot therapy for elders affected by dementia. *IEEE Engineering in medicine and biology magazine*, 27(4):53–60, 2008.

[139] M. Wahlström and A. Törnberg. Social media mechanisms for right-wing political violence in the 21st century: Discursive opportunities, group dynamics, and co-ordination. *Terrorism and Political Violence*, 33(4):766–787, 2021.

[140] H. Wang, M. Olhofer, and Y. Jin. A mini-review on preference modeling and articulation in multi-objective optimization: current status and challenges. *Complex & Intelligent Systems*, 3(4):233–245, 2017.

[141] R. H. Wang, A. Sudhama, M. Begum, R. Huq, and A. Mihailidis. Robots to assist daily activities: views of older adults with alzheimer's disease and their caregivers. *International psychogeriatrics*, 29(1):67–79, 2017.

[142] Z. Wang, M. K. Singh, C. Zhang, L. D. Riek, and K. Chaudhuri. Stochastic multi-player bandit learning from player-dependent feedback. In *ICML Workshop on Real World Experiment Design and Active Learning*, 2020.

[143] Z. Wang, C. Zhang, M. K. Singh, L. Riek, and K. Chaudhuri. Multitask bandit learning through heterogeneous feedback aggregation. In *International Conference on Artificial Intelligence and Statistics*, pages 1531–1539. PMLR, 2021.

[144] B. Woodworth, F. Ferrari, T. E. Zosa, and L. D. Riek. Preference learning in assistive robotics: Observational repeated inverse reinforcement learning. In *Machine Learning for Healthcare Conference*, pages 420–439. PMLR, 2018.

[145] J. Xu, G. B. De'Aira, and A. Howard. Would you trust a robot therapist? validating the equivalency of trust in human-robot healthcare scenarios. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 442–447. IEEE, 2018.

[146] G. C. K. Yew. Trust in and ethical design of carebots: The case for ethics of care. *International Journal of Social Robotics*, 13(4):629–645, 2021.