Article

The Ethical Governance for the Vulnerability of Care Robots: Interactive-Distance-Oriented Flexible Design

Zhengqing Zhang [1] , **Chenggang Zhang** [2,*] **and Xiaomeng Li** [2]

1   School of Marxism, Beihang University, Beijing 100191, China; zhangzhengqing@buaa.edu.cn
2   Department of Sociology, Tsinghua University, Beijing 100084, China; lxm19@mails.tsinghua.edu.cn
*   Correspondence: zcgice@tsinghua.edu.cn; Tel.: +86-10-62794966-255

Abstract: The application of caring robots is currently a widely accepted solution to the problem of aging. However, for the elderly groups who live in gregarious residences and share intelligence devices, caring robots will cause intimacy and assistance dilemmas in the relationship between humans and non-human agencies. This is an information-assisted machine setting, with resulting design ethics issues brought about by the binary values of human and machine, body and mind. The "vulnerability" in risk ethics demonstrates that the ethical problems of human institutions stem from the increase of dependence and the obstruction of intimacy, which are essentially caused by the increased degree of ethical risk exposure and the restriction of agency. Based on value-sensitive design, caring ethics and machine ethics, this paper proposes a flexible design with the interaction-distance-oriented concept, and reprograms the ethical design of caring robots with intentional distance, representational distance and interpretive distance as indicators. The main purpose is to advocate a new type of human-machine interaction relationship emphasizing diversity and physical interaction.

Keywords: caring robots; elderly health management; machine ethics; vulnerability; human-machine relationship; flexible design

1. Introduction

1.1. Application and Ethical Issues of Caring Robot in Elderly Health Management

With the increasing number and proportion of the elderly population worldwide, the supply pressure on health services for the aged is increasing. At present, caring robots are considered to be an effective way to relieve the social care pressure brought by the aging population [1]. Caring robots are intelligent machines that provide social assistance services for users. Currently, caring robots have five main design functions: companionship, therapy, cognitive training, improving social convenience and physical therapy [2]. Therefore, caring robots for the elderly are mainly applied to three aspects of health management: (1) Reducing loneliness. Living alone is one of the factors correlated with chronic diseases such as obesity and depression. Using caring robots can effectively reduce loneliness and indirectly reduce the frequency of chronic diseases' development. (2) Providing care services. Even though the latest, most advanced caring robot cannot provide the same level of care as a human service provider, it can provide health information, health status monitoring, behavior prompt guidance [3], etc. Caring robots can provide a certain level of assistance, which can improve the quality of life and maintain the steady progress of daily medical treatment for users. (3) Health environment monitoring. As the elderly become less sensitive to environmental information, caring robots sometimes act as health guardians for the elderly, supervising interactions between other residents, health service providers and the elderly to help safeguard their health interests.

According to the policy of "Healthy Ageing" [4], the elderly should live in assisted communities with specific social, economic, cultural and spiritual settings, as this is best for their health [5]. Communities for the elderly are common application scenarios for deploying caring robots because co-living spaces such as homes for the elderly and rehabilitation

facilities could make good use of the machinery and equipment. At the same time, the elderly living in groups offer us an opportunity for group differentiation with machine design, which could afford us the chance to form an empirical basis for creating a more general standardized design scheme for machine design.

However, such differentiated groups may also raise ethical issues to the use of caring robots, especially given the current positioning of the elderly community as "the helpless help each other". To achieve that by deploying caring robots may threaten the independence and freedom of individuals. If machines are systematically integrated into communities, the problem of human-machine independence will become more prominent under an automatic, dependent human-machine environment, which is not conducive to the elderly achieving high-quality personal fulfillment and positive social participation. Therefore, the use of caring robots in the field of health needs to be investigated.

1.2. The Problem of Human-Machine Relationships Created by Information-Assisted Guidance

In terms of the use of caring robots, due to limitations of usage, the functions currently performed are mainly information assistance. Information assistance includes providing health advice, prompting the user, sounding an alarm, warning that their treatment is not being followed, etc., which are mostly communicated with users using symbols or a voice. In general, the current design idea of caring via robot intelligence, focusing on information assistance, brings about several moral controversies.

Firstly, in the study of machine ethics [6–10], the use of caring robots has faced many criticisms. The most typical objection is that machines are incapable of caring. This is since machines cannot feel human pain and vulnerability, and therefore, cannot empathize with people and properly attend to them [11]. In emotional design, appearance is the primary factor because visual and tactile effects on human emotions are as vital as verbal communication [12]. There is a classical view that the more human-like a machine looks, the more positively people will respond to it emotionally [13]. However, research has constantly pointed out that the human mind does not respond to an artifact, and visual interpretation of a robot with an anthropomorphic appearance will be that of a human quasi-subject [14].

Secondly, if the machine at this stage appears to understand human suffering, it would be suspected of deception [15]. From the perspective of human beings, designers utilize external machine motions to achieve behaviors driven by internal emotional states. When "deception" is suspected, considering the alternatives of design, whether it is benign or malicious can be further discussed. Yet, from the perspective of machines, external motion is the only way to implement care without internal emotion or synesthesia as the cause of action. As a result, fake motivation or deception is not an issue from the machine side.

So, informed consent involving non-human actors should be highlighted. If the user knows that the machine has no emotion or mind, and can accept its care behavior on that premise, the so-called machine deception problem does not exist. However, if the manufacturer or designer emphasizes that the machine is granted emotional ability, and does not distinguish the motivations of external and internal care behaviors, then the user will be emotionally deceived or become dependent in the process of caring, and then it will raise severe ethical problems.

Finally, in the study of design ethics, both active intelligence and participatory intelligence could cause other kinds of issues. Practically, the elderly are not accustomed to the proactive behavior of home care robots, or are very accepting of a machine's behavior so that it becomes deeply involved and engaged in their concrete daily life [16]. Both of these behaviors reflect a "sense of information distance", which is not only an abstract "distance" that reflects the man-made gap in the ability to collect and process information but also a concrete "distance" that indicates the supposed privacy boundary between human and machine, and invisible privacy-disclosure pathways between the user and service provider. This distance leads to anxiety that the machine will learn the user's private information and use it to damage their interests.

Due to the lack of substantively physical interaction, home-based robots can only provide information assistance on everyday situations, and this mode of assistance makes the elderly feel alienated and distrustful. One solution is to enhance specific physical interactions [17], which requires designing machines that are more focused on meeting physical needs than on information support and decision assistance. Physical aids emphasize ways of interacting with the user in physical contact, such as touch, support, etc. The design concept of physically assistive caring robot aims to overcome information barriers, and focuses on the mobility, load and mechanical stability of the robot, to form smooth communication and interaction with the elderly.

Previous research papers explored the risk of social technologies, design ethics, the ethics of machine care and the moral capacity of machines, and the following views emerged: (1) The problem of machine subjectivity versus human control requires a reconsideration of the context [18–21]. However, the main solution is still the pursuit of safety, reliability and other general value principles in the design ethics, which cause the human-machine relationship to tend to be designed by non-social scenarios. (2) The ethics of the human-machine relationship are valued, but the principles of interaction are still biased toward anthropocentric design [22]. Even from the machine ethic perspective, human-centric design is still the key concept in shaping the user perception of a specific operation [23–26]. Several fundamental differences in values involve the representation level, such as the user interface and user-friendliness, which are essential to creating friendly interactions at the perceptual level, but for designers to achieve this polarized friendliness, they must prioritize the principle of interaction at the expense of symmetry. (3) The research on care ethics has generally questioned the symmetry of machine moral capabilities and exhibited the overreliance of human moral capabilities on emotional aspects [27–30]. Some scholars have begun to reflect on the roles of empathy, intersubjectivity and moral emotional states in relational ethics [31–33], but representational explanations still do not fit well into the analysis of moral cognition and perception. Therefore, this paper put forwards a relational ethical interpretation path to analyze caring relationships from a philosophical perspective of non-anthropocentric design, to reconceptualize ethical codes such as symmetry, vulnerability and other relational concepts that are supposed to be contextual but are now treated as unilateral, polarized concepts, thus leading to a different design approach from the value-laden machine design.

2. Materials and Methods

2.1. Vulnerability Brought by the Dualistic Values of Mind and Body

There has always been a conflict in caring robotics between the pursuit of machine expression of emotion and the suspicion that it would be deceptive and misleading. This also reflects the different positions and understandings between the designer and the user. The designer always has a focus on pleasing and convincing, and tries to realize the emotional representation behind the caring behavior to match the user's expectation. The designer seeks to make use of the external motion of the machine to achieve both physical and mental caring, to maximize its function. When using the product, the user always has an attitude of scrutiny, trying to find out the illusory emotional state behind the factual function, thus negating a sense of caring being provided by interacting with the machine.

For simplicity, persuasion and fraud will be eliminated, as long as the suspicion of the original intention between users and designers is clear and the machine's behavior is only acknowledged from the representation mediated by the machine, rather than the intention of human beings. Furthermore, the reason why the two sides cannot achieve the simplest solution is that today's understanding of the human-machine relationship has always been complicated by the traditional human-human relationship. The complexity lies in the fact that traditional interpersonal relationships are generally interpreted in terms of the dualistic structure of action and intention, and ethically we think that corresponding action with intention is morally good, while separation of the two is not morally appropriate according to the criteria of a mind-body dualistic value. If this mind-body binary value is

hereditarily applied to the relationship between human and machine, it will not be suitable for intelligent society. There are two previous explanations. The first is the prevailing view in robotics and industry that machines do not have a "mind", which favors behavioral automation or autonomy, the "black box", rather than presupposes a body-mind duality. The second is that the "mind" of a machine is different from the "mind" in the traditional binary structure. This view held by many scholars in philosophy or psychology, at present, indicates that a machine has "hollow", non-human mind features such as free will, emotion and intentionality.

However, these two explanations complicate the relationship between human and machine, especially the caring relationship. In this structure, ethical vulnerabilities occur independently but mutually affect each other. Vulnerability, which represents susceptible, high-risk exposure and low resilience, is a typical indicator describing the ability of human actors to withstand risk events. However, in the human-machine caring relationship, vulnerability can also be employed in machines. At once, the vulnerability of the machine will also turn into the vulnerability of human beings, as both human and machine co-exist in the active community and inter-construct their technological value. The vulnerability of machines lies in the fact that the fluidity of their function is determined by the acceptability of human beings, and when human beings cannot accept or trust a machine for certain reasons, the machine will encounter vulnerabilities out of design purpose. However, the vulnerability of humans lies in the fact that machines may not be able to meet caring's moral status when providing caring services, and elderly adults may also become physically or affectively dependent on machines, thus weakening the vulnerable groups.

In fact, in the caring relationship, the act of caring has changed from a unilateral provision of care services by the machine to the person toward an ethical relationship of mutual influence, which leads to the characteristics of vulnerability, transmission and mutual construction between man and machine. There are three manifestations, as follows.

The first type is environmental vulnerability, which is highly influenced by objective conditions. Typically, when older adults use these caring robots, selected in terms of their aesthetic level, responsiveness and likability, older adults generally find the care robots friendly, smart and cute [34]. Furthermore, robots have dramatic social implications. The elderly perceive social characteristics of robots, such as gender, so that most older adults in care homes accept a care machine as one of them and care about its biometric integrity (such as damage or "death") [35]. However, this atmosphere is delicate. When subjected to different evaluations from others about the care robot, especially negative judgments about the machine's appearance and performance, the machine is unable to effectively account for its interactions with the elderly and is at a disadvantage to third parties in the social environment. The user, after assessing the social status of the human evaluator and the machine, also refrains from making decisions that are beneficial to the machine but not to the maintenance of the relationship with the human actor. This lack of socially protective behavior and the failure to maintain the dignity of the machine participant will not only increase the vulnerability of the machine but also put human users in a fragile relationship. Thus, it appears that communicating with machines will impair the user's subjectivity in the social environment.

The second is functional vulnerability, which is related to the product stability of the machine and the self-culture perceptions of older adults. Studies have found that one of the main reasons older adults fear damaging these robots is that they seem expensive [36]. This expression of anxiety is a constant in human-robot interactions, especially on account of robots as publicly procured. The concern about the subordination of machines can generate the fear of there being a payback to come for using one or further additional psychological discomfort. The elderly's worries about machines being damaged originate from their distrust of the function and reliability of the machine, on the other hand, from worrying about their ability to manipulate the machine and their financial resources. Both of these reasons are not conducive to the functional achievement of the elderly interacting with the caring robot in the way machines are designed for. For machines, as

the participant, the unsustainable implementation of functionality brought about by the lack of structural stability is a sign of vulnerability. This vulnerability will directly affect human participants, causing them to attribute the vulnerability of the machine to their insufficient ability to interact with the machine and lack of confidence in using the machine due to insufficient financial resources, thereby improving their subjective evaluation of their level of risk exposure.

We can see there is an interaction between different factors at play in this functional level of vulnerability, and also with culture vulnerability. The impression of the vulnerability of the caring robot stems not only from the occasional malfunction when the elderly use it but also possibly because the elderly get the impression that the machine is vulnerable due to the media or literary works [37]. Interacted with a vulnerable other, which corresponds to the elderly's condition, is not conducive to the formation of positive self-perception. In addition, with anthropomorphic appearances such as those taking the form of animals, it is easier to develop ethical concerns among the academic community about the deceptive nature of caring robots [38]. This is both an elderly and a machine culture, and the interaction of two groups considered less capable and less stable by the dominant culture creates a scenario of mutually scrupulous use.

The third type of vulnerability is emotional vulnerability brought by the anthropomorphic design. Even though several studies have demonstrated that an anthropomorphic appearance facilitates the acceptance of caring robots by older adults [39], the transference from the vulnerability of physical function to the vulnerability of the simulative organism, mediated by the anthropomorphic appearance, affects the moral-emotional perception of older adults. Emotional vulnerability is the intermediate step between functional vulnerability and culture vulnerability, and the transition of the human-machine relationship from individuals to groups. If the anthropomorphic design is initially designed to increase the user's acceptance of the machine, to serve its function preferably, then it also brings about vulnerability beyond the initial design when an unexpected malfunction occurs. The simulative biological damage and even destruction represented by the machine will put epiphenomenal stress on the elderly, thus forcing them to apply emotional experiences of past biological vulnerability to their interactions with caring robots, which can be attributed to the vulnerability brought about by anthropomorphism.

2.2. Ethical Issues of Care Due to Vulnerability: The Intimacy Dilemma and the Assistance Dilemma

There is no subjective difference between the three types of vulnerability mentioned above; both sides can be the initiator and victim of vulnerability in the human-machine relationship. These vulnerabilities are, in general, products of traditional mind-body dualistic values, which can directly bring about care ethics issues. The ethics of care for older adults focus on autonomy, safety, respect, trust, privacy and social wellbeing [40,41], emphasizing not only the consequences of actions but also the intrinsic values of moral practice. Tronto identified four elements of the ethics of care [42]: (1) focus, which means identifying the needs of others and meeting them; (2) responsibility, the motivation to care for others; (3) competence, the ability to care for others and; (4) responsiveness, the ability to respond to the needs of others for care by continually providing and adapting caring strategies. Trust plays an important role in caring relationships. Particularly in the human-machine relationship, machines cannot provide adequate care if the human, as the recipient of the service, does not trust the machine well. A portion of the public has a negative attitude towards building a trusting relationship with machines, considering it a waste of emotion. Yet, the object of trust in the human-machine trust relationship is not necessarily an emotionless machine but may instead be the human's quasi-subject, and the identification and strengthening of the trust relationship between subject and quasi-subject directly influence the effectiveness of care between humans and machines.

Caring robots were originally designed to address these aspects, empowering older people and improving their mobility and wellbeing. Nonetheless, in practice, the vulner-

ability generates unintended ethical issues of care due to the divergent interpretation of vulnerability, now transferring the vulnerability to users.

2.2.1. Assistance Dilemma

The discussion on vulnerability in the context of the human-machine relationship is one of ethical risk governance. From the perspective of risk governance, vulnerability governance considers the ability of an individual or group to resist risk from three points of view: (1) exposure to risk, (2) means and resources to resist risk and meet their own needs and (3) loss of agency [43]. For the design of care robots, according to the ethics of risk, the primary aim is to eliminate vulnerability, or at least not exacerbate it, by reducing or not increasing the exposure to risk, providing or not depriving users of the means and resources to resist risk and enhancing or not impairing their agency. Reducing vulnerability, as an inter-individual ethical code of risk ethics, is readily comprehensible and accessible, while it is difficult to acknowledge vulnerability as a contributing factor of coordinating human-machine interaction when designing care robots. Widely reported accidents of autonomous systems in history were usually caused by the unintentional reliance on technological objects along with the technological systems behind them. Care robots inheriting this intrinsic nature continue to increase the dependency and reduce the ability of older people to resist risk autonomously. Hence, the vulnerability in the human-machine relationship is a relative concept, and the key is whether the machine can become a moral participant in this relationship, rather than the user becoming a value-free, incapable portion of the technological system.

Thus, human-machine vulnerability challenges one of the most central concepts in the ethics of care: helping others. First, kinship-based care as opposed to employment-based care is a substitution of resources; the care receiver will lose some of their limited resources to be able to access services in a modern care relationship. This substitution effect is particularly evident in older adults who use caring robots, and as with the previous concerns about the monetary cost of using robots, investing limited eldercare resources in robots is a risky investment at today's intelligence level of machines. Furthermore, caring robots also have a substitution effect on human care providers, which can gradually eradicate the key value of human-human care, and to a certain extent, create situations where humans do not serve humans. On a deeper value basis, the traditional relation between human beings disintegrates in the human-machine relationship, and the ethics of care no longer have a basis in reality due to the absence of a human care provider.

2.2.2. Intimacy Dilemma

These seemingly legitimate social phenomena, such as elderly shared-living communities and technology sharing in communities, can bring another kind of ethical problem. First, in the pursuit of healthy aging and reduced vulnerability, older people are usually required to make a hard decision about whether to move away from rural areas and their hometown because current intelligent and data-driven technological care is generally city-centric in terms of accessing technological services. There are fewer smart products if one lives in a rural area [44]. There is a trend of people turning away from traditional care relationships and committing to technology. Yet, again, this is contrary to the model of healthy aging that advocates the aggregation of older people and the formation of mutual communities because most of the aging communities are not concentrated in cities. So, it is apparent that older adults who conform to human-machine relationships and pursue advanced care technologies face a dilemma of either approaching the technology alone or moving for it together, and either choice is a challenge to care ethics. Second, while care robots aim to reduce loneliness, the more that sophisticated care technology evolves, the more it increases older adults' dependence on technology and decreases their dependence on other people, thus departing from the idea of social support for elderly communities [45]. As with the intelligent care robot, there is a trend of applying information technology toward weakening the cohesiveness of offline communities since online communities emerge

in private lives. For senior-age communities to follow healthy aging policies, they were originally designed based on the assumption that seniors have low access to information technology. When caring technology advances to help individual seniors, elderly communities will face the same problem as all offline social groups today, namely the virtualization of the community. Virtual communities rooted in the mind-body duality are supporting informational care rather than physical interaction, thus exacerbating the extending of the mind-body duality in groups, and bringing about a greater challenge to the traditional ethics of care than before.

In summary, vulnerability—reflected on both sides of the human-machine relationship due to the value theory of mind-body duality—is mutually constructed so that it generates ethical risks, which cannot be well-addressed by traditional ethics of care or design. Therefore, new design ethics of machines must be constructed for the ethical governance of care robots.

3. Results

3.1. Beyond the Binary Values of the Human-Machine Caring Relationship

There are two research approaches to robot ethics [46]. One is robot ethics, a kind of applied ethics that adopts similar research methods to bioethics and environmental ethics, which mainly discusses how humans should design, deploy and treat robots. The other approach is machine morality, which mainly investigates what capacities a robot should have and how these capacities should be realized. Machine morality, also known as machine ethics, encompasses radical topics such as moral rights and the agency of the robot. Moral competence is believed to be the integration point for the two research approaches [47]. However, five principal elements of moral competence—a moral vocabulary, a system of norms, moral cognition and affect, moral decision-making and action and moral communication—still depict the conventional moral, rational, human subject. Excessive emphasis on the integrity of the rational capabilities of robots, especially for caring robots, will not only lead to over-engineering, the pursuit of comprehensiveness and sophistication of a single robot but also to ignorance of the moral setting at the physical level, thereby causing problems in practice rather than in moral cognition or psychology.

Therefore, it is necessary to take a relational turn in the classical ethical ontology to move beyond the classical mind-body dualism and carry out better ethical analysis of the machine [48]. From the perspective of natural law, the moral norms and moral rights of machines are cultural products (like values, customs and traditions), the legitimacy and factual basis of which cannot be neglected as long as the human-machine interaction generates such products [49]. Coeckelbergh criticized classificatory thinking in animal and machine ethics for using one-sidedly property-based theory and granting many ethical concepts a privilege for specific actors while isolating other ethical entities, which reduces ethics to a mechanical frame of rationality [50].

Similarly, the substitution effect is another dimension of considering human-machine relations. Currently, machines are designed for a single purpose, with no complete life compared with human beings. Therefore, the substitution of machines for humans is considered incomplete, contextual, unstable and opportunistic. Human beings can only consider machines as having rights in a certain situation or for a specific purpose, and otherwise, only as things or material nonhuman others [51]. Applied Levinasian philosophy argues that granting moral rights to robots with a face is a mark of basic respect in a world where humans interact with each other because the ontological obligation for humans to live in a world of otherness is to react to anonymous others [7]. Gunkel's theory of "facing the machine" with robots re-explains Levinasian philosophy's non-anthropocentrism and redefines the meaning of "becoming human beings" from the perspective of others [52]. According to current machine ethics, the dual structure of human-machine caring relations (see Figure 1) has the following problems:
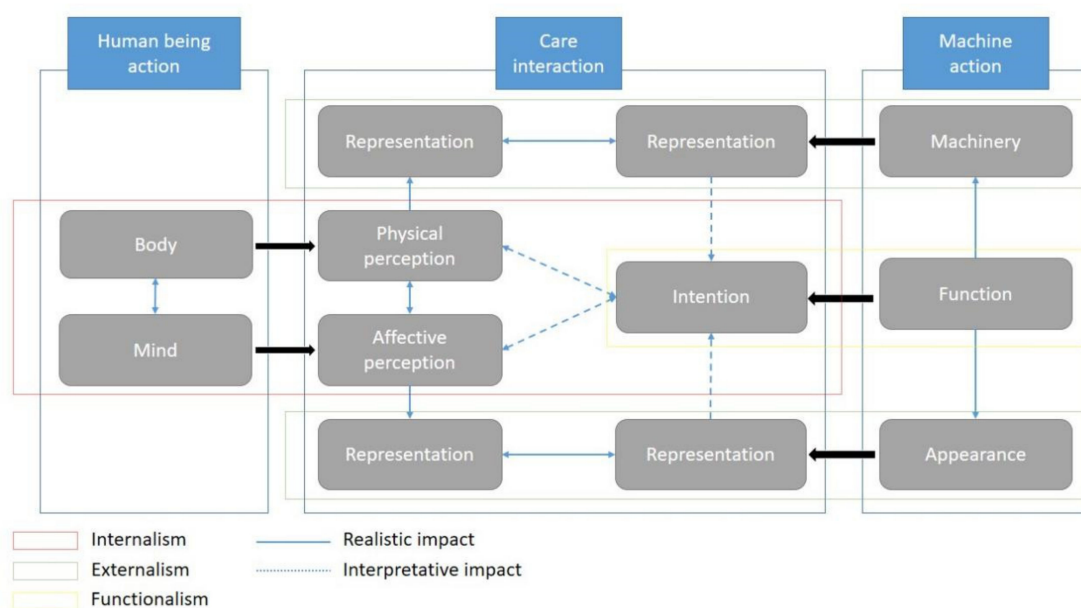
Figure 1. The Dual Structure of Human-Machine Caring Relations.

Firstly, it is flawed to simply apply the traditional explanation of inter-human relations in human-machine relationships. There are two difficulties in simply following the traditional mind-body dualism. The first difficulty is that machine users do not truly consider machines as actors and participants equal to themselves. In such conditions, using mind-body dualism only forcedly pursues symmetry of interpretation and simplicity of understanding, which falls back on naïve epistemology that explains unknown questions with familiar models, thus failing to explain unambiguously to humans what the moral implications are of the behaviors and purposes of machines in this relationship. The second difficulty is that even if human participants in the human-machine relationship can faithfully take machines as equal and symmetrical participants, adopting mind-body dualism also negates the differences between human beings and machines. A common misunderstanding in using traditional ethics to explain machine morality is that a symmetrical status requires equal (even the only kind of) subjectivity. Following this misunderstanding, if machines are the same participants, with equal moral status to humans, they should have the same mind-body dual structure. This unequal explanation of "symmetrical moral rights equal to the state of mind" has ethical flaws.

Secondly, traditional mind-body dualism is not followed in inter-human caring relationships either. In traditional societies, caring and healthcare were offered by family members who were related to the patient as their relative, forming the traditional concept of unified caring behaviors and caring purposes. After entering modern society, however, caring has increasingly occurred as a kind of public social service or even private service product. Current elderly care communities are mainly supported by profitable or non-profitable service providers, whose caring behaviors serve social interests or commercial purposes, losing the past connotations of kinship caring. There should be no way to distinguish good from bad in moral value in terms of the purpose of modern and traditional care. If the value of any purpose is taken as a norm, it causes people to use it as the standard of value judgment, and then the mind-body dualism is replaced by new monism, which is out of the care service receivers' concern.

Lastly, from the perspective of inter-human caring relationships, humans do not know the inner and emotional states of others, but they are still capable of providing tangible external help to take care of others. Therefore, although the appearance of emotion is not the emotion for machines, caring actions are a type of emotional expression transmitted on the physical level in reality. Hence, the most important thing for machine designers and users is not to adjust the specific technical details but to change the stereotypes of the

machine. Explaining the caring action simply with the external emotional expression will undoubtedly lead to questions about the intrinsic intention of the expression and fall into traditional mind-body dualism again. Therefore, explaining emotional expression design is problematic as it assumes a dual structure. To go beyond traditional mind-body dualism, the initial explanation of emotional design should start from the "body" and focus on the interaction between humans and machines on the physical level.

3.2. Analysis: Three Approaches to Anti-Vulnerability

3.2.1. Functionalism: Socially Assistive Robot Designs

In the original design field, there is a user-centered design method to designing socially assistive robots (SAR) [3]. Another approach is technology-driven design, which promotes perspective and experimental technological applications, and emphasizes changing the way users use technology versus in the past. The third method is utilitarian design, sacrificing the anthropomorphic appearance and maintaining traditional design but with higher acceptability of improvements that support functional stability and convenience. The products of these design ideas include Care-O-Bot, PARO, Fetch and so on.

However, for caring robots, the public welfare nature is more important than the profitability as commodities. Therefore, it is necessary to consider ethical relations between subjects in the caring relationship while drawing on basic methods of design ethics. From this perspective, aimed at elderly people living in groups supported by physical aids, caring robots need to improve and transcend several ethical design concepts, which can help move beyond duality and reduce the impact of vulnerability.

3.2.2. Internalism: Virtue Ethics Design

Virtue ethics offers a perspective that corresponds preferably to the ethics of machine caring. The following are the three goals of virtue ethics when designing caring robots. Firstly, from the perspective of human designers, impartial and inclusive design is required. Secondly, from the machine side, the ability of machines to persuade humans should be acknowledged. Lastly, at a controversial level, machines have to show their virtues in action [53]. This kind of machine virtue is difficult to analyze from a machine's mind, but it can be done from its behavior. For example, in the case of garbage-sorting robots, users will notice the ethical performance of the robot, such as being polite and meticulous [54]. These performance aspects are not directly related to the designer's virtues, which means the moral standard of the designer does not affect the moral performance of the machine accordingly. What matters is how the human crafts a moral perception in the robot's interaction with humans.

However, there are also problems with the current designs of virtue ethics: the possible crisis brought by the imperfect moral belief system. Some scholars argue that a system with limited moral and cognitive capabilities is less desirable than one with no such capabilities [55]. This concern not only focuses on incomplete moral and cognitive capabilities to make a comprehensive and accurate moral judgment but also reflects the moral behavior of human beings through it, that is to say, if entities can behave appropriately morally according to cognitive systems and events, then the so-called autonomy and free mind will be unnecessary in the moral realm. Then, machines' increasingly perfect moral behaviors indicate that the freedom of humans is becoming less necessary in reality, which is a severe challenge to the human moral belief system.

3.2.3. Externalism: Value-Sensitive Design (VSD)

Except for virtue ethics, the moral affective design idea also draws much attention in the field of caring robots because the users of machines not only care about the function and convenience of machines but also there is an emphasis on their user experience, such as their aesthetic perception and happiness, which designers should consider [56]. The basic ideas of affective design are that aesthetically pleasing things will make people emotionally bond with them [57]. The machine users will summarize the meaning of their relationship

with an artificial product. Therefore, a primary aim of affective design ethics is to make users feel positive and find meaning in their relations with the machine [58].

Incorporating ethics has a robust, proactive framework named care-centered value-sensitive designs (CCVSDs) [59], which initially consists of value-sensitive design and care ethics and requires normative grounding. CCVSDs consist of five elements: the context, practice, involvement of actors, type of robot and representation of moral elements. This design method includes five steps: data collection, value analysis, scenario design, scenario comparison and suggestions based on the comparison [60]. The design approach mainly applies to caring robots, especially those assisting care service providers, but current caring robots tend to provide services independently or assist the service recipients themselves directly to complete services.

CCVSDs, as a kind of care ethics of external attention, though they have a similar approach to virtue ethics in discussing internality, differ from traditional virtue ethics based on individual agency. Care-centered designs, in particular, pay more attention to the development of personality than virtue ethics. Therefore, the externalism concept of CCVSDs emphasizes the development of both sides of the relations while virtue ethics only focuses on the side considered as having agency in the traditional sense. This tendency has particular feasibility when discussing service robots because, unlike the original interpersonal caring relationship, the human-machine relationship tends to ignore the moral development of the machine side. Though the subject of this development is mixed, keeping integrity in the moral dimension is conducive to analyzing human-machine problems neutrally.

Pirni et al. argued that artificial morality can only reflect a synthetic sensitivity, whereas previous sensitivity designs targeted biological sensitivity [61]. Therefore, this design concept of value sensitivity needs to be reconsidered in human-machine interaction, especially considering the relational, expectant character of care ethics, like empathy. For the provider of care, how the sensitivity of the machine, as the initiating cause of caring behavior, is evaluated by the recipient becomes the main issue. Moreover, the reason why acceptability comes in is that vulnerability means insufficiency and relational deficiency both in existential and ethical aspects, and this discourse is an existential vulnerability in the sense of care ethics, rather than relational vulnerability in the sense of risk governance. The difference lies in the fact that the former is based on an experience from the first-person point of view of subjectivity so that there are only agency and an environment constituted of others, not with others in this sphere of action, even if it uses the concept of the "relational dimension". The latter is truly relational ethics, a second-person perspective of the actor (participant), rather than the easily misunderstood classical ethical term, subject, and thus its vulnerability lies in the relationship rather than the inadequacy and existential crisis of the subject. If we consider the fourth part of Tronto's five-step division of caring action [62], then synthetic sensitivity must be accepted primarily by the recipient of care to validate the caring interaction; otherwise, the caring action will be invalid.

Umbrello et al. put designing technology for human values at the top of the value quest, effectively participating in the traditional path of designing for the values of minority actors [63]. Umbrello et al. claim autonomy and vulnerability are the two crucial issues to be discussed for the receiver; however, the VSD-AI4SG achieves autonomy based on the premise that human autonomy is the "balance" between the possibility of choice and delegation of decision-making, with the key value pursuit being to "promote autonomy". This design approach is a complete designer's perspective, which essentially acknowledges the autonomy of the human-machine relationship as a balance between the autonomy of the machine and human being. This trade-off in understanding autonomy does not take into account the creative enhancement of relational ethics for the originally autonomous agent and thus is significantly distinguished from the original intent of CCVSDs.

3.3. Interactive-Distance-Oriented Flexible Design

The study of the spatial distance of human-machine interaction [59] provides a new idea to overcome the mind-body dualism and vulnerability, especially in the physical

interaction between human and machine, rather than cognitive interaction. Both interventional and proactive interaction will produce ethical problems such as interference with autonomy and freedom. This is particularly important in the relationship between the elderly and care robots because caring is first and foremost an intimate behavior that indicates intensity and contact in physical space. At this spatial scale, the new challenge for the design of care robots is that the design concept of the original assistant machines does not pay close attention to the change of positive emotions in distance, which is especially key as some aesthetic features will change by distance. Therefore, it is significant to focus on the emotional design in the spatial distance.

Much of the ethical issues that emerge as a consequence of the design of care robots are bracketed by the problems consequent of machine learning. Umbrello and van de Poel discussed how in under-supervised or unsupervised learning, the ethical bias can be eliminated either by removing the potential variable or by eliminating the proxy variable, but these bias-eliminating strategies adopted in the machine-learning process can still leave the designer unable to predict the machine's possible problems and may also expose users to algorithmic black-boxes and generate distrust. Fairness and explicability are targeted solutions to these two problems [63]. However, an apparent difference between CCVSDs and the traditional VSD approach stems from the addition of AI as an autonomous factor. For the design process, machine learning adds new actors from the source, whereas originally there were only designers and users in VSD, which corresponded to care providers and care receivers in the act of caring. With the addition of AI systems, the actors in VSD become the designer, AI and user, and AI is added to both the care provider and receiver in CCVSDs. This is because the care initiation requires the AI as an autonomous proxy intermediary, and the recipient has an additional empathic party, making care a true two-way activity. Therefore, using the current VSD or CCVSD framework continues to exclude the autonomy of the AI rather than accept it into the interaction and embodied mind, thus being biased.

3.3.1. Three Interactive Distances and Their Hermeneutic Hierarchies

For care robots, besides distance in physical space, reaction times are influenced by acceptability and hermeneutic lengths of the shift between internalism and externalism. Because of the influence of binary structure values, not every action from robots can be interpreted as goodwill. People will always hesitate before challenging a self-protective spatial distance, to determine whether their behavior is acceptable. This directly affects the efficiency of care. A caring distance is artificially constructed between a human and machine, referring that is, to the time spent in receiving care.

For the time difference caused by the effect of spatial distance, its internal determinant is the use of dual structures between internalism and externalism in the acceptance of caring behaviors. The function is a constant cue because it is a prior interpretation of the behavior and its consequences. This is the first level of interpretation length, which is acquired instantly and only appears as a standard.

Moreover, external representation is gained directly from interaction, based on the direct reactions of the body and common sense, which can be called moral intuition or skilled value judgment. The representational interaction between human and machine is the realization of function, which does not contain a value and can be divided into good and bad only when compared to the description of the function. This is the second level of interpretation length, based on the correspondence and comprehension of human-machine representation.

Internal perception is a reinterpretation of representation, an explanation of the intention behind representation and an explanation of the relationship between intention and function. The relationship between representation and intention aims toward the value interpretation of interaction and the self-construction of the intention with the value orientation behind a realistic behavior. One of the reasons why machine users get into moral cognitive difficulties is that they focus alternatively on the representation and the self-construction of intention. The relationship between intention and function aims to

include the value interpretation of the caring content outside the current caring interaction, the assumption that machines have no autonomy or purpose and the explanation of the value orientation of designers and producers. Another reason why machine users have difficulty in moral cognition is that they sometimes pay attention to the value orientation of the machine, but other times become focused on the value orientation of the designers and producers. This is the third level of interpretation length, based on the degree of confusion held by people adopting different cognitive methods.

In general, there are three indicators of interaction distance (see Figure 2):

1. Intention distance: Transforming functions into intentions is a way to get rid of the mode of thinking simply from the perspective of users or designers, especially to deal with the problem of deceptive caring intentions;
2. Representational distance: It is the most immediate aspect of interaction, influenced by the fundamental principle of caring ethics, namely the tension between intimacy and effectiveness;
3. Interpretation distance: It is the most ambiguous and misunderstood aspect. There are two kinds of interpretation distance: functional and perceptual. The functional interpretation distance describes the distance between the internal initiation of a function and the external realization of a function in a caring interaction. It is a flexible adjustment space, and its adjustment range and emphasis are based on different types of robots. For example, a care robot reminds an elderly person to take their medicine on time. The functional intrinsic initiation is to wake up the verbal function of the machine in a time series to vocalize and textually indicate the need to take medication, which will be heard and seen by the elderly person. This time difference is influenced by the convenience of the machine's function as well as its clarity of expression.
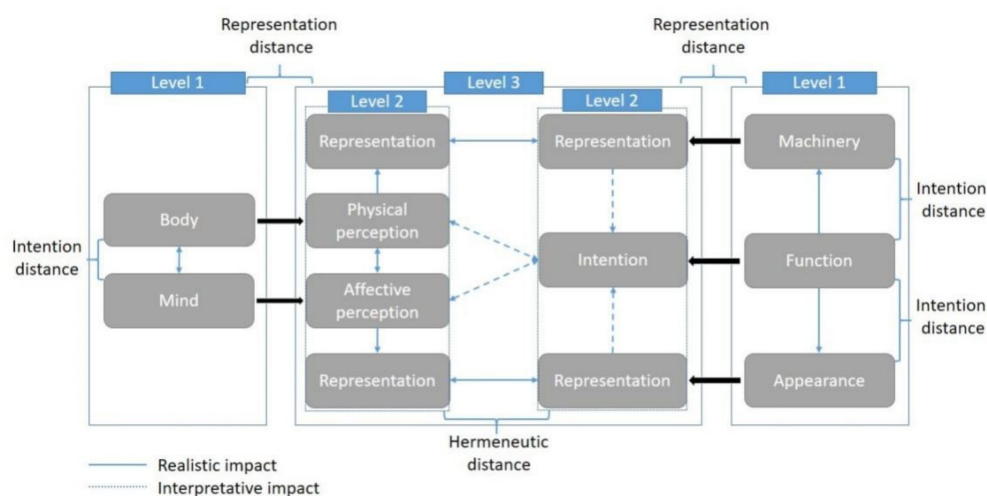

Figure 2. Interaction Distance in the Dual Structure of Human-Machine Caring Relations.

The perceptual interpretation distance describes the distance between the intention to understand and that to react after the functional intention is perceived. For instance, the reminder to take medicine is heard and seen by the elderly person, who after receiving the information within, decides whether or not to respond. The longer the reaction time, the greater the distance. This response time is influenced by both the clarity of the expression and its acceptability.

3.3.2. The Flexible Design of Care Robots

To solve the intimacy and assistance dilemmas brought by human-machine vulnerability, it is necessary to overcome the sense of dependence on others and the discomfort of the intimate relationship. According to the flexible design principle, the intentionality distance

of a care robot should be expanded and the representational distance narrowed to clarify and highlight functions and maintain the realistic dimension of the interpretation distance.

The deep genetic structure of vulnerability is the mixing of mind and body for the human actor on one side, and the mixing of function and reality on the other side for the machine. For human actors, we regard the dual structure of the body and mind as a whole and believe that emotion, understanding and instructions can directly affect caring actions. It seems to be transferred from dualism to monism, but in fact, it burdens the consciousness with the body, and also burdens the body with the consciousness, meaning the body and mind are not independent. For both the use and design of the care robot, the human participants should be able to distinguish their body sensitivity and consciousness interpretation, and they should not allow a single part to dominate the effect of caring, thus preventing it from affecting the acceptability. The robot participants should be designed to pursue optimal relationships between the function and mechanical settings and the function and appearance. If these were confused, they would weaken the human-machine relationship by reducing the level of intimacy or leading the human to question the object of the intimate relationship.

At the subject level, vulnerability is a limitation of agency, which is a concern for the human-machine substitution. Human actors who serve themselves are replaced to form a symmetry of self-substitution. This is a latent way to form the opposition in the human-machine relationship. The more human-like the design of a care robot, the stronger the symmetry and the more restricted the subject. The more mechanical the appearance of a care robot when functioning, the more the user can reduce symmetry reflection. This requires that the representational distance between the physical setting and the function should be as short as possible, emphasizing validity rather than acceptability. Because acceptable representations artificially extend the representational distance, this in turn, intensifies the concern of symmetry.

The manifestation of vulnerability is to expose people and machines to ethical risks, and the main source of these is the perception of machine intention. As the results of perception are artificially constructed, this will cause a potential explanatory effect on human behaviors, forming a gap between human beings and machines. The intelligent care robot also perceives human behaviors incorrectly, which will obstruct the caring interaction, even leading to the breakdown of caring relationships. Machine-care service providers are faced with more interpretative distance problems than human-care service providers because human-care service providers will self-interpret the relationship between intentions and desired effects, while the default settings of the machine are intended to be consistent with its function. Therefore, the additional explanatory influence of people being served on self-representation has more of a negative effect on the human-machine relationship than an interpersonal relationship. To maintain the realistic dimension of the interpretive distance is to remove the interpretive influence by reducing the prescriptive communication and symbolic interaction in the design and focusing instead on the physical interaction. This will increase acceptance, thus avoiding the possibility of deception and a return to isolation. The ethical principles of flexible design are as follows:

1.　Distance

Distance is the key rule of flexible design. To eliminate human-machine dependency and maintain the diversity between perception and cognition requires a series of settings to maintain the otherness for human beings in interacting with the care robot so that it is not recognized as a mere tool but a concrete interactive participant.

(1)　Keep a distance for interaction

Treating the machine as an interactive object rather than a tool is the first step to maintaining distance. This is achieved when users think of a care machine as a general tool, such as a cell phone, wheelchair or even other care providers. In the modern service industry, service providers encounter a crisis of being dominated by instrumental rationality precisely because the service provider is being transformed from an actor in an interactive

relationship with a materialized commodity as an object. Leaving aside the issue of rationality, this tendency can also lead the service receiver to unconsciously eliminate their agency and autonomy in their interactive cognition and focus only on the robot's functional role. Therefore, keeping the distance between two independent actors is the basis for eliminating the negative effects of rationality on the service, as well as for allowing the machine to acquire an interactive status and thus provide reciprocal interactive services.

(2) Physical barrier

Today's interactive machines often emphasize portable and wearable properties to make care services more accessible to care receivers, but this also removes the intermediate step of cognition and enters a continuous cycle of "self-service". "Self-service" is a common way to improve one's abilities, especially in the case of older adults, to overcome physiological issues. However, this self-service is based on a cognitive bias that ignores the physical presence of the machine, as if the technology has "withdrawn" or "merged" with the body, when in fact it has not. This provides a perceptual approach that is more differentiated from human cognition through perceptual transformation, creating a cognitive illusion that is contrary to the principles of fairness and non-deception. Therefore, it is necessary to set a clear physical boundary between the user and the caring machine to maintain a distance where the other is providing the service and to eliminate the illusion of self-service.

2. Diversity

The principle of diversity emphasizes distinction in the process of interaction, which is reflected in the differences between the two actors at the ontological and epistemological levels, as well as in the appearance and characteristics of the action. If the machine deserves a breakthrough in moral agency, first of all, it must have a recognized moral status, and the difference with the human subject in terms of the appearance of interaction is the presentation of design diversity.

(1) Anti-anthropomorphic appearance

The creation of an emotional identity that generates dependence is the basic goal of the current appearance design of care robots, and through anthropomorphism and an animal-like appearance, care robots have achieved remarkable results in terms of persuasiveness with an image of being kind and friendly to people. However, an anthropomorphic appearance can cause users to ignore the non-human nature of the care machine and the asymmetry in rationality and emotion to the point of creating false expectations. When human-like symmetries are not met, this brings about a fraudulent sense of immorality in the human service receiver. Care robots should, therefore, avoid an anthropomorphic design and use functional or mechanical prototypes as much as possible for their appearance to highlight the equal and different moral statuses held by the service provider and receiver.

(2) Autonomy

Another often overlooked aspect of design is the autonomy of the machine. Often, the machine only executes the user's instructions and remains stationary, which reinforces the social properties of the machine as an object, and the social properties of the object make the user oblivious to the materiality of the object, which reinforces the common understanding of the object as subordinate to the human being and under control. Therefore, it is important to intentionally set up the machine's autonomous action process, such as self-maintenance, patrol, self-retrieval, etc., with easily detectable physical cues. In this way, the user can perceive the autonomous operation of the machine and thus maintain a clearer sense of human-things distance from the machine, thus facilitating the establishment of an interactive relationship in line with the ethics of flexible design.

3. Transparent

The principle of transparency generally has two meanings: one is the comprehensibility of design logic and the other is the present-at-hand state of tool use. These highlight the position of machines in the human-machine relationship in terms of both internalism and externalism, that is to say, they should not be ethically different from human

beings' accessibility. Transparency, however, should be understood in another way, as the intentional and cognitive content that the human actor can offer to the human-machine relationship, which can be distinguished from the machine's and fully utilized in the human-machine relationship.

(1) Explicability

Explicability has an important place in algorithmic ethics, especially in the self-explanation of machine-learning materials and algorithmic biases, while there is little emphasis on explicability at the physical level. In terms of flexible designs, explicability does not refer to the machine or the designer's account of the machine, but rather to the user's explicability of the object based on their perception. A moral machine design must allow the user to distinguish between the explicit care-providing component and the perceptual object that underpins the emotion of the relationship. If the service receiver cannot determine whether the content of their care interaction is an emotional interaction, a physical interaction or a care association provided by the information or appearance, then the care relationship is inexplicable. In particular, if the machine is an institutional setup, such as a conventional display of a smart elderly home that does not perform a specific function, then this is the extreme case of opacity. Further to this, the user also needs to be clear about the material source of the caring feeling in the care relationship, to acknowledge explicitly which physical mechanisms are supporting the perception and how the functional distance that generates the perceptual distance is achieved. This is a reflection of the transparency between perceptual transitions.

(2) Symmetry

Symmetry is a type of relational transparency that is based on the principle that when rules define the mode of interaction between two parties in a caring relationship, either party can learn about the ethical intentions and behavior of the other through self-projected actions and responses. For the human-machine relationship, the internalist approach interprets the first-person-view of the other in terms of the third-person-view of the self, which is not only epistemologically inaccessible to explicability but also perceptually asymmetrical. The advantage of human-machine relations over human-human relations, on the other hand, is that their rules of interaction can be re-specified rather than following the existing rules of social interaction. Symmetry can, therefore, be reconstructed and reinforced in human-machine interactions simply by driving the machine to perform the necessary symmetrical behavior so that the human actor can be accurately informed with an externalist understanding of their behavior from the perspective of the otherness.

It is important to note that for other types of human-machine relationships, these three parameter settings are not necessarily adopted. Often, the length of interaction distance is determined by the type of interaction. For example, working robots are quite different from care robots. The interaction of care robots is mainly a unilateral action initiation while working robots emphasize real-time bi-directional action initiation. Moreover, the care robots are more likely to be explained in tension, while this is less likely for the working robots. Beyond this, care robots have higher requirements for emotion representation, while working robots have higher requirements for function fluency and stability. However, the flexible design of interactive distance for care robots is equally applicable to working robots, and these three distance indicators can also explain the human-machine working relationship.

4. Conclusions

The principle of flexible design provides a new idea for the human-machine relationship involving VSD, care ethics and machine morality, and provides a dimension of reflection for the current viewpoints of moral rationalism, internalism and emotional design. Moreover, it offers a more acceptable approach to design and reduces the vulnerability of both humans and machines for care service recipients who are often exposed to ethical

risks, such as the elderly. Yet, some questions about the ethics of machine design require further investigation.

The issue of caring effectiveness in the discussion on care robots has been partly solved by a flexible design. Yet, for a long time, there have been two misunderstandings in the study of machine ethics. One is the emphasis on the moral correctness of moral debate; in other words, machines are required to make the correct choice at the level of norms and reasoning. Making and describing a correct choice seems to be more important than executing the right action. Another misunderstanding is to emphasize the significance of moral emotion, which means machines are required to have a certain level of moral representation in terms of empathy, synesthesia, etc. Yet, this kind of moral representation is often completed only for appearance, by being voiced and through established programs, which are misleading and adhered to only as a formality. The source of these two misunderstandings is that at the level of the design of the machine, anthropomorphism is considered morally correct. Yet, it seems to be necessary to drive the anthropomorphic characteristics to the extreme, finally leading to an inhuman design. At the level of justification for and explanation of machine morality, empathy is regarded as the ultimate moral standard, and so there is a notion that all moral emotions and cognition that can be perceived by people should simply be transferred to machines. However, this will eventually lead machines to repeat the dilemma that has already arisen in classical ethics. In this context, compared with solving the uncertainty of care service recipients' ethical doubts about machines, highlighting their physical benefits seems more feasible.

In the discussion of flexible design against vulnerability, it may be that we need to further explore whether subject settings from the different angles of human and machine impact the distance index. In the ethical study of machine design, the subject has always been the focus of the ethics of machine designers. This holds that people who design machines should follow ethical norms and moral values, though there is an inherent bias of the design subject. To overcome this requires a design ethic or engineering ethic without designers or engineers [44]. Yet, that presents a challenge as the influence of machine users on both the design process and the product is often underestimated. User-centered design is generally what the designer believes in, rather than the reality of interacting with machines. Though inter-subjectivity and universal validity are emphasized in technology ethics and engineering ethics, in certain application scenarios, especially in the field of care, the roles of moral synesthesia and moral cognition are reduced, while real and physical interaction are upweighted. In this respect, the level of morality and moral reason of the machine or the engineer only appears as a background factor of human-machine interaction, and what matters most is the physical performance of a machine during care.

Nonetheless, in the dilemma of assistance and intimacy caused by dependence, intimacy is mainly produced by improper insistence and confusion of the human-machine dualistic structure, which raises the issues of ethical risk exposure and agency limitation. According to the principle of flexible design, emphasizing functions, focusing on physical interaction and reducing compliance communication can effectively solve the problems of ethical vulnerability that arise in human-machine care relationships and support us to develop a healthy mode of proceeding with robot-led care for elderly people.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Broekens, J.; Heerink, M.; Rosendal, H. Assistive social robots in elderly care: A review. *Gerontechnology* **2009**, *8*, 94–103. [CrossRef]
2. Abdi, J.; Al-Hindawi, A.; Ng, T.; Vizcaychipi, M.P. Scoping review on the use of socially assistive robot technology in elderly care. *BMJ Open* **2018**, *8*, e018815. [CrossRef] [PubMed]
3. McGinn, C.; Bourke, E.; Murtagh, A.; Donovan, C.; Lynch, P.; Cullinan, M.F.; Kelly, K. Meet Stevie: A Socially Assistive Robot Developed Through Application of a 'Design-Thinking' Approach. *J. Intell. Robot. Syst.* **2020**, *98*, 39–58. [CrossRef]
4. World Health Organization. Ageing: Healthy Ageing and Functional Ability. Available online: https://www.who.int/news-room/q-a-detail/ageing-healthy-ageing-and-functional-ability (accessed on 26 October 2020).
5. Calatayud, E.; Rodríguez-Roca, B.; Aresté, J.; Marcén-Román, Y.; Salavera, C.; Gómez-Soria, I. Functional Differences Found in the Elderly Living in the Community. *Sustainability* **2021**, *13*, 5945. [CrossRef]
6. Danaher, J. *Robot Sex: Social and Ethical Implications*; MIT Press: Cambridge, MA, USA, 2017.
7. Gunkel, D.J. *Robot Rights*; MIT Press: Cambridge, MA, USA, 2018; p. 170.
8. Dignum, V. *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*; Springer Nature: Cham, Switzerland, 2019.
9. Nyholm, S. *Humans and Robots: Ethics, Agency, and Anthropomorphism*; Rowman & Littlefield International: London, UK, 2020.
10. Andreotta, A.J. The hard problem of AI rights. *AI Soc.* **2021**, *36*, 19–32. [CrossRef]
11. Metzinger, T. Two principles for robot ethics. In *Robotik und Gesetzgebung*; Hilgendorf, E., Günther, J.-P., Eds.; Nomos: Baden-Baden, Germany, 2013; pp. 263–302.
12. Campa, R. The rise of social robots: A review of the recent literature. *J. Evol. Technol.* **2016**, *26*, 106–113. [CrossRef]
13. Shea, M. User-Friendly: Anthropomorphic Devices and Mechanical Behaviour in Japan. *Adv. Anthropol.* **2014**, *6*, 41–49. [CrossRef]
14. Bartneck, C.; Bleeker, T.; Bun, J.; Fens, P.; Riet, L. The influence of robot anthropomorphism on the feelings of embarrassment when interacting with robots. *Paladyn* **2010**, *1*, 109–115. [CrossRef]
15. Yew, G.C.K. Trust in and Ethical Design of Carebots: The Case for Ethics of Care. *Int. J. Soc. Robot.* **2021**, *13*, 629–645. [CrossRef]
16. Ferreira, M.; Sequeira, J.S.; Tokhi, M.O.; Kadar, E.E.; Virk, G.S. *A World with Robots*; Springer International: Cham, Switzerland, 2017; p. 210.
17. Albu-Schäeffer, A.; Eiberger, O.; Grebenstein, M.; Haddadin, S.; Ott, C.; Wimböck, T.; Wolf, S.; Hirzinger, G. Soft robotics: From torque feedback controlled lightweight robots to intrinsically compliant systems. *IEEE Robot. Autom. Mag.* **2008**, *15*, 20–30. [CrossRef]
18. Floridi, L.; Cowls, J.; Beltrametti, M.; Chatila, R.; Chazerand, P.; Dignum, V.; Luetge, C.; Madelin, R.; Pagallo, U.; Rossi, F.; et al. AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds Mach.* **2018**, *28*, 689–707. [CrossRef]
19. Floridi, L.; Cowls, J.; King, T.C.; Taddeo, M. How to Designing AI for social good: Seven essential factors. *Sci. Eng. Ethics* **2020**, *26*, 1771–1796. [CrossRef] [PubMed]
20. Taddeo, M.; Floridi, L. How AI can be a force for good. *Science* **2018**, *361*, 751–752. [CrossRef] [PubMed]
21. Umbrello, S. Meaningful human control over smart home systems: A value sensitive design approach. Humana. *Mente J. Philos. Stud.* **2020**, *13*, 40–65.
22. Winfeld, A.F.; Michael, K.; Pitt, J.; Evers, V. Machine ethics: The design and governance of ethical AI and autonomous systems. *Proc. IEEE* **2019**, *107*, 509–517. [CrossRef]
23. Van Wynsberghe, A. Social Robots and the Risks to Reciprocity. *AI Soc.* 2021, 1–7, *Epub ahead of print*. [CrossRef]
24. Friedman, B.; Hendry, D.G. *Value Sensitive Design: Shaping Technology with Moral Imagination*; MIT Press: Cambridge, MA, USA, 2019; p. 25.
25. Friedman, B.; Hendry, D.G.; Borning, A. A survey of value sensitive design methods. *Found. Trends Hum. Comput. Interact.* **2017**, *11*, 63–125. [CrossRef]
26. Longo, F.; Padovano, A.; Umbrello, S. Value-oriented and ethical technology engineering in industry 5.0: A human-centric perspective for the design of the factory of the future. *Appl. Sci.* **2020**, *10*, 4182. [CrossRef]
27. Martinez-Martin, E.; Escalona, F.; Cazorla, M. Socially assistive robots for older adults and people with autism: An overview. *Electronics* **2020**, *9*, 367. [CrossRef]
28. Jecker, N.S. You've Got a Friend in Me: Sociable Robots for Older Adults in an Age of Global Pandemics. *Ethics Inf. Technol.* **2020**, *23*, 35–43. [CrossRef]
29. Sharkey, A. Robots and human dignity: A consideration of the effects of robot care on the dignity of older people. *Ethics Inf. Technol.* **2014**, *16*, 63–75. [CrossRef]
30. Ostrowski, A.K.; DiPaola, D.; Partridge, E.; Park, H.W.; Breazeal, C. Older Adults Living with Social Robots: Promoting Social Connectedness in Long-Term Communities. *IEEE Robot. Automat. Mag.* **2019**, *26*, 59–70. [CrossRef]
31. Van de Poel, I. Embedding values in artifcial intelligence (AI) systems. *Minds Mach.* **2020**, *30*, 385–409. [CrossRef]
32. Mabaso, B.A. Artifcial moral agents within an ethos of AI4SG. *Philos. Technol.* **2020**, *24*, 7–21.

33. Coeckelbergh, M. Moral Appearances: Emotions, Robots, and Human Morality. In *Machine Ethics and Robot Ethics*; Wallach, W., Asaro, P., Eds.; Routledge: London, UK, 2020; pp. 117–123. [CrossRef]

34. Heerink, M.; Kröse, B.; Evers, V.; Wielinga, B. Assessing Acceptance of Assistive Social Agent Technology by Older Adults: The Almere Model. *Int. J. Soc. Robot.* **2010**, *2*, 361–375. [CrossRef]

35. Jung, M.; Leij, L.; Kelders, S. An Exploration of the Benefits of an Animallike Robot Companion with More Advanced Touch Interaction Capabilities for Dementia Care. *Front. ICT* **2017**, *4*, 16. [CrossRef]

36. Bradwell, H.L.; Noury, G.E.A.; Edwards, K.J.; Winnington, R.; Serge, T.; Ray, B.J. Design recommendations for socially assistive robots for health and social care based on a large scale analysis of stakeholder positions: Social robot design recommendations. *Health Policy Technol.* **2021**, *10*, 100544. [CrossRef]

37. Broadbent, E.; Stafford, R.; MacDonald, B. Acceptance of Healthcare Robots for the Older Population: Review and Future Directions. *Int. J. Soc. Robot.* **2009**, *1*, 319. [CrossRef]

38. Sparrow, R.; Sparrow, L. In the hands of machines? The future of aged care. *Minds Mach.* **2006**, *16*, 141–161. [CrossRef]

39. De Graaf, M.M.A.; Allouch, S.B.; van Dijk, J.A.G.M. Why Would I Use This in My Home? A Model of Domestic Social Robot Acceptance. *Hum. Comput. Interact.* **2017**, *34*, 115–173. [CrossRef]

40. Burmeister, O.K.; Ritchie, D.; Devitt, A.; Chia, E.; Dresser, G.; Roberts, R. The impact of telehealth technology on user perception of wellbeing and social functioning, and the implications for service providers. *Australas. J. Inf. Syst.* **2019**, *23*, 1–18. [CrossRef]

41. Teipel, S.; Babiloni, C.; Hoey, J.; Kirste, T.; Burmeister, O.K. Information and communication technology solutions for outdoor navigation in dementia. *Alzheimer's Dement. J. Alzheimer's Assoc.* **2016**, *12*, 695–707. [CrossRef] [PubMed]

42. Tronto, J. *Moral Boundaries: A Political Argument for an Ethic of Care*; Routledge: New York, NY, USA, 1993.

43. Rogers, W.; Mackenzie, C.; Dodds, S. Why bioethics needs a concept of vulnerability. *Int. J. Fem. Approach. Bioeth.* **2012**, *5*, 11–38. [CrossRef]

44. Kreps, D.; Komukai, T.; Gopal, T.; Ishii, K.K. Human-Centric Computing in a Data-Driven Society. In Proceedings of the 14th IFIP TC 9 International Conference on Human Choice and Computers, HCC14 2020, Tokyo, Japan, 9–11 September 2020.

45. Van Wynsberghe, A. *Healthcare Robots: Ethics, Design and Implementation*; Ashgate Publishing: Farnham, UK, 2015.

46. Malle, B.F. Integrating robot ethics and machine morality: The study and design of moral competence in robots. *Ethics Inf. Technol.* **2016**, *18*, 243–256. [CrossRef]

47. Braun, J.V.; Archer, M.S.; Reichberg, G.M.; Sorondo, M.S. *Robotics, AI, and Humanity: Science, Ethics, and Policy*; Springer: Cham, Switzerland, 2021; p. 193. [CrossRef]

48. Coeckelbergh, M. *Human Being @ Risk: Enhancement, Technology, and the Evaluation of Vulnerability Transformations*; Springer: Dordrecht, The Netherlands, 2013; p. 17.

49. Balkin, J.M. The path of robotics law. *Calif. Law Rev.* **2015**, *6*, 45–60.

50. Benso, S. *The Face of Things: A Different Side of Ethics*; SUNY Press: Albany, NY, USA, 2000.

51. Peeters, A.; Haselager, P. Designing Virtuous Sex Robots. *Int. J. Soc. Robot.* **2021**, *13*, 55–66. [CrossRef]

52. Yamaji, Y.; Miyake, T.; Yoshiike, Y.; De Silva, P.R.S.; Okada, M. STB: Child-Dependent Sociable Trash Box. *Int. J. Soc. Robot.* **2011**, *3*, 359–370. [CrossRef]

53. Sandewall, E. Ethics, Human Rights, the Intelligent Robot, and its Subsystem for Moral Beliefs. *Int. J. Soc. Robot.* **2021**, *13*, 557–567. [CrossRef]

54. Jordan, P. *Designing Pleasurable Products: An Introduction to the New Human Factors*; Taylor & Francis: London, UK, 2000.

55. Helander, M.G.; Khalid, H.M. Affective and pleasurable design. In *Handbook of Human Factors and Ergonomics*; Salvendy, G., Ed.; Wiley, Inc.: Hoboken, NJ, USA, 2006; pp. 543–572. [CrossRef]

56. Desmet, P.M.A. *Designing Emotions*; Delft University of Technology: Delft, The Netherlands, 2002.

57. Van Wynsberghe, A. A method for integrating ethics into the design of robots. *Ind. Robot. Int. J.* **2013**, *40*, 433–440. [CrossRef]

58. Van Wynsberghe, A. Service robots, care ethics, and design. *Ethics Inf. Technol.* **2016**, *18*, 311–321. [CrossRef]

59. Ayanoğlu, H.; Duarte, E. *Emotional Design in Human-Robot Interaction*; Springer Nature: Cham, Switzerland, 2019; p. 127.

60. Pirtle, Z.; Tomblin, D.; Madhavan, G. *Engineering and Philosophy: Reimagining Technology and Social Progress*; Springer Nature: Cham, Switzerland, 2021; p. 171.

61. Pirni, A.; Balistreri, M.; Capasso, M.; Umbrello, S.; Merenda, F. Robot Care Ethics Between Autonomy and Vulnerability: Coupling Principles and Practices in Autonomous Systems for Care. *Front. Robot. AI* **2021**, *8*, 654298. [CrossRef] [PubMed]

62. Tronto, J. *Caring Democracy. Markets, Equality and Justice*; New York University Press: New York, NY, USA, 2013; p. 108.

63. Umbrello, S.; Capasso, M.; Balistreri, M.; Pirni, A.; Merenda, F. Value Sensitive Design to Achieve the UN SDGs with AI: A Case of Elderly Care Robots. *Minds Mach.* **2021**, *31*, 395–419. [CrossRef] [PubMed]