Designing Robots With Care

Creating an Ethical Framework for the Future Design and Implementation of Care Robots

Aimee van Wynsberghe

Promotion Committee:

Rector Magnificus, voorzitter
Prof. dr. P.A.E Brey, University of Twente, promoter

Prof. dr. i.r. P.P.C.C. Verbeek, University of Twente
Prof. dr. N. Sharky, University of Sheffield
Prof. dr. J. Tronto, University of Minnesota
Prof. dr. M. Verkerk, University of Groningen
Prof. dr. V. Evers, University of Twente
Mr. W. Wallach, Yale University
Dr. M. Coekelbergh, University of Twente

Printed by: CPI Wöhrmann Print Service, Zutphen, The Netherlands
Cover image: Inspired by strk3; http://www.zazzle.nl/de_robot_van_da_vinci_vitruvian_kaart-137468253664862439 .
Cover design by: Patrick Cook

DESIGNING ROBOTS WITH CARE

CREATING AN ETHICAL FRAMEWORK FOR THE FUTURE DESIGN AND IMPLEMENTATION OF CARE ROBOTS

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
Prof. dr. H. Brinksma,
on account of the decision of the graduation committee,
to be publicly defended
on Wednesday the 18th July 2012 at 14:45 hrs

by

Aimee van Wynsberghe
Born on 6th May, 1981
in London, Canada

This dissertation has been approved by promoter:

Prof. dr. P.A.E. Brey

ISBN: 978-90-365-3391-1

Contents

Acknowledgements

I thought writing the thesis was hard until I had to write the acknowledgments! There are so many people to thank. First and foremost I would like to thank my supervisor Philip Brey. Thank you for encouraging me to continue with this idea from the very beginning, for giving me the structure I needed to function and the freedom I needed to flourish. I would also like to thank my committee members who contributed to this thesis through their own work. Joan Tronto, thank you for the inspiration. Your work has been fundamental for my own. Wendell Wallach and Noel Sharkey have provided much inspiration through their own writings as well as the fruitful discussions and presentations I have been present for. Thank you both for your encouragement and guidance in the last year. To Peter Asaro, much of your work has also helped to shape the foundational ideas of my writing. As a scared PhD, thank you for making it so easy for me to reach out to a rockstar in robotics (what I refer to you as) so I could then continue to reach out to others.

Thank you to all my friends and family in Canada, the Netherlands and all over the place. Mom, I can't thank you enough for your support for every min you were needed. Most of all, thank you for encouraging me to follow my passions and dreams even when they keep me on another continent. Erinn, thank you for your wisdom and guidance during my big decision making moments. Dr. B, thank you for your continued mentoring and support beginning back in high school. I'm so happy you are here to participate in another graduation! Patrick um, you're an idiot. And I love you so much. Thank you for coming to visit and for being present for every moment of academic and personal significance (and let's face it there were many over the last four years!). Danielle, thank you for being a patient and committed friend for the last 21 years. Here's to another 21! Thank you to Sarah and Nema for being my muses and letting me drill you on questions about your work. Thank you to all the care workers in the nursing home and hospitals that I visited for talking to me and letting me

observe. Thank you to all the roboticists who sat and talked with me in the labs and at conferences.

And to the Enschede crew! Thank you to all the PhDs, post docs and Profs at Twente for a memorable four years. Thank you to Liska for saving me with last minute adjustments and for reading so much of my material. Thank you to Asle, Johnny, Pak and Scott for your feedback. To Johnny and Pak: so many conferences, so many talks, so much fun! Thank you for introducing me to PhD life and for helping me navigate my way at conferences. To Giovane, my brother and my amazing friend, I look forward to the cracking of more shells and peaches!! To Josine, the dinners and talks, beers and patios I cannot wait for our annual trips! To Scott, thank you for introducing me to Latex, for helping me with references, and for proofing. Above all that, thank you for being there to support me and to make me laugh when it was all so stressful. To Lucie, I'm so happy you came to Enschede (and brought speculaas mousse into my world!) on that ridiculous night and agreed to spend the next four years with us crazy people! And to Fede, my office-mate, my house-mate, my sister and my best friend. You have taught me so much on this journey; you have been my voice of reason, my support system and my greatest fan. So much love. Last but certainly not least, thank you to the Faculty club for the memories!

Introduction

The possibility of being cared for exclusively by robots is no longer science fiction. [*Sharkey and Sharkey*, 2011, p. 267]

A Revolution in Healthcare

WELCOME to a revolution in healthcare. As we come into the 21st century, the ageing population is already a major demographic worldwide and will continue to increase dramatically. The care-givers available to care for this large segment of the population are woefully outnumbered by this 'boomer' generation. According to the World Health Organization (WHO), while life expectancy is increasing, fertility rates are declining around the world (WHO 2010). The continued anticipated increase in this population group is reason for concern, as is the challenge of providing care for the population is general. This will be hampered by a lack of resources, a competition for healthcare services, shortages of personnel and care providers, and a changing pattern of need (redirected resources). It will be a test for healthcare systems around the world. How are such setbacks to be mitigated? Increasingly, policy makers and healthcare providers are turning their attention to robots as a solution among others. Interaction with robotic pets, such as Sony's AIBO or the robot seal Paro, are shown to have positive physiological benefits on elderly people. Service robots, such as Aethon's TUG robot or the HelpMate, are currently used in hospitals across the United States for the delivery of sheets and medications. With the widespread introduction of robots used in healthcare, the 'robot revolution' has spawned what can only be referred to as *a revolution in healthcare*. This thesis addresses the initiative to create and use care robots and the many questions surrounding their design and use. Specifically, my aim is to translate ethics into a tangible tool to be used by designers in the design of future robots used in

healthcare.

Currently, in healthcare applications, robots are now available to help in surgical tasks that a surgeon couldn't otherwise complete with the same precision. Although the influence of popular culture conjures images of human-like robots, such as Star Wars' C3PO, performing a surgery on a human, this not the case. Such robots are big and bulky, machine-like in appearance and require the direct input of a human user in order to execute an action. Hospitals and healthcare facilities are using robots in rehabilitation treatments, the sorting of medications, delivery of food, and as a communication platform between patients and physicians when geographical boundaries separate the two. These robots are already commercially available and used in hospitals in the US, Canada, Europe, and Japan. The latest developments in healthcare robotics are those intended to assist the nurse in his/her daily tasks. These robots, now referred to as care robots, may be used by the care-giver and/or the care-receiver for a variety of tasks from bathing, lifting, feeding, to playing games. They may have any range of robot capabilities and features and may be used in a variety of contexts, i.e., nursing home, hospital or home setting. They are integrated in the therapeutic relationship between care-giver and care-receiver and aim at meeting a range of care needs of users. Consequently, they are expected to mitigate the foreseen lack of healthcare personnel and resources or in specific instances to allow persons to stay in their home without having to live in a care institute (as in the care of elderly or rehabilitative persons) [*Tamura et al.*, 2004, p. 85].

I do not claim that these robots should be made or used for all care activities, nor that they should be used for any and every care practice. This standpoint is grounded in the potential benefits a care robot can provide as well as the potential ethical problems that may arise with the use of a care robot. In terms of the first point – that care robots can provide a benefit – a care robot presents the option of providing impartial care 24/7. It cannot be denied that care is required 24/7 and it is not possible for one care provider to provide this kind of assistance, in a hospital or home setting. Thus, care robots also hold the promise of allowing (elderly or rehabilitative) persons to remain in their home longer. This is of course a benefit for persons who wish to remain in familiar surroundings but it should also be recognized that in practice many patients in home care settings may not receive a high quality of care. In a homecare setting there is the risk of maltreatment of patients; "a relationship with a care recipient can evoke a multitude of attitudes and behaviours. At times, deplorable traits can emerge. In fact, individuals suffering from debilitating illnesses such as dementia are sometimes mistreated by family members" [*Cooper et al.*, 2009 from *Borenstein and Pearson*, 2011, p. 257]. The fear of such treatment is

not exclusive to a home setting: "at the present moment when the costliness of labour-intensive care is foremost in the minds of citizens" [*Razavi*, 2007], we frequently hear about abusive or inadequate forms of care [*Tronto*, 2010, p. 163]. In the nursing home patients are often reported to be abused physically, emotionally or psychologically [*Pillemer and Moore*, 1990; *Payne and Cikovic*, 1995; *Podnieks*, 1990][1]. Moreover, in practice many nurses in the hospital feel an affinity for some patients over others, especially when patients themselves are abusive. In summary, each patient is treated differently for a variety of reasons. Consequently, a robot in place of a nurse for certain tasks or at certain times in the night/day presents the potential to overcome concerns of impartiality and abuse as well as providing care at all times of the day. Most importantly, a care robot presents a benefit in terms of relieving certain burdens of care workers but may also be used as a way of regulating the behaviour of human care-workers to avoid any risk of patient abuse or maltreatment. While the care robot movement is pressing forward at an incredible rate in Japan where the gap between care workers and those in need of care is greatest, "Europe and the US are facing similar ageing population problems over a slightly longer time scale" [*Sharkey and Sharkey*, 2011, p. 267] and are expected to follow suit in the robotics trend.

Currently, in elderly care facilities in Japan robot teddy bears monitor and assess the functioning of patients and report back to staff [*Sharkey and Sharkey*, 2011]. The PaPeRo robot is used in a similar way for childcare or monitoring [*Sharkey and Sharkey*, 2011]. Such trends are expected to continue in order to facilitate remote monitoring of patients and are even thought to be used in the event that patients are quarantined. In the more futuristic visions roboticists have, robots are used in a variety of care applications for a variety of tasks. The hopes for future robots include providing companionship, completing multiple tasks required for daily living (assistance with dressing, cooking and feeding), surveillance of one's home, assistance with grocery shopping, assistance with household cleaning and beyond.

In some instances, such as search and rescue robots or robots in outer space, the benefits of using robots are immediately evident. Conversely, in care applications, the presumed benefits may come at the expense of cultural traditions and values. In popular culture discourse, the issue of using robots is fuelled through movies, literature and science fiction writing. Isaac Asimov is the most well-known of science fiction writers addressing the ethical issues to pertaining to robots. In his series of short stories he tackled the rules by which robots

[1] For an exhaustive overview and study of elder care abuse see the Journal of Elder Abuse and Neglect: www.tandfonline.com/loi/wean20

ought to be programmed according to (the ethical principles, if you will) and at the same time showed the impossibility of robots functioning according to such programming. In movies, Western societies are presented with a multitude of dystopian futures in which humans become lazy and completely dependent on robots for their well-being [*Morris et al.*, 2008], or humans suffer at the hands of robots when the robots override the decisions of humans [*Kubrick et al.*, 2001]. Although entertaining, literature and movies fuel the views and beliefs of popular discourse and have left society in fear of these anticipated future visions when living with robots.

The question of robots is also addressed in academic domains. Ethicists are now grappling with the evaluation of the use of robots in these applications. Some studies deal with the questions pertaining to the safety issues, or the issues of human-robot interactions exclusively [*Breazeal*, 2004] while others look at the broader societal questions pertaining to the initiatives to use such technologies. Some of those questions being considered, for example, are: whether robots will cause human societies to decline [*Mowshowitz*, 2008]; whether people will lose a sense of judgement with potentially fatal consequences [*Cooley*, 2007]; whether the use of robots will become a dependency inviting "empty brains" [*Maurer*, 2007]; why we are creating these intelligent systems and for whom [*Capurro*, 2009]; whether the use of robots will result in responsibility gaps [*Gill*, 2008], what Tamburrini refers to as the "responsibility ascription problem" [*Tamburrini*, 2009]; or, the replacement issues pertaining to robots [*Decker*, 2008]. The last point deserves a great deal of attention given the displacement of industrial workers and the systematic de-valuing of their tasks and roles following the implementation of industrial robots in the 1960s [*Moravec*, 1999]. This concern is quite problematic when we consider robots entering into care contexts and the role of women in these contexts. Historically, the skills attributed to women seen as necessary for care – empathy, compassion, ability to connect interpersonally – have been undervalued [*Tronto*, 1993]. Accordingly, one must ask whether care robots reflect and/or propagate such a de-valuation.

Care robots, in particular, pose certain ethical concerns specific to the tradition of care. Scholars have written about the potential for social isolation when a robot is used in place of a human for social and emotional caring tasks [*Sparrow and Sparrow*, 2006]; that a care robot in elderly care has the potential to threaten the rights of elderly persons [*Sharkey and Sharkey*, 2012, 2011]; that a care robot takes away the opportunity for (self) growth of the care-giver [*Vallor*, 2011]; that a robot has the potential to threaten the privacy, security and confidentiality of a patient when the robot is used to communicate information from one setting to another; that the robot has the potential to threaten privacy

when outside parties can contact a person in their home without permission; that the robot may pose safety concerns in terms of the physical well-being of patients; that the robot has the potential to threaten the quality of care of patients [*Coeckelbergh*, 2010]; that a robot used exclusively in the care of elderly persons, children or other marginalized demographics presents a risk of ageist discrimination; and that the use of robots in care may present a risk in terms of distributive justice or health equity. For the last point, the question concerns whether or not developing countries will have access to the technology or, whether the use of robots will be directed towards those who lack a certain social status (robots used for the care of prisoners, elderly persons, children, handicapped persons, etc.).

A large portion of the difficulty in ethically assessing the design, development and use of care robots has to do with knowing what questions to ask. In other words, should the ethical evaluation of care robots focus on how their introduction will impact the organization and provision of care? Or, should the ethical evaluation of care robots address the initiative to use such robots and the assumptions leading to such an initiative? Or, perhaps the most appropriate course of action would be to ethically steer the design and development? Such steering may be in terms of what behaviours the robot elicits from the users, referred to as nudging [*Thaler and Sunstein*, 2008]. In the same vein, perhaps the ethics of care robots ought to centre on the domestication or implementation of the robot. Each of these questions starts at a different point in the design process of a care robot thus appealing to a different set of stakeholders (designers vs. users) or a different context (the lab vs. the hospital/home setting). It follows that the evaluation of care robots ought to encompass all of the aforementioned questions.

Ethics and Care

If we take the starting point to be that the initiative to create and use care robots rests on the belief that care robots will maintain a high standard of care, or perhaps even improve care, then the main question has to do with how care is understood; what is care, what is good care, and how is this achieved and/or evaluated? At the root of all the ethical issues addressed to date appears to be an ambiguity of what care is, how it is structured, what it involves and what it means. This only adds to the problem of articulating when care is good, for whom and what elements make it good. As psychoanalyst Sherry Turkle eloquently points out in her book "Alone Together" [2011], with the current

generation of robots, we as a society are afforded the opportunity to reflect on the values of societal importance and to safeguard their place or alternatively allow for a trade-off between values. This opportunity is what Turkle refers to as "the robotic moment" and is the situation we are currently in. But this is more than an opportunity claims Turkle, it is a necessity. Care robots offer us the opportunity to reflect on care – what it is, how it is achieved – and to tailor the design of the robot accordingly. For authors like Shannon Vallor, this reflection involves paying special attention to the goods at stake for the care-giver when a robot is used [2011]. For Sparrow and Sparrow this involves recognition of the significance of the component of human presence in care [2006]. For Sharkey and Sharkey, this involves recognition of the rights of vulnerable demographics and how a care robot may impact such rights [*Sharkey and Sharkey*, 2012].

I would like to go even further than this and examine the very root(s) of care. Such a feat demands an understanding of care conceptually as well as understanding care in context, in terms of the actions and interactions between care providers and care-receivers. In the care ethics tradition, the many actions in care that make up the overall process of care are referred to as care practices. Consequently, care as a concept is distinguished from contextualized care. The former I refer to as 'care': the conceptual dimension of care that centres on a valuation of another, concepts like dignity and a relationship with the good life. The latter, contextualized care, I refer to as care practices, which provide meaning to abstract values such as human dignity. Understanding the many practices that comprise 'care' will allow me to uncover the fundamental values that make up care. To do so we need a framework for understanding a care practice; how care values are made real, how roles and responsibilities are distributed, and how meanings are established. Only by understanding these variables can we come to understand: the role the care robot will play once introduced; the responsibility and meaning the robot will have; whether or not the robot preserves the expression of values or alters them, and if so in what way. When we understand what is happening in care at the contextualized level (what I will also refer to as the micro level), we may begin to understand the significance of the robot at the same level.

Thus, I formulate my research question as follows: *how can care robots used in care practices be designed and implemented in a way that supports and promotes the fundamental values in care?* This central question takes into consideration all of the aforementioned questions; how will the care robot impact the expression of care values, how will the care robot impact the distribution of roles and responsibilities, and what meaning will the care robot take on? To facilitate this kind of ethical evaluation of care robots I will create a framework for

understanding the web within which the care robot will enter and the potential impact the robot might have.

Creating a Framework for Evaluating Care Robots

How will such a framework be created, how will it be used and what is its purpose? Chapter 1, 'Creating a Framework for the Ethical Evaluation of Care Robots', will explain in detail how I have chosen to address these questions. The framework is both conceptual, in that it allows for an understanding of how values are manifest in care practices among actors (human and non-human), and normative in that it allows for the analysis and evaluation of the impact a robot may have on the promotion and expression of care values in context. In my work, I draw upon a number of theoretical approaches and methodologies, and this chapter aims to explore many of these approaches and concepts. I draw on elements of Actor-network theory [*Latour*, 1992; *Callon*, 1986], script theory [*Akrich*, 1992], the concept of embedded values [*Nissenbaum*, 1998], Value-Sensitive Design (VSD) [*Friedman et al.*, 2003, 2006], and the care ethics tradition [*Tronto*, 1993, 2010; *Little*, 1998]. Chapter 1 addresses in great detail how the framework will be created as well as its strength and utility while chapter 2, 'Values and Assumptions Embedded in Technology', embarks on a conceptual investigation of important concepts like values, assumptions and norms, and how they come to be embedded in a technology.

In order to address the relationship between a care robot and contextualized care, we must first understand what care values are and how they come into being. Chapter 3, 'Understanding Care in Context', goes into a conceptual analysis of the dominant values of the care ethics tradition. Special attention is paid here to the description of a care practice and the significance of understanding care tasks as practices rather than as tasks. I explore the fundamental values in care from a top down approach beginning with the abstract values articulated by the World Health Organization and how they become concrete when understood in context. This chapter reveals three important findings: One, values are manifest (or co-produced) through the actions and interactions among actors (human and non-human) in a network for a particular practice in a specific context; two, a care practice is a small piece in the holistic vision of care as a process [*Tronto*, 1993]; and three, the therapeutic relationship is the vehicle for the manifestation of care values.

Current philosophers of technology argue in favour of addressing the technical details of a technology in order to adequately address the associated ethical

issues [*Verbeek*, 2011; *Nordmann and Rip*, 2009; *Brey*, 2012]. So what is the technology that I am talking about? Chapter 4, 'Care Robots and Robot Capabilities', deals with the definition of a robot, the variety of robot capabilities and features as well as presents existing care robot prototypes currently in use or still in the developmental stages. This chapter reveals the impossibility of translating human capabilities into robot capabilities independent of contextual variables (the care practice and the actors involved). From this I conclude that without an understanding of the context within which the care robot will be applied or the practice for which it will be used, one is not capable of truly understanding the impact the robot may have. Consequently, I begin to set the stage for the various components of the framework, namely, that context and practice must be made explicit if one is to understand the impact the care robot will have.

Here we are faced with the question: how will all of this be used? Chapter 5, 'A Framework for Evaluating the Design of Care Robots', outlines and describes the components of the framework and the justification for their place within the framework. I refer to the framework as the *Care-Centered* (CC) *framework* given the focal role the care perspective plays in its creation and usage. As such, chapter 5 also explores care as a concept in relation to the care ethics tradition and how these insights are integrated into the framework. The CC framework is then used for two types of value-based analyses: (1) for retrospective evaluations of current care robots, and (2) in the prospective design of future care robots. I refer to the first methodology as *Evaluating Care Robots* (ECR) and the second as the *Care-Centered Value-Sensitive Design Approach* (CCVSD).

The ECR approach holds the potential to be used in the evaluation of *any* care robot. For my analysis, I have chosen to address specific care practices for which care robots are currently in the design and development stages, and in some cases commercially available. Chapter 6, 'Care Robots and the Practice of Lifting', investigates both the practice of lifting, and the current robots delegated for such a practice. Two care robot designs used for the lifting of patients are compared with each other to illustrate how differing robot capabilities arise from varying assumptions about the ideal care-giver and care practice, and consequently result in divergent visions of the resulting care practice. Each robot is examined using the current practice of lifting to understand the way in which a care robot might be used to re-integrate values lost in the first wave of automation (i.e., the mechanical lift for lifting) as well as how the robot may impede the promotion of necessary values. The aim of this chapter is to make clear the relationship between the technical capabilities of the robot and its impact on the resulting care practice.

Chapter 7, 'Care Robots and the Practice of Feeding', explores another dimension of analysing practices - the difficulty in understanding the holistic nature of practices, their interconnectedness and their relationship to the overall care process. The relevance of this for care robots stems from the robot's potential to not only impact one moment in the practice but to unintentionally impact a moment in another practice. Through an analysis of the practice of feeding, I explore three moments that fall under the umbrella of 'the practice of feeding': the dietician's assessment and creation of a nutrition plan for the patient, feeding the patient, and the removal of trays from the patient's room. Each of these moments is described in terms of the manifestation of values throughout the practice in both mechanical terms (describing the elements as they relate to a specific practice) and their relationship with the overall process of care. At each of these moments in the practice of feeding, there are care robots under development and commercially available which enter the equation. The goal of this chapter is to assess these robots according to their potential impact on the manifestation of moral elements within the particular practice for which they are developed and to observe and evaluate their impact on the manifestation of values with respect to the therapeutic relationship and the overall care process.

Chapter 8, 'Designing Moral Factors With Care', investigates the moral impact of a care robot in terms of moral agency. Although the moral agency has been addressed implicitly throughout the preceding chapters, my aim in this chapter is to explicitly discuss the moral status of robots and the consequences such a discussion has on the design of future care robots. To take an example of a type of robot that brings this question to the fore, I turn to social robots and care robots with social capabilities. To be clear, I do not categorize social robots as care robots. This is directly related to a difference in the ends that each robot serves. Social robots have as their aim the formation of a relationship between human user and robot – the end being companionship. Alternatively, a care robot aims at meeting the care needs of individuals, and the therapeutic relationship is a means to that end. The relationship in care is not one of companionship but rather, of a therapeutic nature. Care robots may have social capabilities but the goal of the care robot with social capabilities is not to establish a relationship, but rather to fulfill a role within a care practice, to be integrated within the conclave of the therapeutic relationship. When the robot is endowed with sophisticated intelligence to such an extent that it may interact in a human-like manner, as is the case with social robots, the question of whether or not the robot is a moral agent becomes quite important. The answer to this question determines the kinds of roles and responsibilities delegated to

the robot. This chapter outlines the critical questions pertaining to a robot's moral status and how such insights should be addressed through the CCVSD approach.

Up to this point I have illustrated the utility of the framework in understanding the role of the robot once it has been integrated into a context, practice and network of actors. This is a valuable tool for understanding the inscribed script of a current care robot. There is still an element that is missing – the element of prospective analysis. By prospective analysis, I aim to show how ethics can accompany the development of care robots . Chapter 9, 'Designing and Implementing Robots With Care', is meant as the apex of this work. It is the moment in which I show the benefit of the CCVSD approach in the overall design process and implementation of future care robots. The CCVSD approach mirrors that of traditional VSD: it is a methodology for the design of a future system in which values of ethical importance are systematically explored throughout the design process to be included in the technical content of the system. It differs from traditional VSD in that I have selected the values of ethical importance from the care ethics tradition, and care contexts, and have translated this into a tool for designers. With this, there is a kind of built-in technology assessment component: the CCVSD approach is about the design of the system but is also about the development and implementation of the system. When using the CCVSD approach for prospective analysis, the point at which the evaluation or analysis begins differs from retrospective evaluations. Analysis begins at the point of idea generation, when the use of the robot and the capabilities of the robot are first discussed. This means that within the prospective methodology, the fears related to the use of care robots must be addressed. The de-valuation of the role of the nurse if replaced with a robot, the de-valuation of care roles when fulfilled by a robot, and the robot's potential to undermine the cultivation of care skills of the care-giver, are three significant fears expressed in the current academic and popular discourse. I do not wish to undermine these potential risks, but rather wish to show how the framework acts to mitigate these risks, and further how the framework acts to systematically take these risks into consideration through the development of the care robot. Articulating and understanding these fears helps us to uncover and identify the values at stake, and the values that must be protected.

For this, I propose the creation of two novel care robots; a robot for the testing of urine in paediatric oncology, the "wee-bot", and a robot used for waste removal (waste referring not to garbage from a persons room but to excretions of the patient), the "roaming toilet". Neither of these two robots is being developed at this time which provides an opportunity to steer the development of

a care robot, according to the framework, beginning with the moment of idea generation. The ideas for the robots came from observations in the hospital and interviews with healthcare workers. The prospective methodology, however, does not end with the resulting artefact. When we take into consideration the idea that the robot is being designed according to a specific use, one that acts to promote care values and one which determines a particular distribution of roles and responsibilities, we must also consider how the care robot will be introduced, or rather, how the care robot *ought* to be introduced. To this end I will examine domestication studies along with design studies for insights into what "ethical implementation" of the care robot should consist of. The methodology for implementing the care robot is then presented as a way of showing the holistic nature of the CCVSD approach.

Chapter 9 ends with a conclusion section summarizing the main findings and benefits of this work. With the CCVSD approach my goal is to foster an interdisciplinary approach, a division in moral labour, in the design, development, and implementation of care robots. Given the initiative to bridge disciplines, this book is intended to be read by individuals/scholars from a variety of fields. As such, each field of study that I draw upon is presented in the most straightforward manner possible. The creation of the CCVSD approach is meant to mark the 'robotic moment', coined by Sherry Turkle. This 'robotic moment' that Turkle speaks of demands that care robots undergo meticulous ethical evaluation. This 'robotic moment' also demands that our traditional conceptions of relationships, of the meaning of care and of what it means to be human are questioned and subject to re-interpretation. My response to the claim of Turkle is to structure both this revolutionary technology in healthcare applications as well as structuring healthcare institutions in a way that supports the introduction of the robot and supports the roles, responsibilities and valuation of healthcare workers. Thus, not only can one consider the technology of care robots as a revolution in healthcare, but designing and implementing them according to the CCVSD approach is also a revolution; one that re-affirms and supports the values of the healthcare tradition along with the roles of healthcare providers.

Chapter 1

Creating a Framework for the Ethical Evaluation of Care Robots

Designers cannot but help to shape moral decisions and practices. Designing is "materializing morality." [*Verbeek*, 2011, p. 90]

1.1 Introduction

The morally charged contexts into which care robots will be included, and their future role in the moral decision making of humans, demand that they undergo rigorous ethical reflection. Evaluating care robots is complicated for a multitude of reasons; the difficulty in knowing *how* to evaluate (which ethical theory to apply or indeed if there is one theory that is sufficient), the difficulty in knowing *what* to evaluate (the initiative to use care robots, their design, their introduction) or, overall the difficulty in untangling the ethically good from the ethically bad uses. The introduction of this work gave an overview of how care robots are seen to be beneficial in care as well as how they are wrought with ethical concerns. Accordingly, the question to ask is not whether or not we should make them but how they should be made and what they ought to be used for. Based on this, I do not deny the development of this technology; rather, I am seeking a way in which the technology can be made in support of

widely held cultural values. Accordingly, this chapter explains in detail how I have chosen to address the research question presented in the introduction; *how can care robots used in care practices be designed and implemented in a way that supports and promotes the fundamental values in care?* I will do this through the creation of a normative framework to be included in the design process of a care robot. But how is such a framework created and what will it target?

In the following chapter I outline the concepts used to create the proposed framework combining approaches from the computer ethics domain (the embedded values approach and Value-Sensitive Design), STS studies (actor-network theory, script theory and domestication studies) and the philosophy of technology (structural ethics, technology mediation) and the care ethics tradition. All theories are related in that they address the relationship between artefacts and humans in a network and the co-creation/production of values and norms. The approaches from the computer ethics domain emphasize the relationship between the technical content of an artefact, its use and the resulting expression of values. The approaches from the STS domain emphasize the actions and interactions of actors being both human and non-human, and the resulting production of meaning, norms and values. The approaches from the philosophy of technology domain emphasize the moral impact of an artefact not only on the immediate network into which it exerts an influence but also on the associated micro networks and the overall macro network (the institution). And, the care ethics tradition provides the lens through which all of the above traditions are analysed and given a place in the evaluation of a care robot.

This chapter begins by discussing the issues related to robot ethics; what the predominant questions to address are according to robot researcher Peter Asaro. While Asaro presents a compelling case for the need of such a robot ethic, he stops short of presenting the methodology to accomplish this. My aim is to incorporate his insights into an approach for the evaluation of care robots but also as a way of steering the design of future care robots. I use the approach known as VSD as a blue-print for creating my own framework specific to the design and development of care robots and in so doing conclude with an approach that I refer to as *Care-Centered Value-Sensitive Design Approach.*

1.2 Distribution of Responsibilities

Aside from the questions pertaining directly to care and its understanding, how do we make sense of the care robot before and after it is introduced into the care context? In other words, what are the ethics related to the robot? I do not

mean to look at the ethical issues pertaining to a specific concern like privacy with respect to robots, but rather the ethics related to robots in general. For robot ethicist Peter Asaro, a framework for addressing the ethical considerations pertaining to robots – robot ethics - ought to first and foremost recognize a robot as a socio-technical system [*Asaro*, 2009]. Recognizing a robot as a socio-technical system, a common theme in Science and Technology Studies, presupposes an understanding of the complex, dynamic and reciprocal interaction between society and the development of technologies. With this in mind, Asaro then identifies the three dimensions for structuring a robot ethic. These three dimensions structure the variety of questions a robot ethicist should ask as well as the questions which a robot ethic should be able to answer. Accordingly, the three dimensions are: "1. the ethical systems built into robots, 2. the ethics of people who design and use robots and 3. the ethics of how people treat robots" [*Asaro*, 2009, p. 1]. We can conclude from this that Asaro agrees with the view of authors like Swierstra and Rip who claim that paying attention to the technical content of a technology (in this care a robot) is indispensable to the ethical reflection of such systems. The overarching question that each of the three dimensions stem from has to do with the re-distribution of moral responsibility in the social-technical network once the robot has been added [*Asaro*, 2009, p. 1]. Consequently, the distribution of responsibilities is, and ought to be, positioned at the heart of ethical reflections on robots; however, the ethical agent, or subject, in question differs depending on the dimension one is working within. In the first dimension the ethical agent is the robot whereas in the second dimension the ethical agent is the designer and in the third dimension the ethical agent is the user and/or society at large. Thus, all actors involved in the process of designing, developing, implementing and using the robotic system have a role in determining the ethical outcome of the robot.

Any shift in the distribution of responsibilities, when a new technology has been integrated into healthcare settings, is important for a variety of reasons. Take the introduction of surgical robots used for long distance surgery, what is referred to as telesurgery. In these instances, the surgeon and the patient are geographically separated. They may be in different cities, countries or continents. The surgeon performs the surgery from a console on their side (known as the surgeon's side), the tactile information from the console is sent via a telecommunications network or satellite to the patient's side where the robot interprets the tactile movements of the surgeon into robotic movements inside the patient. Thus, the surgeon is performing the surgery through the robotic apparatus. The question then is, if something were to go wrong, who is responsible? Is it the fault of the surgeon's performance, the robotic system (thus making the distrib-

utor or manufacturer liable) or the telecommunications network? Without the robotic system, the surgeon would be responsible; however, in the case of telesurgery, the robotic system coupled with the telecommunications network blurs the lines of responsibility [*van Wynsberghe and Gastmans*, 2008]. Moreover, without international guidelines and standards, it isn't clear who is financially responsible for the procedure; the patient in their home country/hospital or the surgeon in their country/hospital? Consequently, understanding and articulating responsibilities in healthcare scenarios is important for the safety of the patient as well as the healthcare workers.

Thus, articulating the distribution of responsibilities helps to ensure good care of the patient. In healthcare, the distribution of responsibilities in the care of a patient is of crucial consequence given that a range of healthcare professionals are required to meet the multifaceted needs of one patient. The doctor or surgeon is most often responsible for the physical intervention portion of care while the nurse is often times responsible for the activities of daily living (ADLs)[1] of the patient and of course a range of professionals are required for cleaning the facilities, preparing and serving meals along with a host of administrative tasks. In this sense then, clarifying responsibility helps to ensure that all the needs of the patient are met and further that healthcare workers understand which needs are their responsibility. This is not always the case, however, as there are certain wards in which the nurse is responsible for reading and distributing the state of the art in research protocols and treatment options to the patient in addition to the daily needs of the patient like the creation of care plans, bathing, feeding and administering of medications (as is the care in paediatric oncology, personal communication). Regardless, the nurse, in conjunction with a range of additional healthcare professionals, has a variety of responsibilities that, when met, come together to fulfil the range of needs the patient has. As such, the distribution of responsibilities upon the addition of a care robot is also at the heart of this ethical reflection. The question then is where to find a framework that can address each of the dimensions proposed by Asaro, with particular attention to the distribution of responsibilities relevant to a healthcare context, using the technical content of the robot as the foundation for analysis? Consequently, it appears as though no such framework exists and it is this challenge that I will take up.

[1] Activities of daily living refer to the daily self-care activities of individuals. For example, bathing, dressing, feeding, movements from one location to another, bowl and bladder management.

1.3 Value-Sensitive Design

Computer ethics, although dedicated to the reflection of computer systems and software, provides a good starting point when addressing the technical content of robots. In particular, the embedded values approach (EVA) proposed by Helen Nissenbaum [1998]. This concept refutes the neutrality thesis of computer systems and software programs and claims that instead, it is possible to identify tendencies within a computer system or software to promote or demote particular moral values and norms [*Brey*, 2010, p. 1]. These tendencies manifest themselves through the consequences of using the object. When said technology is capable of imposing a behaviour on a user, or consequence to using it, the imposing force within the technology is considered a "built-in" or "embedded" value (or alternatively a disvalue if the computer system hinders the promotion of a value).

Given this view of computer systems and software, the consequences of using them demand ethical attention and thus the computer system or software requires ethical reflection during its design process. This is not to say that the computer is morally responsible in any way but rather, awareness of the force within the computer to impose actions and roles on the future users and non-users, requires ethical reflection. This approach addresses the first dimension of ethical importance proposed by Asaro – the ethical systems built into robots. The belief that a system or technical artefact carries this force without rendering it morally responsible is also something supported by other roboticists [*Asaro*, 2009; *Tamburrini*, 2009; *Floridi and Sanders*, 2004]. Technologically speaking, current robots are not sophisticated enough to render them responsible for their own actions as they cannot recognize the implications of their actions. For others, a distinction between responsibility and accountability renders the robot accountable but not responsible. Questions pertaining to whether or not we can create ethical decision making robots requires further attention and will be taken up in a chapter 8.

This approach also addresses the second dimension proposed by Asaro – the ethics of the designers. EVA may be considered a concept within contemporary computer ethics studies and it is this concept which forms the groundwork for various computer ethics methodologies regarding the design of future systems. Engineers have used this concept coupled with various methodologies for uncovering the embedded values, as the foundation for designing technologies in a way that supports the promotion of certain values [*Brey*, 2010; *Introna*, 2005]. Value-Sensitive Design (VSD) is a well-known approach of this kind that aims at the creation of technical artefacts in a way that encourages the realization

of values. In short, the methodology of VSD requires the concept of EVA as part of its own methodology; however, the VSD approach addresses the issue pertaining to design. For Brey, the link between the concept of EVA and the methodology of VSD can be summarized as follows;

> *If designers are aware of the way in which values are embedded into artefacts, and if they can sufficiently anticipate future uses of an artefact and its future context(s) of use, then they are in a position to intentionally design artefacts to support particular values* [*Brey*, 2010, p. 9].

As such, VSD presents the potential for the creation of future care robots that promote the realization of care values thereby preserving the tradition of care. Value-Sensitive Design as a design process is then a means for steering the design and development of care robots in an ethical manner.

1.4 Why Design?

Discussing robots in terms of their "design" and the "design process" from which they result, demands an understanding of what I mean by both *design* and *design process*. For starters by *design* I neither refer exclusively to the external appearance of the robot nor exclusively to the software programming of the robot; rather, to a combination of the appearance and capabilities of the robot. Of course the capabilities of the robot result from the programmed computer code and thus programming is subsumed within the element of capabilities. Appearance refers to the robot being humanoid, machine-like and/or creature-like as well as the morphology of the robot – the form and structure of the robot. In contrast, Feng and Feenburg describe 'design' as a "process of consciously shaping an artefact to adapt to its goals and environments" [*Feng and Feenberg*, 2008, p. 105]. This process of shaping the artefact is what I refer to here as the *design process*. My insistence to focus on design and the design process rests predominantly on the relationship between artefacts and morality conceptualized in the philosophy of technology and STS domains.

1.4.1 Design and Morality

For some, artefacts are believed to have a kind of morality. Oosterlaken conceptualizes this morality in terms of a technology's ability to 'expand capabilities' [*Oosterlaken*, 2009]. This morality, or moral impact if you will, is a result both

of the designers' intentional decisions as well as the technologies place within a network. I reference the term 'network' intentionally to relate to Latour's approach known as Actor-Network Theory (ANT). For Latour, a network describes an amalgamation of human and non-human actors which interact together for moral decision-making, for establishing norms and meanings and for determining outcomes. Actors are both human and non-human, thus a robot may also be considered an actor. For Verbeek, artefacts have moral relevance given their role in mediating one's experiences and practices. Technological mediation refers to the phenomenon whereby a technology helps to "shape human actions and perceptions and create new practices and ways of living" [*Verbeek*, 2011, p. 92]. Intentionality and freedom – two necessary components for granting moral agency/responsibility – are hybrid affairs between technologies and humans – technologies are intimately involved in the directing of human actions as well as the decision making of humans. It is this aspect that bequeaths a type of moral relevance to the technology. It follows then that "designers materialize morality" [*Verbeek*, 2006], and thus "technology design is inherently a moral activity" [*Verbeek*, 2008]. Consequently, "an engagement in the development of the material environments that help to form moral action and decision making" is called for [*Verbeek*, 2008].

Recent work in Science and Technology Studies shows how technologies can be used to steer the behaviour of users. For Philosopher Bruno Latour this is known as prescription. Thaler and Sunstein build on this idea and claim that technologies can be used to nudge users to behave (or refrain from behaviours) in a variety of ways. The type of behaviours these authors refer to has to do with producing behavioural effects without the user knowing it. This is not a new idea, many authors have argued in favour of such technologies; persuasive technologies, seductive technologies, coercive technologies or decisive technologies. Each of these technologies prompt the user to engage with the technology, and what it demands of the user or the environment, in a different way. Technologies can be used to stimulate reflection, to prompt moral decision-making or to provide feedback about a user's behaviour. Most importantly, the act of engagement is a result of the design of the technology. In a morally delicate situation as care, engagement is significant for meeting the needs of the patient (i.e., good care). Care robots will invariably be programmed with any number of steering capabilities for the care-giver and/or care-receiver which are decided during the design process. It is therefore crucial to address these types of capabilities and their moral implications *before* they become standard capabilities of a care robot. By using such steering capabilities, the care robot is a manifestation of the intention of the designers.

Other scholars in the field of STS study the phenomenon known as domestication. In short, the impact the technology has once it becomes an actor in a network of other human and non-human actors. Hence, domestication studies build on the concept of the network and the interactions between human and non-human actors (the material environment). This impact is observed/studied in terms of the meaning the technology takes on, how this meaning is established, how the technology propagates or alters existing norms, etiquette, prioritization and interpretation of values, etc. Given the technology's propensity to maintain or shift an established morality, the artefact itself is said to be an actor for its role. ANT, however, insists on a lack of subjectivity or a homogenizing of the responsibility attributed to actors in a network whether they be human or non-human (technologies, the material environment, etc.).

Structural ethics, on the other hand, maintains the concept of the network and the emphasis on the interactions between actors in a network but adds the interactions among different networks on both the micro level as well as the macro level (the macro level referring to the overall institution or structure within which other networks exist) as well as giving the issue of responsibility attribution high priority. For the latter, responsibility remains in the exclusive domain of the human actors. non-human or material actors are recognized as having a moral impact on the network and for this reason are referred to as moral factors. They factor into the moral decision making of humans, they are a factor in the establishment of traditional and/or new norms and values and they are a factor in the establishment of the meaning attributed to a practice. A factor because the artefact bears an impact on the decisions as well as the outcome of those decisions; however, *not* an actor because technologies are not capable of being 'responsible' for their moral impact. Placing blame and/or praise on the 'responsible' agent is a necessary condition for attributing responsibility to an actor. This is of no consequence to a robot and thus it is not possible to proclaim the robot responsible. Accordingly, the structural ethics approach concludes that a technology is still recognized as having an impact but in light of it not being able to take responsibility the technology remains a moral factor and the full moral agents, capable of taking responsibility, are the human actors. More on the topic of robots and moral agency to come in chapter 8.

Thus, through design, a kind of morality is manifest, a morality decided by the designers and embedded into the robot. The care robot will invariably shape the decision-making and actions of nurses, patients and other healthcare workers and thus establishes a new morality within the network or reinforces an existing one. It is for these reasons that the design of care robots is the starting point in their ethical evaluation. Acknowledging design as a moral

activity addresses Asaro's second dimension of ethics in robotics. Deciding on the values of importance, the trade-offs made between values, and how values are manifest through the use of a technology are all decisions that make up the design process of an artefact.

1.4.2 The Design Process

For Vincenti, a design process may be divided into either a normal or a radical one. A normal design process is one for which the "operational principle" and "normal configuration" are known and employed. The operational principle refers to how the device works (for example fluorescent vs. incandescent light bulbs have different operational principles). Alternatively, in radical design processes, "the operational principle and/or normal configuration are unknown or a decision has been made not to use the conventional operation principle and/or normal configuration" [*Van Gorp and Van de Poel*, 2008, p. 79]. For example, battery operated cars in contrast with traditional cars. Within a normal design process are regulative frameworks based on the operational principle and normal configuration. Such a framework describes "the system of norms and rules that apply to a class of technical products with a specific function" [*Van Gorp and Van de Poel*, 2008, p. 79-80]. The framework "consists of all relevant regulations, national and international legislation, technical standards and rules for controlling and certifying products. It is socially sanctioned, for example by national or supra-national parliament such as the European Parliament, or by organizations that approve standards" [*Van Gorp and Van de Poel*, 2008, p. 80]. In a random design process no such framework exists.

For robots outside of the factory, no regulatory frameworks exist at present and thus designers resort to radical design processes. Such design processes are radical given the differences between robots in the factory and robots outside the factory. Firstly, the difference in performance environment - the factory is predictable and structured while the hospital or home is not (as) structured or predictable. Secondly, the difference in human contact - robots in the factory remain somewhat isolated while robots in the hospital will inevitably come into direct and indirect contact with humans on a day-to-day basis. Thirdly, the size and capabilities of the robots – robots in the home or hospital are on average smaller than those used in the factory, with a wider range of capabilities and sophistication. And lastly, the materials used to create the robots – robots in the hospital will need to be sterile, for example. Given that robots outside the factory will come into contact with humans much more often and in an unpredictable manner, the same safety standards cannot apply for both. Accordingly,

since industrial robots are used for different tasks than robots in the home, the same ethical considerations cannot apply for both. Normal design processes follow socially and legally sanctioned ethical standards, and therefore the public is inclined to put their trust in designers and the resulting technical artefacts. Alternatively, in radical design processes, the basis for trust may be lacking - designers may not explicitly pay attention to ethical criteria. Then again, with greater freedom in design, designers may pay greater attention to the ethical considerations at stake. The context within which the care robots will be situated (home, hospital and nursing home) and their potential role in the ethical enterprise of care bestows a need for greater attention to ethical considerations. Through systematic and rigorous design processes, greater focus is obtained.

Aside from the distinction between normal and radical design processes, there are hundreds of known processes. In his book, "How do you design" [*Dubberly*, Dubberly Design Process], Hugh Dubberly presents over a hundred known processes. Essentially, a design process is a way of designing, of learning what the problem is, breaking it down into manageable fractions and deciding from this the best way to resolve the problem. Design processes typically involve a series of stages or phases during which the problem is deconstructed and the potential solution is proposed and worked into a prototype. Through each process, values are selected (both explicitly and implicitly) for embedding in the system. When regulatory frameworks aren't available, design teams refer to internal design team norms, context, users, or the ergonomics of use, depending on the design process's and design team's objective (referring to contextual design, user-centered design, use-centered design respectively). Designer Bryan Lawson notes that "many models of design processes are theoretical and prescriptive rather than descriptions of actual behaviour" [*Dubberly*, p. 28]. In other words, although designers ought to observe and address the needs of stakeholders in context this is not always what happens in practice. Such a line of thinking reaffirms that work of Akrich who claims that designs are the result of assumptions an engineer has of a context rather than an understanding of the context in real life. It is for this reason that designers of late have embarked on understanding practices in context as a way of overcoming this discrepancy. For VSD, Nathan et al. make the suggestion to understand values in context – thus the values are conceptually understood from a philosophical perspective but are also understood in terms of their manifestation in context.

Given the nascent stage of the development of robots, explicitly addressing the design and design process of care robots is called for. In the specific case of care robots, the question is how the design process ought to proceed, given it is a radical one. Without a regulatory framework to guide the design of

care robots, VSD presents a (radical) design process of sorts for an enhanced ethical focus. Design processes – deciding what and how to program capabilities and appearance - adheres to Asaro's second dimension of ethics in robotics as well as the first given that the resulting care robot will contain the agreed upon capabilities as a reflection of the intentions of the designers with an in-depth understanding of the values at stake and their interpretation in a specific context.

1.4.3 Design and Empirical Research

Questions revolving around the design of a care robot also address how users will treat the robot once it has been introduced into a socio-technical network. Empirical research of opinions concerning the design of robots indicate that design will play a central role in how humans treat a robot as well as the expectations humans will have of the robot. Taking these insights into consideration through the design process of the care robot addresses the third dimension proposed by Asaro. This aspect also as to do with the domestication of the robot: how the robot will be accepted and used. In a study done by Dautenhahn et al, funded by the European Project COGNIRON ("The Cognitive Robot Companion"), the authors show how participants want a robot as an assistant, a machine or appliance, over having a robot as a friend or mate. The study also shows how participants prefer robots to communicate with them in a human-like manner but do not find human-like behaviour or appearance desirable. What's more, the appearance of the robot plays a crucial role in the interaction between the human and the robot; people expect a robot to look and act appropriately for different tasks [*Goetz and Kiesler*, 2002]. If people believe a robot's appearance ought to correspond appropriately with their assigned tasks and they also believe that robots should not fulfil roles traditionally considered within the human domain [*Dautenhahn and Werry*, 2004a], then one may conclude that robots ought never be designed to resemble a human. This conclusion corresponds with a Swiss survey which reported that only 19% of its participants (n=2000) preferred a human-like appearance [*Arras and Cerqui*, 2005]. In the prospective design of robots, such a response may be seen as a motivation to prevent the creation of humanoid robots and instead search for alternative designs. Maintaining realistic expectations of current robot capabilities is integral for the future success of robots. If users have higher expectations of a robot's capabilities when the robot has a humanoid appearance and these capabilities are not technologically feasible, users may be less inclined to support the future development and use of robots or, users may become overly frustrated with the

robot. This dimension becomes increasingly significant as we embark on discussions of care. Care, at the very least, is a relational activity. Therefore, how people treat robots, their expectations of the robot and their comfort with the appearance and capabilities of the robot will play a pivotal role in the quality of care achieved.

On a deeper level for this dimension of ethical consideration is what to do when a robot steps into a moral setting like that of healthcare. This placement does not necessarily render the robot a moral agent – moral agency may be considered a shared operation between robot and human – but robots will be engaged in activities in which their actions have moral consequences. For example, a pharmaceutical robot to dispense medications, a surgical robot performing surgery, or a robot for the lifting of patients. In other words, "the robot is required to make decisions with significant consequences – decisions which humans would consider value-based, ethical or moral in nature" [*Asaro*, 2009, p. 3]. Will this ultimately result in a need to treat robots as moral agents? If one were to believe that robots are never moral agents on their own, one might be left wondering if it is ethical to create robot soldiers or robot nurses in the first place if we are not to treat them as moral agents. Furthermore, what implications might the exploitation of robots have on other human practices and/or values? These questions pertain directly to the amount of responsibility delegated to the robot, which is ultimately decided through design and the design process (i.e., deciding which capabilities to program, etc.).

1.5 The Care-Centered Framework

Addressing or beginning with issues of design does not presume that care robots ought to be designed for any and every task. Rather, the framework I am creating allows for a critical reflection of current care practices, coupled with an investigation of real world capabilities of robots, to determine if and where care robots may be a benefit without threatening the fundamentals of care. With this in mind, designers are then free to begin exploring the ways in which the values in care may be promoted through the use of a care robot. This does not presuppose an instrumentalist view of technology – that technology is neutral and its impact is a result of usage – but instead relies on the belief that the robot can be created in a way that adheres to the values in care and further promotes them through its usage. Adhering to the methodology of VSD allows designers to take into account all three dimensions proposed by Asaro. Furthermore, such a standpoint parallels the idea that the ethics of technology should aim to

accompany technological developments rather than merely rejecting or accepting their development [*Verbeek*, 2008, 2011].

In order to create a framework for the ethical evaluation of care robots, I use the concepts and methodology of VSD. In short, the idea is to begin with value constructs relevant for a technology in question (ex. safety); to de-construct these concepts in terms of their meaning in context (ex. speed at which the robot moves and stops when a human is nearby); and to continue to program/design the technology accordingly. The interpretation of a value is encoded into the system such that when using the system the value is expressed. The values chosen are those that pertain to the technology in question. For example, a surgical robot is programmed to scale the surgeon's movements to the micro scale. This allows the surgeon to perform in a minimally invasive manner with a plethora of benefits to both the patient and the surgeon; the patient's risk of infection and scarring is reduced along with the recovery time while the surgeon is able to perform in a manner that is ergonomically beneficial for the surgeon. Overall, one can observe how the robot (through its design/capabilities) promotes certain core medical values like that of non-maleficence and beneficence. It follows then that the values to embed in a system are directly related to the system's goal and its context of use. Accordingly, this work uses care ethics for identifying the ethical values of import in the care practices of the nursing home and hospital; for deconstructing these values; and for creating a (normative) framework to analyse and evaluate care robots. Given that the values in care are the focal point for the creation of the framework, I refer to the framework as the "care-centered" (CC) framework. Creating the CC framework follows the methodology of VSD (using core care values as its starting point) and results in operational guidelines, indicating the values of ethical import in care. Putting the framework to use reveals how the values become manifest, their interpretation and meaning in context as well as their ranking in context. The framework is then used for two types of value-based analysis of care robots:

1. for retrospective evaluations of current care robots prototypes in combination with script theory [*Akrich*, 1992; *Latour*, 1992] and,

2. for prospective design and implementation of future care robot designs in combination with the structural ethics approach and domestication studies

1.5.1 Value-Sensitive Design Methodology

Value-Sensitive design (VSD) has been praised by computer ethicists and designers for its success in incorporating ethics in the overall design process of

computer systems of Information and Communication Technologies [*van den Hoven*, 2007] but is also advantageous to guide the design process of a variety of technologies **[*Cummings*, 2006]**. As mentioned, VSD begins from the perspective that technologies, through their use, promote (or demote) certain values. Applying this to robots in care contexts (i.e., care robots), the goal is to design robots in a way that through their use the values in care are promoted. The tri-partite methodology of VSD consists of a conceptual investigation, empirical investigation and a technological investigation. The conceptual investigation is the portion in which the value constructs are chosen and their interpretation and meaning from a philosophical perspective are explored. In traditional VSD approaches, the empirical portion consists of testing the designed system in its context of use. For my work, the empirical investigation makes up the component in which real world variables of the context in terms of the interpretation of the values in context as well as the users and their needs, are taken into consideration for the design of a system. Hence, I am not testing an artefact once it has been designed. This is so given my own inability to create a robot; however, this should ultimately be incorporated within the design process and will be the goal of future work. Finally, the technological investigation is the section in which trade-offs between usability and values is explored. Again, this is done through an analysis of the technical components and their relationship to the expression of values as opposed to observing the utility of the robot in practice. This is the aim of future work once I have proven the function of the CC framework and methodologies for use in this text. Although the three are listed as separate portions of the overall methodology, in practice they are overlapping and cannot be separated. Moreover, relating components to one another, as I will do, only strengthens the fluidity and consistency of the approach.

1.5.2 Creating the Care-Centered Framework

I begin this work with an investigation of the important concepts – a conceptual investigation of certain concepts, terms and ideas that play an integral role in this work. Specifically, a discussion of values, care values and how a value is embedded is necessary (chapter #2) to ground the central idea of this thesis – that a care robot holds the potential to promote the expression of values through its design and usage. I continue with a technological investigation (chapter #3), outlining a variety of robot capabilities and current robot prototypes (some which are commercially available at the moment). This is done to ground the reader in the current state-of-the-art for a realistic assessment

thereby avoiding evaluations of speculative care robots. I then proceed with an additional conceptual investigation of the 'values of ethical importance' according to the VSD tri-partite methodology. Friedman et al suggest a list of twelve values of ethical importance that are often implicated in the design of computer systems. For Lucas Introna this is an authoritative way of dictating the values at stake. Instead, Introna suggests the approach of disclosive computer ethics [*Brey*, 2010] as a way of uncovering the values at stake which pertain to the specific technical details of the system. This method is an exhaustive one identifying and analysing every value possible but one is still left wondering how to decide which of the values are relevant and further who the values are relevant for? Given the institutional context of care, I begin with the more abstract values at the organizational level articulated by the WHO as well as guidelines and ethical frameworks for hospitals – these are the values of ethical importance for healthcare providers to ensure that good care is achieved. To understand the meanings of these values from a philosophical perspective, I use the concepts and insights of care ethics. Thus, care ethics is used to select the values of ethical importance (at an organizational level) as well as de-constructing their meaning meta-physically and meta-ethically.

For the next step in deconstructing care values, I use observations from fieldwork experience in a variety of care settings to articulate how the care values are interpreted and expressed in context (chapter #4). This comes from the noteworthy indication of Le Dantec et al. [2009] with regards to the VSD methodology; that values ought to be understood in situ, in context. This also adheres to the suggestion of care ethicist Joan Tronto in terms of understanding the particularized nature, and necessity thereof, of care institutions [*Tronto*, 2010]. The idea here is to specify how care values are interpreted not in a meta-ethical sense but in a practical sense. In other words, to observe the manifestation of values through care tasks among the network of actors (human and non-human) involved. Furthermore, how values are prioritized throughout a task. This prioritization has to do with the context within which the care is happening as well as the demographic for whom the care is being delivered. For example, the prioritization of values expressed when a nurse is bathing a patient differs when a nurse is bathing a patient in the complex care ward of the hospital versus in the nursing home. In the hospital, greater attention is paid to the safety of cleaning given the fragile state of patients. Compassion is also highly valued throughout the provision of the task given that that may be the only time in the day that the patient has (intimate) contact with another human. In the nursing home, focus is directed towards time efficiency. Nurses are on a tight schedule and are aware that the resident will have many interactions with

other people throughout the day.

Understanding the multiple dimensions of care tasks also means understanding that many tasks are about more than just that specific task and further, that roles and responsibilities are allocated through the provision of said task. This knowledge is crucial before one can undertake an assessment of the impact robots will have on the distribution of roles and responsibilities. This is also pertinent information for understanding exactly what it is the nurse does in order to create a robot that does the same or helps to meet the patient's needs in one dimension without compromising the needs of another dimension. What's more, understanding how values may be prioritized differently for the same task depending on where and for whom this task is being performed, draws our attention to context. It is my aim to show how attention to context is necessary in the evaluation of current care robots as well as the design of future care robots. Context differs from application domain in that it is much more specific. In one domain (ex. healthcare) there may be a variety of types of robots and a variety of uses for these robots. Ethically assessing all of these robots according to the overall domain fails to take into account the range of users and scenarios present. Although I use the ethical values from one domain (healthcare), I aim to show how their interpretation and ranking differs between contexts/hospital wards. This, in addition to the capabilities and appearance of the robot, the contextual practice of care is integral for design considerations of the robot.

Through this analysis, we come to understand that care tasks are not tasks but are rather care practices. They are practices involving human and non-human actors and the interplay among and between actors. This rests on the assumption that values are an interpretation of what is happening between all the elements, of how the elements work together to bring values into existence. As such, the CC framework provides a normalized account of the different kinds of values important in different care practices and the ranking of those values, normatively understood. The additional value of the CC framework is to make the designer understand the complexity of care practices as well as the complexity of designing for these practices. This work (chapters 3 and 4) is a combination of both conceptual and empirical investigations: care ethics for deconstructing value constructs and field work experience for empirical knowledge to align real world practices with theoretical values. The result is the CC framework tailored to the ethical evaluation of care robots with design issues at the core and a particular emphasis on how roles and responsibilities are distributed in a care practice (network). How the framework is then used will be described in detail in chapter 5 once I have presented the necessity of each component of the framework and its relationship to the moral landscape of the care practice.

1.5.3 Criticisms of Value-Sensitive Design

Peter Asaro disregards the methodology of VSD for robots in the military claiming that what is needed instead is: "an understanding of the kinds of information people use to make various sorts of ethical decisions, how they process that information, and how the presentation and representation of that information influences their performance in ethical decision-making tasks" [*Asaro*, 2009, p. 22]. This point presupposes two things: the first being that ethical decision-making in care can be broken down in a systematic manner and programmed into a system used for any care task, and two, that robots ought to be delegated tasks in which they are required to formulate their actions following ethical deliberation. The domain of healthcare is one in which the ethical decisions made by healthcare workers involve life and death issues. Decisions in the hospital concerning patient care are made using real-time variables concerning the patient's unique status physiologically, emotionally, culturally and otherwise. It is fair to say that such information would be beneficial in the development of care robots; however, it is another thing entirely to say that robots ought to be placed in positions in which they are required to engage in ethical decision-making in a healthcare setting . Consequently, Asaro's reasoning cannot be applied for robots in care.

Asaro further discounts VSD claiming that it addresses robot design on a case-by-case or design-by-design basis which merely demonstrates the superiority of one abstract value over another. However, by grounding the value constructs in empirical research (fieldwork experience), the values no longer remain in their abstract form but rather are interpreted in their context of use. Moreover, in light of the intricate distribution of roles and responsibilities and how dramatically they (and the values in a network) may shift depending on the capabilities or appearance of the robot, assessing care robots on a design-by-design basis is the precise manner in which designers ought to proceed. In one design, the robot's capabilities may fundamentally change the distribution of responsibilities whereas in another design, for the same task, the robot's capabilities may not have much of an impact at all or, may re-introduce values overlooked with other technologies. For example, an autonomous robot designed to cue the operator (nurse or patient) for or during care tasks implies that a certain amount of attentiveness is delegated to the robot. Attentiveness, on the part of the care-receiver, in care is a fundamental and necessary ethical component for the provision of good care [*Tronto*, 1993, 2010]. Consequently, such a care robot may enhance the promotion of attentiveness and strengthen the care process. Alternatively, a human-operated robot requiring feedback after cueing

the operator, may take the nurse's attention away from the patient entirely, ultimately sacrificing attentiveness of the nurse. Consequently, this kind of care robot demotes the value of attentiveness, potentially sacrificing the care of the patient, and as such may be morally problematic. These questions pertain specifically to capabilities and features of a given design, and are only identifiable through analysis on a design-by-design basis.

In line with this view, some have criticized VSD for its lack of alignment with one particular ethical theory or approach, that at best it is a descriptive enterprise [*Albrechtslund*, 2007; *Manders-Huits*, 2011]. VSD is thus thought to lack some standardization. This problem does not apply to my work as I have chosen care ethics as the basis for analysis and evaluation. The question then remains whether researchers like myself are free to decide the ethical approach most suitable for the technology in question. To counter this claim, I respond with the insight that care robots are to be integrated into the therapeutic relationship in a context governed by care values and thus there appears to be no other approach that can accommodate such variables.

Lastly, for some, using the methodology of VSD presupposes an instrumentalist view of technology – that technologies are created as neutral and their ethical impact is dependent on the human user. In other words, what counts is not the technology itself but how we choose to use it. Although VSD asserts that the value or disvalue of a technology is manifest through its use, this differs from the idea that the technology is inherently neutral. Instead, VSD relies on the belief that a technology brings with it an embedded or built-in value and it is through the use of the artefact that this value is observable. VSD does, however, rest on the belief that designers are powerful [*Feng and Feenberg*, 2008] and further that "their intentions are expressed through design" [*Feng and Feenberg*, 2008, p. 106]. This belief might appear to place VSD in the realm of technological determinism whereby technology is the means for steering social change and it is through the intentions of the designers that said change comes to be. Nevertheless, the role of care robots cannot be entirely reduced to the intentions of designers. As we have seen, human actions are an interplay between human actors and technologies in their environments – technologies help to shape actions and decision making processes [*Verbeek*, 2005, 2011]. My conception of VSD is dependent on this last point, what Verbeek refers to as technological mediation, and a belief in the co-evolution of technology and morality – neither is independent of the other, the two evolve in tandem. Care robots will be created against a pre-existing background of values and norms and with the introduction of this technology comes new practices and perceptions. When we explore the assumptions, beliefs and/or intentions of designers, we find they are

not exclusive to designers or isolated from public opinion. The choices made by designers are value-laden according to culture and time. As such, ethical reflection of a care robot begins with design as a way to reflect on the quality of caring practices resulting from its integration into a network.

1.6 Conclusion

The projected lack of healthcare workers and resources has invited a range of potential solutions to mitigate the anticipated consequences. As I have stated, one of the proposed solutions is to introduce robots into care contexts (care robots) as a way of aiding care workers in institutions and/or as a way of helping care-receivers remain in the comfort of their own home for a longer period of time before entering into a care facility. Such a technology, given its foreseen place in the moral practice of care, demands appraisal from an ethics perspective. From a care ethics perspective, the main question at stake has to do with the shift in roles and responsibilities following the introduction of the robot. Similarly, for those from the field of ethics and technology, the link between design and morality demands that design be the place to begin an assessment of the impact of said robots. Given the lack of socially and politically sanctioned design standards, ethical criteria in the design process of care robots may be jeopardized. Moreover, once a care robot has been created, no ethical guidelines for its retrospective evaluation exist to date.

In response to the lack of tools for the ethical evaluation of care robots both retrospectively and prospectively, I am creating a framework for such kinds of evaluations. Combining the methodology of VSD with the value constructs and their interpretation from a care ethics perspective, I am creating an approach to the ethical evaluation and assessment of care robots. This approach will result in a normative framework, the Care-Centered framework, which will serve two value-based design analyses. The framework provides a normalized account of the values in care, their interpretation in context and their ranking. Firstly, the framework will be used in combination with script theory [*Akrich*, 1992; *Latour*, 1992] for an ethical evaluation of current care robots. Thus, such a retrospective evaluation may be used for recommendations pertaining to the improved future design and/or the implementation of the care robot. Secondly, the framework will be used along with the structural ethics approach as a tool for designers in the ethical design of future robots in care contexts. The CC framework is meant as an evaluative framework incorporated into (radical) design process of care robots.

The focus on design does not presuppose a blind acceptance of all care robots for all tasks but is meant to reiterate the significance of design and the design process in the resulting artefact. As Verbeek points out, "designers materialize morality" [*Verbeek*, 2008, p. 99] when we consider how artefacts alter experiences and practices. Accordingly, consideration of how a care robot might alter experiences and practices through its design, is placed at the centre of this work. Based on the suggestions of Asaro, we see how design is integral in the ethical evaluation of robots on multiple levels. Each of the levels discussed by Asaro refers back to the central concern being the distribution of responsibilities. In care too, this is a central concern. I now turn to a conceptual investigation of what a value is and how it comes to be embedded in a technology.

Chapter 2

Values and Assumptions Embedded in Technology

2.1 Introduction

The underlying aim at the core of this work is to articulate a series of values which are to be embedded in the design of a care robot in such a way that, through its use, the core values of the healthcare tradition are manifest. To do this, we must understand how values are expressed and ranked within a care network (Chapter 3). But first what is needed is an understanding of what a value is, and how values are embedded into a technology. Thus, to begin the conceptual investigation, I take as my starting point a clarification of terms and concepts. This chapter is descriptive in that it seeks to outline what a value is and how it comes to be embedded in a technology. This chapter also deals with the embedding of assumptions and norms. Assumptions may pertain to users, biases, and norms that may or may not reflect a value but are all components that are integrated into the design of artefacts.

2.2 What are Values?

The discussion of values is multi-faceted and lacking universal consensus. Scholars continue to struggle with the variety of questions surrounding the concept of a value: what is a value, what are the different kinds of values, how they are

achieved, how they are given meaning, where does the worth of a value come from, etc.? The aim of this chapter is not to engage in this discussion, given the lack of agreement, but rather to outline the concept of values used for this work. In the Value-Sensitive Design literature, Batya Friedman and colleagues, opt for the more open definition of a value to refer to "what a person or group of people consider important in life" [*Friedman et al.*, 2003, p. 2]. This is in contrast to the more confined definition of a value as referring to the monetary value of an object. Friedman et al point out the history of discourse pertaining to values and the range that this discourse takes to include "the good, the end, the right, obligation, virtue, moral judgement, aesthetic judgement, the beautiful, truth and validity" [*Friedman et al.*, 2003]. Friedman et al identify twelve human values with ethical import which are often implicated in system design. The reason for this selection is based on deontological and consequentialist perspectives as well as those values which they claim are "related to system design" [*Friedman et al.*, 2003, p. 6]. These values were chosen not only for the impact that systems render on them but also because of their universal importance and recognition. The authors indicate that the list is not comprehensive and that the values should not be considered distinct from one another. The former point means that the scope of values varies depending on the technology, the users, the culture, the time period and the application domain. It is important to note here that the list of values to be observed in care contexts will differ from the values observed in a military or industrial context and further that "each value has its own language and conceptualization within its respective field" [*Friedman et al.*, 2003, p. 6]. This implies then that all the values are not interpreted in the same way. Nathan et al illustrate this with the value of privacy and its divergent ways of being interpreted and therefore protected. Le Dantec et al reinforce the idea that values may be universal or generally accepted but differ in their interpretation. Because of this, Le Dantec et al suggest a way in which the methodology of VSD may be strengthened, through an uncovering of values in situ, or discovering values through experiencing the practice [*Le Dantec et al.*, 2009]. This is of course due to the idea that differences exist between designers' values and users' values [*Nathan et al.*, 2008].

2.3 What are Values in the Embedded Sense?

With an idea of what a value is or what it means to value something, how then are values embedded in a technology? Brey provides a straightforward way to envision this:

> *The idea of embedded values is best understood as a claim that technological artefacts have built-in tendencies to promote or demote the realization of particular values. These tendencies are manifest through consequences and thus a built-in value is a special sort of built-in consequence* [*Brey*, 2010, p. 3].

In other words, the consequences of using the technical artefact may promote the realization of a value. This is referred to as a causal relationship between the artefact, the built-in value and the consequences of using the artefact; "values are related to causal capacities of artefacts to affect their environment" [*Brey*, 2010, p. 6]. For example, let us look to the healthcare context and surgical tools, specifically endoscopic tools used in minimally invasive surgeries, as the technical artefact in question. The most significant value in healthcare, universally recognized, is that of human dignity and thus all practices and technologies in healthcare aim at preserving and promoting human dignity. Although human dignity is somewhat of an abstract value, in the healthcare context it may be promoted through curing and caring tasks, essentially addressing both the physiological and emotional/social needs of patients. How can a technical artefact embed the value of human dignity, or promote the realization of this value? We may suggest that the size of the tool promotes the realization of the value of human dignity by preserving the physical integrity of the patient. In other words, the small size of the tools coupled with the use of an endoscopic camera allows the surgeon to operate successfully with great physiological advantages for the patient – a reduction in pain, scarring, risk of infection and recovery time. Thus, the consequences of using the tools result in an improved quality of care and quality of life in general. Such consequences are built-in to the technology and may be brought about only through the use of the technology.

Alternative to the causal conception of built-in values is the expressive vision. According to this vision, "artefacts are considered expressive of values in that they incorporate or contain symbolic meanings that refer to values" [*Brey*, 2010, p. 6]. What is meant here is that a user will buy a certain product because of the status or success attributed to that product. The problem with this line of thinking is that it does not imply that the technology also functions to realize these values. If someone buys a product because it represents a certain value, the value is not necessarily realized for the individual in his or her life. Id, for example a consumer buys a product from an "ethical company" or buys a product considered ethical (free trade coffee for example), does that make the consumer an ethical person? What is the link between the symbolic value of the product and the status acquired by the user? Or, if someone plays video games

that discriminate against one gender or race, will the individual then take on this set of values as their own? It is possible to suggest that certain symbolic values expressed through an artefact have a causal role – meaning they do in fact bring about the manifestation of the value. In fact, one might suggest that the causal relationship involves, and further relies on, a symbolic one. Energy reading devices in the home exemplify this idea; the consumer believes they are making an ethical choice by buying from the ethical company making environmentally sound technologies and in turn, the technology, through its use, promotes the value of sustainability and minimal or reduced energy consumption.

On the other hand, certain symbolic values will carry no casual role – meaning, you can read or see values in an artefact but when it is used, it might not have a causal role in promoting the symbolic value. For example, Peter-Paul Verbeek speaks of the development of many kitchen appliances and how they came to be designed in a particular way [*Verbeek*, 2005]. Dishwashers, fridges, microwaves all took on a 'clean' look through their colour and style. Thus, the symbolic value of cleanliness (along with the assumption of how cleanliness was represented) was built-in to the technology, but this did not translate into an actual clean kitchen space. In other words, the consequences of using the clean technology did not bring about a clean kitchen but merely symbolically denoted what a clean kitchen should contain.

Both the causal as well as the symbolic embedding of values into a technology will be addressed in this work. Often times the symbolic or expressive vision of built-in values has to do with the appearance of, or the materials used to make, the artefact. Symbolically expressed values may carry significant influence; however, the more interesting and critical embedded values pertain to those that are causal. Causally expressed values will play a decisive role in the distribution of roles and responsibilities.

2.4 Are Values Enough?

The VSD approach addresses issues pertaining to design. The question of whether values are enough has to do with my overall query concerning the most important question pertaining to robots – the distribution of responsibilities in the socio-technical network into which the care robot has been added. More specifically, what is the extent of the responsibility of a robot and what kind of responsibility are we delegating to robots? By type of responsibility, I am referring to whether the consequences of the actions of the robot are considered moral in the eyes of humans and by extension, I am referring to how much

responsibility we actually delegate to the robot. For example, a surgical robot is operated by a human; thus, a human is required to carry out the action and ultimately the human bears the responsibility for the physical well-being of the patient. Because the human is responsible for deciding whether the continued use of the robot is the best option the moral responsibility remains with the human. The type of responsibility delegated to the robot is not moral given that the robot is only delegated the role of visualizing the surgical field and scaling the movements of the surgeon. Again, the surgeon is responsible for deciding what her movements will be and when they will be made. To be clear, I am not referring to telesurgery (which introduces a telecommunications network) but to real-time robotic surgery during which the patient and surgeon are in the same room.

Addressing the values in question, and how those values are promoted or demoted through the design of the care robot, does not explicitly invoke a discussion of the distribution or delegation of roles and responsibilities imposed within the socio-technical network of actors – human and non-human – into which the robot is stepping. When an artefact imposes certain roles, we may consider this a special kind of consequence, one that may or may not carry with it a valuation of said roles and responsibilities. If this does not pertain to a built-in value then what is it? Perhaps a norm?

For Brey, norms are "rules that prescribe which kinds of actions or state-of-affairs are forbidden, obligatory, or allowed" [*Brey*, 2010, p. 6]. Given this definition, we can see how a norm may be based on a value and as such endorse the promotion of said value by encouraging certain behaviours. It is also important to distinguish between norms considered moral and those considered non-moral. Moral norms prescribe or prohibit certain behaviours from a moral perspective. For example, "do not kill humans". Non-moral norms do not rely on the same moral foundation for support. An example of a non-moral norm in the healthcare context may be 'standing out of the way of patients and healthcare personnel as they proceed through the hospital hallways'. The former is based on a moral value of human dignity and life in general while the latter is based on the value of social etiquette or behaviour in the hospital. Brey puts forward the idea that norms may be embedded into a technical artefact in much the same way as values are thought to: "embedded norms are a special kind of built-in consequence they are tendencies to effectuate norms by bringing it about that the environment behaves or is organized according to the norm" [*Brey*, 2010, p. 6].

When we link the concept of values to a specific application domain like that of healthcare for example, we may observe in detail how values are related to

norms. Firstly, I have claimed that the healthcare context supports the promotion of distinct values whose definition and interpretation rely on the details of (specific) healthcare scenarios. Next, I suggest that the norms within healthcare scenarios are based on the continual expression of the predominant value within healthcare namely that of preserving human dignity. How such a value is interpreted depends on a further classification in healthcare. In surgical applications, human dignity is promoted through multiple activities on the part of the surgeon; the acquisition of expert surgical skills, the performance of surgery with careful attention to detail, the ability to perform multiple types/styles of surgery (i.e., conventional, laparoscopic or robotic). How the technical artefact (surgical tools) might embed the value of human dignity, or promote the realization of this value has already been discussed in terms of the size of endoscopic tools. The norm may then be thought of as an obligation to perform surgeries laparoscopically if the patient meets the physiological criteria and if the surgeon is capable. This norm results from the statistically significant benefits of laparoscopy over conventional surgeries (for certain surgical domains and procedures) and is based on the valuation of these benefits as a promotion of the value of human dignity interpreted in terms of physical integrity.

In the nursing context, the value of human dignity may be interpreted using the concepts of caring and touch. The predominant threat in the healthcare setting is that of reducing the patient to an object, which is considered a violation of one's human dignity – namely through failing to recognize the patient as a person [*Gadow*, 2002, p. 35 - 36], or failing to *care about* the patient. Healthcare practitioners have suggested two ways in which the dignity of persons can be upheld in the healthcare context; through caring and through touch. The process of care outlined by Tronto serves to mitigate this threat by paying meticulous attention to the relationship established between the patient and the physician. Touch is the symbol of vulnerability, which invokes bonds and subjectivity [*Gadow*, 2002, p. 40]. Touch acts to mitigate the temptation for objectification. Thus, each is considered a value in the healthcare domain and the outcome of each practice results in the preservation of the value of human dignity. Consequently, the norm, or standard of care, incorporates these two values. Specific guidelines are provided for the element of touch to ensure appropriate boundaries within the therapeutic relationship. The question then becomes how these values can be embedded into a technical artefact, or in other words how can the use of a technical artefact encourage the realization of these values in practice? And further, are they conducive to each other or mutually exclusive (values in care and the technical system)? Melanie Wilson illustrated how a particular computer system implemented in the field of nursing was rejec-

ted as it prevented nurses from "hands-on care", the cornerstone of the nursing practice [*Wilson*, 2002]. This particular system was rejected due to the fact that assumptions made about the care-givers and the care process were embedded in - meaning already there - the system; however, these assumptions did not match the real world practice. This demands that we expand our discussion of values to include a discussion of assumptions on the part of designers.

2.5 Uncovering Values and Norms Through Assumptions

At this point I am addressing the presence of built-in values and norms in a technology; however, the difficulty lies in understanding how these embedded values and norms came to be. Nissenbaum outlines a linear relationship of biases the engineers or designers may have which result in a variety of kinds of built-in biases. Another way of articulating this idea is found in the script theory proposed by Madeline Akrich and Bruno Latour. This theory illustrates how engineers have assumptions about a user group, as well as the world within which the users reside, and these assumptions effectively shape the design of the technology.

> *Designers define actors with specific tastes, competencies, motives, aspirations, political prejudices, and the rest, and they assume that morality, technology, science, and economy will evolve in particular ways. A large part of the work of innovators is that of 'inscribing' this vision of (or prediction about) the world in the technical content of the new object* [*Akrich*, 1992, p. 207].

For Akrich, the assumptions pertaining to users and assumptions pertaining to the future context of users and their environment are embedded into the technological artefact such that they impose a particular role on the user. Akrich claims in the above statement that the assumptions about users are multi-dimensional and encompass things such as users' political views, motives, tastes, aspirations. It follows then that assumptions are made about values and norms in a user context. According to Introna, "through these user inscriptions, those that encounter and use these inscribed artefacts become enrolled into particular programs or scripts for actions" [*Introna*, 2005, p. 76]. What is also assumed here is the role the technical artefact will take on and the roles of the human users involved. The way these roles are ascribed to the actors (both human and

non-human) has to do with the responsibilities dictated to those actors, what Akrich refers to as "a geography of delegation" [1992, p. 206]. In other words, embedded in the technology is a presupposition about the responsibilities of the technical artefact as well as the human actors involved. By responsibilities, Akrich refers to which tasks and roles are delegated to which actors; "technical objects participate in building heterogeneous networks that bring together actants of all types and sizes, where human or non-human". From these responsibilities, roles are assumed. This collection of roles and responsibilities that is exerted through the artefact is referred to as the **script**, the process through which these assumptions are integrated into the technology is known as **inscription** and the process through which the assumptions are revealed is known as **de-scription**. For Latour, the actions, dictated by the technological artefact, of the human and non-human actors is known as **prescription**.

As mentioned, for Nissenbaum, the route for visualizing the progression to built-in values begins not with assumptions but with biases. A bias is a belief about something from an impartial or prejudiced stance or a belief based on information that is "untrue". If an assumption is defined as a belief about something, a belief that one presumes to be true, then when this belief is based on information that is "untrue", impartial or prejudiced the assumption may then be considered biased. With this in mind, we can see a clear link between Nissenbaum's embedded biases and Akrich's notion of scripts. This is not to say that all assumptions are biased. On the contrary, although an assumption may be considered a bias in certain instances, or biased – in the case that it is impartial – a bias is always defined as impartial. In the theory of scripts, it is suggested that the assumptions inscribed in the technology may also pertain to an ideal [*Wilson*, 2002], for example, the technical content of a care robot is dependent on the vision of the ideal care-worker and the ideal care practice. This may be considered a bias on the part of the designer or it may be considered a normative claim made by the designer.

Nissenbaum nicely illustrates the three origins of biases which Brey then applies to values [*Brey*, 2010, p. 8]. According to Nissenbaum, there are three different types of biases; *pre-existing biases*, *technical biases* and *emergent biases*. *Pre-existing biases* refer to the pre-existing values and attitudes at both the individual as well as the societal level *Brey* [2010, p. 8]. An example of this may be the societal attitude that women are traditional care-givers. If a care robot is designed to look female this may be seen as a built-in valuation of care-givers as female based on a pre-existing bias. *Technical biases* are what Nissenbaum refers to as a bias resulting from technical constraints, considerations or limitations. When such a bias is reflected in a technology it may also

make a statement about the prioritization of values. For example, in the case of surgical robots, incorporating tactile feedback (touch) into the technical content of the robot is not yet possible. As a result, one may conclude that the designers were limited by a technical constraint which leaves one wondering if touch was not placed as a high priority or was sacrificed for other values like accuracy or comfort of the attending surgeon. An *emergent bias*, may come about when the system is used in an unintended way or by an unintended group of users. This novel way of utilizing the system may not support the values, interests or capabilities of the unintended users. Again referring to surgical robots as an example, these robots were originally advocated for use in the military via telesurgery (remote robotic surgery); they were thought of as a way to keep surgeons off the battlefield while at the same time providing the capability to save soldiers on the field. When telesurgery was first introduced into traditional healthcare settings, the idea of separating the patient from the surgeon in this way did not have the same appeal. It presents ethical considerations pertaining to the establishment of the surgeon-patient relationship as well as issues of trust, liability, responsibility and a further objectification of the patient [*van Wynsberghe and Gastmans*, 2008].

The relationship between biases and assumptions then has to do with the idea that assumptions are somewhat neutral (unless pertaining to a value in which case there may be a normative claim being made) while biases carry a negative connotation. This work will rely on assumptions to allow for both positive and negative assumptions being made, but will also reference Nissenbaum's illustration of biases.

In contrast to the EVA, Akrich proposes that the force within the technology exerts its effect in terms of roles in the resulting network, or scenario as Akrich calls it. This means that the assumptions embedded into a technical artefact exert a force on the users and this force imposes a particular geography of responsibilities and as such a particular geography of roles. The imposing force, which gives rise to the resulting geography of responsibilities, comes neither from society nor from institutions but rather from the artefact itself. For Nissenbaum the built-in force in the artefact is referred to as the embedded value while for Akrich, the built-in force in the artefact is known as the script. Accordingly, we can uncover what the script is (and the assumptions leading up to the script) by examining the roles imposed on the users through the technical artefact. By addressing values and norms in this way – as assumptions about values and norms and how they are to be promoted – I will take a more critical, normative stance. Consequently, script theory provides the missing link for uncovering what values and norms are assumed in the design process by addressing how

assumptions about users, user context, and the norms and values pertaining to both, impose a specific 'geography of delegation' in the resulting socio-technical network.

To illustrate the changes in responsibilities and roles observed through the consequences of using a technology (by uncovering the embedded script) take we will examine the surgical tools mentioned earlier. Laparoscopic tools shifted the practice of surgery from conventional methods that required a greater amount of recovery time to highly specialized methods that reduced the pain, scarring and overall recovery time of the patient. Patients require considerably less post-operative recovery time in the hospital and may go return home in a matter of days, when previously it may have taken weeks or months for a full recovery. Although there are numerous benefits to the patient as well as to the hospital (which can function more efficiently by serving more patients in less time due to shorter stays) with the introduction of laparoscopy, there are added responsibilities and roles taken on by the home-care providers who are usually family members untrained in medical care-giving and procedures. This can be quite the undertaking for family members who lack knowledge, training, expertise and time. The underlying assumptions here have to do with the role of healthcare in the care of persons – that care is reduced to surgical intervention (curing) and that this is the role and responsibility of healthcare institutions but post-operative care is not. This may be thought of as a pre-existing bias. Or, perhaps this handing over of responsibility to the patient and their family (if they have one) to arrange for their own care may be an emergent bias only perceivable once the practice of laparoscopy was well established. In terms of values, one may perceive this as a prioritization of curing means over caring means. Alternatively, it may be the result of economic concerns from a utilitarian perspective – how can we provide care for as many people as possible?

The idea of prescription, or an imposing force of the technology on behaviour, is what is more recently referred to as engaging technologies or choice architecture [*Verbeek*, 2011]. Designers have picked up this idea and have begun to create technologies that stimulate reflection or ethical decision-making of the user, for example, energy recording devices to be used in homes which indicate how much energy you have consumed in a day. This type of technology is not meant to enforce a certain behaviour but is meant to stimulate reflection of the household such that they are free to decide if and how they will lower their energy consumption. This is after all the aim of ethics – to empower individuals to make choices for their own life based on a rational conception of good. This is an example of a persuasive technology which directly interacts with the user. Other technologies that aim to engage the user do so in a seductive way, a

coercive way or a decisive way. The significance of this concept for my work has to do with the eventual design of the care robot and the imposed behaviour or engagement it elicits from the user. Once I have created the normative framework for evaluating current care robot designs and proscribing future designs, the aim is to understand how and when the care robot ought to engage the nurse and/or the patient and in what way. Engaging and/or prescribing actions is a striking ingredient for the design of future care robots and will be given attention later on.

2.6 The Relationship Between Assumptions, Values and Norms

As I have now identified, there are numerous concepts built-in to the technical content of an artefact – values, norms, assumptions pertaining to these as well as assumptions pertaining to the roles and responsibilities of the actors in the resulting socio-technical network. All of these may be observed through the consequences of using the technical artefact. For Verbeek, mediation is the phenomena to articulate the consequences of using a technical artefact. Using a technology changes our experience and interpretation of the world and it is these consequences that we ought to observe and reflect on throughout the design and development of a technology and continue to do so after the technology has already been established in a context.

But what is the link between all of these elements? An assumption may be defined as "a supposition on the current situation or a presupposition on the future course of events, either or both assumed to be true in the absence of positive proof" . Thus, an assumption has to do with asserting a claim without proof of the validity of said claim. As assumption does not carry with it positive or negative connotations. According to the Collins English dictionary [2003] an assumption is quite similar to a belief; "a principle, proposition or idea accepted as true". The common theme here is that some thing, principle or course of events is believed to be true, factual, or correct, and is in some way reinforced. Thus, assumptions may be made about values, norms roles and/or responsibilities. As indicated above, the assumptions may pertain to how the current situation works or they may be assumptions about how morality and technology will co-evolve. Whether or not assumptions about a user, role or value come to fruition depends on the real world conditions. If the artefact is integrated into the existing socio-technical network, one may conclude that

the embedded assumptions of said artefact were accurate or came to be seen as accurate. If the artefact is rejected from a network, one may conclude that the assumptions were not accurate or conflicted with existing norms and/or values. Melanie Wilson considers the second instance to be a conflict of scripts – when the embedded assumptions about the practice of nursing do not match the real world practice and as a result the technology cannot be integrated [2002].

An important distinction must be made here pertaining to the difference between assumptions and the concept of values and norms. Assumptions are more about the real world; they are descriptive in a certain sense, while norms and values are more about what the real world ought to be like; they are normative. One could surmise then that perhaps the embedding of norms and values based on assumptions of those norms and values is a kind of embedded naturalistic fallacy; that engineers are in some way claiming that what *is*, is what *ought* to be. In contrast, when an assumption is made about a value to be embedded, it does not have to be an exclusive description about what *is* expressed, but could also be a claim about what values *ought* to be expressed, how they ought to be expressed, or possibly even what priority they ought to be given. In others words, when the built-in assumption pertains to a value, or when a valuation is being made, the result is a normative claim about what the values should be, or what should be valued. For Akrich, "many of the choices made by designers can be seen as decisions about what should be delegated to a machine and what should be left to the initiative of human actors" [1992, p. 216]. By making choices about what should and should not be delegated to certain actors (human or non-human), engineers may change the distribution of responsibilities in a network. Understanding this relationship demands that we make assumptions explicit to identify when an assumption is descriptive in nature and alternatively, when an assumption is prescriptive in nature, and then reflect on these assumptions in a critical manner, prior to their manifestation.

Introducing the concept of technological mediation here allows us to link all of these elements in a cohesive manner. If mediation is the phenomenon whereby the technology influences our decision-making and experiences with the world in a way that shifts practices and traditions, and if it is the design of the technology that encourages or forces this shift, then it is that which is embedded in the technology that bears the force for shifting. It is the values, assumptions and norms embedded in the technology that effectively result in the re-shaping of practices and traditions. It is this re-shaping potential that we want to investigate in the design of current and future care robots so the concept of mediation is important here.

When we say that something 'improves' care, we are referring to a type of

mediation. There is a shift in our experience or perception for the better. As Verbeek describes, mediation can be along two dimensions; pragmatic and/or hermeneutic. The first refers to how technologies change our actual practices and the latter refers to how technologies change our interpretations of the world [*Verbeek*, 2006]. Akrich and Latour focus on the pragmatic mediation of technologies: how they enforce certain behaviours or actions, using the speed bump as an example (which says to users "slow down before getting to me"). Idhe, on the other hand, illustrates the hermeneutic dimension: how technologies change our experience and interpretation of the world, using ultra-sound technology as an example (which shifts the ontological status of the fetus and places in a medical context for which medical decisions are to be made about the now patient and not the fetus). If we use the example of endoscopic surgical tools, one can observe their potential for mediation along both dimensions. Their design commanded the surgeon to perform in the least invasive manner possible. This then became a norm, thus in hermeneutic terms, a pragmatic mediation (a change in the way surgery was performed) led to a hermeneutic one (a change in the way conventional surgical methods were perceived) which resulted in a new norm, a new standard of care.

In hermeneutic terms, the tools change the experience of surgery for the surgeon. The surgeon no longer touches the patient or has any kind of tactile feedback from the patient. This may create a new ontological status of the patient in a more profound way; the patient's corporeality is removed from the surgeon in such a way that they become a mock-up (or representation) on which the surgeon performs a sophisticated type of procedure. For Sally Gadow, the hermeneutic mediation is the one to pay particular attention to as it focuses on the objectification of the patient [*Gadow*, 2002]. By limiting or eliminating contact between the surgeon and the patient (the surgeon no longer is touching the patient with her/his hands during the surgery, nor does the surgeon have tactile feedback), the patient is further objectified by the surgeon. The patient is no longer a person with needs to be met but is a 'diseased' object which the surgeon's tools will rectify. In short, the tools symbolically and causally embed a disvalue – an objectification of the patient.

2.7 Conclusion

With an understanding of what a value is and how values come to be embedded in a technology alongside norms and assumptions, what comes next? An interesting connection between these elements is that values, norms, and as-

sumptions are all made real upon examination of the consequences of an action. It is through the consequences of using an artefact that we perceive the built-in value. It is through the consequences of using a technology that we perceive its impact on the distribution of roles and responsibilities and the assumptions leading up to such a distribution. In the same vein, one might suspect that care values too are made real through the actions of actors in a network: that care becomes care through the actions of the nurse, without or without the assistance of a technology. Thus, the element that links all of the aforementioned concepts is that of actions (the actions and interactions between human and non-human actors) and the consequences of an action (in terms of promoting a value or disvalue). The next step towards understanding if, and how, a robot may fulfil these tasks in terms of the promotion of care values is to understand care values in their totality: what they are, how they are understood and interpreted in context and, how they are manifest through care practices.

Chapter 3

Understanding Care in Context

The human act of caring is the recognition of the intrinsic value of each person and the response to that value. [*Schoenhofer*, 2001]

3.1 Introduction

The task for which a robot will be used is an integral component for designing and programming the robot. It follows then that understanding and describing precisely the task for which the robot is to be designed is of crucial significance [*Asaro*, 2006; *Engelberger*, 1989]. In the late 1960's, roboticist Joseph Engelberger proposed the idea of using robots in multiple domains outside of the factory [1989]. The way Engelberger's book proceeded, and the manner in which he suggested all robotics applications ought to proceed, was concerned with outlining the capabilities of robots needed in order to fulfil the range of tasks to be delegated to the robot. For example, if a roboticist envisioned a robot gas attendant then according to Engelberger the next phase is to outline what is required to fulfil this task: manual dexterity for handling the gas pump and planar locomotion to travel from the car to the pump. With this in mind, the roboticist then continues to design/program the robot with the capabilities necessary to meet this mechanistic description of the task. In this way, roboticists could clearly envision the capabilities required to fulfil certain tasks and accordingly what the design of the robot must entail in order to make this a

reality. Engelberger's view presupposes that tasks can be understood as, broken down into, a linear commodified process.

Other tasks, like "caring for the elderly", are not as straightforward as Engelberger may have thought. Care for the elderly could mean anything from having coffee one-on-one, helping an individual put on their shoes, bathing, lifting or feeding (eating assistance). What's more, the idea of *understanding* care tasks precisely, involves more than a description in mechanistic terms (the exact actions the robot will have to perform) as Engelberger suggested. Understanding the tasks for which a care robot is to be designed must also acknowledge the relationship such actions have within the holistic picture of care. From the care ethics perspective, this holistic vision of care is of crucial importance when analysing a care institution and its ability to provide good care. Conceptualizing care as a process helps to overcome the threat of viewing care as a commodity or as a series of unlinked actions done to meet standardized needs. These two threats are, for care ethicist Joan Tronto, among the seven warning signs of bad institutional care and thus ought to be avoided at all costs [*Tronto*, 2010]. An additional threat to good institutional care is the taking of needs for granted; how to determine needs and who is responsible for determining needs. Thus, in order to provide good care there are certain elements that must be valued and in so doing the values are made real. The question then becomes whether or not it is possible for the design of a care robot to assist in the 'making real' of said care values.

The following chapter begins by exploring care as a value on its own, with insights from the care ethics tradition. Most notably, I take the contributions of Joan Tronto in terms of conceptualizing care as a practice with phases and corresponding moral elements. Added to this, I use Joan Tronto's analysis of good institutional care in order to conceptualize the values steering the context. Using a top-down approach I explore the realm of institutional care and the values of ethical importance within. This is done by examining and relating the abstract values of the WHO with the institutional values found in hospital policies and guidelines. The chapter proceeds with an exploration of care practices to elaborate on the meaning of a care practice. This is done by presenting an example of a care task, outlining the range of actors (human and non-human) involved and how they interact, to show how this task is viewed as a practice rather than as a 'task'. This is so for multiple reasons; given their place in the institutional setting against a background of values, norms and assumptions; given their role in meeting the multi-layered needs of patients and nurses; given the range of abilities the nurse must embody; and, given the entanglement of meeting social and physical care needs through one practice. Understanding

the meaning attributed to, and the complexity of, care practices may be the most significant finding for roboticists to understand the gravity of their design choices.

3.2 Unpacking the Concept of Care

Care may be one of the most difficult concepts to articulate. This is in part due to the ubiquity of the word but is also largely a consequence of the fact that one is assumed to know what care means given its revered place in many cultures. The work of Warren T. Reich nicely outlines the broad range of meanings and connotations care has embodied going back as early as Ancient Greece [*Reich*, 1995]: care considered an essential ingredient for the preservation of humanity, "the key to the process of becoming truly human"; care consisting of helping acts directed toward healing; care referring to the bearing of responsibility or a burden, "worrisome or anxious care"; and care as a subjective experience, "the capacity to feel that something matters" [*Reich*, 1995]. Regardless, of how one perceives or defines care, care is still valued as something above and beyond simple care-giving tasks. It has a central role in the history of human kind. It is linked with concepts of good and theories of the good life. Recognizing the needs of another and acting on those needs is what we may call care in its most rudimentary sense, and it is this series of events that bestows a valuation on the care-receiver. Thus, care is a value as it signifies the value of others. It is all of the mentioned aspects or components of care that are valued on an individual level as well as a societal level in and of themselves or as a means to something else.

Care is also conceptualized as a verb. Taking a closer look at the characterization of the verb "to care" one finds that caring may actually be divided into the idea of *caring about* and *caring for*. The dimension of *caring about* in the medical field implies a mental capacity or a subjective state of concern. On the other hand, *caring for* implies an activity for safeguarding the interests of the patient. In other words, it is a distinction between an attitude, feeling or state of mind vs. the exercise of a skill with or without a particular attitude or feeling toward the object upon which this skill is exercised . *Caring for* is unique in that it (often) requires the physical presence of the one exercising the skill for the benefit of the patient. *Caring about* does not place this requirement on the individuals involved in the scenario. "

The idea of caring as a reciprocal activity was introduced as early as 1937 by Martin Buber; "the embodiment of a commitment to protect, enhance, and

preserve human dignity" [*Buber*, 1958]. Care is further described as "the humanity in one being reflected in the other" [*Buber*, 1958]. The prevailing idea of these two descriptions is that of preserving human dignity. Following this idea was the adaptation by Simone Roach that caring is "the human mode of being" [*Roach*, 1999]. More specifically, "the human act of caring is the recognition of the intrinsic value of each person and the response to that value" [*Schoenhofer*, 2001]. As such, "care is not merely warm feelings or positive regard but is the outward expression or communication of those feelings" [*Schoenhofer*, 2001] and presupposes the recognition of the inherent worth of another. Roach's version also pays tribute to the preservation of human dignity through caring, but incorporates the idea that the ability to care is inherent to humans and in broad terms is characteristic of being. Strengthening the significance of care as an act, Leininger [1988] adds that the act of caring is "essential for human survival, growth and development".

As Le Dantec has rightly pointed out it is important to understand values in a given context [2009]. In the healthcare context, care is directly associated with the good of the patient – satisfying the good of the patient is considered a corollary of good care. The good of the patient, in the healthcare context, is teleologically structured with respect to the patient's health as the end. Although the patient may enter the physician-patient relationship with an exclusive clinical need, there are numerous needs which must be met within a healthcare setting. These needs are identified as goods of the patient according to Pellegrino and may be articulated as four components:

1. Clinical or biomedical good
2. The good as perceived by the patient
3. The good of the patient as a human person
4. The Good, or ultimate good [*Pellegrino*, 1985]

In a healthcare setting, each of these components must be served; however, their hierarchical order differs depending on the situation, i.e., for example the disease to be treated, its severity, and whether or not a cure exists. Clinical or biomedical good of the patient is straight forward in that it refers to the physiological effects on the patient. If a disease is the issue here, then the alleviation of the disease or a minimizing of symptoms would meet the requirement of biomedical good. Easing one's suffering is also most often attributed to the good as perceived by the patient.

While the first two conceptualizations of the good of the patient are clear the third and fourth may require further discussion. The good of the patient as a human person refers to the ability of the patient to continue with his/her life plan but could also be interpreted as the good of the patient as a human being capable of partaking in human activities like love, companionship, relationships, etc. Here, the initiative to use social robots [*Breazeal*, 2004] is conceived of as a means of fulfilling this good. Therefore, the good the patient seeks when entering a hospital is to be restored to a prior, or better, state of functioning. If this is not possible, the patient will then wish to be assisted in coping with the pain, disability or dying which the illness may cause . In this way, the patient seeks a good as a human person who can continue on with their life in whatever way possible. The Good also described as the ultimate good, refers to a spiritual good of the patient. If the patient has certain beliefs regarding their spiritual destiny, then the good of the patient is to have such beliefs supported.

In 1982, the work of Carol Gilligan brought the significance of care in ethical decision making to the fore. As a result, care has been given special attention and value in the forum of ethical deliberation and decision making – nurses, bioethicists and care ethicists have become the stewards of care in the healthcare domain [*Vanlaere and Gastmans*, 2011; *Pellegrino*, 1985; *Gilligan*, 1982; *Wilson*, 2002]. In fact, in many healthcare settings, reflecting on the process of care, or appealing to a notion of care, is the norm for ethical decision making. When introducing a new technology into healthcare, claims are often made about improving the level or quality of care. Accordingly, when evaluating the introduction of a new technology in healthcare, it is often done in terms of the impact on the provision of care [*Wilson*, 2002; *van Wynsberghe and Gastmans*, 2008; *van Wynsberghe A and C*, 2009]. From this recognition of care and its lack of acknowledgement in traditional ethical theory, the care ethics tradition evolved. Care ethics is thought of, not as a pre-packaged ethical theory ready to use in a given context, but as a perspective or stance from which one can theorize ethically; a lens from which one may begin moral deliberation in the care of others [*Verkerk*, 2001; *Little*, 1998]. This lens from which one begins to theorize is coupled with the direction in which such theorizing ought to take place. For many care ethicists, the care ethics stance demands a recognition of the relational status of persons along with a focus on responsibilities rather than rights [*Tronto*, 1993, 2010; *Little*, 1998; *Noddings*, 1984; *Buber*, 1958]. This focus on responsibilities over rights does not exclude a discussion of care values, rather care values are understood as being responsibilities of the care institution as well as the care provider.

3.3 The Values in Care

Alternative to the idea that care in itself is a value, linked with the good life and with a valuation of another, is the idea that beneath the umbrella concept of care exist many other values. These values are given importance for their role in care – their role in giving significance to care, in making it what it is. It is through the manifestation of these values that one comes to understand what care really is in practice. It is therefore fruitful for the topic of embedding values, and their relationship with consequences or outcomes, to understand these values and their link with outcomes in a health care context. Thus, to begin from a top-down approach, I look to the values articulated by the governing body of healthcare, namely the World Health Organization (WHO). The WHO framework for people-centered health[1] narrows in on the values in healthcare stemming from the patient's perspective; **patient safety, patient satisfaction, responsiveness to care, human dignity, physical well-being** and **psychological well-being**. This is not to say that other values like innovation or physician autonomy are not valued but rather from the patient's perspective, the listed values are the ones with the greatest ethical importance and will thus be used in my evaluation of implementing robots in the care of persons.

The above values are meant to structure and guide the overall tradition of healthcare. But such values must also be specified when understood in a more specified context – at the institutional level. In support of the values identified through the WHO, the guidelines and mission statements of hospitals and nursing homes in London Ontario[2] include additional values like: **compassion, integrity, dedication, respect** and **accountability**. According to the mission of St. Joe's hospital in London, Ontario, "values are the means by which [we] work together and serve others. [We] do [our] best to demonstrate these values in all actions and decisions, no matter how large or how small. Performance is measured according to; **respect, excellence** and **compassion**"[3]. Here, again, respect and compassion are primary values in the care of patients but

[1] WHO (2010) Health topics: Ageing. Available from: http://www.who.int/topics/ageing/en/

[2] Guidelines and Mission Statements were selected based on the hospitals and nursing home were field work was completed. All hospitals were guided by "The Ethical Framework for Nurses in Ontario" and the "Standard for the Nurse-Client Relationship" developed by the College of Nurses of Ontario.

[3] Information obtained by the "Mission, Vision and Values" document of St. Joseph's Health Care Institute, guided by the Health Ethics Guide of the catholic Health Alliance of Canada.

the hospital guideline has also included the value of excellence. While this may be considered a manifestation of the value of safety, it reinforces the idea that care given in a skilled manner by experts is a valued component in the provision of care. In short, the guidelines for a hospital ward or a nursing home aim to specify the more abstract values presented by the WHO; however, there is still room for interpretation depending again on the healthcare context within which one is working and the demographic one is working with.

As we can see, the abstract values articulated by the WHO are used as a reference for creating guidelines used in individual hospitals and/or nursing homes. For Tronto, good institutional care is has three main foci: "first, a clear account of power in the care relationship and thus a recognition of the need for a politics of care at every level; second, a way for care to remain particularistic and pluralistic; and third, that care should have clear, defined, acceptable purposes" [*Tronto*, 2010, p.162]. Thus, for Tronto, good care begins with hospital policies and guidelines like the ones just discussed. Ones that are agreed upon and not thought of as a hindrance to workers in the institution [*Tronto*, 2010, p.165].

To further specify the values expressed in hospital policies and guidelines, the College of Nurses (in Ontario) (CNO) has created an ethical framework . The framework "describes the ethical values that are most important to the nursing profession (in Ontario)" [*Nurses of Ontario*, 1999a, p. 3]. In support of these values, the CNO points out that the values listed are "shared by society and upheld by law". Values are not listed in any kind of priority; however, the CNO recognizes that client well being and client choice are primary values. As such, the value or insistence on particularistic care is embedded within the guidelines of the nurse's role/profession. The values of ethical importance for nurses, listed in the framework, are as follows: **client well-being, client choice, privacy and confidentiality, respect for life, maintaining commitments** (to clients, oneself, nursing colleagues, nursing profession, health team and quality practice settings), **truthfulness** and **fairness**. These are the values that nurses strive to uphold in their daily work; they are exhibited through the actions of nurses as well as the attitudes nurses embody. Another significant aspect of the framework is the recognition that nurses need to be aware of their own personal values, through articulation and reflection, to ensure that they are not acting to promote their own values but rather the values of their clients.

But listing these values as guide posts for nurses still requires that they be defined and furthermore how they too relate to the values listed by the governing body of the WHO (**see Table 3.1 for a list of values in nursing and their definitions**). In observing the definitions of nursing values we may link these contextualized values with those provided by the WHO guideline. For example,

client well-being is intrinsically linked with physical well-being and psychological well-being. Client well-being assumes that patient safety is adhered to for all activities in meeting physical and psychological well-being. What's more, client well-being presumes that the nurse is attentive to and cognizant of patient preferences in terms of the 'good' of a particular patient. Client choice, when seen in relationship to the WHO values, may be seen as a manifestation of the value of patient satisfaction – a client will be satisfied with their care when they are free to make choices about their care. Privacy and confidentiality may be considered a manifestation of the value of human dignity as well as a respect for persons but are also linked with client satisfaction. Respect for life also includes considerations of the quality of life and with this brings considerations of the quality of care and understanding what good care is in terms of what is good for a particular individual at a particular time. Hence, a recognition of human dignity. The interpretation of trust has changed with the changes in medicine and technology. Whereas before doctors thought it best for the patient's well-being to withhold information pertaining to the severity of their condition, now doctors and nurses believe that full disclosure is best for maintaining trust in the relationship.

Table 3.1: Values According to the Ethical Framework for Nurses (in Ontario)

Value	**Definition**
Well-Being	Facilitating someones good or welfare and preventing or removing harm.
Client Choice	Client choice means self-determination and includes the right to the information necessary to make choices and to consent or to refuse care.
Privacy	Limited access to a person, the persons body, conversations, bodily functions, or objects immediately associated with the person.
Confidentiality	Involves keeping personal information private. All information relating to the physical, psychological, and social health of clients is confidential; as is any information collected through nursing services. Clients have the right to confidentiality, and nurses make an implicit promise to maintain confidentiality.

Respect for Life	Human life is precious and needs to be respected, protected, and treated with consideration. Respect for life also includes considerations of the quality of life.
Truthfulness	Speaking or acting without intending to deceive. Truthfulness also refers to providing enough information to ensure the client is informed. Omissions are as untruthful as false information.
Fairness	Allocating health care resources on the basis of objective health related factors.

3.3.1 Care Values and the Therapeutic Relationship

I pause here to take note of and elaborate on the therapeutic relationship and its significance. The care ethics tradition builds on a recognition of the relational human as opposed to the human in the autonomous, atomistic, sense. Care is fundamentally about relations between people and things [*Tronto*, 1993, 2010; *Koggel*, 1998; *Nedelsky*, 2008; *Noddings*, 1984]. Although the foci of a good care institution, according to Tronto, "grow out of an understanding that takes care as a relational practice"[*Tronto*, 2010, p. 159], one foci in particular brings attention to the asymmetry of power in the relationship between care workers and clients. So far, all of the values central to the nursing practice are observable within the relationship between the nurse and client. Adding weight to the relationship formed between the nurse and client, nursing is actually defined in terms of the relationship as; "the therapeutic relationship which enables the client to attain, maintain, or regain optimal function by promoting the client's health through assessing, providing care for, and treating the client's health conditions. This is achieved by supportive, preventative, therapeutic, palliative, and rehabilitative means. The relationship with an individual client may be a direct practical role or it may be indirect, by means of management education, and research roles" [*Nurses of Ontario*, 1999b, p. 3].

When we examine closely the difference in power and how it must be recognized at the institutional level, it follows too that the framework would specify the boundaries and conditions of this unique relationship; "the therapeutic relationship is established and maintained by the nurse and the client, through the use of professional nursing knowledge, skill and caring attitudes and behaviours in order to provide nursing services that contribute to the client's health and well being. The relationship is based on trust, respect and intimacy and

requires the appropriate use of the power inherent in the care provider's role" [*Nurses of Ontario*, 1999b, p. 4]. Accordingly, care is expressed through and within the relationship between the nurse and client and therefore that which is valued in care is also expressed through and within the relationship between nurse and client. Thus, the relationship is valued on its own but also as a way of manifesting many of the other values central to health care and to care in general.

Not only is the relationship the place where values are expressed and promoted but there exist certain components of the relationship which together define it as a *therapeutic* one. Thus, the relationship between care-givers and care-receivers in a care institution is distinguished from the relationship formed between care-givers and care-receivers in a home setting, classroom, or between friends: "the therapeutic relationship differs from a social relationship or friendship in that the needs of the client always come first" (page 8). Labelling the relationship in this way serves to protect the rights of both the nurse and client involved by demarcating specific roles and responsibilities assigned to each. The components that facilitate such a recognition are: power, trust, respect and intimacy[4]. These components may not all be considered values so to speak, but are *valued* given that their expression symbolizes a manifestation of another value. For example, power is not a value in the same way as safety or client choice, but sensitivity to the unequal relationship between the nurse and client and (i.e., the power of the nurse) is valued; "the appropriate use of power in a caring manner enables the nurse to work with the client towards the client's goals, and to ensure that the client's vulnerable position in the nurse-client relationship is not taken advantage of" [*Nurses of Ontario*, 1999b, p. 5]. Recognition of the power the nurse holds is linked with the component of trust in the relationship; "the client expects the nurse to possess knowledge and skill and to demonstrate caring attitudes and behaviours, and so entrusts his or her care to the nurse. Trust is critical, as the client is in a vulnerable position in the relationship. Because the initial trust in the relationship is quite fragile, it is important that the nurse keep promises to clients. "If trust is breached, then it becomes very difficult to re-establish" [*Nurses of Ontario*, 1999b, p. 5].

Again, care is referred to in the sense that it encompasses behaviours, attitudes and skills, that they are valued in the relationship but are also necessary elements in the promotion of good care. The element of respect has appeared in a more abstract sense but here is presented in a more specific milieu; "part

[4]Taken from the Standard for the Therapeutic Nurse-Client Relationship, College of Nurses in Ontario. ISBN 0-921127-66-9

of respect is being non-judgemental of the client, and seeking to discover the meaning behind certain behaviours of the client" [*Nurses of Ontario*, 1999b, p. 5]. Intimacy is a component of the therapeutic relationship as well but is not interpreted in a sexual sense but rather refers to "the kinds of activities the nurses perform for and with the client which create personal and private closeness on many levels. This can involve physical, psychological, spiritual and social elements" [*Nurses of Ontario*, 1999b, p. 5].

Of significant importance is that the nurse must be sensitive to the context in which care is provided: "the client's home may feel like an informal environment in which to provide care, making the boundary between professional and social relationships less clear". The nurse may be tempted to fulfil tasks outside the boundaries of their role as nurse engaged in a therapeutic relationship; however, it is the responsibility of the nurse to make their role and the boundaries of this role clear.

3.3.2 The Complexity of Care Values in an Institution

Deconstructing or defining the abstract values presented by the WHO in this top-down manner, was intended to show how values are further described once the setting has been articulated. What's more, not only are there values structuring the institution, there are also values structuring the vocations of the multiple disciplines within one institution. Alongside this insight, we must pay tribute to the range and variety of care values but also to the interconnectedness of values. We observed that the WHO value of patient satisfaction and human dignity, when specified at the institutional level were expressed through attention to privacy (limited access to a person's body), confidentiality (keeping personal information private), respect for life (respect and protection of human life), and truthfulness (speaking or acting without intending to deceive) and so on. Other values like patient safety may be understood in terms of competence or excellence of care providers within the institution.

Interpretation of the values should take into consideration the institutional setting but must also take into consideration the demographic for whom care is being delivered to, the therapeutic context within which care is being delivered (meaning if the institution is the hospital, the therapeutic context refers to the specific hospital ward), and by whom the care is being delivered. "Indeed, thinking about the organization's purpose quickly requires us to notice the complexity of care, and that of all those people involved in the organization of care" [*Tronto*, 2010, p. 162]. For example, if we take the value of human dignity as an example we can see how its interpretation is also dependent on

demographic, context and care provider. In the context of paediatric oncology (hospital), the demographic being children, human dignity is manifest through safety precautions like the avoidance of cross contamination. This is the responsibility of the nurses and physicians directly involved in patient care but is also a responsibility for cleaning and other support staff of the ward. Alternatively, in complex care – another context in the same hospital – the demographic is elderly persons with complex issues pertaining to their care (meaning, most are immobile, some are blind and deaf, some are waiting for places in palliative care or nursing homes given that there is no "cure" for their condition), the value of human dignity is manifest through the compassionate completion of care tasks (like bathing and feeding). Safety is also valued and is a priority but safety pertains to the tasks nurses do in this ward which are different from the tasks nurses do in the paediatric oncology ward. Distinguishing between hospital wards and/or demographics and observing how the values are further deconstructed brings us to the concept of a care practice.

3.4 When Care Tasks Become Practices

3.4.1 A Care Task

A predominant threat in institutional care is the conception of care as an isolated action or task without recognition of its place within the process of care or its relationship to other actions/tasks. "Care institutions have to think about the nature of the caring process as a whole in order to guide their actions" [*Tronto*, 2010, p. 162]. Thus, the concept of a care task for which the robot is to be designed must be abandoned and in its place I introduce the concept of a care practice. To exemplify the difference between a care task and care practice I take the example of bathing, carried out in the hospital or nursing home . I have specifically chosen to look at this practice given its recognition as being a moment in care much more significant than the mere bathing of the corporeal dimension. I use this practice and the description thereof to set the stage for further analyses of practices using the Care-Centered framework in future chapters. What's more, it is also a practice for which care robots are currently being designed (Sanyo bathtub), adding weight to the understanding of the practice of bathing in its totality.

#1 Bathing a patient in their hospital room

The human actors are the nurse and the patient. The non-human actors are the mechanical bed, the sink in the room, the cleaning supplies, the curtain to

enclose the patient, the window in the corner of the room, and the door (the list is exhaustive but we understand the range of non-human actors). The practice begins with the nurse entering the patient's room, indicating what she will be doing (cleaning the patient), asking if that is ok and beginning to organize the cleaning supplies in close proximity to the patient. When things are set-up, the nurse encloses the curtain around the patient and begins to undress the patient. Using a wash cloth, the nurse cleans all orifices of the patient's body asking if the temperature of the water is fine and other personal items like how the patient is feeling. The nurse is also checking the patient's skin colour and temperature as indicators of well-being.

#2 Bathing a patient in a "tub room" (hospital or nursing home)

The human actors are the nurse and the patient. The non-human actors are the wheelchair to deliver the patient, the hospital tub room, the large tub for bathing, and the curtain around the tub. The practice begins when the nurse retrieves the patient from their room and escorts them to the tub room. The tub room is a special room in the hospital or nursing home where bathing takes place in a large bathtub of sorts. The tub is equipped with railings on the side to support the patient entering. The nurse lifts or helps to lift the patient from their chair and into the tub. If this is too much effort for the nurse alone they may call for assistance or may use the mechanical life (if available in the room). Once the patient is in the tub, the nurse closes the curtain and proceeds to clean the patient, again asking if the water temperature is satisfactory. In the nursing home, the nurse will be on a time schedule for bathing a long list of patients, there may not be ample time for discussions and instead the nurse focuses on cleaning and assessing the patient's skin colour, temperature and whether the patient has any abnormal markings.

3.5 The Meaning of a Care Practice

Within the field of care ethics, the idea of conceptualizing care in terms of a practice is not new:

> *The notion of a care practice is complex; it is an alternative to conceiving of care as a principle or as an emotion. To call care a practice implies that it involves both thought and action, that thought and action are interrelated, and that they are directed toward some end. The activity, and its end, set the boundaries as to what appears reasonable within the framework of the practice* [*Tronto*, 1993, p. 108].

In a similar manner, I refer to *care practices* to indicate the complexity of actors, needs and values involved. I define a care practice as the attitudes, actions and interactions between actors (human and non-human) in a care context that work together in a way that manifests care values. Thus, a care practice is defined by the interactions between actors but also **how** these interactions take place; the practices are values working together. Aligning with the care value discussion above one might consider that the value of safety is the result of interactions between actors whereas the value of compassion is a result of the manner in which a care –giver speaks or touches a patient. There are human actors, the nurse and patient predominantly, but at times there may be a host of other staff contributing to the practice. There are also non-human actors in the network; the hospital or nursing home room, the TV, the mechanical bed, the mechanical lift, the telephone, the respirator, the window, the door, and so on and so forth. Each of the care values are realized through the interactions of all the elements in the practice, the humans and the material environment. Some elements will embody a value on their own while others require an interaction with another actor to manifest a value.

For instance, in example #1, all actors, human and non-human, interacted and were important for decision making throughout the task. Reflecting on this example, one may presume that the curtain enclosed around the patient is a manifestation of the value of privacy. But, the door to the patient's room was open and the window was not closed. Thus, partial privacy is achieved, but not full. If partial privacy is bestowed, than trust may be affected which shows again the intertwining of values. To compensate for this potential, we may look to the manner in which bathing takes place. This is a manifestation of the range of human attributes on the part of the nurse which are valued, but it also facilitates a building of trust between the patient and the nurse. Consequently, whatever feelings of trust that may have been inhibited through the door or window being open, were renewed through the disposition and manner in which the practice was carried out. The questions asked by the nurse as to whether or not it was ok for her to bathe the patient reflect a valuing of the patient's choice and satisfaction (WHO values). It is also the nurse acting with a conscious respect for life and being truthful (Nurse's framework, College of Nurses). Hence, the values of the hospital/nursing home guide the actions of the nurse. The way in which care was provided along with paying attention to details like closing the curtain are all reflections of the moral responsibility on the part of the nurse. By complying with the sanctioned and agreed upon institutional values, the nurse behaves as a morally responsible individual. Of equal importance, the choices made by the nurse (to enclose the curtain but not close the door or window)

were made in association with the non-human actors in the practice. Therefore, mediation already occurs in current care practices in the hospital, nursing home and homes of care-receivers – the choices made by nurses and care-receivers are a hybrid affair between human and non-human actors in the context. Thus, not only are values the result of interactions between human and non-human actors but so too are the choices made by the human actors.

For nurses, the practice of bathing is highlighted as a moment in which the nurse gains additional information as to the patient's medical status [*Pols*, 2004]. The nurse assesses the patient's physiological status, neurological status, temperature of the skin, and overall sense of the patient's mood and/or recovery. The nurse learns of the patient's preferences for types of soaps or temperature of water. What's more, it is an opportunity for the nurse and patient to engage in social conversation and essentially to build the relationship between the two thereby establishing and/or maintaining trust through the nurse's commitment to privacy, confidentiality and compassion. Building a trusting relationship is important not only for the next time the nurse needs to bathe the patient but also for additional practices the nurse will be engaged in [*Nurses of Ontario*, 1999a]. Later on in the care of the same patient the nurse requires that the patient be honest about their symptoms, comply with their care plan and take their medications. Without trust between nurse and patient, the patient is not as likely to meet these needs of the nurse. Of course I see the irony in referring to these as being needs of the nurse when they are really needs of the patient in order to provide good care for that patient; however, they are needs that the nurse has in order to fulfil his/her role and responsibility in the care process.

In addition to the values already discussed there are other values of equal importance that have not yet been listed. These are values that become apparent when the care practice is described as it occurs in context. Values like eye contact when the nurse is talking to the patient. Values like human presence when the nurse enters the room for the duration of the practice. These are both values not listed in any of the guidelines presented here but are nonetheless essential in the provision of good care. Eye contact plays a role in a patient's assessment of the truthfulness of the nurse. Human presence plays a role in conveying the worth and dignity of the care-receiver. There is also the value of 'tinkering' coined by Annemarie Mol [2010]. Such a value refers to the nurse's ability to alter his/her behaviours and actions based on the changing needs of the patient and the nurse's perception of such changes. This demands a sense of attentiveness on the part of the nurse and presupposes the nurse is physically present to perceive such changes. The nurse's responsiveness will then take into consideration any change in the patient's physiological status

such as temperature of the skin, etc.

All of these additional aspects bestow meaning to the practice and illustrate how the practice of bathing is much more than the corporeal dimension of cleaning a patient's body. Jeanette Pols speaks to this in great detail and discusses the bathing of patients on psychiatric wards and how the practice reflects a conception of the patient as a citizen [*Pols*, 2004]. Pols identifies four different washing repertoires, the first repertoire is labelled, 'washing as part of individual privacy'. In this repertoire patients choose how often and when washing occurs and thus individual preferences and taste's are key; "by relating to a patient's history of washing, the activity can be tailored to what he or she is used to and prefers" [*Pols*, 2004, p. 50]. In the second, 'washing as a basic skill', the bathing of an individual is used as a moment in which they are empowered to keep their own basic motor skills in tact [*Pols*, 2004, p. 55]. The role of the care-giver in this instance is to encourage and motivate the patient to fulfil certain portions of bathing on their own and to assist in the parts of washing that the patient cannot do themselves. The third repertoire, 'washing as a precondition', outlines the practice of bathing as a step towards the self-actualization of the patient [*Pols*, 2004, p. 60]. The practice of bathing is used to develop the patient's autonomy; "autonomy in care for self-actualization is a matter of developing and choosing rather than doing things oneself (basic skills) or being free to decide on one's individual life (privacy)" [*Pols*, 2004, p. 61]. Repertoire four, 'washing as a relational activity', illustrates the practice of bathing as an intentional moment for establishing relations as the goal of the care. Thus, learning basic skills is not the primary aim but rather being connected with others is. As such, the nurse becomes a part of the social network of the patient. In describing this repertoire, using an example from her own fieldwork experience, Pols addresses the aspect of communication: "communication is not always easiest by 'talking'. On the contrary, [he perceives] verbal communication is often difficult. Washing creates an alternative situation for communication. There is a clear task at hand and 'the rest comes with it'" [*Pols*, 2004, p. 67].

Thus, the end goal of bathing may differ depending on the patient and their particularistic condition and history. In line with Tronto's thinking, the care institution must be flexible to allow for different manifestations of the same practice. As seen through Pols' work, when the end goal of the practice changes (developing basic skills vs. developing a relationship between care-receiver and care-provider), the prioritization of values changes along with it. Thus, not only is a practice like bathing recognized for its utility in meeting the corporeal needs of a patient but it is also granted a deeper meaning when we observe

the expression of values and the meaning the practice takes on as a result of such expressions. This understanding must be made explicit in order to begin to understand how a care robot might impact the provision of care when introduced into a practice like bathing.

Naturally the context in which the practice takes place plays a role in the transition from task to practice. As was just shown, all of the actions of the nurse are aimed at fulfilling the socially and legally sanctioned values of the tradition of healthcare found in the guidelines of the institution and the professional codes of conduct. Added to this, the ends that the practice serves, not only in terms of manifesting values but in the meeting of needs, adds a further complexity to the concept of a care practice and it is the element of 'needs' that I will explore next.

3.6 The Multi-Layered Needs of Patients

Care is thought of first and foremost as a response to the needs of another. For Tronto, this marks the starting point of care. Additionally, needs act as the thread linking all actors and elements of the care practice. As humans we are constantly in a state of need, the differences occur in the amount of need one individual requires and the manner in which these needs may be met; some individuals require more assistance than others for their whole life while others require additional assistance for only a short period of time. Thus, to begin, we must first understand needs in a cyclical and fluctuating way.

It follows then that the response to needs is also a process vs. a single act in time. Responding to needs in the process of care articulated by Tronto refers to the following: attentiveness to the needs of the other, taking responsibility for the needs of the other, embarking on an action to meet the needs of the other and engaging in reciprocity with the patient to assure their needs have been satisfied. Maintaining the needs of the other as the central focus of the care process reflects the concept of "ethical sensitivity", a skill of nurses essential for the provision of high quality care [*Weaver et al.*, 2008].

This last point, directs our attention to the idea that not only do nurses, and other healthcare workers, have needs in terms of resources but there of needs of the nurse fulfilling their role as nurse; skills or attributes of the nurse that are needed in order to provide good care. The remainder of this chapter is meant to outline the concept of needs within the healthcare tradition. This is not an easy task, however, determining needs is described as "one of the foremost political struggles of any account of care" [*Fraser*, 1989]. In fact, "any agency

or institution that presumes that needs are fixed is likely to be mistaken and to inflict harm in trying to meet such needs" [*Tronto*, 2010, p. 164]. Consequently, needs must first and foremost be recognized in all their complexity. Added to this is the idea that the complexity of needs comes not only from the way in which they ought to be met but rather from the recognition that needs are dependent on a unique individual at a given point in time.

3.6.1 Conceptualizing Needs

Although need is often spoken of in the economist's sense, in terms of *want*, need ought to be distinguished from want. Want brings with it a connotation of voluntary conduct; I desire something even though I can live without it. Need, however, brings with it a connotation of conditions or things that are necessary but lacking, or conditions requiring relief. Linked with this, is the idea of conditions or things being necessary or vital. Psychologist Abraham Maslow developed a hierarchy of needs that typified, categorized and prioritized needs. Maslow claimed that people progressed from one category of needs to another once their needs were satisfied [*Maslow*, 1970]. In this sense, needs are thought of as a continuum building from one category onto another.

Alternatively, the Chilean economist Manfred Max-Neef proposed a somewhat similar taxonomy to Maslow but instead claims that "all needs must be satisfied at the same time to make a person whole" [*Max-Neef*, 1995]. Thus, he refutes the continuum or hierarchy proposed by Maslow. This fluid or circular way of visualizing needs reflects the vision of Tronto as well as the 'personalist perspective' in care ethics [*Vanlaere and Gastmans*, 2011]. According to personalism, the thread linking all actors and elements in care is a recognition of the patient as a multi-dimensional person. This means, that any decision regarding the treatment of the patient must take into consideration the preferences of that patient recognized as a person with unique spiritual, cultural, educational, and personal factors and any measures taken to meet the needs of a patient must attempt to meet needs along all of these dimensions. As such, each patient is respected for their uniqueness and is treated with dignity. Thus, personalist ethicists conceptualize needs as a product of the multi-dimensional person with a unique history as well, as physical or mental condition, and therefore demand that needs be met through a recognition of the patient as a multi-dimensional person. The work of Annemarie Mol presupposes the personalist recognition of the multi-dimensional person and through her idea of "tinkering" she demands that a care-giver be in-tune with the *changing* needs of the patient [*Mol et al.*, 2010]. Thus, not only are needs specific to an individual but they change from

one moment (month, day, hour) to the next and it is the good care-giver's responsibility to account for this dynamic nature. Tronto adheres to both the personalist perspective as well as the concept of tinkering and purports that needs command the greatest attention given their role in structuring the process of care at both the level of the institution and the level of the practice.

> *What we think of as 'needs' changes. They change over time for particular individuals, they change as techniques of medical intervention change, they change as societies expand their sense of what should be cared for, and they change as groups make new, expanded or diminished demands on the political order. The demands placed upon institutions change. Within institutions, as the particular individuals within the institution change, they have different needs.* [*Tronto*, 2010, p. 168]

Thus, needs in the healthcare tradition must be conceptualized in terms of the meeting of needs of individual patients, the needs of care-givers as well as the needs of the larger institution. For the latter, the institution needs support staff, managerial staff, equipment, insurance and so on and so forth. Needs according to the former conceptualization was discussed above, the needs of an individual patient must be worked out according to that individual person, their medical history and the status of their condition. And needs of the care-giver cannot go unnoticed for without which care could not be provided; the care provider has needs for the institution to meet, needs in terms of the patient they are working with and needs in terms of what is needed of them in order to provide good care (i.e., skill, competence, compassion, empathy). Each care provider cares for different patients, fulfills different roles and responsibilities and is themselves a unique individual.

What all the concepts have in common is the multi-dimensional character of needs based on the multi-dimensional nature of the person, their relationships, and the institution they're in rather than the objectified patient/care provider or the standardized institute respectively. Knowing that everyone has needs and that these needs will change depending on the context and situation is not a new phenomenon, but how are needs actually met? *Meeting* needs isn't quite as clear.

3.6.2 Meeting Needs

When needs are thought to be met through care actions alone they are conceptualized as a commodity; needs met through the fulfillment of tasks without

a sense of the overall process of care or a concern for the disposition of the care-giver. Care-receivers are conceived of as clients and the market model is instilled. But such a commodification of needs often leads to alienation; who will have access to the care for their needs [*Tronto*, 2010]? Often times marginalized demographics will not. To overcome such health equity issues Amartya Sen and Martha Nussbaum propose the capabilities approach [*Verkerk et al.*, 2001; *Nussbaum*, 2000; *Sen*, 1985]. That needs are conceived of in terms of capabilities; the capability for play, the capability for relations; the capability of a healthy life as examples. When such capabilities are met, said person is said to have a functioning. Needs are not defined in terms of majorities but take into consideration those marginalized demographics without the same baseline capabilities.

For Tronto, being in need is considered a fact of life, an inevitable existential condition. All species and technologies are in need at some point. Even baby animals need their parents to provide them with food at the beginning of their life; one might imagine a bird returning to the nest to feed their babies with worms. And technologies too have needs when they break or are not working. Tronto's conceptualization of the meeting of needs is (intentionally) general enough that it accounts for a diversity in actors, both human and non-human. What's more, Tronto's account of everyone being in need at (more than) one point in their life helps to discount the negative connotation of one in need. The view of the passive care-receiver, vulnerable to a care-giver, ought to be discarded in exchange for the recognition that being in need and being cared for is a product of the relational human. This is not to claim that there is no distinction in power between care-receivers and care-givers, as we saw in the code of conduct for nurses and the criteria of the therapeutic relationship. Such an asymmetry in power demands recognition in order to overcome the potential for paternalism, to allow for the care-receiver to have a voice.

For Tronto, needs are met through a four stage process of care; "*caring about* (recognizing a need for care), *caring for* (taking responsibility to meet that need), *care-giving* (the actual physical work of providing care) and *care-receiving* (the evaluation of how well the care provided had met the caring need)" [*Tronto*, 2010, p. 160, *Tronto*, 1993]. These phases are not in a linear order but are rather intertwined and happen concurrently in many instances. Although one could claim that needs are tangibly met when the care-giver engages in an activity to satisfy the needs of the care-receiver (care-giving), meeting needs cannot be reduced to this phase alone. This again refers to the seven warning signs of bad institutional care when we observe that each phase of the care process revolves around needs. Meeting needs presupposes that needs have been

identified correctly, that someone is taking responsibility for meeting these needs and that someone will assess whether the action taken has met the identified needs. It follows then that many actors (human and non-human) are involved in the meeting of needs, each delegated a certain amount and type of responsibility. In accordance, alongside the four phases of a care practice, Tronto identifies four corresponding moral elements. The moral elements act as the buttress of care, the manner in which the ethical nature of care may be evaluated. I claim that through Tronto's account of the moral elements she is ultimately presenting a normative criterion for the evaluation of care practices.

> *Good care requires that the four phases of the care process must fit together into a whole. Similarly, to act properly in accordance with an ethic of care requires that the four moral elements of care (attentiveness, responsibility, competence and responsiveness) be integrated into an appropriate whole. Care as a practice involves more than simply good intentions. It requires a deep and thoughtful knowledge of the situation and of all of the actors' situations, needs and competencies* [*Tronto*, 1993, p. 136].

The element of attentiveness refers to the care-giver's ability to perceive the often changing, unique needs of the patient/person. The care provider in this role need not be the one dispensing care, their role may be only to address the initial needs of a patient to determine which doctor or nurse should attend to them. While their role ends here, their responsibility may not. The nurse in charge of admitting the patient will not follow the progress of the patient but may be responsible to ensure that the patient is directed to the next specialist. This ties in with the element of responsibility. Responsibility as the second moral element means that an individual, or institution, be responsible for the needs of said patient. Thus in the above scenario, the nurse who passed-on the patient to the specialist delegated the responsibility of physiological care to the specialist. In terms of the institution, one might assume that when entering the institution of the hospital, nursing home, hospice, etc., the institution both symbolically and casually takes responsibility for the overall care of the patient. The delegation of responsibility does not necessarily fit the linear model described above; however, it provides a way of visualizing the chain of responsibility among health care professionals, the complex sharing of responsibility and the relationship roles and responsibilities share from one phases/element to another.

The element of competence is translated as skill. The care provider is required to fulfil their role and responsibility of care-taking in a skilful manner. If not, the responsibility assigned to the role requires that they be punished. Reci-

procity refers to the component in which the care-receiver is actively engaged in responding to their care provider and the care provided. This is not a one-time event but rather is continual throughout the multiple phases that may happen any number of times during a patient's time in a care institution. What's more, the phase of responsiveness or engaging in a reciprocal interaction may also be conceived of as a need of the care provider – without a response, positive or negative, from the care-receiver, the care-giver will not know when needs have been met or whether they have been met to the satisfaction of the care-receiver.

Returning once again to discussions rooted in the ethics of technology, I must acknowledge the relationship between needs, care and technology. Such a relationship is neither causal nor linear. One does not begin with needs isolated from technologies and then apply care with or without the use of technologies. Rather, needs are often the result of interactions between humans and existing technologies [*Verbeek*, 2011]. For example, the need for surgical intervention is only recognized as a need when we have the technology to provide such an intervention. Without such technologies, the need would be considered a want. It follows then that the introduction of endoscopic tools has introduced a need for "minimally invasive surgery" whereas prior to the technology, surgery was the only option and so surgery was where the need ended. This may be considered a type of hermeneutic mediation (i.e., the interpretation of what the need is has shifted with the introduction of endoscopic tools) and also ties in with norms.

It is also important to indicate that I do not believe all dimensions of needs ought to be met by human intervention alone. I see that as being neither possible nor desirable. It is not possible because the meeting of needs happens in the context of the hospital which employs a wide spectrum of technologies for its functioning from the mechanical bed to the TV. It is undesirable given that technologies should also be recognized for their benefits to care. For example, sterilization technologies allow for rooms to be kept uncontaminated. Or, technologies used to test the urine of paediatric oncology patients for the presence of chemotherapy chemicals. This provides a significant benefit to the patient and nurse and is not something that could be done without technology (granted the need for it also arises from another technology, chemotherapy and radiation treatment for cancer).

Thus, needs are met through a variety of actors (human and non-human), actions, attitudes, and roles that together express the values in care. Needs are met through care practices which may be broken down into the four phases of care each assigned a corresponding moral element used to evaluate its moral quality. Each of the moral elements must be attended to in striving for good care. The quality of a care practice will suffer if time is not allotted to the development

of attentiveness of the nurse towards a patient or if the necessary skills are not cultivated to ensure the competent completion of an action. Through such an analysis of care, the complexity of identifying purposes, roles and responsibilities is revealed. It is this complexity and intricacies of roles, responsibilities, manifestation of values and interactions between human actors and the material environment that must be made explicit and criticized prior to the design and introduction of a care robot for a given practice. Consequently, this understanding of needs orients the recognition of needs and the multi-dimensional nature of said needs, as a high ranking value. Without this starting point, care cannot proceed or might proceed in a dodgy or hazardous way. Paying tribute to the idea that needs are multi-layered, and thus care practices serve multiple ends, we end up in a discussion of the variety of skills and attributes the nurse ought to embody. The nurse is responsible for understanding needs in this dynamic, holistic sense but is also required to understand a range of communication forms (verbal and non-verbal) and to be able to "tinker" a care practice accordingly. What are also needs in a sense then are the abilities of the nurse (or other care-giver) to bestow care in a compassionate and empathetic manner. This refers to the capabilities of the nurse, as a need.

3.6.3 Abilities of the Nurse – Cultivating Care Skills

"When we think of care as a practice, with all of the necessary component pieces, then we must take into account the full context of caring. We cannot ignore the real needs of all of the parties; we must consider the concerns of the care-receiver as well as the skills of the care-giver, and the role of those who are taking care of" [*Tronto*, 2010, p. 118]. With an idea of what "needs" refers to, we cannot ignore that for good care, another need has to do with the capabilities of the care-giver. Care ethics and bioethics highlight the role of virtues and virtue ethics in the evaluation of care. The "good care-giver" is one that does so in a way that promotes their own moral development while at the same time fulfills the best interests of the care-receiver [*Vanlaere and Gastmans*, 2011]. Each of the moral elements listed above point to a core component of the moral significance of the care-giver – care-givers are (traditionally) moral agents who assume moral responsibility for the care of an individual. For Tronto, all of the moral elements essentially refer to capabilities or necessary skills of the care provider. Specifically, "the second dimension of care, *taking care of*, makes responsibility into a central moral category" [*Tronto*, 1993, p. 121]. As such, moral responsibility is an important attribute for the care-giver to posses. This is because of the types of decisions the care-giver will have to make (those that

carry with them moral consequences) but also because of the manner in which care ought to be provided (a caring disposition and a moral agent is required to comprehend the significance of this). Moral responsibility also assumes that the care-giver ought to exercise impartiality and justice when necessary and not rely on subjective emotions to guide their actions and decisions [*Vanlaere and Gastmans*, 2011].

Being in-tune with the ethical delicacy of the situation, and how to address it, what Tronto refers to as *attentiveness*, has also been referred to as "ethical sensitivity"; "that which enables professionals to recognize, interpret and respond appropriately to the concerns of those receiving professional services [*Weaver et al.*, 2008]. *Attentiveness* also refers to staying in-tune with the dynamic physical demands of the patient. This has also been referred to as "tinkering"; improving care in real-time by asking questions and responding to verbal and non-verbal cues given by the patient [*Mol et al.*, 2010]. The former adheres to the idea of care as *caring about* while the latter is closely linked with care as in *caring for*, albeit they are not mutually exclusive. *Caring for* is the dimension in which skill is valued above many other values and needs. If the caring action is not carried out in a skilled manner, it may do more harm than good to the patient. Consequently, there are skills of the nurse that are valued in care and necessary for the assuring a high quality of care. These may even be the skills the nurse has for operating machinery or tools.

There are other human attributes of the care-giver which are also necessary for promoting many of the values in care, namely the ability to empathize and portray compassion for the care-receiver. These are important emotions for the care-giver to portray; however, the care-giver must always be in a state of balancing the portrayal of emotions with an objective stance pertaining to the good care of the patient. The care-giver cannot allow emotions to interfere with what is best for the patient medically speaking. The care-giver must take a distance from the patient in this respect, while at the same time allowing the patient to understand that they take this stance from a disposition of care – of beneficence for the patient's best interests. The way in which this is done is most often through small tasks like small talk during meetings or assessments or through the way in which the care-giver touches the care-receiver. The latter refers to touching in a way that conveys compassion and empathy for the pain the patient may be feeling or speaking in a voice that conveys empathy for the patient's condition. This idea introduces the prospect that emotional or social needs of the patient can (and are) met through the provision of physical care tasks. This again, reinforces the idea that a task is no longer a task but is a practice aimed at meeting the needs of a patient across a broad spectrum.

It must be stated here that for some, the use of a care robot, or another technology for that matter, poses a threat along the lines of what it does to the care-giver in their role as carer as well as what the impact may be on the care recipient. "Critics fear, perhaps justifiably that care-givers might be less attuned to the specific needs of care recipients because a technological crutch is available" [*Borenstein and Pearson*, 2011, p. 260]. For Shannon Vallor, this aspect presents the main concern when discussing the prospect of robots in care [*Vallor*, 2011]. Vallor claims that the use of care robots prohibits the cultivation of the necessary care skills mentioned above. What's more, that removing the care-giver from their role prevents the cultivation of certain human capabilities like empathy and compassion, necessary for the flourishing of the care-giver as a person engaged in relationships of their own. But I have already outlined the relationship that care and technology already share – that technologies are currently and have been for quite some time, a dominant factor in the provision of care. Thus, what new challenges will care robots present for the provision of good are. Inevitably this leaves us asking the general question about the relationship between care and technology.

3.7 Care and Technology

The idea that technology introduced into care ought to be assessed for its impact on care presumes that care and technology may be at odds. On the one hand, there is a school of thought that believes care and technology to be mutually exclusive; that the two represent different spheres of meaning (also referred to as "gendered spheres") [*Wilson*, 2002]. One sphere conjuring images of subjectivity and emotions (care) and the other conjuring images of objectivity and standardization (technology) [*Wilson*, 2002]. On the other hand, there are also those who believe that the two are intertwined; that care in a hospital today requires technology, that technology allows for a greater level of care and that skill of technology in care represents a manifestation of care. Referring again to surgical robots, the benefits to patients are a concrete example of when a technology can significantly improve care if we consider care in the *caring for* sense (meaning a decrease in scarring, risk of infection, recovery time and pain). If, however, one were to think in terms of *caring about*, robotic surgery may be thought of as an extension of the tendency to objectify patients [*Gadow*, 2002]. The surgeon does not touch the patient, does not even stand over the patient's bed throughout the surgery, but rather sits at a console and performs the surgery in a distant way, much like a video game. The patient then, one

might assume, is represented not as an individual but as they are displayed on the screen, as a model of the human body.

In the nursing context, many technologies are thought to improve care by providing specialized information to the nurse. Take for example, the external foetal monitor used in gynaecology/obstetrics wards to monitor the heart rate of the fetus in utero. The monitor allows the nurse to receive information pertaining to the status of the fetus. This information can be collected even when the nurse is not in the patient's hospital room. In this way, the nurse has a constant flow of information about the status of the fetus recorded and the idea is that the nurse can make more accurate assessments. A drawback, however, to using this technology is that it takes the nurse's attention away from the patient and instead directs the nurse's attention toward the machine. If a technology distracts the nurses attention, the nurse may no longer be attentive to the particular needs of the patient. Consequently, while the machine may provide a way of increasing care in terms of *caring for* it poses a threat to care in terms of *caring about*. Moreover, one might suggest that the nurse is no longer capable of detecting foetal cues in the traditional way once they come to rely on technological methods. This is what is often referred to as de-skilling; when technology introduces a new way to collect and interpret information and one no longer relies on traditional methods. To conclude de-skilling as bad presumes that conventional methods are superior in a deterministic way without acknowledging the potential benefits of the new skills introduced – namely, that the nurse has more accurate information even when he/she must leave the bedside for a time. However, to rejoice in the superiority of the technology and its ability to provide different kinds of information gives technology a "saviour" status [*Borgmann*, 1987].

Another valued aspect to care differs from the conception of care as doing. There are times in a person's treatment when there is nothing to do and care is expressed just by *being there* for the patient – being physically present or conveying a sense of being there (perhaps through pre-preference). This also combines the dimensions of *caring about* and *caring for* but in a very different way. While in normal situations in the hospital *caring for* requires that a task is being done, in this instance *caring for* is achieved through one's presence alone. Meeting this need may be thought of as a manifestation of a caring attitude or *caring about*. And meeting this need may only be achieved through the use of technology if the one *caring about* is at a distance and does so through pre-preference (e.g., Skype). This brings us back to the idea that good care (valued care) is care that is tailored to the dynamic needs of individual persons. This is one of the most crucial points articulated in the care ethics literature; care is

only thought of as "good care" when it is personalized [*Vanlaere and Gastmans*, 2011; *Tronto*, 1993, 2010]. This means understanding the individual needs of the patient and choosing a care plan with the patient based on the preferences and values of the patient as a person with distinct and unique cultural, spiritual and personal values. The most problematic aspect to a technology then would be its prospect for interfering with a nurse understanding the particular needs of a patient.

Thus, care in healthcare, is valued as a means for meeting the physiological, physical, psychological and emotional needs of patients in an individualized manner through the exercise of skill, with a compassionate disposition. Given this description, technologies can be included as aids in the practice of care in a manner that supports this vision. A technology may provide greater insight into the physical problems of the patient and thus aid in understanding the patient's individual ailment. Or, a technology may present a way to target the emotional needs of a patient. One must understand what the values are in the network into which the technology is being introduced to uncover the impact a technology may impose.

3.8 Conclusion

The goal of this chapter was to outline the vast amount of values within the care ethics tradition and the complexity of understanding their relationship to one another, as well as how they are expressed in a care institution. The noteworthy contributions of care ethicist Joan Tronto become indispensable for understanding care as a practice with corresponding moral elements. Only by understanding values at the level of the care context and care practice, can we begin to speculate on the program requirements of a care robot. An additional goal of this chapter was to demarcate the significance of the care practice, both in terms of understanding the meaning of the care practice as the forum in which values are made real as well as understanding the relationship one practice shares with another and with the overall process of care. For these reasons I insist on referring to care practices as such rather than as mere tasks. The language of 'task' fails to pay tribute to the meaning of the practice. This meaning that I speak of is brought to the fore with the recognition of the significance of the practice as well as a recognition of the practice's role in meeting needs; both the needs of care-receivers as well as care-givers.

By understanding that many actions are done in a way that preserves the dignity of the care-receiver, the practice is seen as a means for the valuation of

the care-receiver which adds an ineffable dimension to the carrying-out of tasks. Added to this is the recognition that care practices serve more than one purpose or need. The simple act of asking 'how one is doing today' means a great deal to the patient on a personal level; it reflects a valuation of the patient through their eyes. The act of bathing a patient with gentleness and compassion while asking personal questions and staying true to the preferences of the patient reflects again the valuation of the patient's dignity, integrity and respect for life. The act of telling jokes while one is serving a meal to elderly patients in a nursing home adds a light-hearted social dimension to a resident's day adding to patient satisfaction. Asking residents of the nursing home 'what they would prefer to do today' grants them choice and preference in their care – a manifestation of their worth once again. These practices add meaning to the care practice above and beyond the expression of institutional values, they emphasize a valuation of the client as a person rather than as merely a patient. In this way, we observe the link that care shares with the provision of well-being, of a good life. The values in care mirror the values which serve as the foundation for theories of the good life. From a care perspective, the good life is one in which persons are engaged in relationships and feel connections with others. In other words, care practices make real/tangible the values in care and thus take on a deep meaning.

One might wonder then about the meaning of the care robot once integrated into a care practice and suggest that its meaning arises within a care practice from the interactions with nurses and patients, in the value-laden context of a hospital or nursing home. Given the intertwining of actions between actors and the expression of values within this network, the meaning of the care robot has to do with its embedded assumptions and norms as well as its ability to promote the realization of values when introduced into the network of actors interacting in a care practice. The questions we are left asking now are multiple; will care practices still be considered such if the robot providing care does not possess the same attributes as a human care-giver, does not fulfil social care in tandem with physical care, does not address the multi-dimensional conception of needs and does not project empathy and compassion throughout the care practice? Alternatively, we may find that a care robot has the potential to re-introduce values into a care practice that were lost at some point.

But 'meaning' requires further analysis. The questions to address now have to do with how a robot will shift the expression of these values, how the robot does this through its interactions with the other actors in the network, how the robot alters the conventional distribution of roles and responsibilities and how the robot can embody the values identified here: being attentive to, responding to and taking responsibility for the multi-dimensional needs of the patient as a

unique person; fulfilling practices as a morally responsible agent; fulfilling practices while conveying human attributes like compassion and empathy thereby meeting the physical and social needs of the patient in tandem; and fulfilling actions with skill (in partnership with non-human actors). The Care-Centered (CC) framework and methodologies for use are intended to assess all of these questions/dimensions. Before engaging in evaluations using the CC framework, however, I now direct our attention to understanding the technology in question: robots.

Chapter 4

Care Robots and Robot Capabilities

Welcome to the Robot Revolution. [*Lin et al.*, 2011, p. 3]

4.1 Introduction

KNOWING what values I aim to discuss and how they come to be embedded in a system, the goal now is to understand the technology in question. The current popular discourse on robots creates ambiguity surrounding different classes and/or features of a robot. The confusion surrounding robots begins with its definition. In the *Springer Handbook of Robotics* [2008], it is acknowledged that there is no consensus on what a robot is today; however, divergent definitions are given depending on the class of robots discussed or their application. Because of the wide variety of robots currently available, some authors discuss particular prototypes while others leave out a definition of robots altogether and instead focus on robotics – the study of robot foundations and methods, or the science and technology of robots.

In terms of defining care robots, there is not one capability, appearance or function that is exclusive to a care robot. It may have any number or combination of capabilities and appearances. For this reason, I use the concept of interpretive flexibility [*Howcroft et al.*, 2004] to discuss their classification. According to interpretive flexibility, a care robot is classified as such based on: its context of use, the function for which it is used and the user. Thus, one robot

may be referred to as a care robot when used by a nurse in a hospital setting to lift patients but the same robot may also be classified as an industrial robot when used in an industrial setting by factory workers for lifting heavy objects (the robot I am referring to here is an exoskeleton). Care robots may be used by the care-giver and/or the care-receiver directly, again dependent on the task and context of use. They may be used for physical care tasks, activities of daily living (like lifting, bathing, feeding) or to provide companionship. Accordingly, the capabilities of the robot are dependent on this distinction as well.

The goal of this chapter is to introduce the reader to the current state-of-the-art in robotics research. By presenting current robot capabilities and prototypes, the aim is two-fold: 1. to prepare the reader for evaluations of real world robots rather than speculative robots [*Smits et al.*, 1995; *Nordmann and Rip*, 2009] and, 2. to educate the reader on the technical capacities of a care robot as a pre-requisite for plausible evaluations [*Swierstra and Rip*, 2007; *Lucivero et al.*, 2011]. I outline what a care robot is, the functions it serves, and the technical capabilities a care robot may posses now and in the foreseeable future. To do this, I must also clarify key terms with respect to robots. I begin this chapter with an introduction into what robots are, the difficulty in their definition, the variety of types and the range of capabilities a robot may posses. I continue with what a care robot is in terms of its functions and the applications in which it may be employed and I conclude with an outline of the range of capabilities a care robot may posses now and in the foreseeable future. The chapter concludes with an attempt to match robot capabilities with values in an attempt to translate certain values in care into capabilities of a care robot. The information used for this chapter is a product of: scholarly articles and books in the many fields associated with 'robotics' (e.g., human-computer interaction, human-robot interaction, social robots, service robots, artificial intelligence, computer science, robots and automation, among others); conference attendance of technical conferences (e.g., The International conference on Robots and systems, The International Conference on Social Robots, the International Conference on Human-Robot Interaction, among others); work experience in a robotic's institute (Canadian Surgical Technologies and Advanced robotics, i.e., CSTAR); and, field work experiences visiting a variety of robotics labs (CoTeSys lab, Technical University of Munich; University of Hertfordshire, University of Twente). It should be noted that the aim of this book is to present information relevant to a discussion on care robots and to translate this into a usable language for readers coming from multiple disciplines. As such, there are specific issues in robotics research that I do not take up (e.g., the creation and use of unmanned drones in military applications).

4.2 Defining a Robot

Just as care is an incredibly difficult concept to define, robots may be one of the most difficult technological innovations to define. This is in part due to the immense technical knowledge required to understand their functioning but also to the role media has played in shaping the image of a robot in the minds of society. The image given by the media, represented by – Star Wars' C-3PO, Star Trek's Data, Pixar's WALL-E – all represent a class of robots not yet realized by today's technology. These futuristic human-like robots may be part of the future or may never be realized.

The word *robot* was coined by Karel Capek (1890-1938) in his play R.U.R. (Rossum's Universal Robots) where he used it to refer to a race of manufactured humanoid slaves [*Capek and Selver*, 1923]. Thus, robots were conceived of as machines that can do the work of humans. The term robot essentially replaced the terms android and automaton which had been traditionally used until that time. Although the word robot was introduced in the last 100 years, society's fascination with robots, or artificial man-made creations resembling the biological, can be traced back to Ancient times including the clay Golems of Jewish legend or Talos, the man of bronze guarding the Cretian island of Europa from pirates in Ancient Greek mythology. In Renaissance Italy, daVinci®sketched plans for a humanoid robot [*Lin et al.*, 2011]. In the 17th and 19th century Japan, automatons like mechanized puppets were made. In 18th century France, Jacques de Vaucanson made several life-sized automatons: a pipe player, a flute player and a duck. The mechanical duck could flap its wings, crane its neck, swallow and give the illusion of excreting food (the excretions were actually stored in a hidden compartment). The first modern robots capable of fulfilling simple tasks were those in industrial applications. George Devol and Joseph Engelberger are recognized as having made the first industrial robot, Unimate, sold to General Motors in 1960. Thus, although robots are thought to be a modern fascination history reveals this to be far from the truth.

When defining a robot, the "sense-think-act paradigm [*Borenstein and Pearson*, 2011, p. 259] is as close to consensus as one might find in terms of a robot differentiated from a computer system. Alternatively, some believe that robots are similar to computers up until the point of physical agency – "a robot physically embodies the link between perception and action" [*Franklin and Graesser*, 1997]. Others believe the distinction between robots and other appliances or devices is the element of autonomy – that robots are capable of completing tasks without direct human input [*Thrun*, 2004]. Others believe the definition of a robot changes as the technology for creating robots develops. The term robot

differs from robotics; "the study and use of robots." [*Mitcham*, 2005, p. 1654], or "the science and technology of robots" [*Siciliano and Khatib*, 2008, p. 1]. The term robotics has also undergone changes depending on the state of the art. In the 1980's robotics was defined as "the study of intelligent connection between perception and action". This more sophisticated description of robotics indicates the underlying core definition of robots as being capable of both perceiving their environment and executing some type of action.

Robots come in all shapes and sizes intended for performing a variety of tasks; however this does not mean there are no defining characteristics shared by all robots. In short, all robots are man-made or *artificial* intended for the purpose of *fulfilling a task* for a human. All robots can *sense* their environment and can *manipulate* and *interact* with things in this environment. A robot is *programmable* and re-programmable in many instances. A robot must have a high level of *intelligence* which affords them the ability to make choices based on the environment or a set of pre-programmed sequences of action. This high level of intelligence comes in different forms and in varying degrees. The kind of intelligence endowed to social robots (robots that interact in a human-like manner) leaves one questioning whether or not the robot has intent of agency of some sort. This is a highly controversial point and will be discussed in greater detail in chapter 8, 'Designing Moral Factors With Care'. The ambiguity of intent or agency also results from its physical *embodiment* and level of functioning. Many robots can also move on one or more axes of rotation or translation; however, many robots may reside at fixed workstations.

Added to the confusion of defining a robot is the variety of types of robots and the ambiguity surrounding such classifications. According to the most recent survey made by the United Nations (UN) and the International Federation for Robotics Research (IFRR)[1] in 2002, robots were grouped into three major categories, primarily through their application domain but also through the sophistication of the technology required for that domain. The three groups of robots are; industrial, professional service and personal service. The main distinction between such classes of robots refers to the context of operation, the factory vs. outside the factory, the former being a structured and predictable environment, the latter unstructured and unpredictable. This does not, however, say much about the capabilities, the appearance or the tasks of any of the classes of robots and thus offers nothing in the classification and definition of care robots aside from the fact that they may be considered professional service or personal service robots.

[1] See www.ifr.org/

Another class, or type, of robot that is often confusing is that of humanoid robots. For some, humanoid robots refers to those robots which "selectively emulate aspects of human form and behaviour and may come in a variety of forms from complete human-sized legged robots to isolated robotic heads with human-like sensing and expression" [*Kemp et al.*, 2008]. The latter description – with human-like sensing and expression – is often referred to as social or sociable robots. These robots are designed "to engage people in an interpersonal manner, often as partners, in order to achieve social or emotional goals" [*Breazeal et al.*, 2008, p. 1349]. A common misunderstanding with humanoid robots is that they also possess the capabilities referred to for social robots [*Kemp et al.*, 2008; *Veruggio and Operto*, 2008; *Ng-Thow-Hing et al.*, 2009]. An assumption is being made here that a human-like appearance presumes a set of social capabilities the robot must have. Kiesler and Goetz [2002] found in experiments that "the presence and absence of humanoid features and the behaviour of the robot influences people's assumptions about its capabilities and social inclinations". In contrast to the more mechanical looking humanoid robots are androids; "android robots are designed to have a very human-like appearance with skin, teeth, hair and clothes" [*Breazeal et al.*, 2008, p. 1351]. These robots are most commonly used to test Masahiro Mori's theory of the uncanny valley [*Mori*, 1970] – when robots resemble humans to a close degree it elicits feelings of revulsion and discomfort among observers. "Once the robot reaches a point at which their resemblance is close to perfect but eerily dissimilar enough such that we no longer trust them – that sudden shift in our affinity is represented by a dip or valley on the curve. But the trust returns as the anthropomorphism approaches perfect resemblance to human appearances" [*Bekey*, 2011, p. 25].

4.3 Robot Capabilities and Features

The following section is intended to draw the reader's attention to the variety of capabilities a robot may have. Sub-section 4.3.1, 'Capabilities for safe interaction with humans', discusses the feature of robot safety in terms of interacting with humans in a human environment, hence a robot outside the factory. Programming for these conditions may be done in a variety of ways as discussed in the section. This section will also touch on the standards for safety pertaining to industrial robots as they bear weight in the creation of standards for robots outside the factory. The remaining sub-sections discuss additional features of a robot that result in the robot having differing capabilities. The relationship between features and capabilities may be described as follows: the robot is en-

dowed with certain features (ex. auditory, vision or locomotive) which result in the robot having the associated capability.

4.3.1 Capabilities for Safe Interaction with Humans

"An essential component of the duty of care is that a carer must keep their charges safe from physical harm" [*Sharkey and Sharkey*, 2011, p. 268]. In a discussion of robots that may be used for care, I begin with safety to emphasis its significance. The feature of safety renders the robot capable of interacting directly or indirectly with humans in a human environment. Many European initiatives are in place to test the safety standards of the new generation of robots outside the factory . Based on these initiatives, I refer to robots safe for human interaction as human-friendly robots (HFR). The human-robot interaction may be '*hands-on*' [*Bicchi et al.*, 2008, p. 1341] or '*hands-off*' [*Bicchi et al.*, 2008, p. 1337]. The former refers to robots designed intentionally to interact with humans, a robot to assist with feeding for example, while the latter refers to those robots which may accidentally interact – come into contact with – with a human, for example, a Roomba®vacuum cleaner. Others refer to the interactions between humans and robots according to the flow of information. The interaction is *direct* if the flow of information is bidirectional meaning "information is communicated between the robot and people in both directions, and the robot and the person are interacting on 'equal footing'" [*Thrun*, 2004, p. 17]. Alternatively, interactions which are considered *indirect* are referred to as such because "the operator commands the robot, which communicates back to the operator information about its environment, task and behaviour" *Thrun*, 2004, p. 17).

The difficulty with building HFRs, or 'safe' robots, is the trade-off of speed for accuracy. The dilemma now is to design robots that are safe to interact with humans without having to sacrifice performance criteria. One way to design HFR is using the concept of intrinsic safety: "a robot will be safe to humans no matter what failure, malfunctioning, or even misuse might happen" [*Bicchi et al.*, 2008, p. 1337]. One aspect of intrinsically safe robots is to quantitatively assess the risk of injuries in accidents for comparison with other solutions and for optimization of the robot design. For this, the severity of a potential impact is linked with the statistical probability of causing a certain level of injury. Other methods for designing intrinsically safe robots take the hardware of the robot into consideration to increase their sensorial apparatus (ability to sense objects in their environment) or to add protective layers to manipulators (arms) which may come into contact with humans. Other avenues explored look at

introducing mechanical compliance into the design, this means a motor in one area of the robot (i.e., one manipulator) can be decoupled/turned off if an impact has occurred in another area. This design, known as compliant transmission, is thought to diminish performance but this may not be a problem when the robot is used for an entertainment application. In other applications, especially in care contexts, speed and accuracy of task execution are vital.

Due to the length of time in which robots have been employed in industrial applications, standards are already in place . The standards pertain to the use of the technology in the workplace but are also robot-specific standards. The introduction of robots that can interact with humans in the workplace requires revision of these standards. The International Organization for Standardization (ISO) undertook a revision of standards in 2002 from the original standard for robots in 1992, ISO 10218[2]. One part of this revision is meant to address workplace safety for end-users rather than for manufacturers. This revision in standards allows for new modes of operation (simultaneous control of multiple manipulators, mobile robots, collaborative operation) between humans and robots in a defined workspace. Control reliability no longer relies on hard-wired electromechanical components but rather acknowledges the significance of state of the art software, electronic and network based technology for safety-related soft axis (layers on the hardware of the robot) and space limiting control activity (sensors on the robot to determine and control for their space in an environment). Moreover, instead of relying on distance between the human and robot, the new standards recognize that stopping time and distance are more important criteria when robots and humans share a working space.

These standards are particularly relevant when speaking of hands-off robots in an industrial application. Things become more complicated when speaking of hands-on robots which will be applied in a variety of domains outside the factory. For hands-on robots, the T-15 committee of the American National Standards Institute (ANSI) is setting safety standards regarding intelligent assist devices. Although these standards cover a wide range of technologies from assistive devices to mobile autonomous robots, these standards are promising in that they may be translated into policy governing domestic applications of robots. For example, one aspect of the standards involves risk assessment replacing fixed rules: "instead of declarations regarding how to accomplish safe operation, risk assessment procedures are advised for assistive devices and physical human-robot interaction robotic technologies, to identify and mitigate risks in proportion to their seriousness and probability". Another aspect refers to

[2]www.iso.org/iso/catalogue_detail.htm?csnumber=41571

safety-critical software: under any condition that the robot malfunctions, the entire system will shut down in a safe manner. The standards also indicate dynamic limits which restricts the capabilities of robot design such that a human operator must be able to outrun, overpower or turn off the robot.

One application in which safety is of paramount importance is the use of robots for rehabilitation. These robots come in direct physical contact with patients in a variety of ways. For therapy robots, the robot is in direct contact with the disabled patient and the therapist simultaneously. Roboticists in this area must be sure that the robot is designed in such a way that it cannot cause injury by moving a user's limbs outside their range of motion, with too much strength or with too much speed. In addition to this, limits imposed on the robotic apparatus, redundant sensors (additional sensors) are used as back-up so if one sensor malfunctions another can identify the problem and shut down if necessary. Outside all this, rehabilitation robots must also be designed to be intrinsically safe; "from the systems perspective, when all else fails, to actively to protect the user, it must be the design itself that makes the robot inherently unable to injure the user" [*van der Loos and Reinkensmeyer*, 2008, p. 1244].

Thus, for care robots capable of hands-off interactions (between the human and robot), intrinsic safety is required in terms of shutting down/off if a problem occurs. Control reliability via software, electronic and networked (if tele-capable) technology for space limiting control activity (stopping time and distance between human and robot). High levels of sensorial apparatus (redundant sensors) for assisting with control reliability. Decoupling motors to ensure if one part of the robot has a problem the whole robot will stop (this will decrease performance but will increase safety). For robots capable of hands-on interactions the above considerations are required as well as safety-related soft axis – the robot is soft to touch. The remainder of robot capabilities discussed in this chapter are related to the aspect of safety in that each capability renders the robot safer for interaction with a human whether it be the robot's capacity for vision or force feedback.

4.3.2 Robot Vision

Vision for robots may be considered the primary means for sensory input. To achieve vision, visual sensors are required on the outer surface of the robot. This allows the robot to extract a vision from its environment, restore it and enhance the vision (through adjusting pixels) for analysing it. Recognizing the object depends on the stored knowledge of the robot. To do this, however, the robot must be able to accurately define the structure of the objects. Vision allows the

robot to find its way about, to analyse chaotic scenes, to recognize faces and/or environments of a human, to detect its own arms and to determine where it is in a given environment [*Engelberger*, 1989]. Laser range finders for vision were shown to enable a robot to create a 2-D map of a nursing home and navigate its way around in the absence of environmental cues [*Thrun*, 2004]. Once the robot has sensed its environment, what it does with this information is left to the control architecture of the robot. While vision is useful for acquiring information about the robot's environment, it is also useful for acquiring information about the humans present in the robot's environment. Research for detecting people is customary and widespread; "it is common to endow (service) robots with sensors capable of detecting and tracking people" [*Schulz et al.*, 2003; *Pineau et al.*, 2003]. Robots may be capable of detecting and recognizing gestures [*Kahn et al.*, 1996], they may track gazes [*Heinzmann and Zelinsky*, 1998] or they may visually perceive head motions, breath expulsions, and/or eye motions [*Fong et al.*, 2003].

Thus, the feature of robot vision renders the robot capable of perceiving its environment as well as objects and people in its environment. Vision for a care robot is essential for the safety of the humans (provides a means for ensuring the robot can sense its environment to avoid objects) but also for the successful completion of tasks (can accurately locate an object to be moved or a person to bring an item to). Along the same lines, the feature of vision may endow the robot with more sophisticated vision capabilities. Infra-red vision would allow the robot to detect and locate a human in a dark room [*Engelberger*, 1989] which would provide a remarkable capability for the care robot, surpassing the capability of a human care-giver. Researchers are also highly involved in the programming of vision such that the robot is capable of facial recognition [*Jain and Li*, 2005]. Researchers use a variety of techniques to balance light conditions, robot position and human position, but the end result endows the robot with the capabilities for recognizing one face or multiple faces depending on the robot's sophistication and needs. Above facial recognition, researchers are investigating the potential for the robot to recognize certain emotional states [*Kim et al.*, 2004; *Breazeal and Aryananda*, 2002; *Breazeal et al.*, 2008; *Mayer et al.*, 2010]. This is done in a variety of ways from endowing the robot with the capability for perceiving physiological cues and/or bodily movements/gestures as a means for determining emotional state. One promising avenue for emotional recognition can be seen in the work done at the technical university of Munich's CoTeSys (Cognition for Technical Systems) lab. Using the six universally recognized facial expressions corresponding with emotional states the goal is to make the robot capable of recognizing the human user's

emotional state based on its facial expression. This kind of research is meant to facilitate the forming of an empathic bond between the human user and his/her robot [*Mayer et al.*, 2010].

4.3.3 Auditory Capabilities; Dialogue Management, Voice Synthesis and Voice Recognition

This capability refers to the robot's ability to recognize and understand spoken language as well as the robot's ability to communicate using language. The degree to which the robot is capable of any of these tasks may vary depending on the task the robot is required to fulfil. Some robots generate speech but do not understand spoken language [*Thrun et al.*, 2000], others are capable of understanding spoken language [*Bischoff and Graefe*, 2003] or may use keyboard interfaces to communicate using language and bypass speech recognition altogether [*Breazeal et al.*, 2008]. When the robot is capable of understanding spoken language, it is be referred to as "dialogue management" and consists of a "set of procedures and rules designed to ensure that effective two-way communication is maintained between operator and machine in the face of imperfect (error-prone) communication channels" [*Engelberger*, 1989, p. 211]. If imperfect communication refers to non-recognition of spoken commands, the robot must be endowed with sophisticated software designed for those who suffer from a speech impairment (i.e., stroke patients). If, however, imperfect speech communication refers to a tele-capable robot and the telecommunications link connecting the care-giver and the care-receiver, the robot ought to be endowed with redundant (an additional set of) encoders and decoders to avoid a complete break-down in communication on either side.

Speech as a communication modality is easy to control and may be quite effective for human-robot interactions [*Thrun*, 2004]. There are problems, however, when speech is involved. The number of speakers to recognize and the presence or absence of environmental noise play a role in the success of dialogue management. If the system is 'speaker dependant' it will only recognize the voice of a designated speaker [*Engelberger*, 1989]. Programming through demonstration [*Friedrich et al.*, 1996; *Billard et al.*, 2008] creates templates which can be matched to spoken words and subsequent utterances will be matched with remarkable accuracy. If, however, the robot must recognize any number of speakers, the software demands increase and the range of vocabulary recognized inevitably decreases to the use of 'yes' and 'no' in some instances. Environmental noise may also present a problem by masking the sound of the speaker. One further difficulty with the capability of speech is a misunderstanding, or unreal-

istic expectations, of the capabilities of the robot. A speaking robot may create a false perception of the robot's level of intelligence, its social capabilities or its overall capabilities [*Goetz and Kiesler*, 2002; *Fong et al.*, 2003]. One might suggest that maintaining the appearance of the robot as a machine-like artefact may compensate for this tendency.

Auditory capabilities resemble robot vision in the sense that they fulfil requirements for sensory perception and come in a variety of degrees of sophistication. Whether the robot is capable of using this information for the execution of a task or is capable of transferring this information to a human operator who controls the execution of the task is again dependent on the control architecture of the robot. Baseline auditory capabilities, in which the robot can recognize the human voice and match the words spoken with a template held in its memory, may be enough. At other times, sophisticated voice recognition and speech analysis is required (what may be referred to as a social capability).

4.3.4 Mobility/Locomotion

The capability of mobility allows the robot to travel along the x-y planar axis and demands distinct planning and control to achieve this[*Kavraki and LaValle*, 2008; *Chung et al.*, 2008]. Mobility can be achieved through a variety of modalities, for example wheeled mobile robots [*Morin and Samson*, 2008] which in turn demands that the robot in motion is programmed for obstacle avoidance [*Minguez et al.*, 2008] (obstacles being material or, in some cases, people) Mobility is meant to "extend the robotic aide's working volume beyond the desk-top workstation environment. It may also include vertical mobility to facilitate access to floors and shelves" [*Engelberger*, 1989, p. 213]. Thus, robot mobility is meant to distinguish mobile robots from traditional stationary industrial robots with a fixed platform, or other robots that perform a function in the kitchen, on the desktop or by the bed. The Diet-Assist robot developed at MIT is an example of a stationary social robot which resides in a common room of an individual's house and serves the function of providing support and encouragement to an individual on a diet [*Kidd and Breazeal*, 2006; *Kidd et al.*, 2008; *Turkle*, 2011].

Locomotion is different from a robot which is capable of moving an effector or manipulator (arm or hand). For example, the surgical robot daVinci®does not travel as it operates but its robotic arms must be moveable during the course of the surgery. In contrast, examples like In Touch's RP7 robot or the TUG robot, are meant to travel through the hospital to the patient's bedside and therefore must be capable of locomotion. It follows then that depending on

the task for which the care robot is intended, it may or may not be capable of locomotion.

Robot locomotion may be accomplished through a variety of modalities; wheels [*Campion and Chung*, 2008], legs [*Kajita and Espiau*, 2008], wings [*Wood*, 2008], or snake-like movements [*Hirose and Yamada*, 2009] among others. It should be noted that winged locomotion differs from aerial robotics [*Feron and Johnson*, 2008], the latter referring to the development of aerial drones used in military applications [*Singer*, 2009]. Robot locomotion may also be controlled through a variety of modalities; an autonomous mobile robot or a human-operated mobile robot. In many instances, the architecture of the robot is determined according to the chosen means for locomotion. Mechanics for mobility vary depending on the institution or company designing the robot and the terrain which the robot is expected to move on. Researchers at the Tokyo Institute of Technology are creating a locomotive robot that moves in a snake-like manner [*Hirose and Yamada*, 2009]. In contrast, researchers at Honda are designing a bi-pedal humanoid robot, ASIMO, to walk like a human using a zero-moment technique [*Ng-Thow-Hing et al.*, 2009]. This technique means the robot equally balances all forces so there is no point at which the robot would lose balance and fall. Additionally, this type of motion requires that the robot be on a smooth surface – not an optimal condition if the robot is to exist in an unstructured environment where these things cannot be accounted for. Other researchers are exploring the use of gravity to propel the 'legs' for moving, a technique referred to as 'passive dynamics' [*Hosoda et al.*, 2008]. This approach/technique uses little motor power to accomplish walking and is considered a promising, efficient substitute to the zero moment technique used for ASIMO.

Wheels are the most typical means for motion for reasons of simplicity [*Campion and Chung*, 2008]. A Segway is commonly thought of as a mobile robotic platform which uses wheels for motion. Researchers at Carnegie Mellon are also investigating the use of a ball for locomotion [*Lauwers*, 2006]. The "ballbot" is a battery operated, omnidirectional robot that balances on a single urethane-coated metal sphere. Because of the use of the ball it is able to manoeuvre in tight spaces and has the potential to interact in human environments better than wheeled robots.

Control of the robot's mobility may be human-controlled or autonomous. An example of a human-operated mobile robot is In Touch's RP-7. This robot is aimed at facilitating patient-physician communication when the physician cannot be physically present at the bedside of the patient. The physician, seated at a console in another area of the hospital or in another place en-

tirely, guides the robot through the hallways of the hospital to the patient's bedside. Using a video monitor attached to the mobile autonomous robotic platform, the patient and the physician may communicate directly. In contrast, iRobot's®Roomba®vacuum cleaner or iRobot's®Scooba®(pool cleaner) are both mobile robots which operate autonomously; no human manipulation is required to guide the robots locomotion.

In terms of safety, the capability of locomotion requires the use of sensors on the hardware of the robot to indicate if, and when, the robot is approaching an obstacle in its environment [*Minguez et al.*, 2008]. If the mobile robot is autonomous, additional safety considerations must be accounted for; whereas a human-operated mobile robot is less likely to collide with other objects because of the control of the human, an autonomous mobile robot requires redundant (additional) sensors for perceiving their environment. The issue of speed for travelling and stopping in autonomous mobile robots is also significant. Vision capabilities may also be incorporated into autonomous mobile robots to provide the robot with the ability to recognize environmental cues like landscapes or faces. The capability of locomotion, in most cases, will be a pre-requisite if the care robot is to fulfil certain physical caring tasks.

4.3.5 Grasping and Manipulating Objects

The earliest robots were praised for their capabilities to grasp and manipulate objects. In fact, these were the only capabilities endowed to traditional industrial robots. For Engelberger, grasping refers to "the property of a robot that allows specific objects to be selected, positioned and oriented. It is typically associated with 'hands' and includes a variety of functional attributes such as detection of slippage and evaluation of object geometry for stable holding" [*Engelberger*, 1989, p. 212]. Robot grasping[*Prattichizzo and Trinkle*, 2008] is often, although not necessarily, accomplished through robot 'hands' [*Melchiorri and Kaneko*, 2008]. Although Engelberger makes reference to 'hands', these are also often referred to as 'end effectors' while the arm is referred to as a manipulator. The end effector can take the form of a humanoid hand but can also be a simple gripper consisting of two fingers. These simple grippers can open and close to pick up and let go of a range of small objects. Vacuum grippers are used for heavy lifting of objects with a smooth surface. These types of grippers should not be used for the lifting of humans. Some of the more sophisticated robots have effectors in the form of humanoid hands with up to five fingers (four is more

common). Examples include ASIMO, the Shadow Hand[3] or the Schunk hand[4]. The latter two are highly dexterous manipulators (20 degrees of freedom) with an incredible number of tactile sensors. The feedback from the tactile sensors allows the robot to apply the correct amount of pressure so as not to break the egg or light bulb it is holding.

Closely in line with the capability of grasping is that of manipulation [*Brock et al.*, 2008]. Manipulation refers to "the capability to move objects from one place to another while maintaining a correct orientation of the hand and avoiding collisions with stationary objects" [*Engelberger*, 1989, p. 212]. Consequently, manipulation requires knowledge of where the robot is in its environment. Moreover, the robot must have hand-hand coordination to understand when one hand has picked something up and placed it down. In short, we may suggest that the capability of manipulation requires the capability of grasping which presupposes the presence of end effectors (grippers, hands, etc.). This does not, however, imply that if a robot has grippers or hands they will be used for grasping or manipulating. For example, the robot for assisting with elderly care developed at Carnegie Mellon University, named PEARL[*Pollack et al.*, 2002], has arms but these are used for communication and not for grasping or manoeuvring objects.

4.3.6 Force Feedback and Tactile Sensation

Force feedback and/or tactile sensation features of the robot fall under the umbrella of haptic research [*Hannaford and Okamura*, 2008]. Although the most significant form of sensory perception for a robot is vision, next in importance may be force feedback and/or tactile sensation, especially in the case of care robots which may have to handle delicate objects like humans, as is the case of lifting or bathing robots. Force feedback refers to the amount of pressure the robot can feel [*Cotin et al.*, 2000]. To accomplish this capability a variety of sensors are required. These 'touch sensors' measure pressures applied to various points on the robot end effector, slip detectors to sense loss of a grip and joint-force sensors that measure forces applied by a robot's hand, wrist and arm joints. Placement of these sensors relies on the area which requires feedback (for example, if the end effector/hand/gripper is meant to perceive the strength with which they are grabbing or manipulating an object, such sensors will be placed on the end effector). These sensors allow the robot to feel whether a grasp

[3]See www.shadowrobot.com/hand/

[4]See europeanrobotics12.eu/media/15066/Tomas%20Berg%20-%20Schunk.pdf

is proper for the task and whether there is interference in fitting one object into another [*Engelberger*, 1989]. In turn, the robot may adjust their behaviour accordingly (in the case of an autonomous robot). Tactile sensation refers to the robot's ability to sense its environment through touch [*Kawasaki et al.*, 1999]. The robot, being able to perceive the variety of textures a surface might have, is then capable of determining whether an object is hard or soft regardless of whether the robot is in the dark or visually obstructed [*Engelberger*, 1989]. Intuitively, one need only imagine the necessity for force feedback capabilities in any number of care robots. In the field of surgical robots, tactile sensation and force feedback are of incredible importance. The first surgical robots (Computer Motion's Zeus Telesurgical System and the daVinci®robotic system) were not endowed with such capabilities and thus surgeon's had to learn new skills in order to perceive the surgical field properly. The latest in surgical robots, namely the Amadeus Composer®, aim to incorporate haptics into the architecture of the robot. Those that are used for lifting, bathing, or even feeding patients will all require haptic capabilities (force feedback and/or tactile sensation) to some extent.

4.3.7 Social Communication/Capabilities

Social communication may be considered a feature of a robot which is facilitated by a range of social capabilities. Social capabilities of the robot refer to the ways in which a robot can communicate with a human user in an engaging, interpersonal, social manner, essentially in a more human-like manner. The field of social robotics began around the 1940-50's by William Grey Walter and has been developed extensively in the 90's by researchers like Kerstin Daughtenhahn (University of Hertfordshire) and Cynthia Brazeal (MIT) [*Fong et al.*, 2003]. The work of Daughtenhahn has focused on creating a kind of robot etiquette outlining proxemic cues to be programmed in the robot to ensure the robot complies with the preferences of humans for successful interaction [*Walters et al.*, 2007]. Dautenhahn is also engaged in extensive work investigating robots endowed with social capabilities used as teaching tools for children with autism [*Dautenhahn and Werry*, 2004b; *Dautenhahn*, 2003]. The work of Cynthia Brazeal has focused on different aspects. Brazeal aims to create robots that can interact over a long period of time with humans in a meaningful way: "social or sociable robots are designed to engage people in an interpersonal manner, often as partners, in order to achieve social or emotional goals" [*Breazeal et al.*, 2008, p. 1349]. The capabilities the robot must posses in order to communicate in this highly developed manner rely on the mechanics of the vision and

auditory capabilities previously mentioned. For these sensory capabilities to fulfil the requirements of a social robot, however, the level of sophistication of these capabilities differs greatly. The domain of social robotics is motivated by questions like: "how to design for a successful long-term relationship where the robot remains appealing and provides consistent benefit to people over weeks, months and even years"? [*Breazeal et al.*, 2008, p. 1350].

It should be noted also that a robot may have the capability for communication but this does not presuppose the type of communication to be in a social manner. Additionally, a robot may have the capability of vision but it may be used for social interactions or alternatively for sensing objects in its environment. Thus, social capabilities build on the existing capabilities a robot may have and presume autonomous decision-making throughout the human-robot interaction.

Social capabilities are distinct from physical capabilities a robot might also have; "the benefit that social robots provide people extends far beyond strict task performing utility to include educational, health and therapeutic, domestic, social and emotional goals and more" [*Breazeal et al.*, 2008, p. 1350]. Methods for communication are often referred to as the interface and a large portion of research is dedicated to human-robot interactions. At this point in time these robots have their own class because of the technical requirements to program them accordingly. In time, as technology progresses, it is believed that the function of social interaction will be a feature added to the existing technical framework. These robots interact with humans in a social way, meaning they communicate (visually, auditorily or verbally) with humans beyond indicating the initiation or completion of a task. "Social robots use a variety of modalities to communicate from whole-body motion, proxemics (i.e., interpersonal distance), gestures, facial expressions, gaze behaviour, head orientation, linguistic or emotive vocalization, touch-based communication, and an assortment of display technologies" [*Breazeal et al.*, 2008, p. 1350]. The robot must be able to perceive this information, interpret it accurately and respond appropriately. The issue of interpretation is quite complex due to the range of human behaviour and communication modalities and thus social robots are considered among the more sophisticated robots of today. This type of interaction with a human user presupposes that the robot is safe to interact with humans in either a hands-off or a hands-on form.

The interpersonal manner in which these robots are meant to engage people relies on verbal as well as non-verbal cues referred to as paralinguistic cues. The roles of paralinguistic cures are, according to the Springer Handbook of robotics, as follows:

1. *Regulator*: expressions such as gestures, poses, and vocalizations that are used to regulate/control conversational turn-taking.

2. *State displays*: signs of internal state including affect, cognitive, or conversational states that improve interface transparency.

3. *Illustrators*: gestures that supplement information for the utterance. These include pointing gestures, iconic gestures. [*Breazeal et al.*, 2008, p. 1350]

Robots can be placed in sub-categories depending on their ability to express paralinguistic information, to understand paralinguistic information, to communicate in group conversations, to understand expressive emotion-based interaction, to understand and interact in an empathetic way, to mirror emotions and to adapt to a human's changing emotions or goals. All of these aspects require distinct technical programming. The end goal is to amalgamate these mechanics together to create the future social robot – one that is capable of interacting in each of the above mentioned manners. Current robots classified as something other than social robots are equipped with both linguistic and paralinguistic mechanics to varying degrees depending on the task which they are intended to perform.

Social robots may also be used for enjoyment, learning and/or therapy or for personal growth. Social robots are used for teaching/learning. Philips ICat robot is used to understand human-robot interactions in order to program future robots accordingly. In therapeutic instances, a predominant use of social robots is to interact with children with autism. The work of Daughtenhahn shows how social robots can be used as a tool for teaching children skills of interaction. Robots with social capabilities are also used as diet assists; individuals wanting to lose weight use the robot to help motivate, encourage and keep track of progress or lack thereof. The goal is to foster a meaningful bond between the human and the robot in order to achieve weight loss goals with greater success. Aside from these instances, sociable robots, or the capabilities of social robots are also used in service applications. For example, a robot to greet people as they enter an office, a hospital, an airport or an elderly home [*Vongsoasup and Mataric*]. In this case, the robot is not intended to form a long term bond with users; however, the same capabilities for such an initiative are used to facilitate successful human-robot interactions to meet a service goal [*Shieh et al.*, 2007; *Fong et al.*, 2003]. In these scenarios the robot might understand and produce natural language, or may understand and produce proxemic cues. Sociable robots used in these 'service situations' fulfil a short term functional goal rather than social robots used to meet the long term emotional needs of persons.

In terms of care robots, communication in the interpersonal manner for which they are designed, seems quite appropriate for the interaction between a care-receiver and a care robot. This may make the care-receiver more comfortable having the care robot in their personal space, whether it be their home, hospital or nursing home room. Additionally, it may make the care-receiver feel more at ease having a care robot incorporated into a care practice to begin with if the robot is able to communicate in a social manner whether verbally or via paralinguistic cues. One might suggest then, that a care robot intended to provide care for the care-receiver directly ought to be equipped with social capabilities to the extent that they may sense and perceive a range of verbal and non-verbal communication paradigms and may communicate with a human user in a similar fashion. This may also ensure the robot acts politely according to cultural standards, a robot etiquette if you will.

One might also suggest that communicating in an empathic way seems quite appropriate as a capability for a care robot. Such a capability may be facilitated through facial recognition or emotion recognition as mentioned previously. For the delicate and sensitive tasks in care that require a caring, or empathic, disposition we may even go as far as suggesting that a care robot *ought* to be endowed with such features to ensure the capability of forming an empathic relationship. However, is this not simply a recognition that a human ought to be present? Taking a second glance, is it wise to have the robot capable of eliciting this type of communication? If the care-receiver believes the robot to be empathetic could it not interfere with the care-receiver communicating, or wanting to communicate, with the human care-giver? If the robot were endowed with this capability, is it possible that a human care-receiver will find it easier to communicate with the robot than with a human care-giver (Pransky et al, 2004).

4.3.8 Appearance

The appearance of a robot is not necessarily a capability but is referred to here as a feature. It is one of the most important features as it plays a large role in the expectations users have towards the robot as well as their comfort level with the robot. Reasonable expectations are required for a functioning interaction between human and robot. Comfort level is of paramount importance as robots are now entering into our personal spaces. According to a survey done by Fong et al, participants claimed that a robot's appearance ought to correspond with their intended tasks [*Fong et al.*, 2003]. If the robot is meant for cleaning it need not appear human-like but may appear as a machine, like the Roomba®. Much

emphasis is placed on the robot having a humanoid appearance. In fact, many engineers believe it is favourable for the robot to have a humanoid appearance and have been designing according to this assumption. Alternatively, many believe this confuses the expectations of users.

I have already discussed humanoid robots in terms of appearance (resembling a human in form meaning it has arms and/or legs and/or a head, etc.) and social robots in terms of their capabilities (can interact with humans in a more human-like manner regardless of their appearance). Another category of human-like robots is that of androids. In terms of their appearance, they are designed to resemble a human with as much detail as possible, thus including skin, teeth, eyes, etc. Currently, androids are not used in service applications of any kind but rather are used to test human reactions to robots resembling humans to varying degrees [*Minato et al.*, 2004; *Sakamoto et al.*, 2007]. If we recall, the uncanny valley is the phenomenon whereby feelings of disgust are elicited by a human when viewing a robot that nearly resembles a human [*Mori*, 1970]. A US company called Hanson Robotics is currently involved in the design of android robots for entertainment purposes. Their most well known android being the Einstein head [*Oh et al.*, 2006].

As social robots are meant to communicate in a social or human-like way, the embodiment of a social robot is often humanoid or animal-like; however, the class of social robots refers to the capability for social communication and not a presupposition of appearance. There are many examples of social robots which have animal-like features like the baby seal robot, Paro. Paro does not verbally communicate but relies on touch based communication to perceive information with the user but also to communicate to its user (that it wants to be held or petted, etc.) [*Wada et al.*, 2005]. There are mobile social robots fitted with a face to enhance social interaction, like the elder-care robot, Pearle, developed at Carnegie Mellon University. For mobile robots, issues of proxemics, as a modality of social communication, are particularly important and culturally dependent (proximity preferences in communication differ between cultures). Other social robots do not have an animal or humanoid appearance, like MIT's Kismet or the Keepon made by the National Institute of Information and Communications Technologies in Japan. Kismet has a mechanical face with anthropomorphic features (large blue eyes) to enhance social communication, while Keepon, the small dancing robot, has a simple face and expresses itself by squashing or stretching its body. There are also social robots with no face or eyes or any anthropomorphic features. These robots resort to language-based communication and proxemics. Therefore, although the robot is assumed to communicate in a human-like way this does not presume the appearance of

the robot must also be human-like for social robots.

Robots are also very machine-like, like the RP-7 robot, the daVinci®surgical system, the Hybrid Assistive Limb, and the Sanyo bathtub to name a few from the healthcare domain. Alternatively, robots might also be creature-like resembling a known creature like Paro, the white baby seal with fur, or ambiguous like Keepon the small dancing robot, which has a simple face and expresses itself by squashing or stretching its body. For care robots there is once again no defining appearance that renders the robot a care robot. They may have any kind of appearance; however, selecting the robot's appearance will depend greatly on the empirical research indicating that the robot's appearance should fall in line with its capabilities and task for use.

4.4 Modes of Robot Control

The discussion up until this point has addressed the appearance and capabilities a robot may have. These features all make reference to the sensory capabilities of the robot. In other words, the capability of vision, grasping and manipulating are features which provide the robot with capabilities to sense and perceive things in their environment and where they are in their environment. I now turn to how these sensory mechanisms are controlled. With the information obtained through the robot's sensing capabilities, an action or task must then be accomplished. There two options for executing the task; human-operated control or autonomous control. In short, human-operated means a human is responsible for executing a task given the information provided through the robot's sensing capabilities/sensors. In contrast, autonomous control asserts that the robot is responsible for deciding how to accomplish a task given the information it has acquired through its sensing capabilities/sensors.

4.4.1 Human-Operated Robots

Human-operated robots, as a type of robot, have the feature that they must be operated on, controlled, by a human in order to execute their required task or action. Human-operated robots represent one of the earliest classes of robots, one which requires a human to guide the action/movements of the robot; the movements or commands of the human are translated into movements made by the robot. This configuration or interface is often referred to as master-slave: the human operator as master and the robot as slave.

For hands-on human-robot interaction in industrial applications, such robots

are often referred to as "cobots". These are "collaborative robots" designed to relieve humans from fatigue or stress and to prevent injuries; "cobots presume a division of control between human and robot, with a robot perhaps supporting a payload and allowing a human to guide it" [*Bicchi et al.*, 2008, p. 1345]. In this scenario, the operator is in direct physical contact with the payload. This description may also be used to describe exoskeletons used for rehabilitation purposes [*Hayashi*, 2005]; "exoskeletons are also controlled by a human operator, leaving all planning and high-level challenges to the user" [*Niemeyer et al.*, 2008, p. 741]. Again, the user is in direct contact with both the robot and the payload.

Within the class of human-operated robots is a subclass known as telerobots whose infrastructure is designed such that a human operator controls the motion/movement/task execution of the robot in the same way with the added condition that the human operator is at a distance [*Niemeyer et al.*, 2008; *Mitra and Niemeyer*, 2008]. Again, all planning and cognitive decisions are made by the human user and the robot is used strictly for mechanical completion of a task. The use of 'tele' (derived from the Greek word for distant) presumes a geographical separation between the user and the environment in which the task is being performed. The inaccessibility of the environment may be for any number of reasons; the user cannot or will not physically reach the environment (as in robots used to search pipes fro gas leakages), the environment is dangerous (as in search and rescue robots, robots used in marine or extra-terrestrial environments), the environment needs to be scaled (as in surgical robots). The physical distance between the user and the robot varies depending on the application (e.g., for surgical robots the surgeon is often in the same room, for robots in space or underwater the distance is much greater). In most cases there are two sites to speak of; the local site with the human operator and the remote site with the robot. For the information to travel from one side to another, the two sites must be connected. Traditionally this was done through the use of cables; however, recently computer networks have made it possible to transmit this information from one side to another using a telecommunication system [*Rayman et al.*, 2006], a dedicated network or a satellite in some instances [*Rayman et al.*, 2007].

Control of the robot may occur through one of three architectures; direct control, shared control, or supervisory control [*Niemeyer et al.*, 2008]. Direct control assumes no autonomy or intelligence on the part of the robot, thus, all the motions of the robot are directly controlled by the user (ex HAL exoskeleton). Shared control refers to a sharing between local and remote sites whereby the human operator decides what to do and how to act while the robot can autonomously refine the command for the environment. For example,

in the case of the daVinci®surgical platform, the surgeon performs its movements which the robot autonomously scales down to the appropriate size for the surgical field. Supervisory control is described as analogous with supervising a subordinate staff member whereby the supervisor is responsible for giving orders to the subordinate but in turn receives summary information. This approach is compared with direct control or autonomous robot control by Sheridan, who introduced the concept of supervisory control; "human operators are intermittently programming and continually receiving information from a computer that itself closes an autonomous loop through artificial effectors and sensors" [*Niemeyer et al.*, 2008, p. 746]. In other words, "the operator plans activities at a level which can be performed by the robotic system independent of human intervention" [*Niemeyer et al.*, 2008, p. 747]. At all times, the human operator may take over control of the task. All three of these models appear appropriate for the use of care robots which aim to fulfil physical caring tasks. Furthermore, a degree of control by a human seems an intuitive requirement when the physical safety of the care-receiver is at stake.

Telerobotic systems or human-operated robots are unique in that they provide information to, but also require commands from, the user. For Thrun et al, this interaction between the robot and the human is termed indirect according to the flow of information; "the operator commands the robot, which communicates back to the operator information about its environment, task and behaviour" [*Thrun*, 2004, p. 17]. These robots are distinguished from autonomous robots in which "a robot executes a motion or other program without further consultation of a user or operator" [*Niemeyer et al.*, 2008, p. 746]. It may be suggested that autonomous robots evolved from the design of supervisory control robots [*Haselager*, 2005].

4.4.2 Robot Autonomy and Intelligence

In contrast to robots controlled directly through the input of a human user are a class of robots labelled autonomous robots. The issue of robot autonomy is problematic due to the diverse conceptions one may have of the concept of autonomy; "the capacity for independent (unsupervised) action versus the freedom to choose goals" [*Haselager*, 2005, p. 528]. While philosophers approach autonomy from the question of *why* one acts in a certain way, roboticists approach autonomy from the question of *how* the robot fulfills its task (with or without assistance or supervision); "within robotics, the increase in autonomy of a system is related to the reduction of on-line supervision and intervention of the operator, programmer or designer in relation to the robot's operations

in a changing environment" [*Haselager*, 2005, p. 518]. This engineering interpretation of autonomy says nothing of the robot's freedom to choose its actions. Autonomous robots are therefore a class of robots with the capability to fulfil a task without real-time manipulation from a human operator. For Engelberger, "autonomous planning is performed by the machine when sensed data are operated on by application programs with the result that the machine makes navigating (or equivalent) decisions" [*Engelberger*, 1989, p. 211]. These decisions do not require human interaction but are, on the robotic side, subject to human supervision and veto" [*Engelberger*, 1989, p. 211 - 12]. Later visions of autonomous robots claim that such a robot may "operate under all reasonable conditions without recourse to an outside designer, operator or controller while handling unpredictable events in an environment or niche" [*Franklin and Graesser*, 1997]. These two definitions of robot autonomy maintain that the robot is acting according to a pre-programmed set of rules, and the robot is capable of planning their action without referring to a human operator (or designer or controller) *during* execution of the task. What the first definition [*Engelberger*, 1989] allows for is a human to supervise and veto an action. This type of interaction may be referred to as 'supervisory control' of the robot, as discussed in human-operated robots. What the second definition adds is the capability of the autonomous robot to fulfil its task in an environment in which it has not been trained, and/or one that is unpredictable.

More recent definitions, or conceptions, of robot autonomy include the criterion of adapting to their environment; "autonomy refers to a robot's ability to accommodate variations in its environment" [*Thrun*, 2004, p. 14]. Other definitions claim autonomy also includes the capability of the robot to fulfil its task within time constraints and with the added component of potential interference by others. For Bensalem et al, autonomous robots must: "operate in highly variable, uncertain, and time-changing environments; meet real-time constraints to work properly; interact with other agents, both humans and other machines" [*Bensalem et al.*, 2009, p. 67] From the description of an autonomous robot today, we see that some of the challenges for their design relate to challenges posed for other classes of robots; meeting real time constraints while at the same time safely interacting with humans is a similar challenge for HFRs. The difference lies in the robot's capability for autonomous function; predictability of the robot's actions decrease without a human-operator, risk that the robot misinterprets environmental cues and acts improperly increases or, if the robot is unsure how to respond, without the guidance of an operator, it may malfunction or shut down. Thus, safety is significant for this class of robots both in terms of interacting with humans but also in terms of reliability and

efficacy – that the robot is capable of accomplishing its task.

We may infer from the preceding discussion that autonomy refers to the intelligence of the robot and may therefore be spoken in terms of artificial intelligence. Autonomy, or the intelligence of the robot, does not refer to the robot's capability to sense, nor does it refer to the systems which command the robots' motor responses but rather, the robot's capability of interpreting the sensory input in order to make a situational judgement for action [*Engelberger*, 1989, p. 99]. For the average reader perhaps the best way to describe the distinction between sensing and perception is as follows: "the gathering of data is mechanistic, interpreting and understanding this data is intelligence" [*Engelberger*, 1989, p. 100]. To exemplify autonomy as intelligence we may look to the robot guide in a museum. When faced with a large number of visitors, and their unpredictable behaviour, in the museum the robot must alter its path in order to reach its destination [*Thrun*, 2004, p. 15]. Sensing the number of visitors and their behaviour is the mechanistic element but translating this information to alter the robot's path is intelligence, the autonomy of the robot.

In short, we may observe that the feature of robot autonomy renders the robot capable of different things depending on the conception of autonomy one holds. Without prescribing what an autonomous robot *should* refer to, we may suggest that it has the following properties; it can perform its pre-determined task in an unpredictable environment without consulting an outside source for assistance. The task is performed under time constraints in an unstructured and/or dynamic environment. If humans are present in its working environment, depending on the robot's function, the robot will be capable of interacting with humans in a hands-off and/or hands-on manner and as such the associated safety considerations are required.

For Thrun [2004], a robot's autonomy has both types and degrees. Types of autonomy refers to a robot being capable of making decisions about its environment, of making decision about a human's behaviour. Autonomously navigating through an unpredictable, unstructured or hazardous environment presupposes the robot be capable of acquiring environmental models. Alternatively, a robot may be capable of detecting people and their behaviours to autonomously accommodate for them. This means that when referring to a robot's autonomous capabilities, it is important to clarify in which way the robot is autonomous; adapting to its environment or adapting to its user. The degree to which the robot is autonomous has to do with the sophistication of the programming and the amount of tasks a robot may be capable of fulfilling in an autonomous manner. The latest in autonomous robot research refers to robots capable of learning. A learning robot may be considered the most sophisticated form of autonomy to

date.

Robot learning may be used to refer to a feature of a robot – it can adapt by changing its behaviour based on its previous experience [*Franklin and Graesser*, 1997] – or to the way in which the robot is programmed – learning by demonstration [*Friedrich et al.*, 1996; *Billard et al.*, 2008], mimicking [*Mayer et al.*, 2010], or reinforcement [*Billard et al.*, 2008; *Santoro et al.*, 2008] [5]. The concept of robot learning invariably increases the degree of autonomy the robot has and increases the success with which the robot will manoeuvre in a new, unknown environment. This desired way in which the robot will behave invites the concern that a robot is then free to choose a certain course of action. With respect to programming robots by learning, it is thought that robots learn general rules from their experience in order to meet task assignments in highly variable environments (meaning human environments) [*Santoro et al.*, 2008]. There are many ways in which roboticists are exploring how to program learning into the robot. The "Child-robot" developed in Suita, Japan, is said to develop social skills by interacting with humans and watching their facial expressions, mimicking a mother-child relationship [*Minato et al.*, 2007]. The aim of the creators at Osaka University in Japan is to develop this robot to think like a baby, meaning the robot will be able to evaluate facial expressions and cluster them into basic categories like 'happy' or 'sad'. Similar research is investigating sophisticated forms of facial recognition to provide insight into the users' emotional state. Using algorithms to locate features like the corners of the eyes and eyebrows, the nostrils and corners of the mouth, the computer program/robot is able to recognize six basic emotions; disgust, happiness, sadness, anger, fear and surprise. Once referred to as learning-by-demonstration (this approach was targeted towards the use of industrial robots), this term was replaced with 'imitation learning' to reflect the way in which the robot would 'learn' in order to interact with humans in a more natural way (by demonstrating similar skills and processes). These robots are considered the predecessors of more advanced social robots. On the other hand, learning robots do not always have to be safe for human interaction because they may be applied in military,

[5] Learning robots are problematic for many reasons; how do we transfer the human notion of learning to robots, how can we reliably say when a robot has learned, will the robot be able to act in a way not intended by their designers? Various authors have addressed the issue that these robots may be capable of acting in ways not anticipated in their design and as such there exists a problem of responsibility; who is responsible for these robots if the designers cannot with complete confidence predict the robot's behaviour [*Santoro et al.*, 2008]? This question is of paramount importance for the creation of policy concerning such a class of robots. Further research on this sub-class is required as the technology is still in its beginning stages.

surveillance, or search and rescue applications.

Another example of sophisticated robot autonomy is seen in the work done at the Technical University of Munich. In the CoTeSys lab (Cognition for Technical Systems), roboticists are investigating the ability to program a kind of cognition into the architecture of the robot[6] [*Zaeh et al.*, 2010; *Buss and Beetz*, 2010; *Tenorth et al.*, 2010; *Tenorth and Beetz*, 2012; *Tenorth*, 2011]. In this way, the robot is expected to know what it is doing and why, what it sees and saw, what it is capable of doing and what it is not capable of doing, and to predict the consequences of its actions based on such reasoning. The manner in which this is accomplished is through programming semantic links between action codes [*Kunze et al.*, 2011]. To give an example, the CoTeSys lab works with two robots to make pancakes [*Beetz et al.*, 2011]. One robot is mobile, and used to gather the ingredients, and the other is stationary, used to pour the mix on the pan and flip the pancakes. The robots are given verbal commands and respond to the commands verbally. If one were to ask the stationary robot to fetch the milk from the fridge, the robot would respond that it is not capable of moving and thus cannot fulfil the request. As such, the robot is thought to be cognizant of its own capabilities and the limitations thereof. The idea is that such capabilities allow for more human-like interactions as well as building trust between the human and robot counterparts when the robot can account for its own actions. A further discussion of cognition in terms of robots will be the subject of chapter 8 when discussing the moral status of the robot and the link this shares with responsibility. One might wonder whether the robot is capable of being responsible if in fact the robot is cognizant (or aware) of its own actions and the reasoning for such actions[7].

Closely in line with discussion of such sophisticated forms of robot autonomy, and machine learning in particular, is a discussion of robot intelligence, also referred to as Artificial Intelligence (AI). It is beyond the scope of this dissertation to provide a comprehensive overview of the field of AI; however, for the purposes of this book, it is important to introduce this field given its predominant role in the development of current and future (care) robots. The field of AI is concerned with the study of intelligent beings, just as philosophy and/or psychology is; however, AI strives to *build* intelligent entities – hence, artificial intelligence. The study of intelligence demands the question of how one defines intelligence. The Turing Test [*Turing*, 1950], proposed by Alan Turing was in-

[6] See ias.cs.tum.edu

[7] Information pertaining to the research conducted at the CoTeSys lab is a result of both a literature review of scholarly articles as well as fieldwork experience, Oct 2010.

tended to provide an explicit, concrete, definition of intelligence: "the ability to achieve human-level performance in all cognitive tasks, sufficient to fool an interrogator" [*Russell and Norvig*, 1995]. In order for the computer to be 'intelligent', it was/is assumed that the computer "needs to posses the following capabilities: natural language processing, knowledge representation, automated reasoning and machine learning" [*Russell and Norvig*, 1995, p. 5]. There are a range of approaches within the field of AI for programming such capabilities: the cognitive modelling approach, the laws of thought approach and the rational agent approach [*Russell and Norvig*, 1995, p. 6][8]. This high level intelligence renders the robot capable of a sophisticated kind of reasoning. Such reasoning has to do with the way in which information input into the system is processed before giving output. The robot may be mobile and as such taking in information pertaining to their environment. As we saw in the case of mobile robots, path planning and object avoidance are key issues. Using AI, autonomous mobile robots are capable of such tasks without the direct input of a human user.

Alternatively, the robot may be stationary but may still be endowed with sophisticated intelligence for the processing of alternate inputs. To give an example, let us take the APACHE-III database. APACHE-III is currently used in healthcare institutions to aid Intensive Care Unit (ICU) physician's make decisions regarding the treatment of patients [*Wallach and Allen*, 2010]. APACHE-III has over six hundred thousand patient records to draw upon. The physician inputs patient information and APACHE-III processes that individual's criteria against the stored database to provide a recommendation to the physician pertaining to the patient's continued treatment, or the cessation thereof. The APACHE system is considered an intelligent system. It is questionable whether or not this system is, in fact, a robot due to the lack of embodiment and/or engagement with its environment; however, given its sophisticated capabilities, many are inclined to label it a robot (this may change over time). The benefits of such a system are to aid a physician in decision-making as well as aid healthcare administrators for cutting costs and justifying such cuts. The fears associated with the use of such a system have to do with the delegation of such decision-making faculties to a computer. Once used pervasively, will the use of such systems ultimately end with a delegation of morally delicate decisions to computers? In time, will the prudent choice be to defer to the robot given its superior capabilities over humans for processing large amounts of information?

[8] I will not delve into the details of each approach, for a comprehensive discussion see Artificial Intelligence: A Modern Approach [*Russell and Norvig*, 1995].

Such questions are the subject of chapter 8, Designing Moral Factors With Care.

Another recent avenue that AI is taking, is that of Affective Computing (AC) [*Picard*, 2000]. The goal is to endow computers with the capability to perceive emotions of human users in order to improve human-computer interactions (i.e., make these interactions more intuitive) [*Oosterhof*, 2005]. The main questions of AC deal with how a computer can detect emotions and whether a computer should itself have 'emotions' (which of course presupposes the more philosophical question of whether it is possible for computers to have emotions). Recognizing emotions requires that the computer (or robot) have sensors that go beyond those already described. The robot's vision capabilities must be sophisticated in the sense that it can detect facial cues (and not just faces). The robot's auditory capabilities may also allow the robot to perceive emotional information about the user (e.g., the volume the user uses to speak). The robot could also perceive physiological changes in the user as a means for emotion perception. Additional kinds of non-verbal cues the robot has to perceive refer to bodily cues like gestures and or proxemics (i.e., spatial distances) and are described in greater detail in the section on social robots later in this chapter. Accordingly, the class of social robots are a direct result of the field of AC.

The field of AC is of great significance in a discussion of care robots. As we have seen in chapter 3, good care is dependent on the nurse's ability to perceive the emotions of patients (i.e., attentiveness of the nurse) and to tailor their behaviour accordingly (i.e., competence of the nurse). The question then becomes whether or not robots used in care must be endowed with such capabilities given their mere presence in a care context. Such a question is addressed in detail in chapter 8 by asking what assumptions are being made in the endowing of a robot with such capabilities.

With an understanding of the variety of robot capabilities, features, appearances, and modes of control the aim now is to provide a definition of a care robot, current examples of care robots for analysis, and to project future capabilities of a care robot based on the care analysis of the previous chapter.

4.5 What is a Care Robot?

When defining a care robot, there are a few definitions currently available: "Carebots are robots designed for use in home, hospital, or other settings to assist in, support, or provide care for the sick, disabled, young, elderly or otherwise vulnerable persons" [*Vallor*, 2011]. For others, defining a care robot has to do with the anticipated roles it is expected to fulfil. Sharkey and Sharkey

discuss care robots specific for elderly persons and list three main uses of such robots: to assist the elderly, and/or their carers in daily tasks; to help monitor their behaviour and health; and to provide companionship. Alternatively, care giving robots may be distinguished between an "affective robot which refers to if/when a robot is supposed to be a friend or companion to a human being as opposed to a utilitarian robot which refers to if/when a robot is used in a similar manner to a tool or instrument" [*Shaw-Garlock*, 2009, p. 250].

For the purposes of this work, I align myself closely with the first definition of a care robot which does not specify demographic of use. My definition of a care robot takes into account the care ethics perspective, thus, I claim that care robots may be defined as *any robot used in a care practice to meet care needs, used by either or both the care-provider or the care-receiver directly, and used in a care context like the hospital, nursing home, hospice or home setting.* With this definition I aim to make clear that a care robot is one which will be integrated within a care practice and consequently is integrated within the therapeutic relationship between care-giver and care-receiver. As was evident in chapter 3, the expression of care values is dependent upon care practices and the relationship between care-giver and care-receiver. In care contexts, such a relationship is referred to as a therapeutic one and is established through care practices like bathing. This is not to say that a care-receiver using a robot on their own is not using a care robot; the robot has been provided to them by a care-giver through a care institution. Those robots which are commercially bought and used in home settings, I classify as domestic robots. They may serve care purposes but by virtue of their acquirement, and without being integrated into the care relationship, they cannot be referred to as care robots. Thus, the care ethics perspective, and its emphasis on relationships, is also prominent in the very definition of a care robot.

In the current academic discourse, care robots are thought of as a class of robots to be used by nurses, as opposed to surgeons. This is due to the fact that care robots are thought to assist with physical care tasks, ADLs like lifting, bathing and feeding, all of which are fall within the realm of the nurse's responsibilities. Although there are scholars claiming that surgical robots should be categorized separately from care robots given that one is used by the physician and the other by the nurse [*Veruggio and Operto*, 2006; *Vallor*, 2011], I claim that surgical robots too ought to be considered care robots and consequently should also be evaluated according to the framework. This is so for multiple reasons. First, the surgical robot is one that is integrated into a therapeutic relationship. Second, the use of the robot not only changes the way in which the nurse assists in surgery (nurses must undergo training with the robot) but changes also the

ways in which nurses care for the patient post-operatively. When we consider care at the level of the institution, or care as a process, one must recognize the linkages between practices. Consequently, this robot, integrated into the healthcare tradition, although used predominantly by the surgeon, also changes the way in which care is practised by the nurse.

Examples of current care robots include: Secom My Spoon automatic feeding robot; the Sanyo electric bathtub robot that automatically washes and rinses; Mitsubishi's Wakamaru robot for monitoring, delivering messages, and reminding about medicine, and Riken's RI-MAN robot that can pick up and carry people, follow simple voice commands, and even answer them" [*Sharkey and Sharkey*, 2011, p. 267]. The RI-MAN robot has of late been replaced with RIBA which has an animal-like appearance as opposed to RI-MAN's humanoid appearance. Other examples include Paro, the baby seal for companionship; RP-7, the mobile robot for patient contact with a physician not geographically present in the hospital; HelpMate, for delivery of sheets, medication and food tray removal; daVinci®surgical robot; Titan's new Amadeus Composer™surgical robot[9]; the medication reminding robot developed by Susan and Michael Anderson [2010a] (University of Connecticut and University of Hartford respectively); the TUG and Helpmate robots used for the delivery of sheets, medication and/or food tray removal in the hospital. In more sophisticated care institutions, robots are used for security and monitoring purposes; "In the high-tech retirement home run by Matsushita Electrics, robot teddy bears watch over elderly residents, monitoring their response time to spoken questions, and recording how long they take to perform certain tasks. These robots can also alert staff to unexpected changes [*Lytle*, 2002].

With respect to the latter example, Sharkey states that these kinds of robots could have a significant impact on elder care in the home or in care institutions as well as caring for quarantined patients [*Sharkey and Sharkey*, 2011, p. 269]. Along the same lines that are a range of robots used for monitoring, the child-care monitoring robot PaPeRo for example [*Yoshiro et al.*, 2005]. "A range of monitoring robots to pick-up on dangerous activities like leaving on taps or gas cookers [*Orpwood et al.*, 2008], or cameras to determine if any elderly person has fallen over[10]. "Other technologies being developed for security robots, such as fingerprint and retinal recognition, could be useful for monitoring individu-

[9]This robot is the latest in surgical robots with a uniquely designed external robot and flexible instruments. See www.titanmedicalinc.com

[10]From the Toronto Rehabilitation Hospital: Our Journey in 2008-9: Annual Report. See http://www.torontorehab.com/About-Us/Corporate-Publication/2008-2009/hospital.asp

als, for example, visitors or an Alzheimer's sufferer, and helping prevent petty robberies" [*Sharkey and Sharkey*, 2011, p. 270].

As we can see, there is no capability exclusive to all care robots rather; they may have any number and range of capabilities from planar locomotion (vs. stationary) to voice recognition, facial or emotion recognition. Additionally, they may have any degree of autonomy, from human-operated (as in the surgical robot daVinci®) or varying degrees of autonomy (like the TUG robot for deliveries in the hospital which requires minimal human input or the RIBA robot intended for lifting patients without input from a human user). The specific capabilities of a care robot are dependent on the context, actors involved and the action which the robot is intended to fulfil. Thus, the definition of a care robot relies on the idea of *interpretive flexibility*, that a robot is defined by its context, users and task for use [*Howcroft et al.*, 2004]. This means that the same robot might be called by a different name if the robot is used for rehabilitation or for care purposes. The Hybrid Assistive Limb (HAL) is an example of this phenomena; the robot may be used in rehabilitation when worn by the patient or could be used to relieve the stress of lifting on the nurse. For the purposes of this work, a care robot will be defined as such according to its application domain (hospital, nursing home, home setting), its intended use (a care practice deemed as such according to its use domain) and its intended users (care-givers and/or care-receivers), in a care domain for a care practice.

If a robot's role in an ADL, fulfilled within a care institution, demands that the ADL be fulfilled in a certain manner (with compassion and empathy), does this, and more importantly *should* this change the requirements of the robot's capabilities? Should the care robot posses certain social capabilities by virtue of their role in care? This brings us to the question of how to know what capabilities to program the care robot with according to its place in the value-laden milieu of the care institution and its role (and associated responsibility) within a care practice.

4.5.1 Care Robots, Social Robots and Companionship

Given the distinction between affective and utilitarian purposes [*Shaw-Garlock*, 2009] made above, a further clarification must be made with respect to my definition of a care robot: a distinction between social robots and care robots. As was seen in chapter 3, care practices for activities of daily living (ADLs) were done in a way that also met the social and/or emotional needs of the care-receiver. Thus, it is impossible to definitively separate the meeting of needs along the dimension of physical care versus social care. Because of this, can we

claim that a social robot may in fact be a care robot when by definition social robots meet social needs exclusively? I suggest that it is not possible to qualify a social robot as a care robot. Care robots are those intended to meet the care needs of individuals. Such needs are fulfilled through care practices and rely on the therapeutic relationship. Thus, the relationship or bond between care-giver or care-receiver and the care robot is not an end in itself. Rather, is it a means to the end of meeting care needs. Alternatively, with social robots, the end goal is the establishment of a bond between robot and the human user; to make the robot a companion. Thus, the end of the social interaction between human and robot is the formation of a relationship rather than some other end; the social robot is engaged in social practices rather than care practices.

Added to this is the kind of relationship that is formed in the two interactions. The goal of the relationship between a social robot and human user is that of a companion, of companionship [*Breazeal*, 2004]. It is possible to suggest that a specific type of care need is that of companionship but the question remains whether or not this is the kind of need met by the nurse. In response, the relationship between care provider and care-receiver is a therapeutic one, explicitly differentiated from one of companionship: "the therapeutic relationship differs from a social relationship or friendship in that the needs of the client always come first. The nurse is in a privileged position because of the trust the client puts in the nurse and because of the power imbalance" [*Nurses of Ontario*, 1999b, p. 8]. The differences have to do with the explicit recognition of the asymmetry in power and the ends which the relationship serves (care needs along a broad spectrum). Moreover, the institutional context within which the relationship is established plays a role in creating the boundaries of the therapeutic relationship. It should be noted that as health care and nursing practices are changing so too is the context of care along with the roles of the nurse. Care is being provided in the community and in certain instances in the homes of patients which shifts the activities of the nurse and has the potential to blur the boundaries of the relationship:

> *The nurse may be taking on a stronger counselling role with clients and/or focusing or concentrating on psycho-social issues. The nurse may need to clarify the role for him/herself and explain that role to the clients who may be expecting a more traditional role from the nurse.... In some instances the role of the nurse can include teaching clients how to grocery shop or do banking, or a community nurse may be involved in planning meetings in the community. The nurse needs to be clear with the client that this activity is part of the nursing role*

> *and not an extra activity outside of that role.* [*Nurses of Ontario*, 1999b, p. 10 - 11]

From the above, it is quite clear that the boundaries of the therapeutic nurse-patient/client relationship must be managed through all of the actions and attitudes of the nurse regardless of the context within which care is being provided. Another manner in which the boundaries of this relationship are made clear are evident in its termination: "at the beginning of the relationship, the nurse establishes with the client, family and health team an estimated period of time that the relationship will last. The health-related goals and need of the client determine when the relationship will end" **[*Nurses of Ontario*, 1999b, p. 11].** Thus, we can see clearly at this point the many ways in which a therapeutic relationship differs from a companionship and as such a social robot defined by its initiative to foster companionship cannot be considered a care robot.

This is not to say that care robots will not or cannot have social capabilities. Take the example of Cody, the diet assist robot developed at MIT, it is a robot with the end of keeping the user motivated towards their goal of weight loss and the bond formed is an integral component of this. Cody is a care robot with social capabilities. Its end-goal is weight loss and not strictly becoming a companion. Alternatively, Paro aims to be a companion for users. Empirical evidence shows the use of Paro with elderly residents in a nursing home to have outcomes that are beneficial for the patient in terms of a reduction in stress and anxiety. The same results one would observe through animal therapy. Added to this is a boost in self esteem that comes from the feeling of being needed, something that is (questionably) generated by the robot. But such a reduction mirrors the effects of animal therapy so perhaps the problem here is the way in which Paro is framed, as a medical device rather than as merely a means for meeting social needs of persons. The same feelings of self-esteem could be observed when a patient interacts with a human, an animal, reads a good book or even does a yoga class. In these instances, Paro is an alternative to a real animal if one is not able to take care of 'living' one. It is an alternative to doing a yoga class when a human is not physically capable of doing so. It provides the same benefits but is not intended to be anything besides a companion.

Discussing the boundaries of the therapeutic relationship reminds us of the many attributes of the nurse and the many values expressed in, and through, a care practice. This now brings us to an important question: how do we translate values and human attributes into capabilities of a care robot?

4.6 Designing a Care Robot According to the Values in Care

If we consider that a care robot will be used to fully or partially replace a human care provider in a given care practice then the robot ought to be evaluated along the same lines as the human care-giver. Not only this, but for the prospective design of care robots, designers should aim to endow the care robot with those capabilities and attributes that make a human care-giver a good one. In terms of design, for Engelberger, "the better a robotic aid can match tasks with capabilities, the more useful it will become" [*Engelberger*, 1989, p. 211]. Thus, one must understand the human capability in order to program the robot accordingly. Table 4.1 provides an outline of the task of strictly translating certain human capabilities into robot capabilities independent of context, practice and actors. I have chosen to list the human capabilities according to the necessary elements indicated by care ethicist Joan Tronto: attentiveness, responsibility, competence, and responsiveness. As we may recall from the previous chapter, these elements provide the criteria for ethically evaluating whether or not care is *good* care. What's more, they may loosely be considered attributes of the *good* care-giver and as such provide a starting point for the evaluation of a care robot.

Table 4.1: The moral elements of a care practice (Joan Tronto) aligned with corresponding robot capabilities.

Moral Element	**Translated into Human Capability**	**Translating Human Capability into Robot Capability**
Attentiveness	Capability and capacity of recognizing the dynamic needs of a patient	Robot Vision; Facial recognition, Emotion perception and recognition.

Responsibility	Closely aligned with trust; requires an understanding of what one is doing and why. Capability of identifying adequate response to needs and delegation to meet them. Presumes individual will be held accountable and liable in the case that something goes wrong.	Emotion recognition as a means to establish trust (mimicry). Robot knows what it's doing, how and why it's doing it . Knows what it sees, saw and can see. Can predict the consequences of its actions. Know what it can and cannot do. Can acquire new knowledge (learning robots/algorithms)
Competence	Capability and capacity of executing an action to fulfil the identified needs in a skilled manner	Safety, efficiency and quality of task execution (speed of robot, stopping distance, emergency shut off/power down, materials used for robot). Force feedback and tactile perception.
Responsiveness	Capacity and capability to engage with the care-receiver regarding the meeting of their needs (can be physiological, verbal or other cues given by the patient to the care-giver)	Multi-modal communication platforms: verbal and non-verbal/paralinguistic communication paradigms, hand gestures, proxemics, head gestures, lights, eye gaze, facial expressions, force feedback and tactile sensation.

According to Table 4.1, one might suggest that in order for a care robot to be considered attentive, it must be able to, at the very least recognize the face of the user (care-giver and/or care-receiver). When discussing such a task, however, using examples we can see how a difference in the context, practice and/or actors changes whether or not the robot capability conforms with the requirements of the care practice or if it presents itself to be ethically problematic. For example, let us continue to discuss the element of attentiveness and endowing the robot with facial recognition (so that the robot recognizes the patient) in order to say that it meets even a portion of the requirement of attentiveness. What happens when we observe this capability in more than one practice, more than one context and more than one robot type? In a practice like lifting, if the robot were autonomous, a capability like facial recognition seems appropriate to guarantee that the robot is aware of the individual patient it is lifting and further that the robot is privy to other information about the patient which it can access once it has the patient's facial information (the context is the hospital or nursing home where the robot would have more than one patient to lift). What if we take the same practice, in the same context but we look at a human-operated robot, the exoskeleton. First, is facial recognition required if the robot is essentially acting as an assistant to the nurse? Second, if the robot were still required to be attentive, perhaps attentiveness in this practice refers not to the robot's need to recognize the patient but perhaps the robot would enhance the attentiveness of the care provider – attentiveness as defined for the network and for the practice. If the robot has facial recognition capabilities, it could indicate pertinent information to the nurse prior to lifting, it could collect relevant information pertaining to that patient at that time and could save it for future use. Switching contexts, if the autonomous robot were in the home where only one person is being lifted, there is no need for facial recognition nor the information (or attentiveness) acquired through such facial recognition.

The example with the human operated robot introduces the idea that a robot need not be endowed with certain capabilities if their role does not demand it. Would a surgical robot need to be capable of facial recognition of the patient? If the robot were autonomous one wouldn't hesitate to agree; however, when the robot is human-operated the element of attentiveness is always in the realm of the human's role and responsibility. This is not to say that attentiveness may not be distributed throughout the network. The robot is delegated a role in which its responsibility is to enhance the capabilities for attentiveness of the human care-giver rather than take-on the role of attentive care-giver. This means also that attentiveness for a particular practice must be understood as it differs between practices (which can be seen through the discussion of the

various care practices in chapters 6 and 7). Attentiveness for lifting refers to the mechanics of lifting as well as its affiliation with the establishment of a relationship between care provider and care-receiver, and the overall process of care (as was discussed in chapter 3). Alternatively, attentiveness as it pertains to surgical intervention is specific for the type of surgery but is nonetheless defined by the surgeon's perception of physiological cues and bodily responses to surgical intervention rather than in the establishment of a bond. What's more, the meaning that is attributed to the endowing of facial recognition on the autonomous robot makes a statement about the robot's moral status. The very idea of endowing the autonomous robot with facial recognition entails that the robot is also delegated a role that requires such a level of attentiveness. When the robot is delegated the task of being attentive without passing along that information to the human actor, the robot is then *responsible* for what is to be done with that information. This is counter to when the human actor is responsible for being attentive with or without the assistance of the robot.

Thus, the interpretations of the moral elements themselves are dependent on the context, the practice, and the actors involved. Isolating the moral elements from these components is logically incoherent given that the meaning and prioritization of moral elements is dependent on all the other factors. In short, we are talking about an entire network of human and non-human actors which all contribute in some way to the manifestation of values. Thus, all are delegated a role and responsibility. Which of these roles and responsibilities can be aligned with a moral status within the network? What's more, assigning robot capabilities without understanding how a capability is linked with the assigned role and responsibility the robot will take on then becomes quite problematic. In other words, the care robot is not evaluated against the criteria of a human care-giver necessarily, but is evaluated according to a broader set of criteria, one that encompasses the role and impact it may bear on the network of actors and the ethical nature of the care practice.

When it comes to design recommendations for robots based on ethical criteria, Sharkey and Sharkey are pioneering such a feat [*Sharkey and Sharkey*, 2011]. Based on an analysis of human rights, the authors focused on balancing the right of safety and the right of privacy/confidentiality with users and robots used in care applications. They conclude that a robot ought to indicate its presence when entering or exiting a room and further that the robot ask permission to enter, a kind of robot etiquette. Additionally, they recommend that when the robot is engaged in the recording of information the robot should indicate this through the use of lights: "a robot should always have an indicator when it is recording or transmitting images" [*Sharkey and Sharkey*, 2011, p.

278]. One might suggest such communication between user and robot manifests the element (or value) of reciprocity – that the robot actively engages with the user such that the user can respond to the action of the robot. If the user is uncomfortable with the robot's presence, perhaps they can leave the room. In the same vein, if the user would prefer not to have their confidential information recorded, they may refrain from speaking. But this leaves a great deal to be decided by the user, the user is delegated the responsibility of deciding what they want to share and once they have decided so, the information is out of their hands and into the infrastructure of the robot. Perhaps in care contexts delegating additional responsibilities to the user is not always the best idea, elderly or child patients may not be the best judge or may not understand the technology and its limitations or powers to be able to adequately decide. Thus, the demographic for such design recommendations and the practice in which such designs are used again play a dominant role.

Sharkey and Sharkey illustrate through their exploration of design recommendations the need for taking the larger picture into analysis; what is the practice, who is the demographic, what kind of information are we talking about, etc.? I conclude from this as well as the above analysis of translating the moral elements into robot capabilities independent of external criteria, that design recommendations must take into account not only the values at stake but an in-depth study of the practice into which the robot is stepping. Such a study is dependent on the context, the actors, the distribution of roles and responsibilities, and the resulting values manifest through the interactions and actions of all actors. Thus, the framework for evaluating a care robot must incorporate all such values in order to evaluate the robot's capabilities in ethical terms.

4.7 Conclusion

This chapter was meant to introduce the field of robotics and to outline the definitions, capabilities and features of robots in general and care robots in particular. The main goal of this chapter was to illustrate the difficulty in translating human capabilities into capabilities of the robot. The capabilities of a human care-giver cannot be translated into robot capabilities independent of an understanding of the task for which the robot will be used and the context within which the robot is used. This is currently, for my own research, the main problem in the ethical evaluation of care robots: without understanding the practice within which the care robot will be applied or the context within which it will be used, one is neither capable of evaluating the ethical implications of

the use of a care robot for a specific care practice nor of assigning capabilities to a future care robot. It follows that, one is not capable of truly understanding the effect the robot may have. Consequently, I begin to set the stage for the various components of the framework, namely that context and practice must be made explicit if one is to understand the impact the care robot will have. With this in mind, I now turn to demarcating the components of the care-centered framework used for the evaluation of current care robot designs as well as in the prospective design of future care robots.

Chapter 5

A Framework for Evaluating the Design of Care Robots

Together, technology and care form a praxis in the Aristotelian sense, i.e., a shared process of action, structured by the orientation towards a common end: the good of man. [*Widdershoven*, 2002]

5.1 Introduction

The research question guiding this work addresses the ethical issues pertaining to the design of care robots, namely *how care robots may be designed and implemented in care contexts in a way that supports and promotes the fundamental values in care?* The reasons for addressing issues of design are manifold, as has been discussed in chapter 1. To target the design process of care robots I use the blue-print of the Value-Sensitive Design approach – I adhere to certain components and digress in others. Namely, I adhere to the conceptual investigation of the values of ethical importance, an analysis of the technical content of the system, and incorporating certain empirical insights. The values chosen, however, are not those of Batya Friedman and colleagues [*Friedman et al.*, 2006] but rather are those values which form the buttress of the healthcare tradition (in Western cultures) interpreted through the lens of the care ethics tradition [*Tronto*, 1993]. What's more, the empirical insights are those taken from the

literature rather than results from contextual experiments with a robot prototype. The task now is to outline the culmination of these investigations in the form of a framework to ethically evaluate care robots according to their impact on the expression of care values and their potential to shift the distribution of roles and responsibilities within a care network/care practice.

By summarizing and synthesizing the findings from the previous chapters I arrive at a conceptual framework for evaluating artefacts and practices from the perspective of good care. I refer to this framework as the *Care-Centered* (CC) framework given the focal role the care perspective plays in both its creation and methods of use. This framework is then used in two separate methodologies to accomplish divergent goals; 1. to evaluate the design of current care robots and, 2. to steer the design of future care robots. Each methodology differs in multiple respects. The former, which I refer to as "retrospective evaluations of care robots" (REC), is used to evaluate both a current care robot and a current care practice. Thus, the technology has already been made and the evaluation occurs downstream in the design process. The recommendations resulting from such an analysis may be used for the improvement of future designs of the care robot in question as well as other robots that may be made. What's more, the evaluation also allows for a critical analysis of current care practices and whether or not a robot will in fact maintain the same standard of care, will minimize the current standard, or will enhance the current standard. Alternatively, I refer to the latter methodology as the "Care-Centered Value-Sensitive Design" (CCVSD) Approach. It is a prospective methodology beginning further upstream in the design process, at the moment of idea generation. Hence, there is no artefact made and thus the methodology is intended to shape the future design, design process and implementation of a care robot. Although each methodology is presented as being separate from the other; however, the two are inter-related in that the findings from the REC inform and substantiate the CCVSD approach.

The care centred framework and the methodologies for using the framework, in either a retrospective or prospective manner, both pay tribute to the central thesis in care ethics, namely that the care perspective provides an orientation from which to begin theorizing as opposed to a pre-packaged ethical theory. The framework articulates the components which require attention for analysis from a care perspective while the methodology indicates how these components are to be dealt with. The framework consists of five components: context, practice, actors involved, type of robot, and manifestation of moral elements. Each of these components will be described in detail for understanding their place within the framework from the care ethics stance. This chapter acts as a 'user manual' for designers and ethicists alike who wish to engage in the

retrospective evaluation of a current care robot. The same kind of 'user manual' will be presented for prospective analysis in chapter 9. The driving force behind the creation of both the care centred framework and the methodologies for its use, are to stimulate and guide the reflection of ethicists and engineers in a deliberative manner through the lens of the care ethics perspective in which context, practice, responsibilities and roles of all actors in a care practice are brought to the fore.

5.2 The Care-Centered Framework

To reiterate, the Care-Centered (CC) framework responds to the three rationales mentioned at the beginning of this work pertaining to why one ought to target the design of care robots: the early stage of development of care robots, the lack of regulatory frameworks and the relationship between designers' assumptions pertaining to norms and values and the resulting care robot. More specifically, the framework aims to translate ethical requirements for engineers and designers into a tangible format for inclusion in the design process of care robots. The hope is to foster trust in the design process of these robots as well as the resulting care robots. Thus, it is important to clarify here that the framework is intended to foster ethical reflection of the human actors involved in the design process of the robot, the designer, engineers and ethicists. I am therefore not making any kind of claim about the internal ethical deliberation or reasoning of the robot per say, this will be taken up in chapter 8. What's more, the framework addresses the belief that assumptions, norms, values and biases are built-in to the robot and attempts to explicitly and systematically account for this.

The care orientation structures both the framework and the methodologies for its use retrospectively and prospectively. To claim that the framework adheres to the care orientation refers to the idea posited by Gilligan and reinforced by Little that care is an orientation from which one begins to theorize [*Gilligan*, 1982; *Little*, 1998]. The difference between an orientation and a theory comes from the interplay between propositions and the execution of said propositions. The care stance does not assert that any one proposition remain at the top of a hierarchy for any given ethical problem. Instead the care orientation asserts that one begins with a certain perspective both for deliberation as well as uncovering the moral problem. This perspective or stance is described in terms of an:

> Emphasis of concern and discernment (*to notice and worry more, say, about the dangers or interference rather than the dangers of*

> *abandonment*), habits and proclivities of interpretation (*the proclivity, say, to read 'the' moral question presented by a situation in terms of rights rather than responsibilities*), and selectivity of skills (*to have developed, say, an ease of abstraction more than an attunement to difference*). [*Little*, 1998, p. 195]

An emphasis on concern for abandonment rather than interference aims towards directing one's attention to the ways in which people can relate to each other rather than be left free [*Verkerk*, 2001]. Interpretation in terms of responsibilities rather than rights recognizes the commitment on the part of both the care-giver and the care-receiver to act together for attaining the desired outcome rather than the obligation of a care-giver based on the rights of a care-receiver . Interpreting in this way, draws our attention to the significance of the therapeutic relationship in care. For Tronto, the relationship between care-giver and care-receiver acts as a buttress for the provision of good care. It is the vehicle through which the care-giver comes to know what the needs of the care-receiver are, the forum where trust is established, and the conclave in which both the care-receiver and care-giver learn about themselves. As such, it is a value in care but also acts to safeguard may of the other values in care like human dignity, self-sufficiency, autonomy, freedom from bias, trust, connection, reciprocity and human emotions like compassion and empathy.

The CC framework is a conceptual framework used for evaluating both the care robot and the care practice into which the robot will enter. It emphasizes the components demanding attention from the care ethics perspective, namely the context, the practice, the actors involved and how care values are manifest. From the analysis of care values and the requirements of a good care institution (according to care ethicist Joan Tronto, 1993) in chapter 3 it has become apparent that the standards for evaluating good care are dependent on: context, the individual care-receiver and care-giver, fulfillment of care values, and the meeting of needs through care practices. It follows that a framework for the ethical assessment of care robots ought to address these dimensions. As such, one performing an analysis according to the CC framework has a clear picture for emphasizing concern and discernment with respect to: the robot's capability to interfere with the manifestation of values or the forming/strengthening of a relationship; interpreting and re-interpreting the moral question in terms of responsibilities rather than rights; and, develop the necessary skills for continued analysis along these lines.

The framework is general in that it cannot standardize the creation of care robots in terms of the technical content every care robot ought to embody. This

is not possible for a variety of reasons. The first being that the capabilities of the robot will differ depending on the practice for which the robot is intended. A robot designed for delivery of sheets will have distinctly different capabilities from a robot designed for feeding and will be evaluated accordingly. Second, the capabilities of the robot and the robot's control will also differ depending on who the robot is intended to be used by (the care-giver, the care-receiver or a combination of the two). In other words, not every care robot will have the same capabilities or end-goals which make standardization of the care robot's design problematic. My goal with the framework is to standardize along a different line – to ensure that every care robot is designed according to its potential impact on the values and components of care that render a care practice 'good' regardless of the robot's capabilities and/or the end-goals of the practice.

Table 5.1 outlines the CC framework. The components comprising the framework are then discussed in greater detail. Added to this, it is of the utmost importance to recognize that care must be understood in its totality, as a practice integrated into a holistic process rather than an unlinked series of actions or tasks to be fulfilled. This means that the framework and methodologies must point towards an understanding of care in this sense. Accordingly, the framework demarcates the components demanding attention and through the methodologies for its use the relationship between these components and the overall process of care are made explicit for analysis. This aspect will be given greater attention in the discussion "Applying the Framework" (section 5.3).

Table 5.1: The Care-Centered Framework

Context – hospital (and ward) vs. nursing home vs. home
Practice – lifting, bathing, feeding, delivery of food and/or sheets and/or medications,
Actors involved – nurse and patient and robot vs. patient and robot vs. nurse and robot
Type of robot and robot capabilities – assistive vs. enabling vs. replacement
Manifestation of care values – Attentiveness, responsibility, competence, responsiveness

5.2.1 Context as a Component

Firstly, one must identify the context within which the care practice is taking place. For example, the specific hospital and the ward vs. a nursing home vs. a

home setting. The context within which the care practice takes place is important for a variety of reasons. Recent research indicates a relationship between religious beliefs and one's acceptance of using robots in care-taking roles [*Metzler and Lewis*, 2008]. Metzler and Lewis are investigating the hypothesis that when one believes in "a god" they may not be as inclined to accept human-robot interaction with life-like robots at an intimate level. Thus, the design of a robot for a Catholic hospital ought to take this kind of research into consideration for the appearance of the robot. Similarly, the context in terms of one hospital ward or another is also of great importance when designing the robot. Research done by Bilge Mutlu of the University of Wisconsin, Madison [*Barras*, 2009] shows how the same robot (the TUG robot) used in one hospital was accepted differently depending on the ward. Workers in the post-natal ward loved the robot, while workers in the oncology ward found the robot to be rude, socially inappropriate and annoying. The same workers even kicked the robot when they reached maximum frustration.

Specifying context in terms of a nursing home vs. a home setting is also of importance given that the prioritization of values differs. For example, certain practices, like lifting, in the nursing home place efficiency as a high priority (and even more so in the hospital) while in the home setting there may not be the same time constraints. In addition, certain practices, like bathing, in a home setting may not require the same demand for privacy as the hospital or nursing home setting given the lack of other patients around. What's more, context plays an integral role when we consider the need for establishing and/or maintaining the relationship/bond between care-giver and care-receiver. In a home setting, the relationship has already been formed between care-giver and care-receiver (they are often family members or spouses who already have a deep understanding of the preferences and routines of each other), thus a robot may not pose the same ethical concerns. Alternatively, in a hospital setting where daily practices are intended to establish and/or strengthen the bond between care-giver and care-receiver as well as to learn about preferences and styles, each practice serves a pivotal role in this process.

5.2.2 Practice as a Component of the Framework

A care practice is defined here as an identifiable moment in which the actions and interactions between and among actors (human and non-human) result in the manifestation of values. The carrying out of the practice is also how we come to understand the distribution of roles and responsibilities and thus the practice is one of the central foci in the evaluation of a care robots's impact.

Examples of practices are lifting, bathing, feeding, fetching items, delivery of medications/food/x-rays/sheets to the room or to the nurse, personal communication, social interaction, games and activities like singing songs or painting. Each care practice results in the manifestation of care values; however, they mean very different things depending on the type of practice. For example, the value of privacy when discussing the practice of bathing refers to maintaining the privacy of the body of the individual – not allowing external parties to view the patient's body. Privacy for the practice of personal communication when thinking of a vulnerable patient sharing their feelings and fears amounts to keeping this information secret between the individuals privy to the discussion . Thus, privacy in both instances refers to a respect for the dignity of the care-receiver, through the non-disclosure of information, bodily or verbal. The detailed description of each practice is further dependent on the context within which care occurs as well as the actors or demographic involved.

5.2.3 Actors Involved

Authors Borenstein and Pearson point to the differences in needs between demographics and thus how a care robot's capabilities ought to reflect this – that certain demographics will have needs that differ from others. Take robots in child care and robots in elderly care as an example. One may suggest that the element of 'play' is crucial in the design of a robot in the first instance while natural language communication is of crucial importance in the design of a care robot for the second instance [*Borenstein and Pearson*, 2011]. In the same vein, authors Sharkey and Sharkey also point to differences between these two demographics only this time they stress that significance of human contact in the cognitive development of children and the significance of human contact in the reduction of stress, increase in global cognitive functioning, and decrease in risk of developing dementia in elderly populations [*Sharkey and Sharkey*, 2011]. Hence, human contact is important in both instances, but for very different reasons.

From the care orientation, the actors involved are of great significance for structuring moral deliberation. One of the most important findings to come from the care ethics perspective is the ontological status of humans as relational. Its significance for this work lies in recognizing that the care practice which a robot will enter involves a network of human (and non-human) actors in relationship. The robot then has the potential to shift the roles and responsibilities distributed within these relationships as has been stressed already. What's more the robot will engage in relationships with any number of actors in the network. The

actors involved in a care practice will differ between a hospital, nursing home, hospice, children's hospital, or home context. In any setting, the patient may be completely dependent on others for certain practices. In a hospital setting, the actors involve the patient and any number of healthcare personnel. If the patient is receiving care in their home perhaps the actors involved are family members or a visiting nurse who is not present on a daily basis. In a home or hospital setting a patient may fulfil certain practices on their own prior to a robot assisting. This does not mean the care-receiver is entirely on their own, in the atomistic sense, but rather that the robot may be delegated a certain portion of the role of the care-receiver (as is the case with a feeding robot like Secom's MySpoon). In each instance care-giver and/or care-receiver enters into (a) relationship with the robot. By entering into a relationship with the robot I do not equate human-robot relationships with human-human social relationships where empathetic and trusting bonds are formed. Rather, I am referring to a superficial/simple interaction between robot and human, which over a long period of time, and with the added dimension of expressing values, can be referred to as a relationship more so than an interaction. It is true that certain robots may have social capabilities and thus convey feelings of empathy but I hesitate to refer to such interactions as forming a relationship. I defer to the work of Turkle here and claim that the one-sidedness of the relationship makes it possible to equate it to that of a human-human one. Regardless, this component is meant to highlight the roles and responsibilities attributed to actors prior to the robot entering the scene.

It is important to remember too that the human actors are not acting alone to manifest values. They work together with each other but also with technologies already in use in the healthcare system. In nursing and technology studies, technologies have often been considered extensions of the nurse's body or self [*Sandelowski*, 1997]. Nurses become so skilled at using the technology they do so without being distracted by the technology's presence. Of course, we must recognize there will always be a time in which the nurse learns to use the technology and appropriates it into his/her daily routine; the technology being an extension of the nurse does not happen automatically. What's more is that the nurse's role is one that incorporates the use of technologies in a variety of ways from the mechanical bed to heart monitoring devices. Thus, technologies are not only extensions of the nurse but they also mediate the relationship between the nurse and the patient shifting both the role and the responsibility of the patient and nurse in order to include the technology in the equation.

In other words, we are not speaking of interactions or relationships that occur without the use of technologies. Therefore, the question is not what happens

when a care robot enters the nurse-patient relationship that is devoid of any technologies. Rather, we are speaking of a context within which technologies are already employed to a high level and the question is how will a care robot alter the existing relationships; what relationships will it form and what relationships might it interfere with?

5.2.4 Type of Robot as a Component of the Framework

The typifying of robots is done in many different ways. Some consider a type of robot according to the domain for which it is used; industrial vs. rehabilitation vs. military vs. search and rescue robots [*Veruggio and Operto*, 2006]. For others, types of robots may be in terms of industrial robots vs. service robots vs. personal robots. This classification of robots is dependent on the amount of human interaction the robot will have and the predictability or structuring of the environment within which the robot is working. To specify for the purposes of the framework discussed here, the manner in which I classify 'type of robot' has to do with the how the robot will be used among the human actors – how a role and responsibility is delegated to a robot. For example, an enabling robot is one which enables a human to perform an action previously not possible without the robot or, the robot enhances the human's performance during a task – the robot and human are working together toward a goal but the human is in control of both him/herself as well as the robot. Thus, the responsibility for accomplishing that role is a shared effort with the robot perceived in an instrumentalist way, as a tool. Robots of this type are tele-presence robots like the RP7, surgical robots like Intuitive Surgical's daVinci®, or exoskeletons like the Hybrid Assistive Limb (HAL). A replacement robot is one that fulfills a practice in place of the human. The role of the human and the associated responsibilities are delegated fully to the robot. An example of this type of robot is the RI-MAN autonomous robot for lifting, Secom's MySpoon automatic feeding robot, or the Sanyo electric bathtub robot that automatically washes and rinses. An assistive robot is one which aids a human in performing an action by providing a portion of the practice without the direct input of a human operator and is thus delegated a partial role and a partial responsibility. This robot differs from an enabling robot in that it does not require consistent input from a human but rather can execute a practice once given its command. Examples of this kind of robot are the TUG or HelpMate robot used for deliveries in hospitals or the Mitsubishi Wakamaru robot for monitoring, delivering messages, and reminding about medicine. In the case of the delivery robots, the role and responsibility of the delivery is shared between the robot and the human deliverer/receiver;

however, the robot fulfills many steps without any input from a human user (e.g., navigating through hallways and corridors).

In addition to the type of robot identified, it is important to list the robot's capabilities as they too play a role in the ethical analysis. Will the robot be collecting information? Will this be stored? If so, for how long, how is it encoded and who will have access to it? Will the robot be linked to a telecommunications network during its functioning and thus be subject to hackers? Will the robot be engaging in physical interaction with a human (as in the case of lifting or bathing robots)? Will the robot be speaking and if so at what volume will be acceptable in a given hospital ward (or nursing home)? Will the robot autonomously return to its battery station and if so what happens when it cannot make it there on time, what de-fault mechanisms will be in place if the robot breaks down in the middle of a hallway? Many capabilities of the robot will demand specific ethical treatment to comply with the ethical standards of the healthcare tradition.

Alongside robot type and robot capabilities, the appearance of the robot is of significant importance. Arras and Cerqui report on their Swiss survey that only 19% (n=2000) of the participants preferred a human-like appearance [*Arras and Cerqui*, 2005]. Dautenhahn [2005] report that although human-like communication is desirable for a robot companion (a personal service robot), human-like behaviour and appearance are less important. The appearance of the robot is of paramount importance for the expectations of users as well as their comfort when interacting with the robot. Moreover, the appearance of the robot may provide insight into the task it is intended for; "people expect a robot to look and act appropriately or different tasks"[*Goetz and Kiesler*, 2002]. Thus, the appearance of the robot must also be taken into consideration in the discussion of the robot's capabilities and features as it too will impact both the acceptability of the robot (the domestication of the robot) as well as role of the robot as perceived by the human users.

In general, this component serves to prepare the reader for evaluations of real world robots rather than speculative robots [*Smits et al.*, 1995; *Nordmann and Rip*, 2009] thereby grounding the reader in the technical capacities of a care robot as a pre-requisite for plausible evaluations [*Swierstra and Rip*, 2007; *Lucivero et al.*, 2011].

5.2.5 Manifestation of Care Values

While many care ethicists make clear the range of values and principles that provide a normative account for care [*Vanlaere and Gastmans*, 2011; *Little*, 1998; *Ruddick*, 1995; *Noddings*, 1984], they fall short by providing a systematic

way to visualize and evaluate these principles and values. The vision presented by Tronto allows for a perception of care as a practice with stages [*Tronto*, 1993], which provides the most enticing conceptualization for engineers to work with [*van Wynsberghe*, 2012]. This component of the framework refers to how the values are manifest within such a defined care practice; in a given context, with the specified actors involved. This proves quite difficult when we are aware of the many values which are expressed at any given time during any given care practice. Thus, before I define the manifestation of values, certain values need to be selected.

As I have argued thus far, uncovering the many layers at work in a care task leaves us with an enriched vision of said task rendering it a care practice. This is so for a variety of reasons. The first has to do with the context within which care is happening – placed in a context structured by values, care takes on a deeper meaning. Second, care is fulfilled through a variety of skilled actors (human and non-human), each delegated a specific amount and type of responsibility. Third, the care practice is an expression of attitudes and values resulting from the actions and interactions of actors in the network. Fourth, good care requires an understanding of the multi-dimensional needs of care-givers and care-receivers and fifth, it is all of these elements and variables that guide the ends of the care practice as well as forming a means for the evaluation of care.

The linkages between values and the manner in which they are intertwined with actions is most evident when speaking of the value of trust. Trust is a need for both the care-giver and care-receiver in order to fulfil care practices and yet it comes into being through care practices. This was noticeable in the care practice example of bathing in chapter 3. Through bathing, trust was established and yet without trust bathing might not occur in the first place. This also points our attention to the holistic vision of care. Certain components of a care practice, like trust, are required further along in the process of care. While establishing trust through bathing, the element of trust comes into play later on such that the care-receiver will be more honest and open about their symptoms and will comply with their care plan IF they trust their care-giver.

The next task is to articulate the fundamental values that ought to comprise the framework for engineers. Essentially, care is a cluster of values that come into being through the actions and interactions of actors in a care context (see chapter 3 for a list of care values). Creating a standardized framework to guide the promotion of these values which applies to any care context, task, care-receiver or care-giver reveals itself to be quite problematic given the range and variety of care values discussed. Moreover, their ranking and prioritization is dependent on the context (e.g., one hospital domain or another vs. a nurs-

ing home) and practice (e.g., lifting vs. bathing). What's more, to claim that human dignity, compassion or respect for power are values to be embedded in a care robot offers nothing for the designer in terms of the robot's capabilities. However, in the care ethics literature, alongside values, needs play a central and crucial role in the provision of good care. The needs of the patient mark the starting point of the care process and the process then revolves around a care-giver taking steps to meet these needs. Understanding the multiple layers of needs, the many ways in which they might be fulfilled, the preferences for one way over another, and the divergent needs between individuals adds a further complexity to the meeting of needs. Moreover, the care-giver has needs too! Needs in terms of resources, skills, responsiveness from the care-receiver to understand when needs have been met as well as their own personal needs.

Given the central role of needs in a care context, what might the relationship be between needs and values? Although many authors have written on the subject, little consensus can be found. I suggest then that the values in healthcare are given their importance for their role in meeting needs. This corresponds with Super's conceptualization of the relationship between needs and values: "values are objectives that one seeks to attain to satisfy a need" [*Super*, 1968, p. 189-190]. This means that, the value is the goal one strives towards and in so doing, intentionally meets a need. In other words, we begin with needs, and the values represent the abstract ideals which, when manifest, account for the needs of individuals. It follows then that a framework for designing care robots ought to address the meeting of needs. Unfortunately, I've just shown how multifaceted and intricate needs are for the care-giver and care-receiver. What's more, according to the field of care ethics, it is neither possible nor advisable to outline a series of needs which pertain to all care-givers, care-receivers or care institutions [*Vanlaere and Gastmans*, 2011; *Tronto*, 2010]. While useful for policy, it goes against the vital element in care – that of the individual and their unique, dynamic needs. In other words, care is only thought of as good care when it is personalized (Tronto, 1993). There is, however, a solution to this barrier. It is possible to delineate a set of needs for **every care practice**. To that end, I propose using the criteria from the care ethics literature explored in chapter 3 to uncover the specific needs in any care context which ought to be met in order to ensure good care. More specifically, the moral elements given by Tronto: *attentiveness, responsibility, competence* and *responsiveness*. (See Table 5.1 for the elements and their definition). These may be considered general needs of any care practice which must be fulfilled in order to meet the requirements of good care. To recapitulate, together the phases and the moral elements make up a care practice. The practices are values working together

and the vehicle for this is the moral elements. If we assume a care practice ought to proceed according to Tronto's phases than the needs for every care practice are the corresponding moral elements. It is therefore these elements that ensure the promotion of care values. Consequently, it is these elements – attentiveness, responsibility, competence, responsiveness – that make up the value-base portion of the framework.

Table 5.2: The Moral Elements and their definitions according to Joan Tronto.

Attentiveness	capability of recognizing the dynamic needs of a patient
Responsibility	closely aligned with trust; requires an understanding of what one is doing and why; capability of identifying response to needs and delegation to meet them
Competence	capability of executing a means/action to fulfil the identified needs in a skilled manner
Responsiveness	capability to engage with the care-receiver regarding the meeting of their needs

With this suggestion, there are two assumptions being made; that every care practice will ALWAYS have the moral elements as needs, independent of the care-giver and care-receiver, and that all values are subsumed within the moral elements. Using the practice of bathing as an example to illustrate the first assumption, I am making the claim that this practice will ALWAYS require attentiveness, responsibility and competence on the part of the care-giver and will ALWAYS require a reciprocal interaction between care-receiver and care-giver for determining whether or not the needs have been met, no matter what the context is. I cannot make the same claim if I took the valued action of touch as an example. Touch is important in and for bathing; however, I cannot say that touch is a valued action that ought to be present in every care practice. For the delivery of sheets or food to the room, there is no reason to assume that touch is required. Or, for social interaction in which a care-giver sits with the care-receiver to talk, there is no grounds on which we can say that touch is required. I can, however, make the claim that attentiveness is required and refers to recognizing the preference of the care-receiver in terms of touch. Thus,

the moral elements are needs which are independent of the care-giver and the care-receiver but which are the necessary and sufficient criteria of a good care practice.

The moral elements are, however, dependent on the context and the specific practice for their interpretation and prioritization. If we were to compare the practice of lifting with the practice of feeding we would see how the element of competence is uniquely interpreted in each practice – skilfully bearing the weight of another without dropping or causing pain vs. skilfully coordinating timing with placement of food and utensils. In terms of context, the practice of lifting in the hospital requires greater efficiency than the practice of lifting in a home setting where time may not be as much of an issue. Thus, although the moral elements must always be present, context and practice still play a crucial role in their interpretation, prioritization and manifestation. This recognition again reinforces the care ethics stance.

For the second assumption – that all the values are subsumed within the moral elements – the values are often analogous to a phase or moral element or are expressed through the manner in which an action takes place. The value of patient safety is fulfilled through the competent completion of a practice (the phase being care giving and the moral element being competence). The valued action of touch requires attentiveness on the part of the care-giver for determining when and to what degree touch is considered necessary. The manner in which care practices take place is tailored to the specific likes of one care-receiver or another and again requires attentiveness to those preferences and competence in meeting them. What's more, paying attention to those unique preferences is the vehicle for establishing trust and allowing for successful reciprocal interaction.

In short, the moral elements may thus be referred to as "the fundamental values in care". They represent the values needed for each and every care practice. What's more, they subsume the variety and diversity of care values discussed in chapter 3. As such, ensuring the elements are present or strengthened through the design and introduction of a care robot, ultimately results in a manifestation of the core care values. The differences in prioritization and manifestation of elements between practices and/or contexts is something that the care ethicist may draw the attention of the designer to. But, the designer must first be aware of the necessary elements and their manner of manifestation.

5.2.6 In summary

The first part of this chapter demarcated the components comprising the CC framework and their relevance for being included within the framework. The

framework is distinguished from its methods for use in that it is intended as a point from which one begins to theorize: the point from which one is directed to the components of ethical relevance requiring further analysis and understanding. The methodologies for use act as a 'user manual' and differ in multiple respects: at which point in the design process evaluation takes place, the kinds of recommendations that result from the differing evaluations and the intended audience of those recommendations (designers vs. policy makers). Each methodology for use, or user manual, will now be discussed in detail.

One point of clarification that should be made explicit at this time has to do with the normative criteria for evaluating the care practice and the care robot. Many design studies (e.g., user-centered design and certain versions of scenario-based design) address the preferences of users to mark both the starting point for design as well as the evaluative criteria for the design of an artefact. It should be clear from the description of the values of the CC framework (see section 5.2.5) that in fact the preferences of patients are not represented as part of the evaluative criteria. This is so for a variety of reasons. First, a distinction in preferences and needs must be made clear; needs represent the essentials for functioning. In care terms, needs are defined as the essentials for the provision of good care. In contrast, preferences assume there is a choice between one alternative and another, and an individual having an affinity for one or the other is a preference. The use of patient preferences to steer the design of the robot does not recognize the pivotal role played by the needs of the institution and the needs of care providers. It is true that healthcare institutions revolve around meeting the needs of patients, (i.e., the needs of patients delineate the actions and attitudes taken in order to meet those needs); however, in order to meet the needs of patients, the needs of the institution and the needs of care providers must come first. Care-givers must be afforded the opportunity to form a therapeutic relationship in order to meet the needs of patients. Care institutions must allocate resources efficiently in order to meet the needs of all. Therefore, by designing care robots in a way that facilitates the meeting of needs of care-givers as well as the needs of the institution, the care robot is ultimately meeting the needs of patients. This is not to undermine the preferences of patients but rather the assist care providers in their routines as a means for enhancing their capabilities for tailoring care to patient preferences.

Second, the impossibility at this time of creating one robot that is capable of adjusting its behaviour based on individual patient preferences, for every patient that it encounters[1]. Of course in the coming decades this is subject to

[1]This finding is based on recent reports of work done with the Casero robot in a nursing

change, at which point my claim may have to be adjusted, but the work of this book is an explicit avoidance of speculative ethics and as such I do not go into much detail in speculating possible future robots that may or may not come into fruition. For now, adjusting one's behaviour based on individual patient preferences remains a capability exclusive to human care-givers. This point also ties in with the assumptions and meanings associated with the delegation of certain roles and responsibilities to care robots, the work of chapter 8. Third, it is fiscally impossible to incorporate robots in a healthcare institution and to leave the use of the robots up to individual patient preferences. This would be analogous to the claim that all patients should have the choice as to a private room versus a room with other patients. Of course this may be an option in a private or two-tiered healthcare system but the same usage with a care robot would rightly encounter questions of health equity and distributive justice.

5.3 Applying the Framework

Referring back to Hubberly's discussion of design process models, he distinguished the model from the actual method of using the process. I too distinguish here between the framework and the method for using the framework. The framework points to the components demanding ethical attention and applying the framework allows me to analyse the components of good care practices with and without the presence of a care robot. Thus, I take the stance that the initial investigation of a care practice must be criticized with respect to the moral elements rather than taken as the gold standard or carrying any kind of normative force. This is also done to avoid a kind of built-in naturalistic fallacy. In other words, that how care 'is' currently practices becomes the normative standard for how care 'ought' to be practised and thus built-into the care robot. The

home. Robot designers used the method of scenario-based design in the design process of the Casero robot. When the robot prototype was tested in context it became quite clear how impossible it was to incorporate user preferences into the design architecture. For example, when Casero was bringing a glass of water to give to residents, some residents would not take the glass if it was a certain colour. The yellow cup was used in the care of demented residents and thus other residents did not want to be associated with such cups. The robot, however, could only recognize that colour of cup. Added to this, the robot was only used for testing the delivery of water. Residents did not want water but preferred "root beer" and again did not want to take the cup from the robot. Thus, aiming to meet the variety of preferences of patients is not feasible considering the current state-of-the-art in robotics. Information regarding the Casero robot was obtained through personal communication with social scientists collaborating in the design process of the robot.

framework is meant to be used by both ethicists and designers of care robots: to divide the moral labour between the two disciplines and to encourage interdisciplinary work of this kind. When used retrospectively, such an evaluation is integrated into the design process further downstream and as such will yield recommendations for an improved design of the care robot (according to the demands of the care ethics tradition) and/or for guidelines pertaining to its implementation. When used prospectively the framework is intended to shape the entire design process of the artefact. Therefore, the CC framework is used at the moment of idea generation and steers the design of the resulting artefact and its implementation.

5.3.1 Retrospective Evaluations of Care Robots (REC)

For retrospective evaluations using the CC framework, one identifies the context, practice, actors and the manifestation of moral elements for the practice, prior to the inclusion of a care robot. This is necessary in order to understand who is acting when (the care-attentives and the care-responsibles), for what reason, and how values are manifest through such actions. The manner in which the care practice is described draws on the Actor-Network Theory of Bruno Latour which asserts that relations are simultaneously material (between things) and semiotic (between concepts) and together form a network. Verbeek uses this to discuss the concept of mediation – that over time the presence and use of a technology shapes and changes the way we practice in, and view, the world [*Verbeek*, 2006]. Mediation does not assume that a technology is made in isolation from the world and then asserts an impact but rather this happens in a manner of co-production – technologies develop as our norms change and our norms change as technologies develop. Thus, the concept of the network helps to clarify that through the interactions of the material and the semiotic, both evolve. For Pols, the network is used to describe how the network acts to shape or to co-produce subjects (as patients): "instead of being an active, autonomous and authentic individual with a perspective on the world, the subject becomes a co-production, a result of interactions with others and a material world" [*Pols*, 2004, p. 136]. The subject in question may be the care-giver and/or the care-receiver. For the purposes of this work, I align myself closely with the insights of Pols and claim that in the actions and interactions between and among actors (the material) so too are there interactions between concepts, values and norms (the semiotic). I have already shown how care values are the result of such actions and interactions and thus the section of manifestation of moral elements aims to make explicit how the moral elements come to fruition in a network of

actors (human and non-human) in a given context for a specific practice.

Of equal importance in this description is making explicit the roles and responsibilities of all actors (human and non-human) in the network (care practice) as well as how the practice relates to other practices and to the overall process of care. This is done to ensure a fluid description of the care practice meeting needs understood as fluid and dynamic as well as making explicit the roles and responsibilities that often go unnoticed. For Tronto, "any account of institutional care that fails to name explicitly the 'care-attentives' and the 'care-responsibles' allows those people, and their roles in caring, to pass unnoticed. Such not-naming contributes to the process of 'naturalizing' care relations and to blaming the care-givers who may have inadequate resources" [*Tronto*, 2010, p. 165]. Is it then of crucial importance that the distribution of roles and responsibilities prior to the inclusion of the robot be made explicit in order to understand any shifts that may occur following the introduction of the robot as well as to understand the chain of responsibility within the overall process (who is responsible for what action/role).

This process of naturalizing care is also of significant interest in the discussion of the care centred framework and its use. By naturalizing care, Tronto makes reference to the model of care in a family or household that is often referred to as the exemplar by other care ethicists, namely Nel Noddings [2002]. One of the problems with this has to do with the establishment of certain elements implicit in a home care setting that need to be worked out consciously in any other institutional setting. Although Tronto speaks of these elements in terms of the asymmetry in power relations, a clear vision of the purpose of care and a means for maintaining the particularities of care, I aim to emphasize other elements as well, namely the relationship and/or bond between care-giver and care-receiver. Within a home setting, such a bond "evolves out of ongoing interactions among the personalities in the household" [*Tronto*, 2010, p. 159]. This natural evolution means that such relationships and the mechanisms for their establishment could be taken for granted; however, in any other institutional setting such aspects of care need to be made explicit, to be consciously dealt with. "This does not make these elements less achievable, but it does mean that they become more visible and require a deliberate, political process to enact them" [*Tronto*, 2010, p. 159]. It follows then that the establishment and maintenance of the relationship between care-giver and care-receiver requires attention when discussing the introduction of a technology that has the potential to alter this relationship in any way.

Following this meticulous illustration of the care practice, one then discusses the type of robot (assistive vs. enabling vs. replacement), the capabilities,

features and appearance of the robot, and the manner in which the proposed care robot is assumed to fulfil its role. One describes the practice again in great detail, only this time with the introduction of the care robot. Thus, one must indicate what role and responsibility the robot has been delegated as well as how the moral elements are manifest throughout the care practice with the inclusion of the robot. With this, one can make clear the impact of the robot's capabilities on the framework of values (i.e., the manifestation of moral elements). Relating this to the care stance and its implications, one now has the potential to uncover if, when, and how a robot interferes with the provision of good care understood in terms of the manifestation of moral elements. Evaluating the robot against a human's capabilities is done in the case of a replacement robot, but in the case of an enhancement or assistive robot the robot is evaluated according to the role it has been delegated. This evaluation then addresses the robot's impact on enhancing or assisting the human care-giver's capabilities for meeting the requirements of the moral elements.

Of particular interest is the impact on the overall process of care, the shifts in role and responsibility distribution, the variety of assumptions embedded in the design and the prioritization of needs/values that the robot assumes. Thus, the ethical assessment begins with the manifestation of moral elements and incorporates additional reflective tools to uncover a range of alternative ethical considerations. Akrich discusses the embedding of elements in terms of assumptions made about user preferences and competencies [*Akrich*, 1992]. Placed in context, each care robot will take on a distinctive meaning and the meaning of the robot has to do with the assumptions embedded within. This description is quite useful for my reflection but an important distinction must be reiterated here, one that pertains to the difference between assumptions and the concept of values and norms. Assumptions are more about the real word, they are descriptive in a sense while values are more about what the real world ought to be like, they are normative in a sense. When an assumption is made about a value to be embedded, it does not have to be a description about what is, but could also be a claim about what values ought to be expressed, how they ought to be expressed, or what priority they ought to be given. In others words, when the built-in assumption pertains to a value, or when a valuation is being made, the result is a normative claim about what the values should be, what should be valued, or what the ideal is. For Akrich, "many of the choices made by designers can be seen as decisions about what should be delegated to a machine and what should be left to the initiative of human actors" [*Akrich*, 1992, p. 216]. By making choices about what should and should not be delegated to certain actors (human or non-human), engineers may change the distribution of

responsibilities in a network according to a pre-conceived ideal.

Consequently, each robot will reflect divergent assumptions pertaining to the understanding of a care practice, the aim of the care practice and the prioritization of values manifest through a care practice. It is for this reason that I continue to insist on evaluations of care robots on a design-by-design or practice-by-practice basis. Making potential assumptions explicit helps to identify additional ethical issues to be addressed. From such an analysis the goal is to provide normative suggestions for either the implementation of the robot or an improvement in design. As such, the retrospective evaluation addresses certain prospective suggestions based on the analysis but differs from the prospective evaluation proposed here.

5.3.1.1 Retrospective Evaluations of Care Robots in Practice

To give an example, let us look at InTouch Health's RP-7 robot. For the purposes of this retrospective analysis then let us say that the robot is in a hospital context, for the practice of stroke care consultation (after an individual has had a stroke and requires check-ups as well as observations of exercises) and the actors are the patient and physician. Traditionally, the physician would visit the patient's room directly, often with a resident (or multiple residents) for teaching purposes. The physician enters the room and begins discussing the patient's status with them before beginning the necessary exercises in order to test the recovery of the patient. In these instances, attentiveness is often split between the patient and other patients in the room, other residents or care workers in the room, or perhaps other patients the physician will see afterwards. The physician, however, is physically present and capable of picking up on multiple cues given by the patient; physical presence thus enables both attentiveness and reciprocity. The responsibility of the physician comes in the form of being physically present and the physician's competence comes in the form of their expertise in assessing the patient's well-being.

Enter the robot. The RP-7 robot "is a mobile robotic platform that enables the physician to be remotely present". Remote presence is used in a variety of instances, for: visitation to patient's bed side, training between an expert and novice physician, collaboration between physicians, care management, critical care experts remotely present, and stroke care consultation. The robot is equipped with a "Panoramic virtually there visualization system, Holonomic Drive system, SenseArray system 360, an easy-to-use control interface and enhanced audio capabilities which allow the user to focus in on a specific conversation, similar to using a person's own two ears". The robot allows for "direct

connection to Class II medical devices like electronic stethoscopes, otoscopes (a medical device used to look inside the ear) and ultrasound" in order to transmit medical data to a remote physician. The robot also has "RP Dock, which allows the robot to autonomously dock"[2].

With the robot in the picture, the robot is partially human-operated and partially autonomous – the robot requires real-time input from the physician in order to drive around the hospital and arrive at the patient's bedside. The robot is autonomous in that it can dock itself when its battery is low. The robot is typified as an assistive robot for this case. The physician may be somewhere else in the hospital, in a different building, city or country but can still operate the RP-7 platform remotely. The physician can also have access to patients from their home during the evening and weekends when they are not on-call or in the hospital. In terms of attentiveness, the physician is able to both see and hear the patient and can respond to cues accordingly. In terms of responsibility, the physician is still taking responsibility for this specific care practice rather than delegating it to an autonomous robot or another healthcare worker. In terms of competence, the physician is without access to touch, smell and taste; however, for this particular practice those senses are not as important as sight and sound. In terms of reciprocity, the patient and physician are able to communicate verbally and with visual cues in real-time. What's more, the physician is able to observe the patient as they perform exercises – thus reciprocity does not demand that the patient be speaking to the physician but rather that the physician can competently assess the recovery of the patient through visual and auditory observations.

In short, with respect to the manifestation of moral elements, the RP-7 robot has the potential to meet the needs of both the patient and the physician. In fact, the robot actually enhances the ability of the physician to take responsibility for the needs of the patient as well as facilitating the physician's capability for engaging in the reciprocity of the relationship. If the robot was not available the physician may not otherwise have the chance to consult with the patient and thus would not be available for the consult. This line of thinking is somewhat dangerous and invites the discussion of whether the bare minimum of care is better than no care at all. When that becomes the sole criterion for evaluating care, the line between providing good care and providing care by any means becomes blurred. This is when the policies governing the use of the robot play an integral role. In terms of the robot's capacity to mediate interactions or practices, one

[2]Information obtained by InTouch Health's website: http://www.intouchhealth.com/products_remote_presence_endpoint_devices.html

might consider that in the future, the responsible decision of a physician at a distance may be translated into an obligation to tele-communicate. Or perhaps, given the initial results of pilot studies and the high acceptability of patients (granted, a small cohort), perhaps the responsible thing for the physician to do in *any* circumstance will be to engage in such tele-interactions. Again, the policies outlining when the technology should be used and when it ought not be used will play a role here; however, such policies will be (or at the very least ought to be) taking such insights into consideration.

This brings us to a discussion of the link between this practice and other practices as well as with the overall process of care. Post-operative consultations with the operating surgeon, or the predominant physician, are important both for the care provider and the care-receiver. For the former, the physician needs to be aware of the success, or lack thereof, of the medical intervention. If the medical treatment has not been successful, the physician will need to intervene. If the medical treatment has been successful, this will determine the length of stay in the hospital. In both cases, the relationship between the patient and the nurse care-givers will be impacted – the nurse will need to tailor his/her care based on the physician's assessment. Thus, the practice of post-operative assessment is linked with the initial medical intervention as well as later interventions. Of equal importance is the impact the consultation (and results thereof) will have on the future treatment of the patient by the nurses. What they must be attentive to, responsible for, competent of, and responsive to will be dependent on the post-operative consultation. For the latter, the patient needs to be informed of their own status as well as feeling connected with the individual involved in their treatment. Thus, communication post medical intervention structures the daily rehabilitation/care plan of the nurse. If the robot is not available and the surgeon could not complete the assessment, further care may be negatively impacted.

But what about values like human presence, of tinkering (a kind of attentiveness), of dignity? Without the use of the robot the value of human presence is fulfilled by the very presence of the physician. Tinkering is then possible because of human presence – tinkering here refers to adjustments made during the consult based on the observed recovery of a particular patient. The expression of dignity in this practice is seen in terms of closing the curtain around the patient's bed (privacy), eye contact with the patient to ensure that they are the priority, and speaking in a certain volume to avoid sharing information with others in the room or outside the room (confidentiality). Without the robot there is no guarantee that eye contact will be ensured (the physician may be occupied or distracted) or that information will remain between the patient

and physician (other patients and their family may be in the room). With the inclusion of the robot, eye contact (or direct and exclusive attention towards the patient) is guaranteed – the physician's field of vision is determined by the robot. The robot, however, cannot close the curtain around the patient and thus another care worker is required for this (if the curtain isn't already closed). In terms of confidentiality, the same risks are present (meaning, if others are in the same room there is the potential that they will hear) but what's more is the question of whether the information transmitted through the telecommunications network will remain completely confidential, whether it will be stored and whether others may have access to it.

In conclusion thus far, the robot does not threaten the relationship between patient and surgeon/physician nor does the robot undermine the responsibility of actors. In fact, the robot appears to enhance such a relationship by allowing for a moment of reciprocity, by allowing the surgeon to take responsibility for their act of care-giving. The robot does change the traditional method of stroke consultation given that the surgeon is not physically present; however, not in a way that would render the robot unethical to use. What's more, the robot ensures that a practice integral for the provision of future care remains in tact. I should include here that the analysis was restricted to stroke consultations and not for basic rounds of a physician. It is not possible to say that every practice for which the robot might be used would uncover the same results from the analysis. For physician rounds it is recommended that the physician have access to all senses rather than to visual and auditory only. For patient's suffering from dementia or some kind of mental disorder it is questionable whether the robot would bring more fear and confusion than good.

Given that the robot is already available, the evaluation takes place further downstream in the overall design process (almost to the point of being a kind of health technology assessment). Thus, from the above analysis I may make tentative recommendations regarding the policy of implementation within a hospital. Firstly, that the information acquired by the robot is not stored for a long period of time (this time period is dependent on the expected length of recovery and the stage of recovery of the patient) and secondly that a dedicated network is used for the transmission of information, thus safeguarding the value of confidentiality. In order to promote the value of privacy perhaps there is an option for headphones which the patient could wear to avoid others from listening. This, however presents a problem if the patient is not capable of putting them on themselves. One might also conclude that when the physician has multiple patients and emergencies in one day that they be recommended to consult with their stroke patient in the comfort of their own office or home

where they will not be subject to outside distractions.

Of course this is not an exhaustive analysis of the RP-7, the goal rather was to outline in general how the framework is used. In the proceeding chapters I will complete a more thorough evaluation of a variety of care robots while at the same time addressing the built-in assumptions and/or biases, norms and values (according to script theory).

5.3.2 Retrospective Evaluations of Surgical Robots

In section 4.5, 'What is a Care Robot?', care robots were defined in a manner that included surgical robots. Setting aside the specialized type of surgery when using the robot, we may still briefly evaluate the robot according to the components of the framework. To begin, one might say that the robot enhances the competence of the surgeon. The surgeon, seated at a console to perform the surgery, manipulates hand controllers while viewing the surgical field (via an endoscopic camera inserted inside the patient). The movements of the surgeon are then translated into movements of the robotic arms which have been inserted inside the patient. The movements of the surgeon are scaled for the surgical environment it is in. In this way, the surgeon can perform to a degree of accuracy not possible before the robot [*van Wynsberghe and Gastmans*, 2008]. In consideration of the attentiveness of the surgeon, one might question whether or not the loss of force feedback and tactile sensation will restrict this element of the surgeon: when compared with traditional or conventional surgical methods, the surgeon no longer has any 'feeling' of the surgical environment. This is not a new phenomenon for surgeons, however, as laparoscopic surgical methods presented the same difficulty. Added to the discussion of attentiveness, one should also consider the ergonomic factors related to the procedure. When using the robot, the surgeon is seated comfortably at the console. Prior to the robot the surgeon would have to stand in an uncomfortable position for any number of hours (depending on the surgery) which had the potential to limit the attentiveness of the surgeon. Altering the ergonomics for the betterment of the surgeon may be considered a benefit in terms of enhancing the attentiveness the surgeon. In terms of responsibility, one would suggest that responsibility is a shared endeavour between human and robot. With this in mind, however, if the robot were deemed accountable for a failure in the surgery, it is still unclear where liability will lie; in the hands of the robot manufacturer, the distributor, the hospital purchasing the robot or the surgeon performing the surgery [*van Wynsberghe and Gastmans*, 2008]. These are still questions left unanswered in the area of surgical robots. In terms of reciprocity, the robot may diminish

the surgeon's capacity for reciprocity again in consideration of the loss of force feedback and tactile sensation. In preliminary studies assessing the subjective experience of surgeons using the robot in comparison with laparoscopy and conventional methods have shown no significant difference in their own perception of the surgery [*Van Koughnett et al.*, 2009]. Thus, concerns for attentiveness and reciprocity may not be as problematic as I am presenting.

The robot's impact on the overall process of care introduces some interesting questions pertaining to the shift in roles and responsibilities. The surgical robot is included within a therapeutic relationship (granted a patient-physician therapeutic relationship) and has a direct impact on the distribution of roles and responsibilities of all actors, both directly and indirectly, involved in the practice. Surgical robots, like the daVinci®, demand that both the surgeon and the nurses receive additional training with the introduction of the robot. Thus, roles remain the same but responsibilities of both the surgeon and the nurses have increased. Added to this, the surgery performed using the robot falls under the umbrella of minimally invasive surgery resulting in a decrease in scarring, a decrease in infection rates, a decrease in post-surgical recovery time and therefore a decrease in hospital stay. In this way, the nurse's role and responsibility in the daily care of the patient post-surgery is changed: the patient will not be under the monitoring of the nurse for the same amount of time. Instead, the patient will be sent home and come under the care of a family member (who now becomes an indirect user of of the robot). Thus, the actors now involved in the post-surgical care of the patient are enlarged. The delegation of responsibilities is also enlarged and now encompasses persons not trained for such activities. This may be daunting for family members who now have to take care of a patient at home but also blurs the lines of who is responsible if something were to go awry post-surgery.

In short, during the practice of surgical intervention, the robot provides the means to enhance many of the moral elements. I have written elsewhere about the robot's potential to shift the patient-physician relationship and for this reason guidelines pertaining to the robot's use must enforce the significance of this relationship by ensuring that pre- and post-surgery the patient and physician are still granted the forum for establishing a therapeutic relationship [*van Wynsberghe and Gastmans*, 2008]. However, when considering the extended impact the robot bears on indirect users, one might suggest that in order to ensure the robot as ethically sound, inclusion of indirect users be taken into account in the policies guiding the use of the robot. To do so, nurses, or another healthcare providers, may be required to make home visits or to conduct post-surgical assessments via phone and/or video-conferencing to ensure that

the responsibility of the institution and its staff ultimately remain responsible for the outcome of the patient's surgery.

5.3.3 Prospective Analysis: Care-Centered Value-Sensitive Design Approach

The *Care-Centered Value-Sensitive Design* (CCVSD) approach provides the answer to the research question guiding this work. It is meant as the climax of this dissertation. When contemplating the nature of a 'good' care institution, Tronto claims that "care institutions have to think about the nature of the caring process as a whole in order to guide their actions. This requirement does not only demand that the 'needs' of 'customers' come first but also that the needs of care-workers, the allocation of responsibility and proper assessment also happen within the organization" [*Tronto*, 2010, p. 162]. Accordingly, the CCVSD approach does just this. It is a prospective methodology that reflects the overall methodology of Value-Sensitive Design. It begins the analysis of the ethical dimensions of the care robot further upstream in the design process, at the moment of idea generation (meaning, at the point in which designers begin to think about developing a robot to be used in a care context). Hence, there is no artefact made and the methodology is intended to shape the future design of a care robot. The approach begins by taking into consideration the fears associated with the use of care robots in general. These fears mark the starting point and are used as a tool in the design process: the care robot is designed in a way that intentionally avoids the manifestation of such fears. A practice is then selected for discussion. The selection of the practice is a result of the ethicist and engineer's experience observing in the healthcare context in question. The manner in which a practice is described prior to the introduction of a care robot is the same in the CCVSD approach as it was in the REC approach: the practice is described in meticulous detail to understand when and how values come into play along with the distribution of roles and responsibilities. The practice is criticized for ways in which any of the moral elements have gone unnoticed or are threatened in a sense. Through the analysis designers become aware of the necessary capabilities the care robot ought to have in order to preserve the current standards of care for a given practice and/or to enhance the current level. As such, the REC methodology works to inform and substantiate the methodology of the CCVSD approach.

In deciding the capabilities the care robot ought to have, careful attention is paid to the relationship between such capabilities and the resulting role and responsibility the care robot will take-on. This insight comes from the work in

chapter 8 discussing how the robot's role is often based on a presumed moral status of the robot. This discussion results in the conclusion that based on the necessary element of responsibility in care institutions, the robot must be designed as an intentional positive moral factor (see chapter 8 for a discussion of the robot's moral status). As such, the robot cannot be delegated a role for which it bears any responsibility for actions with life and death consequences. What's more, the care robot may not be delegated a role for which moral-reasoning capabilities are required. Thus, the decided role of the robot has a direct impact on the kinds of capabilities the robot can have. This is not a static moment in the design process but rather is a dynamic dialogue through which the role of the robot may change depending on what this role demands in terms of capabilities: if the robot's capabilities delegate a responsibility to the robot that it cannot bear then the capabilities (and assigned role) must adapt. Thus, recommendations resulting from this portion of the analysis are intended for designers.

Additionally, the CCVSD approach incorporates the aspect of implementing the care robot. For this, recommendations resulting from the analysis of care practices alongside the development of the robot are incorporated into the introduction of the robot in the context of use. In other words, if the robot is designed to be used in a specific way, by a specific user, at a certain time, such recommendations are expressed to the actual users in context as they become acquainted with the technology. This is to avoid the manifestation of an unexpected or random morality within the context of use. While this may seem paternalistic at first glance, the entire aim of the CCVSD approach is to shape the care robot according to the values of the institution, an in-depth understanding of the practice and context of use, and a commitment to the care ethics tradition. Thus, the care robot is designed with a specific use in mind in order to maintain the values of the institution and a high quality of care such that allowing for random or unintended uses threatens the integrity of the CCVSD approach.

5.4 Conclusion

Given the lack of regulatory frameworks in the design process of robots used outside the factory coupled with the need for ethical consideration throughout the design process of care robots, I have created a tool to guide design according to certain fundamental concepts and principles in care. This tool is meant for use in the design and development of any care robot. The focus on care practices and

the therapeutic relationship comes from the context, practice and relationship within which care robots will be placed. Thus, the framework is tailored to the context of a care institution, the complexity of care practices and the central focus of the therapeutic relationship in the provision of care. Beginning with the care orientation and using the blue-print of VSD I arrive at a conceptual framework for evaluating artefacts and practices from the perspective of good care. The framework is heuristic in that each component is subject to description and interpretation (from the care ethics perspective) alone as well as in relation to the other components. Additionally, it is prescriptive in terms of directing one's attention to the necessary components and the means for their analysis. The framework, what I refer to as the *Care-Centered Framework*, may be used as a tool for retrospective ethical evaluation of current care robot designs as well as a tool for use in the prospective design process of care robots.

In short, the methodology for using the framework retrospectively involves three steps: 1. described the current care practice without the robot in detail, 2. describe the robot in detail and, 3. describe the care practice with the introduction of the robot in the same detail as in step 1. The methodology for prospective analysis mimics that of the retrospective methodology but includes two additional steps; it begins by addressing the relevant fears associated with the use of a care robot and finishes with the implementation of the robot.

When conducting a retrospective evaluation of a current care robot the methodology makes clear the link between robot capabilities and their impact on the manifestation of care values. What's more, through an exhaustive analysis and critique of current care practices without the introduction of the robot we are able to see the ways in which one practice is linked with another and further how one practice fits into the overall picture of care. With this, not only do we have a better insight into the broader impacts of the care robot but we are also engaged in a structuring of the care institution according to the criteria of good care institutions proposed by Tronto [2010]. Consequently, we may ask the question; what kind of care do we want to provide and in so doing we may steer the design and development of care robots. Hence, recommendations resulting from the retrospective analysis are intended for an improved future design of the care robot. The next task is to provide examples of the Care Centred Framework used retrospectively in order to understand the relationship between components, to better understand how the retrospective methodology will proceed and to set the stage for the prospective methodology.

Chapter 6

Care Robots and the Practice of Lifting

6.1 Introduction

Institutional care is comprised of a variety of care practices. Many of the practices can be labelled "activities of daily living" (ADLs). ADLs is a terms used in healthcare referring to practices like personal hygiene and grooming, dressing and undressing, self-feeding, self-transfer such as getting into and out of bed, bowel and bladder management. They are practices considered essential to one's daily functioning and are therefore routinely used to test the functional status of a patient [*Krapp*, 2002]. The majority of current care robot initiatives aim at developing a robot to assist with ADLs. Some good examples are: Secom's My Spoon, an assistive robot for feeding, RI-MAN, or RIBA, a robot for lifting and the Sanyo electric bathtub for bathing [*Sharkey and Sharkey*, 2012]. ADLs serve a variety of purposes beyond the primary role of the ADL. In chapter 3 with the example of bathing, this ADL was revealed as also a moment in which the nurse's vision of the patient as a citizen was enacted throughout the practice [*Pols*, 2004]. They are practices during which the therapeutic relationship between patient and care-giver is either established or strengthened. They are practices during which social needs are often met as well [*Mol et al.*, 2010; *Tronto*, 1993; *Pols*, 2004]. Consequently, the ethical evaluation of a care robot used in an ADL must take into consideration the multiple ends which an ADL serves in the therapeutic context. Evaluating the robot using the Care-Centered

(CC) framework does just this.

As shown in the example of lifting, the current practice does not always allow for the full manifestation of elements. This adds strength to the idea that robots present the opportunity to re-introduce certain elements that have been neglected or overlooked. The aim of this chapter is to make explicit the relationship between a robot's capabilities and the resulting promotion of care values; a difference in one capability or mode of control is shown to dramatically shift the resulting care practice. Added to this, the chapter aims to make clear the kinds of assumptions that find their way into the architecture of a care robot.

The following chapter uses the CC framework for the retrospective evaluation of current care robots used in the practice of lifting. This is done according to the 'user manual' described in chapter 5. To recapitulate, the practice is described in detail with careful attention to how and when the moral elements are manifest. The distribution of roles and responsibilities are also clearly articulated, along with the relationship between this care practice and others (the overall care of the patient). Following this, the robots are described and once again the practice of lifting is presented in the same manner as before, only this time with the inclusion of the robot. To be clear, I am not describing the current practices and concluding from this their normative force, rather I am describing the current practice to articulate where and how the moral elements are manifest according to their interpretation from the care ethics perspective along with where and how certain values may have taken precedence at the expense of others. With the addition of the robot, I then analyse the technical content of the robot using script theory [*Akrich*, 1992; *Latour*, 1992] – the script of the care practice inscribed in the robot is revealed through its technical content.

6.2 The Practice of Lifting Using Human Actors

In the home, hospital and/or nursing home setting, the practice of lifting may involve only human actors for the act of lifting. The description of the practice is as follows: the human actors are the patient and the nurse (or nurses if more than one is needed, or porters if nurses are not available) and the non-human actors are the chair or mechanical bed which the patient is getting out of, the curtain (if the patient is in bed) and the hospital room . The nurse approaches the patient and asks if they are ready to get up. With agreement from the patient, the nurse places his/her arms at the patient's waist and waits for the patient to place their arms around her neck or on her shoulders. The nurse

makes eye contact with the patient at all times to cue the patient. Together, they work to lift the patient from the chair and onto the bed or into another chair (from wheelchair to dinner chair for example). If the patient requires more assistance and is confined to their bed, often times porters are called upon for assistance. In this case, the nurse and/or porters enter the patient's room, speak to the patient to let them know they will be moved, enclose the curtain around the patient's bed and either the nurse begins to lift or one porter will lift from each side of the patient.

This is the kind of description of a practice that leaves one with the image of a task – that lifting is a moment in the care process in which the only goal of the task is to lift. As we have seen from the previous discussion of care practices this is not so. The practice of lifting is about much more than the exclusive act of lifting when we observe the presence of the moral elements. For this practice, attentiveness may be considered in terms of the nurse's perception of how the patient is doing on that day, at that time. Perhaps the patient's condition is worsening and the nurse can observe this through their physical presence during the practice of lifting. Perhaps the patient is in a greater level of discomfort or is recovering quite quickly. Or, perhaps the patient's neurological status is worsening. Thus, not only is the nurse responsible for being attentive to the mechanistic criteria defined by the practice of lifting (lifting at a certain speed, applying the appropriate amount of pressure, etc.) but the nurse is also responsible for being attentive to the unique state of the patient on that day at that time – these considerations have to do with the overall process of care. The nurse will use this information and tailor other care practices of that patient's overall care accordingly. Consequently, lifting becomes not only a moment for establishing and maintaining the trusting bond between nurse and patient but it is also a moment for incorporating the element of attentiveness as it refers to the patient's overall state of being.

The moral element of responsibility is closely aligned with the element of competence and both rely on the concept of trust. Responsibility refers to the capability of the nurse to be accountable and liable for accomplishing the lift with competence. Responsibility in this sense demands that the nurse and/or porters be blamed in the case that something were to go wrong. Responsibility also refers to a capacity in terms of the nurse's role in the overall provision of care. For the first point, the nurse (or porters) must competently lift the patient. This refers to lifting the patient at an appropriate speed, using an appropriate amount of force and lowering the patient according to the same criteria. The nurse or porters must be attentive during the practice of lifting to ascertain whether or not the patient would prefer to be lifted at a slower pace

(or a faster pace) or using more (or less) force. The nurse is also responsible for fulfilling the practice competently as it applies to the overall provision of care. By understanding what the patient prefers in this practice the nurse begins to understand the unique details of one particular patient with particular needs and preferences (i.e., having the capacity for being attentive to and understanding the personalized patient). It is this piece of the practice that strengthens the therapeutic bond, helping to encourage the formation of a trusting therapeutic relationship between care-receiver and care-giver(s). This relationship is necessary in the provision of good care throughout the rest of the care process. This bond is required for the patient to be honest about their symptoms, to take their medications, to comply with their care plan and in general to follow the advice of the nurse.

The trusting bond is achieved not only through attention to, and an understanding of, the unique patient and their needs. The *value* of trust is maintained or promoted throughout this practice in additional ways as well – through the interactions between not only the human actors but the material world as well (i.e., the practice and the meaning of the practice is a co-production of actions and interactions between actors and material objects). One might assume that enclosing the curtain acts to encourage trust, through a respect for privacy. Privacy, if we recall from chapter 3, refers to a non-disclosure of the corporeal dimension of an individual. One might also assume that the eye contact made between nurse and patient also acts to promote trust. Or, one might suggest that having the nurse physically present in the room encourages the manifestation of trust. Accordingly, the nurse's presence is a causal and symbolic representation that they are responsible for the well-being of the patient and further, that this is a necessary element for trust in this practice. It is symbolic in that the nurse is a representation of the legally and culturally sanctioned policies of the hospital. It is causal in that the physical presence of the nurse demands that the nurse be liable in the case of something going wrong, according to the sanctioned policies.

This last point, the presence of the human actor for lifting seems to be the necessary criterion when it comes to the moral element of reciprocity when the actors for lifting are human. Reciprocity in this practice refers to the patient giving cues as to their own well-being – these may be verbal or non-verbal. The patient may make verbal recommendations as to their discomfort or satisfaction and the nurse (or porters) can respond accordingly. Alternatively, if the patient cannot speak, or chooses not to, the nurse is still capable of picking up on subtle cues given by the patient like wincing or a look of fear. Reciprocity involves an interaction between nurse and patient and adjustments being made according

to the attentiveness of the nurse regarding this reciprocal interaction. This has been discussed previously as 'tinkering' [*Mol et al.*, 2010] or 'ethical sensitivity' [*Weaver et al.*, 2008].

The above was a description of how lifting happens in an institutional setting currently. When understood in this way, the expression of the moral elements is clear and coincides with the recommendations from the care ethics tradition. The attentiveness of the nurse, as described here, facilitates an understanding of the patient as a unique person with dynamic needs and preferences. The human presence of the nurse/care-giver encourages the formation of a trusting bond. And lastly, the interactions with the material world – namely enclosing the curtain around the patient for privacy – also helps to express core care values. But, this is not always how the practice of lifting occurs in the hospital or nursing home setting. Interestingly, although the moral elements are in tact, they seem to come at the price of the nurse's physical well-being. Often times the nurse does not have the physical strength for lifting multiple patients in a given day and many times there are no porters around for assistance. In these cases, to ensure efficiency within the institutional setting, nurses rely on a mechanical lift.

6.3 The Practice of Lifting Using a Mechanical Lift

In the previous description of lifting, I investigated the expression of values such as trust, human presence and privacy that are co-produced through interactions between human actors (the nurse and patient) and the material world (enclosing the curtain). This was done through the analysis of the moral elements, their interpretation and how they become real, for this particular practice. Moreover, I made explicit how the moral elements, and their expression in this vision of lifting, contributed to the overall care process. In this picture of lifting, using the mechanical lift, the expression of values and their ranking differs once the inclusion of the mechanical lift is understood according to the CC framework.

The lifting of patients has proven to be quite a challenging feat for nurses. Many elderly patients in the hospital or nursing home require partial assistance for lifting themselves out of bed, or out of a chair. Alternatively, many are not capable of supporting their own weight at all and require complete assistance of a nurse to get out of bed or out of a chair. In addition, many post-operative patients require partial or full assistance for transferring themselves from the

bed to a wheelchair, etc. Given that the nurse must do this for any number of patients, there is a risk to the nurse's physical safety if she/he is required to lift multiple patients in a day. What's more, many nurses are not physically strong enough to do this. As a result, nurses have opted to use mechanical lifts on the many occasions that patients need to be lifted [*J. Li*, 2004].

In the context of the hospital's critical care ward the practice of lifting involves the following actors: the nurse and patient, the mechanical lift and its remote control, the mechanical bed, the curtain to enclose the patient and the room. To provide complete lifting assistance for raising the patient out of their bed, a slip is placed underneath the bottom of the patient while they are lying in bed (the patient is raised to an appropriate degree using the mechanical bed) . Each side of the slip is hooked onto ropes hanging from the ceiling. At the press of the remote control the ropes work to lift the patient off the bed and into the air. Using the remote, the nurse moves the patient to position them over their wheelchair and begins to lower them into the wheelchair, again using the remote.

In this picture, attentiveness of the nurse for the patient is directed more towards the mechanical lift and its remote than the human patient. When the patient is being lifted, there is no physical contact with the nurse; although the nurse is physically present there is no chance for eye contact as the patient is raised quite high and the nurse is paying attention to the remote control. Consequently, eye contact and touch are not possible. As I have already shown, these values are integral for establishing and/or maintaining a trusting bond, and this bond is integral for the provision of good care later on in the process. In terms of trust, one might assume that the patient trusts the mechanical lift only because the nurse (whom they trust) is using it, or perhaps because of the trust that they have placed in the institution they are in. The professionalization of medicine and nursing aims at fostering such a trust. If this trust is not there, the patient is in quite a vulnerable position with no other options. In terms of competence, we might say that the nurse and mechanical lift accomplish the lift efficiently without injury; however, this undermines the link between competence and attentiveness – only when a care-giver fulfills the caring actions with attentiveness can they be considered fully competent. With the nurse's attention and focus directed towards the remote, or even shared with the remote, the attentiveness of the nurse towards the dynamic reactions, and/or cues, of the patient is threatened. Reciprocity is observed when the patient is placed in their wheelchair; however, many opportunities for reciprocity are missed during the course of the lift while the nurse is watching the remote and positioning the chair she/he may not catch a sudden wince on the face of the patient indicating

some pain.

Although the mechanical lift promotes the value of efficiency, it also presents the opportunity to threaten other necessary values integral to the formation of the patient-nurse therapeutic relationship, such as eye contact, touch, reciprocity, and attentiveness. This is not to say that efficiency is not a value in healthcare contexts or practices. In fact, the vast needs of the institution demands that it be efficient in order to meet them all. Rather, that efficiency be measured against competing values. This first wave of automation presents a rather flat view of the care practice of lifting. It would appear that this practice has been viewed as a task, as an event that is separate from the process of care and uninvolved in the manifestation of the care values shown to have importance in this work. The current technology involved in the practice of lifting shows us how important it is for designers to understand the holistic vision of a care practice – how it acts as a moment for the promotion of care values. We are reminded that the introduction of care robots may perpetuate the trend to minimize certain care values if we do not question current practices, but it also presents a unique opportunity to re-introduce certain values of ethical importance that may have been lost or overlooked in the previous wave of automation.

6.4 Enter the Robots: Care Robots for the Practice of Lifting

There are two robots which will be used to show the utility of the CC framework in the retrospective ethical evaluation of current care robot designs. The first is the robot previously known as RI-MAN which now goes by the name of "Robot for Interactive Body Assistance", or Riba (Riba is the next generation of the RI-MAN robot). Riba " has a teddy bear face, and can pick up and carry people from a bed to a wheelchair. It can recognize faces and voices and responds to human commands" [*Sharkey and Sharkey*, 2011]. The second is the Hybrid Assistive Limb (HAL) from Cyberdyne [*Hayashi*, 2005]. This robot is most commonly referred to as a "rehabilitation robot" because it is intended for use by patients with severe spinal chord injuries who can no longer walk; wearing the robot, patients are able to walk again . Thus, the robot was not originally intended to be used in the practice of lifting; however, of late, roboticists are exploring the possibility that it be worn by a nurse for assistance in lifting within the practice of bathing where the patient must be transferred from the bed to the bathtub [*Satoh et al.*, 2009]. There are of course other versions of both

types of robots; however, the main point is to compare the distinction in robot capabilities between that of an autonomous/replacement robot and a human-operated/enhancing robot. Both robots can achieve the same task (lifting a patient); however the technical capabilities through which this task is achieved differs and thus changes the way in which the care values are met along with the resulting care practice.

6.4.1 An Autonomous Robot for Lifting

The RI-MAN/RIBA autonomous robot is autonomous in that it is capable of lifting a patient and carrying him/her from one place to another without being controlled by a human operator. Thus, the robot is considered a replacement robot. This robot is designed to work directly with humans and as such is programmed for safety considerations such as speed and the materials which are used for its structure are pre-tested for human contact. The robot has a humanoid appearance, meaning the robot has a head, eyes, a nose and arms. This robot is intended to work in a hospital, a nursing home or in someone's home. Regardless of the actual robot used here the intention is to discuss a completely autonomous robot – one that is delegated the entire role and responsibility of lifting. There may be robots that are delegated a certain portion of the practice, which fulfil that role autonomously, but these will not be the focus of this section . For the purposes of this discussion, the interesting point is the complete delegation of the role and responsibility of lifting to the robot in the context of the nursing home and/or hospital.

The practice of lifting in the nursing home or hospital context using an autonomous robot involves the robot entering the patient's room and fulfilling the practice entirely on its own . The robot must recognize the person who it will lift and respond to a voice command to lift. Such a response may or may not be in the form of verbal communication. This is at least the hope for this robot in the future. At present the robot is still in the early stages of development. This early stage, however is ideal given that the recommendations resulting from the following evaluation may be incorporated into its design process and/or implementation. Requiring that the care-receiver give the command to lift assumes that the person is capable of giving a voice command that the robot can understand; in the case of frail or elderly persons this may not be the case. It may also be possible that the patient is not fluid in the language of the institution [1]. This also assumes that the manner in which the patient is lifted

[1] Imagine a patient from an English speaking country in a country like France or Italy

will not differ between patients but will be standardized; the patient must be lying in a position which the robot can work with. The robot lifts the patient and hopefully may alter the force, speed or angle at which it lifts. These are all current design considerations.

Immediately with this description we can see how the design of the robot denotes a de-valuation of the elements of attentiveness and competence when compared to the value of efficiency – the use of the robot prioritizes the value of efficiency over these elements. The robot is designed to lift a patient in a standardized manner. Consequently, there is no possibility for the robot to be attentive to the individual needs of the patient. Such attentiveness is also a component of competence. The robot may be capable of fulfilling the action of lifting according to a mechanical description of the practice (the angle at which one is lifted, the speed with which one is lifted, the force with which one is lifted); however, even such a mechanical description of the practice is dependent on the unique patient. Not every patient is lifted in the exact same way. Without the ability to be attentive to the changing needs and/or status of the patient, the practice cannot be tailored accordingly and thus the competence of the care-giver is in question.

In terms of responsibility, the robot has been delegated the full role and responsibility of lifter. Thus, if something were to go wrong, the robot would be liable for damages. But if the robot cannot be liable according to the traditional conception of liability (human care-giver would be sued or fired), then is it possible to say it is really liable and as such responsible? More on this in chapter 8. In terms of reciprocity, the patient's placement in a chair, toilet or back on the bed is the only source of reciprocity. At the moment, the robot is not sophisticated enough to acquire cues as to the patient's satisfaction during lifting nor can it tailor its performance accordingly. One might wonder whether this is the kind of recommendation for design or whether this kind of role ought to remain in the domain of a human care-giver.

Not only does the robot impact the manifestation of moral elements when it comes to the practice of lifting when analysed on its own, but it may also impact the manifestation of moral elements as they relate to the overall care

where English is not a national language, how would they ever know what to say to the robot? This may seem too speculative at the moment; however, what if it turns out that robots are revealed to provide a lower quality of care but are still kept around for those people who cannot pay the full price and are consequently used to care for people coming from outside the nation or people in a lower income bracket? Such health equity concerns need to be addressed through the creation of policies for a given care robot. Policies, like design, must also be made according to a robot-by-robot basis.

process (i.e., the multiple practices within care that are linked via the values). For example, human presence was necessary for achieving attentiveness of the nurse to the personal preferences of a particular patient towards the mechanics of the practice of lifting (speed at which one lifts, angle and force) but also for ascertaining the daily changes of a patient. The human presence also creates a moment for establishing and/or maintaining the bond of trust between nurse and patient. What happens then when the nurse is entirely removed from this practice? This is particularly evident when we observe the lack of values like human touch and eye contact. Current research at Georgia Tech University investigates subjective responses to robot touch [*Chen et al.*, 2011]. Results are preliminary but indicate that humans may be in favour of touch by a robot. Although touch may be considered important for the establishment of trust between the human and robot, in this case, one must wonder what impact this will have for future human-human interactions between nurse and patient; will the patient no longer have the same trust for the human nurse if their daily encounters are with the robot?

It is possible to suggest that the nurse will have another opportunity to check-in on the patient; however, the information obtained through the practice of lifting, just as in the practice of bathing, speaks to the neurological, physiological and sociological status of the patient. One might recommend for policy guidelines in a care institution that when a robot is used for lifting there may not be another robot used for the practice of bathing and or feeding. It is possible to suggest that the robot be programmed to ask certain questions, the answers for which can be transmitted back to the nurse taking care of the patient. But this still poses many problems: what if the patient can't answer the robot? What if the patient doesn't trust the robot and doesn't want to answer? What if the patient lies to the robot? What if the patient is worried about who will have access to that information and consequently doesn't respond to the robot? It's not simply a question of information that is verbally given – much of the nurse's work deals with those cues and signals that are not given verbally but that the nurse's training allows him/her to pick up on [*Pols*, 2004]. These are all questions that have to do with the therapeutic relationship between care-giver and care-receiver; the patient would ask the same of the human care-giver and through repeated interactions they would be resolved (hopefully).

As important as the therapeutic relationship is, it is of value to examine what happens to the overall linkage between care practices when the role and responsibility of the human care-giver is delegated exclusively to the care robot. The robot is not making linkages between one care practice and another for a particular patient in its head (aka internal programming) in the sense that

it is not comparing Mrs. Smith's fragility being lifted in the morning with her clammy skin during bathing in the afternoon. The reason for this has to do with the level of sophistication of the robot at this time but one must also consider whether or not this is the kind of responsibility we ought to be delegating to the robot. What does it mean to enforce that the robot, and not a human care-giver, be attentive to a patient in that manner? What are the underlying assumptions?

Not only is there a relation between this kind of attentiveness and how competently the nurse fulfills his/her multiple practices with a patient in a day, but there is also a relationship between this kind of attentiveness (presupposing or demanding human presence) and the interpretation of, or feeling of, responsibility on the part of the care-giver. When the nurse is delegated a role and responsibility in the multiple practices of a patient's day, they are tacitly reminded of their overall responsibility for the patient. Moreover, when the patient interacts with the same nurse on a given day or at the same time from one day to the next, they begin to sense the role and responsibility of the nurse which encourages a trust in the nurse, a bond between the two and an increased likelihood for moments of reciprocity. One might say that there will be persons who prefer to talk to the lifting robot or who prefer to have the robot lift them in the hospital or nursing home. Given the requirements for good care in care institutions, namely the formation of the therapeutic relationship and the fulfillment of the moral elements, it is of paramount importance that the elements remain in tact regardless of the patient's preferences. The moment of pleasure that the patient feels sharing a conversation with the robot may have a detrimental impact on the nurse's ability to provide good care at another moment. Of course one may respond and say the good of the patient is the goal of institutional care [*Pellegrino*, 1985; *Vanlaere and Gastmans*, 2011] and therefore that patient should be able to share a conversation with the robot if they so desire. I will not deny a person a conversation with a robot but I do claim that the needs of the care practice (and ultimately the care-giver) must be placed at the fore to ensure that the system of the institution is functioning efficiently as well as in accordance with the values of the institution. Thus, I too prioritize efficiency above other values in the healthcare tradition, but how I interpret and arrive at an efficient system is based on the entire system and the interconnections within the system rather than from one moment to the next. Accordingly, my vision of a care institute is a reflection of a system conceptualized according to the structural ethics approach [*Brey*, 2012].

The example of an autonomous robot used for the practice of lifting in the hospital or nursing home context leaves much to be desired in terms of the

manifestation of moral elements and the link of this practice with the overall care process. This is also a moment in which we see the significant role that context plays in the evaluation of a care robot; when evaluating the same robot using the same evaluative criteria with only a change in context, the results of the evaluation differ. In a home context the reflection may differ. In a home setting in which a close family member or care-giver is the one providing daily care services there may not be the same need to ensure moments for establishing and maintaining the relationship or the trusting bond; these may have already been established through the existing relationship between the care-giver and care-receiver. In this case, there is not the same pressure to ensure the presence of the human care-giver as a means for ensuring a link with the nurse and the overall care of the patient. The care-giver at home is most often already aware of the personal preferences of the care-receiver. What's more, the care-giver may be the spouse, child, or other family member of the care-receiver making the care-receiver feel quite vulnerable and powerless to be in such need. In these cases, perhaps the more dignified means for lifting is in fact the use of an autonomous robot. In the hospital and nursing home context, this is not the case. Hence, the importance of context bears significant weight in terms of the recommendations for implementation. From this analysis, one may conclude that an autonomous robot in a nursing home or hospital setting fails to promote the elements of the care practice of lifting that demand attention and provision. One might also add that this is not necessarily the case in a home care setting.

In short, it appears as though this robot is quite problematic if the idea is to implement the robot in a nursing home or hospital setting. But this does not end the analysis. We saw that in the nursing home and hospital contexts where mechanical lifts are used there are certain aspects linked with attentiveness and competence that were overlooked in favour of the mechanical lift and its efficiency in alleviating a burden of the nurse. The question then becomes, is there an alternative to this? It is possible to suggest that a human escort accompany the autonomous robot in its functioning. The human escort may be a nurse or another care worker/volunteer whose role is to speak to the patient while they are being lifted and to adjust the robot in the case that it lifts too quickly, etc. Added to this, the human escort may also keep track of pertinent patient information. If this escort is the nurse then the robot relieves the burden of lifting but doesn't necessarily free up any of the nurse's time. If the escort is another care worker (like a porter), will this too threaten the development of the therapeutic relationship and the recognition of linkages between practices?

6.4.2 A Human-Operated Robot for the Practice of Lifting

An alternative robot exists which presents the potential to re-introduce certain aspects of the moral elements and at the same time to address the burdens of the nurse associated with lifting patients. The class of robots that fit this bill is known as 'exoskeletons'. One example of this robot is Cyberdyne's "Hybrid Assistive Limb" [*Hayashi*, 2005]. The robot is a weight displacing robot so that the human does not feel the full effects of the weight. Versions of this type of robot exist in factory and military applications. They are used to prevent over-exertion of factory workers or soldiers respectively (in industrial applications they are known as 'cobots'). It is not an autonomous robot, but a human-operated one. Another version of an exoskeleton made by Toyota[2] for example, can be used in place here. It too will interact directly with a human (more than one in most instances) and must be programmed for the appropriate safety considerations. Given that the robot is human-operated, the safety considerations for this robot are slightly different from the autonomous robot. For example, the robot will not have the same sensors for perceiving a wall, person or object in its range. This robot, in contrast with the first, does not have a humanoid appearance, but appears rather machine like. This robot can be used in the hospital, nursing home or home setting. While the previous robot, RI-MAN/RIBA is capable of replacing the human care-giver that would normally lift the patient, this robot is meant to assist the human care-giver with their task. It is an enhancing robot. By reading the biometric signals of the care-giver, the robot is able to bear the burden of the weight of whatever the care-giver is lifting. This could be a patient, a bed, a heavy box, etc. We can see with this robot that if used for the rehabilitation of a patient unable to walk it is a rehabilitative robot [*Kawamoto and Sankai*, 2002], whereas, when it is used in the hospital by a nurse, it is considered a type of care robot [*Satoh et al.*, 2009].

The practice of lifting using the human-operated robot proceeds in a similar manner as the practice of lifting with human actors. The nurse, wearing the robot, enters the room, indicates to the patient that it is time for lifting, encloses the curtain around the patient and begins to lift the patient with careful attention to the speed, angle and force with which lifting occurs for this particular patient. The nurse's attention is not directed towards the suit or a remote control and consequently, she/he is capable of engaging in eye contact with the patient to pick up on any non-verbal cues. Their presence also allows them to

[2]See http://www.technologyreview.com/blog/helloworld/27330/

converse with the patient if desired and/or needed. They touch the patient in a manner of speaking; touch, however, occurs through the robot's apparatus rather than by human-human touch. In the case of HAL, the element of attentiveness is still in the domain of the human as is the element of reciprocity. The care-giver uses his/her faculties to ascertain when the care-receiver needs to be lifted, at what speed, from which angle, and with or without social interaction. For the latter, reciprocity is something that happens between the care-giver and care-receiver in real time by verbal and non-verbal cues which are detected by the care-giver. This means that the nurse can ask the patient how they are doing while they are being lifted. The nurse is present to observe non-verbal cues which supplies them the opportunity to learn about the patient.

As for responsibility and competence, these elements now become shared endeavours between the human and the robot. The robot here has the role of weight bearer and is responsible for carrying the weight of the patient but for nothing else. The care-receiver and care-giver must both trust the technology – responsibility for the safety of the practice becomes a hybrid event between the human care-giver and the robot helper the nurse must be capable of competently using the robot. Thus, a certain amount of competence for the skilful completion of the practice is delegated to the robot. The robot must accurately translate the movements of the care-giver into its own movements with synchronicity just as the surgical robot translates the movements of the surgeon's hands into movements of the robot hands. Therefore, a portion of the responsibility for lifting is delegated to the robot as is a certain level of skill. Such robots are not endowed with tactile sensation or force feedback, at this time, and one might question whether or not this is the kind of information the robot must be capable of acquiring in order to adequately label the robot competent. Alternatively, as in the case of surgical robots, a lack of force feedback has demanded that the surgeon be trained for a new style of surgery. Thus, the surgeon conforms to the technology rather than the alternative . In short, the robot is assistive. Given the role and responsibility it has been delegated, the overall successful completion of the practice of lifting remains the responsibility of the human care-giver such that they are accountable for a failure .

So what does this mean for the overall process of care if the human care-giver and the robot share certain roles and responsibilities rather than delegating all roles and responsibilities to the robot? To begin with, the human care-giver is present throughout the entire care practice and is focused on the needs of the patient and is monitoring the patient's preferences. The main question of concern is whether or not the robot will pose the same problems as the mechanical lift – will it detract the attention of the care-giver from the patient

to the robot? When the robot is worn by the nurse there is no remote for the nurse to focus on. The robot essentially becomes an extension of the nurse in his/her role of lifting the patient just as a stethoscope becomes an extension of the physician in his/her role of assessing the patient. In this way, the assumption (and hope) is that the nurse's attention will be directed entirely on the patient whom she/he is lifting. The robot use of the robot also frees up the nurse's attention. She/he can make eye-contact with the patient and can engage in conversation because she/he is not straining to bear the weight of the patient. By ensuring the presence of the nurse throughout this practice, not only are the moral elements safeguarded but so is the linkage between this practice and other practices the nurse engages in with this patient. Consequently, the robot has the potential to re-integrate certain values such as eye contact and a portion of attentiveness while at the same time relieving the burden of the nurse and maintaining the practice of lifting within the overall process of care.

6.5 Attributing Meaning to Design Through Assumptions

It is only through a deeper understanding of what care values are and how they are manifest throughout a care practice that we come to grasp the impact a design might have on the care practice. Above and beyond the direct relationship one can uncover between the moral elements and the technical capabilities of the care robot, there is greater meaning attributed to these capabilities upon further reflection. I suggest that this meaning is uncovered, or revealed, through an analysis of the technology's script and the assumptions prescribing such a script. Akrich discusses the embedding of elements in terms of assumptions made about user preferences and competencies [*Akrich*, 1992]. Placed in context, each robot takes on a distinctive meaning; hence the meaning of the robot has to do with the assumptions embedded within. This description is therefore quite useful for my reflection.

An important distinction must be reiterated here pertaining to the difference between assumptions and values and/or norms. Assumptions are more about the real world, they are descriptive in a sense, while values are more about what the real world ought to be like, they are normative in a sense. When an assumption is made about a value to be embedded, it does not have to be a description about what is, but could also be a claim about what values ought to be expressed, how they ought to be expressed, or what priority they ought

to be given. In others words, when the built-in assumption pertains to a value, or when a valuation is being made, the result is a normative claim about what the values should be, what should be valued, or what the ideal is. For Akrich, "many of the choices made by designers can be seen as decisions about what should be delegated to a machine and what should be left to the initiative of human actors" [*Akrich*, 1992, p. 216]. By making choices about what should and should not be delegated to certain actors (human or non-human), engineers may re-distribute responsibilities in a network.

Assumptions may pertain to any of the elements in the care network; the care robot, the care-giver, the care-receiver, or the care practice. Additionally, I claim there are a variety of *types* of assumptions; those that pertain to the technical content and its practical implications (for example, assumptions about the appropriate speed of a robot), those that pertain to the distribution of roles and responsibilities (for example, assumptions about what is delegated to the robot and what is delegated to the human), those that bear significance at the cultural level (for example, assumptions about the burdens of a care-giver or care-receiver) and those that bear significance (exist) at the feminist level (for example, assumptions about the value of traditional care-givers and their tasks or the appearance of the care robot as feminine). It is not possible to outline a linear order to content or type of assumption because they all seem to co-exist in a web of beliefs and meanings. Thus, although I offer a sequence to the addressing of assumptions, this by no means implies a sequential order to the assumptions themselves. Instead, the assumptions must be understood in a cohesive manner.

6.5.1 Assumptions Pertaining to the Care-Robot

Firstly, let us begin with the assumptions pertaining to the care robot. In both cases we may assume that the designers are presupposing an overall respect and trust in technology as a means of solving a problem – a belief in "the saving power of technology" as Borgmann might say. Additionally, care robots highlight the belief that care is computational, that care may be translated into a computer algorithm (a robot is essentially a sophisticated computer with the element of embodiment). If we refer to the insights made by Latour concerning what the delegation of tasks to a technological artefact presuppose, one might suggest that the attitude this robot embodies (figuratively or symbolically) is one in which the care-receiver is not valued enough to have a human present for their care. Another attitude that the robot may embody is one in which the robot symbolizes the lack of discipline or trust in today's care-givers. This

brings us to the second point made by Latour which is perhaps, the most interesting assumption embodied in the initiative to use care robots in the first place – humans are unreliable and thus their tasks ought to be delegated to a non-human. We must question then whether this is the assumption or belief underscoring the entire initiative to use care robots? What's more, the claims that such robots are intended particularly for the care of elderly or young persons raises issues of ageism. Why are these robots not well suited for other patients, post-operative patients, for example?

These are some of the assumptions or beliefs that may have given rise to the initiative to use care robots in the first place; however, there are also assumptions we may derive from the robot's script which pertain specifically to robots. RI-MAN, being an autonomous robot, presupposes that a robot alone can imitate care and further that the ideal care robot is one that can fulfil its task without requiring the input of a human operator. The appearance of RI-MAN as humanoid presupposes that if a technology were to be incorporated into a care scenario it ought to resemble a human. Or perhaps, the robot's appearance as humanoid is intended to facilitate a greater trust between care robot and care-receiver. Interestingly, one might assume that this robot would be programmed to have social capabilities because it is replacing a human care-giver. In this case, one might assume that human-like capabilities are necessary capacities of a replacement robot which only goes to show the importance of such capacities in the provision of good care. In short, the belief here is that the ideal care robot is one that can fulfil the task of the human care-giver without requiring the human to be present. Consequently, there is also an assumption made about the kinds of roles and responsibilities that one may delegate to the robot. The autonomous robot, being delegated the entire role and responsibility of care-giver for the practice of lifting, is embedded with a script that claims robots ought to be delegated these high levels of responsibility. Such levels of responsibility are those traditionally attributed to the human care-giver recognized as a moral agent. A moral agent is capable of skills like empathic reasoning and moral deliberation and therefore held liable/accountable for any problems that occur. More on this discussion in chapter 8. Alternatively, assumptions pertaining to HAL appear to be somewhat different. Because HAL is human-operated, the assumption about the ideal care robot is such that they may only be used as a tool in the care process that occurs between two humans (a care-giver and a care-receiver). Thus, the care robot is integrated into the therapeutic relationship without threatening its establishment or the elements which it aims to serve. What's more, the role and associated responsibility delegated to the human-operated robot dictates that the final responsibility remains in the domain of the human

actor. Such an assumption may also be translated as a statement on the kinds of responsibilities, or the limits of responsibilities, a robot ought to have.

6.5.2 Assumptions Pertaining to the Care-Receiver and/or the Care-Giver

There are also a variety of assumptions pertaining to the care-giver and the care-receivers in both robot scenarios; we must first return, however, to the initiative to use a care robot and the ideas of Bruno Latour. Latour claims that "delegation to a non-human presupposes a lack of discipline on the part of the human" [*Latour*, 1992]. Are we to assume here that the actual initiative to use a care robot presupposes some kind of mistrust in the capabilities of human care workers? Or is this rather an assumption pertaining to the plethora of responsibilities and duties of the care-giver? And, given the paucity of available healthcare workers, assistance is sought wherever possible. Here, the point must be made that a lack of healthcare workers is a necessary, but not a sufficient criterion for the initiative to use technology as the solution.

Added to this, as Latour also points out, "the non-humans take over the selective attitudes of the humans" [*Latour*, 1992]. For Turkle, this line of questioning demands we question whether or not human nurses will fulfil their roles in a robotic-like manner when working alongside robotic companions. In the same vein, one might question this idea at a cultural level [*Turkle*, 2011]. It is possible to suggest that at a cultural level there is a negative connotation towards ageing, the perception that as we age we are considered more of a burden and generally undervalued [*Neven*, 2010]. Perhaps such a cultural view is taken over by the care robot when used exclusively in the care of elderly persons; the burdens of the ageing population are much too onerous and troublesome, and as such a robot is the most desirable way to fulfil the mundane, repetitive and arduous tasks required for this demographic. Alternatively, if it is presumed that the women of the family group are the traditional care-givers, and if they are to be encouraged to enter or remain in the workforce, then the robot is seen by society as a potential solution to this problem in particular. Such considerations will only become apparent through the policies guiding the use of the care robot, as well as through the the meanings attributed to the robots with continued use.

Through an analysis of assumptions, we may also investigate the ideal care-giver and/or care-receiver presupposed by the design of the care robot. In the case of RI-MAN the ideal care-giver is one which completes its task without question or complaint in a timely, safe and efficient manner. The ideal care-

giver does not need to engage in small talk with the care-receiver, nor does s/he need to assume responsibility for all the daily details of the care-receiver. Thus, the role of the ideal care-giver, in this scenario, is literally that of a tool. The ideal care-receiver is one who will remain patiently in place until the care robot has reached her, who will position herself accordingly for the care robot to lift her at the programmed angle and who will tolerate the minimal level of discomfort that the robot may cause. This ideal care-receiver appears quite passive and is obliged to alter their behaviour in order for the robot to work. This ideal care-receiver is also given an additional range of responsibilities in this scenario and may be seen more as an independent person rather than as one dependent on others. Both the care-giver and the care-receiver are seen in an atomistic way and not as relational persons.

In the case of HAL, there is an assumption that a human care-giver is available to participate in the practice of lifting the patient and, furthermore, that the ideal care-giver ought to be physically present for all steps of the care practice. The ideal care-giver here is also depicted as technologically competent. In the nursing practice, much has been said about the skills of nurses and the idea that technical skill is now an expression or manifestation of *caring about* patients [*Sandelowski*, 1997]. Alternatively, there is also a large body of literature arguing that technology and care are mutually exclusive [*Widdershoven*, 2002]. Demanding that the care-giver be technologically competent prescribes a role and responsibility to the care-giver; the nurse takes on the role of the technician and is responsible for learning how to use the technology as well as for adjusting in the case of a malfunction. Thus, the nurse is required to be skilled not only in the area of attentiveness to the patient, in terms of bestowing empathy and compassion on his/her patients as well as understanding their wishes and needs, but also in the area technology skills and competencies.

From the perspective of Shannon Vallor and the impact these robots may have on the cultivation of care skills, it is clear that the replacement robot offers no means for cultivating the skills, attributes and relationship described for this practice. For the enhancing robot, there still remains the moments in which such skills may be cultivated; however, the skills of the nurse now include knowledge of the technology in addition to human attributes like empathy and compassion.

6.5.3 Assumptions Pertaining to the Practice of Care

Based on the above assessment of care robots, we arrive at an overall picture of the care practice inscribed in the two care robots. Based on the distinctive

technical details of the robots two divergent visions of care, or ideal types of care, may be drawn out. For the first, the autonomous robot, the vision of the ideal care practice is a standardized one with efficiency as the top priority or value. This one-dimensional view of good care as efficient may have negative implications for the care process. One may presume that the quality of interactions, the number of social interactions, and the presence of a human are threatened by this efficient system. Alternatively, the system may be considered efficient considering that time of the human care-giver is freed up, ultimately improving the number and quality of social interactions. This is a claim made by designers and robot enthusiasts. Of considerable significance is the subsequent implementation of the robot in context. For the second robot, the overall aim is to provide individualized care using a human-operated robot with a human care-giver present at all times for all parts of the care practice. Efficiency is still a priority; however, it is achieved through aiding the care-giver. Furthermore, an ideal care practice is one in which the care-giver is present for all stages of the care process regardless of how mundane or burdensome they may be.

6.6 Conclusions

In response to the burdens associated with the practice of lifting, roboticists are currently designing robots to meet such a need. Of course there are many ways in which a lifting robot may be designed to meet such a goal and as a result, such robots may have any range or variety of capabilities. In the above chapter, my goal was to show how the design of the robot is intricately related to the manifestation of moral elements and how a difference in type of robot (enhancement vs. replacement) and its capabilities changes the overall picture of a care practice. An added benefit is to compare the robots with the current practice of lifting.

When we look at the current practice of lifting in the setting of both the hospital and the nursing home, eye contact and touch are values that are not present with the use of the mechanical lift. This has the potential to threaten the element of trust which has been shown to be of great importance in the carrying out of other care practices. RI-MAN/RIBA does not do anything to resolve this discrepancy in values but HAL has the potential to restore them. Thus, HAL presents the potential to improve the quality of care in terms of this specific practice. Consequently, we may easily ask the question; what kind of care do we want to provide? There need not be a common universal answer to this, one culture may choose one vision and another culture may choose another

vision. The important step is making the assumptions generating each vision of care, and the elements within care, explicit. This will allow us to make decisions about the future of caring practices in a well-informed manner. While I may not be sure that my prediction of the assumptions was the epistemic aim of the engineer, it is hoped that this study will encourage and inform forward thinking about what the technical content reveals about the potential users.

This analysis is also not intended to conclude that RI-MAN ought to be disregarded or labelled as unethical. Illustrating this robot in a different context could conceivably alter the outcome/consequences. For instance, in the home of two elderly persons who may not be equipped for wearing HAL, or who may not want to burden their spouse in regard to the practice of lifting, or whose spouse is not physically capable of lifting another person, RI-MAN may be the more suitable, ethical choice. Clearly, decisions concerning the use of a robot and its ethical implications are multi-faceted and complicated and demand an understanding of the specific context and users in order to anticipate how the elements will be served to their greatest potential.

Based on the above evaluation of these two types of robots, recommendations can be made about potential future designs as well as the implementation of the robot. For the human-operated robot, one may suggest that the robot be endowed with force feedback and/or tactile sensation/perception in order to enhance attentiveness and competence in lifting. The current HAL prototype is also somewhat big and bulky in contrast with the newer models proposed by Toyota. One might also suggest that the exoskeleton be made in a way that its presence is not as noticeable. For the autonomous robot, still in the early stages of development, one might suggest a range of capabilities the robot ought to have: force feedback and tactile sensation, high level of sophistication to pick up the patient regardless of the patient's positioning, and multi-modal communication platforms to communicate with the patient.

In terms of the implementation of the robots, the replacement robot in the hospital explicitly threatened the link between the practice of lifting and the overall process of care (seen when the robot is thought to eliminate a moment for establishing the therapeutic relationship between care-giver and care-receiver). Consequently, in an institutional setting which places a significant weight on the establishing of a bond between care-giver and care-receiver, such a robot would present as ethically problematic. I suggested earlier that it is possible to have a human escort, responsible for enhancing the care-receiver's trust in and comfort with, the robot, but if the main incentive for using care robots is a lack of personnel then one would question whether or not this is realistic. Alternatively, the same robot used in a home setting in which a bond is already formed

between the care-receiver and their predominant care-giver, the same moment for establishing and/or maintaining the trusting bond may not be necessary. It would appear that in a home setting, an autonomous robot does not present itself to be ethically problematic.

Only when we understand the relationship between the robot's capabilities and the resulting care practice can we come to understand in what way a care robot should be designed. What's more, once the robot is integrated as an actor into the network of a care practice, it too takes on a role and a responsibility. With this in mind, I now move on to another practice for the same kind of evaluation.

Chapter 7

Care Robots and the Practice of Feeding

7.1 Introduction

ACCORDING to the Care-Centered (CC) framework, the ethical evaluation of a care robot begins with an in-depth understanding of the care practice for which the robot is intended. In doing so, the evaluation is tailored both to the practice and the care robot understood in context. What's more, the relationship one practice shares with the overall care of an individual is made explicit. Consequently, the care robot is evaluated according to the impact it exhibits not only on the one practice in which it is used but on the many other interconnected practices.

In line with the practice-by-practice design approach I argue in favour of, I now turn my attention to another practice. Another activity of daily living (ADL) for which care robots are currently being developed is the practice of feeding. By 'practice of feeding' I refer not only to the moment of eating assistance but to the additional connected practices also considered part of feeding. While the main goal of chapter 6 was to highlight the relationship between a care robot's capabilities and the resulting expression of values, chapter 7 explores another prominent difficulty associated with the study of practices: knowing when one practice ends and another begins. To do this, I investigate three separate moments subsumed within the practice of feeding using the CC framework and the methodology for retrospective evaluation. Each of these moments may also

be considered a practice in its own right and as such I refer to the practice of feeding as the process of feeding and each of the moments are referred to as practices. These practices are: the dietician's assessment, feeding assistance and, food tray removal. Accordingly, I then present and evaluate three separate care robots intended for each of the three practices presented. The aim is to strengthen the reader's conceptualization of the interconnectedness of practices and as such the care robot's potential to impact other practices outside the one for which it has been designed.

7.2 The Process of Feeding

Similar to the many practices incorporated into care, the practice of feeding is a crucial one, yet a complicated one. Poor nutritional status has been studied extensively and shown to have a negative impact on patient rehabilitation, length of stay in an institution, complication rates, mortality and the cost of health care [*Correia and Waitzberg*, 2003]. One's dietary regime is an integral component to one's care and is often tailored to one's particular background, diagnosis, and preferences. Thus feeding and food play a direct role in the physiological good of the patient. Feeding also provides a moment of social interaction and a moment in which the bond is formed between care-givers and care-receivers. What's more, feeding is a moment, much like that of bathing and lifting, important for gaining understanding of an individual patient's particular history, preferences, life-style and character. Feeding, like bathing and lifting, is not an isolated event, rather, it is connected to the overall picture of care. This moment for social interaction affords the care-giver the opportunity and space to gain a better insight into the patient as a person as well as keeping tabs on the patient's well-being at that moment in time. This information is then applied to the care of that patient in other practices.

But what is meant by the practice of feeding? There is much to be said about the variety of actions involved in the practice of feeding and their interconnectedness, and this variety often blurs the lines of knowing when a practice has begun and when it has ended. Do we consider the practice to begin with the assessment of the patient by the dietician in order to create a meal/care plan? Or does feeding begin when the food has arrived at the patient's room and the patient begins to eat what has been prepared. Alternatively when does the practice of feeding end? When the nurse has assessed what the patient ate and didn't eat in a day? Or does it end when the food tray has been taken away from a patient's room. In order to do justice to the practice of feeding,

the many actions and actors involved, I refer to the practice as a process. Each of the moments within this process are then conceptualized as practices on their own, interconnected through their relationship to the nutritional well-being of a patient. I begin with the practice of the dietician's assessment for the creation of a meal plan, which is integrated into a patient's care plan. I refer to this practice as 'the practice of diet assessment'. I continue with the actual moment of eating and refer to this practice as 'the practice of eating assistance'. I conclude with the removal of the food tray from a patient's room, referring to an assessment of the patient's nutritional intake, and refer to this practice as 'the practice of food tray removal'. For each of these moments in the practice of feeding, there is a care robot already in development to which a role and/or responsibility will be delegated. With the analysis of each moment in the process of feeding I will address the role the care robot is intended to play and its impact on the manifestation of moral elements.

7.3 The Practice of Diet Assessment

It is possible to suggest that the process of feeding, when seen in holistic terms, begins with the dietician's initial assessment of the patient in order to create a nutritional plan [1]. In long-term care facilities in which the context is the hospital or nursing home, the dietician visits the patient in their room for the assessment and passes on the nutrition plan to the kitchen staff. The responsibility for ensuring that the patient complies with the nutrition plan is a shared endeavour amongst the cooking staff, nurses, dieticians, and physicians. The dietician communicates with other healthcare personnel to make sure the patient is eating, drinking and taking their supplements. It is the role of the dietician to stress the significant relationship between the patient's nutritional status and their outcome to patients as well as other healthcare personnel. For example, burn victims need high protein and high caloric intake in order to get better faster: proteins for making the necessary skin cells and calories because of the amount of energy needed to do so.

Alternatively, in a home-care setting, the dietician will visit the home of the patient for the assessment and will give the nutrition plan to the patient themselves. Responsibility for complying with the nutrition plan lies in the hands of the patient at this point with check-ups by the dietician. Autonomy

[1] Information pertaining to the practice of diet assistance comes from both dietician literature as well as personal communication with a dietician in a long-term care facility (in London, Ontario) who has also practised in home-care settings as well.

is very important in the home setting and needs to be respected; "people will lie to keep their autonomy" [2]. Autonomy in this instance refers to not having other people nag at them, not being restricted in their choices and not feeling burdened. It's also important to give reminders if the patient forgets, so there is a fine balance between the two values – the autonomy of the patient and the physiological good of the patient. Autonomy in an institution is also important but differs because patients are comfortable with having others watch over them and are also more used to human interaction. Patients need to feel like they have *some* control over their dietary habits and will often want a different portion size of the food available just to exercise this. The difference in the autonomy of the patient plays a role in the responsibility the patient has for complying with their care plan. In an institutional setting, other healthcare professionals ensure that the patient comply with their nutrition plan; however, this presupposes that the patient trusts their healthcare providers and will comply. In the home-setting, the patient must also trust the dietician in order to comply with their nutrition plan but because there is little interaction between the two, it is crucial for the dietician and patient to establish a trusting rapport during the assessment.

The purpose of the dietician's assessment in both contexts is to understand the details of the patient's life and to tailor a nutrition plan according to this. When in a home setting the dietician must understand aspects like: how much the patient makes, can they access a grocery store easily, do they know what their diet should be, what is their culture, and/or what is their religion? All these details are considered the *determinants of health*. The dietician must also assess the stage the patient is in in terms of making a life change.

> *It's important to understand the relationship between body and mind when we care for the body; recognizing their culture, food preference ability to cook or not, social and economic situation, religious beliefs, and physiological aspects like blood work, etc. We have to take in all that and decide the diet most suitable for that particular patient based on all that information. Then we have to ask ourselves, can they do it (i.e. start the nutrition plan) and if not what steps do I need to take to get them started (instead of eating 10 cookies a day get them to 5 for example).*[3]

For patients in a home-care setting it is vital that a therapeutic relationship is

[2] Personal communication with dietician

[3] Information obtained from interview with a registered dietician (i.e., personal communication) at St. Joseph's Hospital, London, Ontario, Canada

formed during this assessment. Often times the dietician will make more than one house call in order to get the necessary information in the short time slot available; patients don't want to talk about personal information before they trust the dietician. The patient needs to believe the dietician *cares about* the patient before they may *care for* the patient.

> *During the assessment I also get a feel for how much involvement is needed from me and what their relationship to food is. The assessment establishes the bond/relationship and without this you can't guarantee they will be honest about their behaviours or will do what you recommend. People emotionally eat so I need to be aware of this. Lots of times we don't even talk about food at first because they need to know I care. And when they're ready to talk about food they will.*

In an institutional setting, the patient and dietician meet for the same goal but in this context the nutrition plan is heavily reliant on the physiological status of the patient rather than on personal eating habits. The dietician does not need to assess the income or ease of access to a grocery store because this information will not be integrated into the plan. It is of equal importance for the dietician to get to know the details of the patient in order to tailor a nutrition plan and to create a relationship so the patient will do as the dietician has recommended. In both an institutional and a home-care setting, assessments are done within a 15min time slot and the dietician won't see more than six patients in a day. Most often there is one dietician for an institution consisting of anywhere from 100 to 500 patients. The dietician must work within these time constraints because assessing the patient isn't all they have to do in the institution, they must also check up on the other patients as well. Thus, efficiency is a highly valued aspect in the work of the dietician.

In short, care during the assessment is about understanding and getting to know the patient in order to provide future care. After the assessment, care from the dietician shifts: "care also involves coaching and mentoring in not making them feel bad if they get off track but also encouraging them not to get back into bad habits". In translating the role of the dietician during each of these moments – assessment and post assessment – we may observe a shift in their role and responsibilities as carer. During the assessment the dietician is a carer in the sense that they want to contribute to a form of physiological care. They are responsible for acquiring the relevant information which will allow them to target the appropriate means for care in this sense. Post assessment, the dietician still aims to provide physiological care but this responsibility is

a shared effort amongst other healthcare personnel (when in an institution) or the patient themselves (when a home setting) and the dietician then acts as the coach.

When taking into account the manifestation of moral elements we may observe a shift in the interpretation and ranking of elements between assessment and post-assessment. During the assessment, attentiveness refers to the dietician taking in multiple cues from the patient, both verbal and non-verbal. The dietician must know when to ask probing questions and when to wait until the patient feels comfortable. Competence then has to do with the skill with which the dietician asks questions and reads the cues of the patient. Given the necessity for patients to feel *cared about* before *caring for* can happen, the dietician must convey their sense of responsibility for the well-being of the patient and the future success of their care. A successful reciprocal relationship, in which the patient speaks to the dietician openly and honestly and complies with their nutrition plan, is dependent on the other elements. The dietician shows their responsibility for the patient and competently ascertains the relevant information from the patient. The context in which the assessment happens does not change this.

Post-assessment the dietician is responsible for mentoring and coaching the patient to stick with their nutrition plan but at this time the home dieter takes on the role and responsibility of their daily progress. They are often required to complete a daily log to keep track of their daily calories consumed and their daily amount of exercise. In the home setting, the dietician must be attentive to the changing status of the patient – are they struggling with the nutrition plan, are they lying about following it, are they confident about their progress. Based on this the dietician knows how to proceed with coaching, acknowledgements or pep talks. The dietician must also be attentive to the patient's perception of their own autonomy to proceed – does the patient feel burdened by the diet plan or empowered? Competence is understood in terms of how well the dietician can perceive the status of the patient but also how well the dietician responds to the patient's existential state. Again, the dietician must ensure a sense of responsibility on their part for the well-being of the patient and the continued success of their weight loss goals. Successful reciprocity happens when the patient is honest about their status and whether or not they are following the plan which presupposes a trusting relationship between the two. Thus, during the post-assessment phase, the role and responsibility of the dietician has shifted to coach and the patient now bears the majority of the responsibility for their daily actions. Thus, the element of responsibility is now understood in terms of the patient's responsibility for their own actions and is ranked above all other

elements for successful care. As such, the patient may then be praised or blamed for adhering to or straying from the diet/care plan, respectively.

In the hospital or nursing home, the dietician is also responsible for following the progress of the patient but the dietician does not have the same role as coach or mentor. This relates back to the difference in the autonomy as a consequence of the context – the patient in the hospital is not responsible for their own meal preparation. The dietician then is responsible for ensuring the other healthcare personnel are facilitating the patient's nutrition plan. Thus, unlike the home care setting in which the patient acquired a new role and responsibility post assessment, in the institutional setting a certain portion of the role and responsibility of 'supporter' is delegated to other healthcare personnel. In this context, the dietician must be attentive to the physiological status of the patient – is the nutrition plan having the intended/desired effects? Why or why not? The patient's body provides information and thus reciprocity does not rely solely on verbal communication. Competence is linked with how well the dietician can acquire this information about the patient but is also related to how the patient is doing. If the nutrition plan is having the desired effect should there be any adjustments. Alternatively, if the nutrition plan is not having the desired effect, what should the next step be?

Overall, the care ethics tradition as well as others (the psychoanalytic tradition for example), it is the relationship formed between the patient and dietician (or other care-giver) that facilitates 'cure' in the respect described above and the moral elements are the driving force behind the establishment and maintenance of this relationship. Although roles and responsibilities shift at different stages, as do the interpretation of moral elements, the relationship between care-giver and care-receiver remains strong. This ensures a chain of responsibility for the care of the patient. Regardless of the dietician's role shifting from 'curer' to 'coach', the dietician plays an integral role at the starting point and throughout the long term nutritional care of the patient.

7.3.1 A Diet Assist Robot

In certain instances, when the patient is obese, discussions concerning the patient's care plan may take place with a physician rather than a dietician or with both. In other instances, if no dietician is available, the conversations, mentoring and coaching will also take place with the physician. For the purposes of this investigation, it does not matter if it is the physician or dietician performing the assessment, mentoring and coaching. What matters is that it is a human care-giver working within the value-laden milieu of an institution that engages

in a human-human assessment with the patient.

The robot I evaluate here is known as the 'diet assist robot', created by a researcher at MIT Corey Kidd, working with sociable robotics expert Cynthia Breazeal. In the initial stages of planning, the robotic platform included the iRobot's®Aibo robotic dog. The robot was intended to be used by patients, previously obese, who had lost a significant amount of weight and would like to keep this weight off. Kidd and Breazeal discuss the real world problems of these scenarios and how a robot might help to mitigate these [*Kidd and Breazeal*, 2006; *Kidd et al.*, 2008]. For starters, the patient is requested to keep a daily record of their caloric intake and caloric output via. Not only do patients lose interest in keeping track of this information but they are often not accurate in their estimates. The robot was intended to keep an accurate record of the patient's caloric intake and output and provide this information both for the patient's education and the physician's records. The robot was also meant to sustain the patient's interest and motivation through both its physical presence as well as its sociable capabilities. In fact, Kidd and Brazeal discuss the remarkable potential that social robots will have on patients in a home setting requiring reminders and motivation for behavioural change. The sociable robot would provide benefits in terms of social support which is so crucial in the case of weight loss (or quitting smoking). In order to do this the robot would have to establish itself as trustworthy (reliable and consistent), would have to engage the patient (keep the patient's attention over one discussion and over a long period of time) and would have to help motivate the patient (in terms of fulfilling one's goals but also in terms of being motivated to care for and engage with the robot) .

Although the initial stages of the design process involved Aibo, this robotic platform/interface provided too many obstacles and instead Kidd and Brazeal opted for a stationary robotic platform with a humanoid appearance. This robot, named Autom, is 15 inches tall with a touch-screen belly, big eyes and a female voice. It is stationary so it sits in a person's home on the kitchen table, an office desk or somewhere else accessible. The 'assessment', if one may call it that, requires that "the user provide some baseline information about food and exercise and the robot then offers encouragement if people slip up and suggestions for how to better stay on track" [*Turkle*, 2011, p. 114]. According to the project description:

> *The (robot) coach offers feedback on recent behaviour and makes recommendations for near-term behaviour. The feedback is based on comparing recent diet-related behaviour, such as calories consumed*

> *and exercise performed, with goals set by the user. Recommendations come from general information on diet, nutrition, and exercise and are tailored to the individual based on the current stage of the relationship between the coach and the user. The system maintains a database that keeps track of interactions with the user, information gathered from the user, and goals set by the user. This is used in the relationship model and for the feedback to determine how each interaction should occur. The system consists of an interactive robot coach and a computer that maintains the necessary information.* [4]

Using this robot makes for an interesting analysis when we discuss not only the moral elements but also the distribution of roles and responsibilities. Although the robot was originally intended for those patients that had already lost a significant amount of weight and would use the robot for continued success, the robot is now being marketed as a weight loss and diet maintenance robot. Consequently, it can either be used to begin the weight loss or to continue with one's diet plans. Thus, although a care robot has formerly been defined as having to be included within a therapeutic relationship, the therapeutic relationship now occurs between the robot and the human user. For starters, using the robot has removed the moment of dietician or physician assessment. The patient is responsible for inputting their information so the robot can calculate the necessary plan for the patient based on weight loss goals. This reaffirms that the once human-human relationship for diet assessment is now a human-robot relationship for the same goal. There is no recognition here that the patient will need to feel *cared about* before allowing the robot to *care for* it. Also, there is very little information that the robot will take in for its assessment. The robot will not be sensitive to the patient's economic status, ease of accessing a grocery store, religious and/or cultural determinants. The robot will not be able to comprehend what the patient's relationship to food is, existentially speaking, or what the small steps are that the patient will need to take in order to get started. Although a future version of this robot may be able to ask the questions about the patient's relationship to food the robot has no experience in this regard and will not understand the meaning of the patient's stance. In short, the robot's capacity for attentiveness, competence, and reciprocity during the assessment is minimal – the robot is not capable of understanding the multi-dimensions of the patient as a person. What the robot does provide in the moment of assessment is a greater amount of time for the patient and robot to interact.

[4] http://web.media.mit.edu/~coryk/weightloss.html

Post assessment, when the robot takes on the role of coach/mentor, we see the robot's strengths in terms of attentiveness and competence. The robot's competence to obtain accurate information regarding the patient's caloric intake and output is superior to the dietician's. The robot's ability to interact with the patient on a daily basis and provide daily reminders and cues is also superior to the dietician's. The robot's attentiveness to the patient's daily activities and therefore its competence to tailor its daily instructions and goals to the patient trumps the capabilities of the dietician. Of course all this has to do with the amount of time that the dietician has for these kinds of activities. Dieticians, and physician's for that matter, are few and far between and cannot possibly provide daily support in the same way that the robot can. But does this mean that the robot should be used in place of a dietician and/or physician? Is it not possible to have a dietician provide the assessment and program the robot according to the personal details of the patient? This would increase the amount of overall attentiveness of the care-giver (referring to a combination between the dietician and the robot). These questions highlight the potential assumptions embedded in the robot. Given the significance of the assessment for establishing a relationship, should the assessment not be the most integral part of this practice? Or alternatively, is it enough to have the robot fulfil this role exclusively?

In terms of the delegation of roles and responsibilities, whereas without the robot the patient was responsible for tracking their daily caloric intake and output, often inaccurately, the responsibility was still in their hands. Their weight loss regime was within their own control and it is their own internal motivation and consistency that is developed throughout the process. Of course this poses a problem for patient's who require time to develop this but by delegating the role of motivator to the robot one must question whether the patient has the room to develop their own internal motivation. We might assume that the responsibility for weight loss becomes a hybrid affair between the dieter and the robot. But what kind of responsibility could the robot have in this instance? The robot will not be punished or praised if the dieter succeeds so how is one to interpret the element of responsibility in this instance? The robot is neither responsible nor accountable for the consequences or outcome of its actions.

7.3.2 The Meaning of the Diet Assist Robot

Although the robot is capable of enhancing the promotion of certain moral elements, namely a higher level of competence and attentiveness to the physiolo-

gical status of the patient, is this enough?

Or, is this the exact fear that Turkle is talking about – the ability of the robot to create a lasting relationship with its user to encourage long term compliance with a diet plan [*Turkle*, 2011]. The user engages with the robot on a daily basis in a reciprocal manner, the robot's social capabilities encourage this interaction and facilitate a lasting relationship (the robot is programmed to alter its suggestions and behaviours based on the patient's progress). Alongside Turkle, I am left wondering what a reciprocal relationship really is and further what does it mean to encourage this kind of relationship with a robot rather than a person? This is where the assumptions of the designer come into play. Is the relationship based merely on an exchange of information in order to fulfil the goal? Isn't a reciprocal relationship about more than this? Is it not about forming a bond between the dietician/coach and the patient? This bond helps to encourage and motivate the patient to grow and change based on their own internal goals and motivations but feeling the support of the human care-giver. By placing the robot in the home for both motivation and accounting of information, isn't some form of motivation being delegated to the robot instead of encouraging development within the patient? However, does the support need to come from a human care-giver or is it not possible to suggest that given the necessity for internal motivation, as long as the dieter is motivated to achieve his/her goals the source of this motivation is irrelevant. Or, is this ultimately reduced to a discussion of the existential state of the patient? If one were to view the dieter/patient in the atomistic sense, one might conclude that the origin of their motivation need not matter. Alternatively, from the care ethics perspective, the individual is viewed as relational – defined by, dependent on, and a product of their relationship to others and the material world. Thus, forming a bond with another human being, being vulnerable in that bond, is in fact a part of what it means to be the relational being and to cultivate the skills of such a being.

Such an insight reflects certain fears proposed by Vallor [*Vallor*, 2011]. Vallor is concerned with the cultivation of skills of the care-giver, and one might wonder here about the impact on the dietician. One would assume that the profession of the dietician would not be in jeopardy but instead the dietician would be restricted to seeing only a subset of the patients requiring assistance. Would these patients be those who are the worst off thereby limiting the development of certain skills of the dietician? Alternatively, in light of the little time dieticians have for assessments, it is possible that the robot may help to enhance efficiency. The robot could be a benefit in the hospital or nursing home for keeping the records of multiple patients: recording their daily progress and making this

information available to the dietician, nurses and physicians. Perhaps in this way the robot is able to increase the attentiveness of the dietician without demanding that the dietician be present on a daily basis and instead the dietician has more time for longer assessments with patients. In short, this idea supports the claim made by Brey within the realm of the structural ethics approach – that a human must always be responsible for the outcome of actions [*Brey*, 2012]. Thus, the recommendation from this perspective is to create/design the robot in a way that it enhances the capabilities of the human dietician while at the same time enforces the human agent's responsibility.

7.4 The Practice of Eating Assistance

I now shift practices in the process of feeding, to the actual moment in which a patient is consumes food; the practice of eating assistance. Here there are many ways in which the practice of eating takes place. Certain clients may be fed in their room if they are bed-ridden while other clients gather in the dining room for their meals . For feeding in a common dining room, certain clients are capable of feeding themselves, once the meal has been brought to them, while others require direct assistance for bringing the utensils to their mouths. In both cases the actors involved are: the client (the care-receiver), the care-giver (in most instances a nurse but often times a family member may come to help), nurses who distribute medication throughout the meal but who are not involved with the feeding, the cooks who remain somewhat isolated from the care-givers and care-receivers but are still present , the tables, the chairs, the utensils and the plates. Care-receivers often have a choice between two options and are asked what their preference would be. A care-giver then brings their choice to the table and places it in front of them. If there is no choice, a care-giver still brings the meal to the table and places it in front of the care-receiver. For those that are capable of eating on their own, they do so. For those that require assistance, a care-giver is seated beside them and begins to cut the food appropriately and slowly, according to the pace at which the care-receiver is eating, brings the food to their lips using a fork, lets them take the food with their mouth, takes the fork back to the plate, replaces it with more food and repeats the above steps until the food is done or the patient has indicated they are done eating. During this time there is often small dialogue about the temperature of the food, the taste of the food, how the care-receiver's day has been, etc. When the care-receiver decides they are finished eating (regardless of whether or not all the food is done) the care-receiver removes their plate and utensils, brings

them to the kitchen for cleaning and begins to bring the patient back to their room or a common TV room. At the same time there is often a nurse present responsible for distributing medications to care-receivers.

In terms of attentiveness, this element is expressed through the entire practice of eating; the care-giver is required to be attentive to the preferences of the care-receiver regarding of their choice for food but must also be aware of whether the temperature of the food is adequate. This also requires reciprocity, however, as the care-receiver will indicate whether the temperature of the food is satisfactory or not. Thus, although attentiveness is required on the part of the care-giver, there is an element of reciprocity that is required throughout the entire practice, as in most practices. Being honest about these things is something that requires trust on the part of the care-receiver as we saw previously when discussing the practices of lifting and bathing – if they do not trust their care-giver or the cook(s), they may not be as honest or forthcoming with their preferences. Note here, however, that reciprocity in the practice of eating, as in many other practices, does not demand that the care-receiver speaks. In fact, silent patients do not drastically hinder the work of nurses and other carers. Nurses address the needs and preferences of silent patients every day. When they do not know the individual preferences of a patient, through trial and error, they come to acquire this information [*Harbers et al.*, 2002].

The element of attentiveness as described above is consistent with the value of *tinkering* coined by Mol [*Mol et al.*, 2010]. Tinkering refers to the balancing of different goods according to a particular care-receiver with a unique history and preferences, in a given context. We might suggest that tinkering presupposes an understanding of Louvain personalism – understanding the many dimensions of an individual [*Vanlaere and Gastmans*, 2011]. Tinkering also refers to an attentiveness to what does *not* work [*Mol et al.*, 2010]. Thus, there is an active and engaged taking-in of information for use in the future. This is both in terms of the future eating of the care-receiver as well as in other practices. Competence is thus dependent on the attentiveness of the care-giver – only when the care-giver is able to provide individualised care, by knowing the patient, is care competently provided in this instance. It follows then that understanding the likes and dislikes of a patient is the responsibility of the nurse. Of course this is one dimension of the practice of eating, when the nurse is directly feeding the patient, competence also refers to the speed at which the nurse feeds the patient, the amount of food placed on the utensil, the force used by the nurse and the length and amount of pauses in between for chewing and digesting. These are all variables associated with the practice of eating in terms of logistics. Preferences along the lines of these variables will also need to be determined

by the nurse and again the competent nurse feeds according to the patient's individual preferences.

Responsibility in the practice of eating is understood in terms of the nurse's competent completion of the logistics or mechanics of eating (raising the fork to the patient's mouth, acquiring food on the fork, checking the temperature of the food, etc.). The nurse may therefore be help accountable and liable if the patient is injured in this practice (via the fork or temperature of the food). Responsibility in this practice is also related to the nurse's role in knowing what the patient is eating and why. In this sense, the nurse communicated with the dietician and/or physician to discuss the eating habits of the patient with respect to their overall care. For patients that need to take their medications with meals, the nurse is responsible for making sure that patient gets the proper medication and that the patient takes it. In certain instances, a specific nurse is there to distribute the medication and to ensure that each patient receives the medication intended for them; however, it is the nurse assisting with feeding that ensures the medication is taken.

Not only is the nurse's presence one that facilitates attentiveness (i.e., learning about the patient) in order to provide good care, but it is also a moment once again during which the nurse and patient can establish and/or maintain a therapeutic relationship. Research has shown how the relationship between care-giver and care-receiver is linked with nutritional care and focusing on this relationship can enhance nutritional care [Evans et al 2005 from *Wright et al.*, 2006]. What's more, "mealtimes can also be used to practice skills relevant to rehabilitation, such as improving mobility, dexterity and mental state [*Davies and Snaith*, 1980]. "Mealtimes have been highlighted as a good opportunity for social engagement as well as practical assistance with nutrition by nursing staff, which may promote independence and self-care for older people [*Stabell et al.*, 2004]. Consequently, the practice of feeding takes on a much larger role. It is not simply a daily activity/task which must be fulfilled but rather it is a moment in one's care that contains much depth and significance for one's overall care.

It is possible to say that a care-giver has fulfilled the necessary steps of the practice of feeding while paying attention to the moral elements; however, it is also important to draw our attention back to the manner in which a practice is fulfilled. In Annemarie Mol's paper "Care and its values; good food in the nursing home" she eloquently illustrates the complexity of the practice of feeding referring to food, eating and the cosiness of the eating environment [*Mol*, 2010]. In the paper, Mol reflects on her experiences in two care home facilities for elderly persons and draws attention to the plurality of goods and values as-

sociated with nourishing care – the preparing, distributing and eating of food: "while values sometimes go together, on other occasions they clash – giving rise to ongoing tensions or a victory of one alternative over the other" [*Mol*, 2010, p. 216]. Mol claims that condensing all the values associated with *nourishing* care cannot be done on a one-dimensional scale of quality assessment or accountability schemes. Consequently, "the overall quality of nourishing care does not depend on the addition of *bits* of *good* along a single scale, but on tinkering with different goods that map onto different dimensions" [*Mol*, 2010, 216 - 17]. Mol distinguishes between dimensions involved in the practice: nutritional value vs. cosiness; and the taste of the eater vs. the taste of the food. In other words, meals are not just about the substance, they are practices, or events as Mols says, that are 'sociable and cosy'. The subject that concerns me for the purposes of this work is about neither the taste of the eater nor the taste of the food but of the cosiness of the practice and its relationship to good care. Cosiness is accomplished through the human-human interactions of the care-giver and care-receiver. Thus, not only is the practice of feeding about acquiring information pertaining to their preferences in order to tailor care accordingly, but the manner in which this interaction happens contributes to the experience and the meaning of the experience. I do not claim that taste of the food and of the eater is not something important and in need of addressing but it is not something that the robot will have an impact on or at least I am not addressing such a robot in my analysis . By making note of this factor I hope to illustrate once again the multiple dimensions and variables involved with the practice of feeding.

7.4.1 A care robot for eating assistance

There are a few robots currently on the market that assist with the moment of feeding. One such robot is known as Bestic, a battery powered eating aid available on the market. Bestic was created through a process of user-centered development and is tailored to each individual user's abilities. When using the robot, according to the designers:

"The user controls both the settings and the operating cycle when using Bestic.

The user first chooses a piece of food, and then directs Bestic to pick it up by one of the two methods described below. The user then leans forward, takes the food, and gives a new command, which lowers the spoon back down to the plate.

- **Controlled**: Controlled with joystick or multifunction buttons. The user steers the spoon freely over the plate and gives a command when the food shall be lifted to the mouth.

- **Simplified Control**: Controlled with one button. The user turns the plate so that the desired piece of food is located were the spoon can reach it. A command is then given and the spoon moves towards the plate's right edge where it scoops up the food, and lifts it to the individual's mouth."[5]

Perhaps the most widely known of the feeding assist robots is Secom's MySpoon. Developed in Japan, MySpoon is intended for patient's lacking the ability to feed themselves. The robot can be used in a nursing home or hospital setting as well as a home setting. The robot has three modes of operation; manual mode, semi-automatic mode, and automatic mode. Each mode differs in the amount of user input required for the robot to function. In manual mode, the user has maximum control and controls the spoon with a joystick. The steps for manual mode are as follows:

1. Using the joystick, select the compartment from the included tray which contains the desired food item.

2. After the spoon reaches the compartment, use the joystick to fine-tune its position near the item.

3. After the spoon is in place, instruct the spoon to grasp the food by using the joystick.

4. The spoon will grasp the item and automatically approach the mouth. When the mouth comes in contact with the spoon, the fork will automatically retract.

Semi-automatic mode simplifies the users operation but prevents the user from selecting a specific food, instead the user indicates the compartment and My Spoon will pick up whatever food is in that compartment. The steps for semi-automatic mode are as follows:

1. Using the joystick, select the compartment from the included tray which contains the desired food item.

[5] http://www.bestic.se/eng/beskrivning.html

2. The spoon will automatically grasp an item and approach the mouth. When the mouth comes in contact with the spoon, the fork will automatically retract.

Automatic mode reduces the user input once again. This time, the user simply presses the on button and My Spoon will begin to pick up food. The steps for automatic mode are as follows:

1. When the button is pressed, the spoon will automatically grasp a food item from within the included tray and carry it towards the mouth.

2. When the mouth comes in contact with the spoon, the fork will automatically retract.[6]

Both robots are intended for the home setting but also have the potential to be included into a nursing home context when multiple patients exhibit difficulties with eating and not enough care providers are able to assist. I will begin by analysing the use of the robot in a nursing home context. Although both robots are quite similar, I focus my analysis on Secom's MySpoon to investigate its range in modes of operation. Each of these modes of operation results in a different picture in terms of the impact on the moral elements as well as changing the amount of responsibility delegated to the robot as well as the role the robot takes on. For example, using My Spoon in automatic mode renders the robot as a replacement robot – the robot replaces a human (either the care-giver or the care-receiver) from involvement in the practice. Alternatively, when the robot is used in manual mode, it becomes an assistive device. One in which the care-giver and/or the care-receiver still plays an active role in the practice of eating. Thus, the already complex practice of eating is made even more so with the introduction of the robot in terms of delineating the distribution of roles and responsibilities based on its mode of operation.

The same holds for the impact on the manifestation of moral elements – their impact is dependent on the mode of operation. When in automatic mode the robot is delegated numerous roles and responsibilities; from selecting the food in the container to the actual action of feeding. The robot is attentive in that it responds to cues given by the patient (pressing the button to start) and to its physical interaction with the patient (it will automatically retract once it has come into contact with the patient's mouth). The robot has control over which food to select but is not attentive in the sense that it knows the

[6] All descriptions of steps for feeding were taken from the Secom home site describing MY Spoon: http://www.secom.co.jp/english/myspoon/usage.html

preferences of the patient as to what foods he/she would like to eat in what order, etc. Thus, the robot is competent in the action of feeding in the logistical or mechanical sense but with minimal reciprocity and virtually no attentiveness to the preferences of the patient. Alternatively, in manual mode, the care-giver and/or care-receiver has control over which food they would like to select from the choices on the plate. Thus, the robot bears no role in being attentive to the patient's preferences or for tinkering to meet the needs of a particular patient but rather this role remains in the hands of the nurse (when used in the hospital) or the care-receiver (when used in the home setting).

Both the MySpoon and Bestic robot are considered competent when the practice of feeding is understood as the mechanics of eating – the speed and angle at which the spoon is raised to one's mouth and the speed and angle at which the spoon is lowered back to the plate. This is the manner in which Engelberger has described the design of assistive robots – understand the practice and design accordingly. Of course this is necessary in order for the robot to be of any use (thus, in terms of use and/or user-centered design approaches); however, we are now acutely aware that this mechanical description of a practice, while invaluable for the programming of the robot, misses many other dimensions that are integral in a care practice. In the description of the practice of feeding we saw how important this moment was in terms of a patient's rehabilitation, quality of life and successful recovery. The moment of eating is also a valuable moment for social interaction, for attentiveness on the part of the nurse and for cultivating the relationship between nurse and patient.

The analysis thus far is restricted to the use of the robot in an institutional setting like the nursing home. There are certain wards in a hospital that may have use for the robot, those that have patients eating in a common room like psychiatric wards; however, for the most part patients in the hospital are eating in their rooms. The feeding robots described here are mainly intended for use in a home setting . One of the goals made explicit from the creators of MySpoon and Bestic is to allow persons with disabilities to re-join their families for meals rather than having a family member feed them. Again, the context within which the robot will be applied plays a role in the ethical acceptability of the robot. This is, in part, due to the fact that the care-receiver and care-giver will already have established a trusting relationship (in the case of the care-giver being a member of the family.

Furthermore, the moral elements are interpreted somewhat differently in the home setting. As we saw with the diet assist robot, there is a stronger sense of autonomy, on the part of the patient, when care takes place in the home. Thus, attentiveness and competence have to do with how well the care-giver can per-

ceive the patient's state with respect to their sense of autonomy. Values like dignity then take on a new meaning. Whereas in the hospital dignity is accomplished through an understanding of the individual patient as a person, dignity in the home setting is often achieved through cultivating skills and empowering the independence of the patient/care-receiver. This is of course also true in an institutional setting [*Pols*, 2004] but in this setting patients already have a different vision of what their own autonomy is; they are already dependent on the healthcare staff for their own daily activities and social interaction.

7.4.2 The meaning of the eating assist robot

In an institutional context like the nursing home, the practice of eating serves multiple ends, which add a social dimension to the practice and which give it meaning. Using the robot in automatic mode to fulfil the role of feeder removes those elements. If, however, the robot is in manual mode it is possible to suggest that certain elements like cosiness will remain in tact. The dimension of cosiness is of particular interest here. This dimension, or value, is claimed by Mol to be of great importance in feeding, so much so that it shifts the perception of eating from merely an action to a social event. To view eating in terms of cosiness one need only think of the variety of staff involved: the cooking staff who may or may not be nearby and visible to patients, the nurses dispensing medications, the nurses assisting in feeding, the other workers and volunteers assisting in serving. If a robot were there to fulfil the role of eating assistant will there be the same kinds of interactions in the dining room? Meaning, the serving staff who deliver a plate to the table of a patient may also engage in small talk with patients. Will this no longer be the case if the staff are instructed to bring the plate to, and place the plate in front of the robot without any need to engage with patients? Without room for tinkering in terms of cosiness one must wonder about the assumptions on the part of engineers that either cosiness is not a value in the practice of eating or alternatively, that efficiency of eating trumps that of cosiness. What's more, what kind of statement is being made about the value of care-receivers in this context? That the practice of eating assistance is too mundane or burdensome to allocate their time towards? Added to this, one might suggest that a script is embedded in the technology along with an assumption pertaining to the significance of the practice for building on the therapeutic relationship between care-giver and care-receiver.

Switching contexts from the nursing home to a home setting, it is still the case that the practice of eating serves multiple ends beyond just the consumption of food; however, from a care ethics stance the robot appears quite differently.

Instead of negatively impacting cosiness by minimizing social interactions (or human-human interactions), in a home setting the robot may in fact add to a sense of cosiness by allowing for the care-receiver to actively participate in mealtimes. The robot thus appears to provide a more dignified sense of care.

Special attention to thee actors in a home setting must be made here. One must consider what happens when a care worker is normally present for practices like eating assistance and/or lifting. If this care worker is replaced with a robot for eating assistance alone, what impact might this have on the therapeutic relationship but also on the care-giver's knowledge of the patient's overall care? Meaning, how can the care worker assess whether a bowl movement is indicative of what the care-receiver ate (or did not eat) when they weren't present for eating. Should the robot keep a record of such information for the care worker to access when they return? Should the robot only be used with eating assistance when the care-givers are family members? Given that family members are not held to the same responsibility or criteria as an institutional care worker, one might suggest that this be a preference decided upon by the family.

7.5 The Practice of Food Tray Removal

The final moment one might consider in the process of feeding is what happens when the patient has finished eating their meal. In the nursing home setting when patients are seated in a common room, eating together with other patients and nurses, food plates are removed by the nurses and brought to the kitchen. The nurse assisting in eating is privy to what the patient ate, how much they ate, how eating occurred (with great difficulty or with ease) along with other variables like the patients neurological functioning, the patient taking their meds, etc. The patient then returns to their room or to a common room for activities.

In the hospital context, patients do not often eat together in a common room but rather the practice of eating happens in their hospital room (or the cafeteria if they are mobile and cleared by a physician to do so). The food is prepared by staff other than the nurse, but is brought to the hospital room by the nurse or other hospital support staff (much like a porter). The food tray is then placed at the patient's bedside table. If there are other patients in the same room, the nurse may enclose the curtain around the patient. This is based on the preferences of all patients in the same room. If the patient requires assistance with eating a nurse or family member will be present and will assist with this. If, however, the patient does not require assistance, their food tray

will be delivered and removed once they have finished their meal (or what they can eat of their meal).

The practice of food tray removal affords the nurse an additional opportunity for cultivating attentiveness towards the patient as a unique person with unique preferences based on their personality and physiological state. Attentiveness here refers to a general knowledge or assessment of what the patient has eaten in order to pass this information along to either, or both, the dietician and physician. The nurse must be sure that the patient had the correct tray and will take note of what the patient ate on that day. Nurses may come to know certain preferences and moods of patients based on what they have or have not eaten. For example, a patient who normally eats their jello but doesn't on a given day may be indicative of the patient's physiological functioning or their mood. Such information, if the nurse is attentive to it, may be acquired through observing the contents on a patient's tray especially in the case when a nurse is not present to observe the patient during eating. Competence thus refers to the nurse (or alternate care-giver) picking up on verbal and non-verbal cues of the patient as well as being informed as to the dietary requirements of a particular patient and ensuring that the patient receive the food that meets their diet plan (here diet plan does not refer to a weight loss regime but to the dietary requirements of a patient based on their health and any surgical treatment or other medical intervention they are undergoing or have undergone). Responsibility for this role refers to the nurse being accountable to observe the eating habits of the patient and to report or question the patient about abnormalities in their routine. If the nurse is not observant to changes in the patient, or to the patient not complying with their diet plan, they may be accountable to the physician or dietician. Reciprocity is observed in both verbal and non-verbal cues. The nurse and patient may communicate about preferences or the nurse may be observant to the behaviour of a patient.

When we consider these elements in terms of the larger picture – the processes of the institution – we can see how important it is for this practice to happen in a timely and accurate manner in order for the hospital to function. The efficiency of the hospital ward depends on all the small practices that make up the process of care. Thus, the delivery of simple things like sheets, food, meds or x-rays takes on a richer meaning when understood in terms of the institutional context of care. What's more is *who* is delegated the responsibility of food tray remover: it is not just a visitor in the hospital or a family member visiting a patient, it is a healthcare professional, one who is obliged to fulfil their role according to the values that encompass the institution.

But perhaps more important than recognizing food tray removal as part of

the efficiency of the hospital is the relationship between this practice and that of the overall care of the patient. The nurse is engaged in a practice which is dependent on the practice of diet assessment and eating and results in both a deeper understanding of the patient as a unique individual with specific needs and preferences. This practice also allows for additional moments of social interaction and helps to maintain the therapeutic relationship between nurse and patient. We see then that this moment of food tray removal takes on a deeper meaning when linked with the overall process of care – it is a moment in the larger process of feeding that completes a chain of responsibility for the nutritional well-being of the patient but which also helps to build a deeper understanding of the patient as well as maintain the relationship between nurse and patient. Of course, I must acknowledge that this is not how food tray removal will occur in all institutions. In fact when understood in this way, according to the manifestation of moral elements, this may be considered the ideal performance of the practice. In some institutions there will not be time for the nurse to have a conversation with the patient about their food, there may not be a nurse available to remove the food try and a porter will be sent instead. I claim here that in my description of the practice, my aim is not to build normative criteria strictly from a vision of how the current practice is performed but to combine current practice methods with the ideal. Thus, although in many institutions there will not be time for the nurse to engage with the patient at this time, I am making the claim that in order to enhance the attentiveness of the nurse, to facilitate the nutritional guidance of the dietician and to maintain the relationship between the nurse and patient, the nurse ought to be present at the moment of food tray removal and to assess the eating habits of the patient. In the hospital, there is not always time for this but the question then is whether or not a robot could make this ideal a reality.

7.5.1 A food tray removal robot

Currently, in healthcare institutions across the world, the HelpMate robot created by HelpMate Robotics Inc. (formerly known as Transition Research Corporation founded by Joseph Engelberger) is used as a delivery assistant [*Thrun et al.*, 2000]. The HelpMate robot is a fully autonomous mobile robot that can navigate a building's corridors and elevators on its own and tell humans it has arrived with a delivery of meals, X-ray photos, medical records, linens, diagnostic samples, and other loads [*Thrun et al.*, 2000]. A similar robot called TUG, developed by Pittsburgh-based Aethon, is also used for delivery assistance in over 100 hundred hospitals across the US. TUG also navigates through

a hospital's corridors and elevators on its own (using "laser whiskers" to avoid obstacles and humans), wirelessly opens doors and elevators, announces its arrival and delivers medications, medical supplies, blood, meals and/or linens and removes food trays from a patient's room to bring them to the kitchen [*Mutlu et al.*, 2008]. TUG has different applications depending on what it is delivering; the pharmaceutical TUG application is locked while the linen application is unlocked (Mutlu, personal communication). The Casero robot developed in Germany by the Fraunhofer Institute[7] (in conjunction with the University of Duissberg-Essen), is similar in that it is used to deliver items on a ward; however it cannot navigate an elevator and is thus restricted to one floor only. The first two are commercially available robots while the last one, Casero, is still in its prototyping phase. These robots are considered aids to care-givers and thus I label them as assistive robots.

Interestingly, these kinds of robots are integrated into a variety of different care practices depending on what the robot is delivering and/or receiving or taking away. When the robot takes linens from one ward to the laundry facilities, the robot is involved in the practice of cleaning. When the robot delivers medications from the pharmacist to the nurse, the robot is involved in the practice of medicine giving. When the robot removes food trays from the patient's room it is involved in the practice of feeding. When the robot delivers a blood sample from the lab to a doctor it is involved in the practice of diagnosis. Once the robot has been integrated there are subtle differences concerning how the practice is impacted. In the first instance, when the robot removes linens, the shift comes in the additional responsibility delegated to the nurse. When the robot delivers medication, the robot bears a great deal of responsibility for ensuring the timely and appropriate arrival of medications on a ward. When the robot removes food trays it prevents a certain level of attentiveness of the part of the nurse responsible for gathering information pertaining to the patients' eating habits. And when the robot delivers substances from the surgeon to the lab for analysis, it prevents reciprocity between the lab and the surgeon by the fact that there is no communication from the original source when delivery happens. Although each use of the robot differs in terms of the impact on persons and/or the practice, what may be observed through all of these examples is the robot's propensity to detach certain aspects of a practice from the overall picture of care. It is for this reason that we are reminded to evaluate the robot not on its capabilities alone but on its role within a practice. The TUG robot then ought to be evaluated separately for each practice within which it plays a role. For

[7] www.robots.nu/casero-for-robot-service/

the purposes of this work, I look at TUG's role in the practice of feeding when used for food tray removal.

Significant research has been conducted regarding the TUG robot's impact on work flow (in terms of the impact on work distribution); the differing expectations, perceptions, and acceptance of the robot based on one's role in the hospital [*Mutlu et al.*, 2008], and anthropomorphism of the robotic interface [*Zhang et al.*, 2008]. These ethnographic studies concluded with design recommendations taken from observations in the hospital context. For Mutlu, the goal was to understand how organizational factors influenced the design of a technology *for* organizations. Thus, their normative criteria for design comes from a description of behaviour. For the ethicist, that misses a crucial point – namely what principles ought to guide the evaluation of a technology. Although a large focus of my work involves the robot's impact at the institutional level, what the listed research fails to address is the impact this robot will have on the promotion of care values integral to the care context. It is exactly this point that my work aims to attend to.

TUG is custom programmed to fit the work environment within which it is placed:

> *The TUG robot allows for the CAD drawings of the hospital to be converted into a map for the robot. If these drawings don't exist for an area, the installation engineer from Aethon will walk the robot through the space. Once TUG knows its new home, it maps its position from a starting point (a recharge station). An array of laser range finders (the 'whiskers'), ultrasound sensors, and IR monitors help it navigate the hospital and avoid moving obstacles like busy surgeons. WiFi allows the robot to communicate with other bots (for optimal delivery and performance) to receive requests for action, and to open doors and call elevators* [*Saenz*, 2010].

In terms of the human–robot interface:

> *hospital staff can program the TUG for scheduled routes through a touch-screen. WiFi connectivity allows for delivery requests or pick-ups, which can also be sent through VOIP or pagers. When TUG arrives at a destination for pick-up or drop off it makes a verbal request. Two large buttons ('pause' and 'go') allows users to signal when the robot should continue on its route. Depending on the hospital and tasks given to a TUG, the cart may come equipped with a secure door that requires a keycode to open, and RFIDs may track*

> *inventory. Automatic doors and elevators can be equipped with wireless controls for the TUG to operate. Generally TUG will recharge in between each delivery to keep itself ready* [*Saenz*, 2010].

Now let us envision the moment of food tray removal once a care robot has been added. The robot is assistive in that it provides a service to the care-giver (and indirectly the care-receiver) that the care-giver could have otherwise fulfilled but it is also considered a replacement robot as it replaces the human care-giver delegated the role of removing the food tray. The context is the hospital and the actors involved now include the robot in addition to the patient. The robot does not enter the room of the patient but rather the food tray is placed on the robotic platform waiting outside the hospital room. Anyone can place the food tray on the robot. Attentiveness refers to the robot's capabilities for taking and understanding instructions. In the case of routine deliveries and pick-ups, food trays are loaded onto the TUG carts and the robot is programmed to know its final destination. Or, the robot is programmed on the spot in terms of irregular deliveries and pick-ups. Thus, the robot's attentiveness for destination is determined according to its programming. Attentiveness in this respect is a shared endeavour between the robot and human. Attentiveness therefore differs when discussing the capabilities of the robot compared with the humans. While the human's attentive role was in terms of observing verbal and non-verbal cues of the patient, attentiveness for the robot refers to its capability to find its destination and retrieve the food tray.

When speaking of responsibility, the sender and receiver must trust that the materials being sent will arrive at their final destination and furthermore, that they will be received and/or delivered in a timely manner. This presupposes that the robot knows where the final destination is and how to get there. The infrastructure built into and around the robot (the environment is mapped out for the robot ahead of time) is pre-programmed, thus, the 'Aethon installation technician' programming the robot bears a certain amount of the responsibility for this. Of course, the type of materials being delivered changes the seriousness of responsibility. A trusted role and responsibility has been delegated to the robot in this practice given that the robot is assumed to keep the information attained via the food try confidential. In terms of competence, again competence is linked with attentiveness. How competent the robot is, is dependent largely on how attentive the robot is – the deliverer must be attentive to the other people in hallways, corridors and elevators and proceed according to this. The technical mechanisms to ensure this are the robot's pre-programmed map, the robot's connection to WiFi, and the robot's sensing "whiskers". With the use

of the robot what is taken for granted, is *how* the deliverer travels through the hallways and corridors and greets the sender or receiver – the social etiquette implicit in and expected of the deliverer. The robot must have the same social etiquette. What's more, the same robot must be attentive to the differences between one ward and another. For example, with the TUG robot (as seen with other pre-preference robots in human-robot interaction studies), the volume of the robot was fine in the post-natal ward but was inappropriate in the oncology ward rendering the robot "rude". Context in terms of the hospital ward has a huge impact on the interpretation of attentiveness and competence – the robot must be aware that its volume must change depending on the ward it is in.

For reciprocity, the robot is given basic verbal capabilities such that it may indicate when it has arrived with a delivery and when it is leaving. This information provides the sender and receiver with cues about the robot's state of functioning but also encourages a reciprocal interaction in that the sender or receiver must press buttons accordingly. Thus, the robot facilitates a certain type and amount of reciprocity between robot and human but does nothing to facilitate reciprocity among other healthcare personnel or between the patient and care-giver.

When we analyse the moral elements, we can see that the robot is capable of meeting the requirements for attentiveness, responsibility, competence and reciprocity when seen in terms of the mechanics of the practice. What the robot fails to incorporate is the link with the moral elements and the overall practice of care. The robot has the propensity to detach the moment of food tray removal from the overall practice of feeding. The robot fulfills the task of food tray removal without recognizing that this 'task' fulfills additional purposes linked with the bond between nurse and patient as well as the nurse's learning about the patient. In the design of the robot, it is possible to suggest that designers envisioned the robot as relieving a burden of nurses to free up time for other activities. What the designers did not take into consideration was the significance of this practice in the overall practice of feeding and essentially the overall care of the patient.

7.5.2 The meaning of the food tray removal robot

The main assumption of interest in the discussion of TUG has to do with the view that the practice of food tray removal is a separate, unlinked practice in the overall care of a patient. In other words, an assumption about the detached nature of care practices. Evaluating the design and use of TUG makes clear the relationship between such an assumption on the part of a designer and

how it structures the design of the technology and the eventual use. Such an assumption may be categorized as one pertaining to the kinds of care work done; what counts as care work and what doesn't count as care work. Also of significant interest is the assumption embedded in this robot about who actually does care work. Take for example the use of TUG in the removal of linens. While originally linens were removed from a patient's room by support staff, with the use of TUG nurses are now obliged to remove the linens and load them onto the robotic platform. The robot then travels to the cleaning area where support staff unload the robot and wash the linens. Interestingly, there is a displacement or hiding of care workers in this example. Support staff normally present throughout the hospital are now restricted, or confined, to a specific area. They are hidden from view and as such their role in the provision of care is made invisible. Such may be an assumption about the value of their work but also about their role as support staff and accordingly a valuation of their role in the care process of the institution. The same kinds of assumptions may hold when the robot is used for food tray removal. When the robot is takes on this role, what other support staff are being displaced or made invisible?

Tronto discusses such a line of thinking in terms of the marginalization of care workers: "care work is often distributed along lines of class. One of the main ways in which societies are able to distinguish among castes is by the kind of caring work they do. It is difficult to determine whether care work is poorly compensated because its citizens tend to be the less privileged in society or whether, given the relative unattractive nature of care positions, people who face discrimination elsewhere in the workforce become care workers. In either case, regardless of cause, the fact that care is still disproportionately the work of the less well-off and more marginal groups in society reflects care's secondary status in society" [*Tronto*, 2010, p. 166]. From this, one can conclude that the care activities of cleaning the institution already make a statement about the individuals doing such work. Added to this, when such work and workers are made invisible, a further reduction of the work and the work they do is made. The question then is how to maintain the status of such workers and their role in the institution. Certainly, the use of a robot that makes them invisible is counter-intuitive to the goal. Consequently, I would suggest that the care workers in roles which are threatened to be made invisible become the carers of the robots. They are the ones trained to be proficient users of the robots. This does not mean they must be trained as technicians and are responsible for fixing the robots if a problem occurs. Rather, they are responsible for checking in on the robots throughout the hospital to ensure the robot is used in the intended manner and that the robot is doing what it is intended to do. Their role as

support staff is reinforced and they are actively engaged in the addition of the robot as an extension of their previous roles.

7.6 Conclusion

While the insights provided by Engelberger provide a useful starting point in working out the mechanics of a task, and thus programming the robot accordingly, they fail to indicate the impact the robot has on a broader scope. Not only this but listing the mechanics of a practice proves to be more problematic then one might first think given how difficult it proves to articulate when exactly a practice begins and when it ends. Only when one attempts to grasp the ways in which practices are interconnected can one begin to design the robot in a way that both recognizes and implements such a finding. In the practice, or process, of feeding, there are multiple moments during which time the component of food is integral to the care of the patient. One might be the moment during which the patient/person consumes food but this need not be the only moment under the umbrella of feeding. What the patient eats, monitoring of the patient's caloric intake, and attention to patient preferences are all considerations related to the moment of consumption but are also are considerations relevant to the overall care of the patient. When designing a robot for any one of these moments it is important to recognize the impact it will bear on another moment in the process.

Based on the evaluations above, I recommend that the diet assist robot be used *following* an in-person interview with a human dietician. Moreover, that the diet assist robot is programmed by the dietician based on their assessment and that the dietician monitors the progress of the dieter. For the eating assist robot, I suggest that the robot be used in a case in which the patient seeks increased autonomy and not in a situation in which the patient is in need of both feeding as well as monitoring (as seen in the care of elderly persons). In the case of the food try removal robot, I recommend that the robot be programmed in a manner that demands that the nurse be responsible for placing the food tray on the mobile robotic platform and to reaffirm their role as nurse. Thus, the responsibility for observing the habits and proclivities of the patient remain within the realm, and responsibility, of the nurse/human care-giver. These recommendations add strength to the idea that incorporating the ethical reflections according to the CC framework further upstream in the design process would allow the robot to be tailored according to distinct needs of care-givers working within the value-laden milieu of the hospital and/or nursing home.

In addition to providing recommendations pertaining to the future improved design of these care robots or to their implementation, an added goal of this chapter was to train the reader (and the ethicist and/or engineer wishing to use the CC framework) to observe care practices and the addition of the robot from the lens of the care perspective. Both chapters 6 and 7 provided a forum for such analytical training: knowing how to picture a care practice, what to look for in terms of the moral elements and the interactions between actors, how much emphasis to place on the therapeutic relationship, to view the threat of the robot not in terms of rights of individuals but of (re-)distributions of responsibilities, observing the relationship between one practice and another, observing the relationship a practice has on the overall provision of care, observing the care practice in the larger context of the care institution, knowing how to ask questions and knowing which questions to ask. We have seen how a care robot included into a care practice has the tendency to detach one practice from its relationship to other practices as well as from the overall process of care. Such was the conclusion with the autonomous robot for lifting in the hospital or nursing home, the diet assist robot and the feeding robot. It is for this reason that I adamantly argue in favour of the necessity to begin the evaluation of the care robot with an understanding of the practice. It is for this reason that I further insist on beginning the design of a care robot with an in-depth understanding of the care practice which the robot will be included within. Consequently, it is for this reason that I insist on the use of the CC framework further upstream in the design and development of care robots. This will be the work of chapter 9; however, before I continue with this I digress for a moment in chapter 8 in order to discuss in further detail the issue of delegating certain roles and responsibilities to robots and what this presumes in terms of the moral status of the robot.

Chapter 8

Designing Moral Factors With Care

In order to do justice to the profound role of technology in society and in people's everyday lives, technologies need to be approached as morally relevant entities rather than as mere instruments in the hands of moral human beings.

[*Verbeek*, 2011, p. 118]

8.1 Introduction

Up to this point in the book, we have a clear picture of the relationship between a care robot's capabilities and the resulting manifestation of values as has been made clear through evaluations of five current care robots (chapters 6 and 7). We have also seen the relationship between a care robot used in one practice and its impact on any number of related practices (chapter 7). This can take us quite far in terms of evaluating current care robots; however, an additional insight is necessary in order to begin the prospective methodology of a care robot. Namely, a deeper understanding of the relationship between the design of the care robot (i.e., its capabilities) and the robot's moral status.

The moral status of the care robot is something I have implicitly made reference to throughout the previous chapters but would now like to explicitly address and discuss because of its significance in the prospective design of such robots, the subject of the proceeding chapter. 'The moral status of a robot' has

to do with the kinds of roles and responsibilities delegated to the robot. The delegation thereof presupposes a view of the robot's status as a moral agent, a moral actor or a moral factor. When the robot is delegated certain roles in which it is required to make decisions according to the programming of an ethical theory it is presumed to be a candidate for moral agency. The robot is delegated moral responsibility such that it ought to be responsible for the outcome of its decisions. But can technologies, specifically robots, be responsible for their actions? According to the distinctions made by Floridi and Sanders, they can be accountable but not responsible [*Floridi and Sanders*, 2004]. Accordingly, we may be able to identify the robot as having created the resulting problem but we can neither blame the robot nor make it liable. Hence, the robot cannot be fully responsible in the sense that a human is responsible/liable for an outcome. For Floridi and Sanders, this does not exclude the robot from being a moral agent but rather shifts or broadens the definition of what it means to qualify someone or something as a moral agent. Consequently, material objects may be considered moral agents. Other scholars disagree with this characterization of a moral agent provided by Floridi and Sanders [*Soraker*, Forthcoming; *Torrance*, 2008]. According to traditional conceptions of a moral agent, the criteria of agency is specified to include capacities for empathic reasoning, sentience, experiences, consciousness, etc. According to these descriptions of a moral agent, a robot does not fit the bill.

The issue of responsibility is of the utmost importance in healthcare contexts and in the therapeutic relationship. A human care-giver *must* be able to take responsibility for the outcome of actions. The professionalization of medicine and nursing is grounded on this fact. The question then is whether we can, and should, delegate roles and responsibilities attributed to moral agents to care robots. This question is especially important in view of the reality that the care robot cannot be held responsible for a behaviour, or outcome, resulting from its action(s). Many current robotics initiatives discuss robot ethics along Asaro's first dimension; the ethics of the robot itself [*Asaro*, 2009; *Anderson and Anderson*, 2007, 2010b; *Wallach et al.*, 2008, 2010; *Wallach and Allen*, 2010; *Wallach*, 2010; *Moor*, 2006]. This refers to the robot's capability for moral decision-making and reasoning. Roboticists and robot scholars have made the claim that based on the role of the care robot and the context within which it is placed, it ought to be endowed with moral reasoning capabilities. In this chapter I aim to challenge this claim along two dimensions: first, the temporality of the argument and second, the presumption of moral status associated with the claim. The following chapter discusses what a moral agent is both independently of the robot and in terms of a care robot. In order to illustrate

the relationship between the care robot's capabilities and the role and responsibility delegated to the robot, I use the medication reminder robot developed by Susan and Michael Anderson. I then discuss social robots as another example of robots with sophisticated capabilities that endow them with a distinct role and responsibility. As such, the moral status of a care robot becomes a central issue of concern for prospective design and is necessary in order to proceed with the Care-Centered Value-Sensitive Design (CCVSD) Approach.

8.2 Moral Agency

To begin the discussion of moral agency, Author Steven Torrance presents an interesting overview of the organic view of moral status. From the organic view of moral status, only a genuine organism (human or non-human animal) may be considered a candidate for intrinsic moral status. This has to do with the belief that moral thinking, feeling and action arise organically out of the biological history of the human species [*Torrance*, 2008, p. 507]. From this, of course robots cannot be considered to have full moral status. But is it the case that robots will never be granted moral consideration? Will there ever be a time when robots will be granted moral consideration in so much as the interests of keeping the robot functioning will be taken into consideration against the interests of other human actors in a network? Moral consideration refers to an entity deserving of a certain kind of moral appreciation without being capable or required to act in a certain way (infants and animals would fall under this categorization). Torrance questions whether or not moral consideration is the necessary and sufficient criterion for moral agency. If a robot deserves moral consideration can we then conclude it to be a moral agent. Further, what is the criterion or criteria to conclude that a robot is deserving of moral consideration? Is it sentience or some kind of potential as in the case with animals and infants respectively? To tackle this issue Torrance speaks of a distinction in terms of moral producers, those that are a source of moral reflection, and moral consumers, those that deserve moral attention. In the latter category fall infants and non-human animals – groups that we recognize to have a certain degree of sentience but who are not capable of moral reasoning at a high level. Those that fall within the former category are full moral agents capable of sophisticated moral reasoning along the lines of both intellectual and empathic reasoning. Intellectual reasoning has to do with a weighing of the pros and cons of all the details of the situation – a consequentialist approach if you will. Empathic reasoning, according to the organic model of moral agency, refers to the capability and the inevitability to

incorporate a kind of "affective or empathic identification with the experiential states of others, where such empathic identification is integrally available to the agent as a component in its moral deliberations" [*Torrance*, 2008, p. 510].

Robots cannot, at this time, exhibit this kind of empathic rationality which has been shown to be both a value in care and a pre-requisite for the good/ethical care-giver. Of course, roboticists and computer scientists may counter this claim and suggest that at a time in the future it may be possible to outline the process of thinking in order to achieve a kind of empathic reasoning and to program a care robot accordingly. However, this raises the question of why we would bestow such capabilities, and the associated roles and responsibilities that are attributed with such capabilities, in a care robot. To conclude, Torrance suggests that given the organic view of moral agency, natural humans are full moral beings and artificial humanoids may be deemed as "courtesy moral beings" [*Torrance*, 2008, p. 520]. This presents an interesting caveat: robots cannot be considered full moral agents but perhaps there is a sense of moral status that the robot will have once placed in context. Thus, we are left with the idea that care robots may be granted a kind of moral character or status based on the context within which they are placed, but what this entails is not quite clear given that robots do not have the same capabilities as full moral agents (i.e., human beings).

In contrast to the organic view of moral agency, the *standard conception* of a moral agent refers to: "beings who are capable of acting morally and are expected by others to do so." Thus, "moral agents are beings that are 1. capable of reasoning, judging and acting with reference to right and wrong; 2. expected to adhere to standards of morality for their actions; and 3. morally responsible for their actions and accountable for their consequences" [*Brey*, 2012, p. 1]. Here, there is no indication as to the physical make-up of the agent but rather solely to the capabilities, expectations and associated responsibilities of a moral agent.

> *An agent is a moral agent when the intentional states that it cultivates and the subsequent actions it performs are guided by moral considerations. This requires a capacity for moral deliberation, which is reasoning in order to determine what the right thing to do is in a given situation. A capacity for moral deliberation requires a capacity for reasoning and knowledge of right and wrong. Moral deliberation typically results in moral judgements, which are judgements about right and wrong. It also frequently results in intentions to perform certain actions that are held to be moral, and to refrain from per-*

forming actions that are held to be immoral. [*Brey*, 2012, p. 2].

Based on the above definition, one may be inclined to ask what is meant when one considers a moral agent to be "guided by moral considerations'? Furthermore, what could this mean for a robot? Is it a reference to the designers' intentions or the reasoning capabilities of the robot? According to this view, it is once again problematic to include technologies let alone robots within the category of moral agents. This is because moral agency rests on the moral reasoning and thinking capabilities that arise from a human's organic nature, as in the organic view, or a human's expectation of right behaviour and punishment of wrong behaviour, as in the standard view. Thus, what happens when we apply the concept of moral agent to technological artefacts and/or care robots? Currently, there are two schools of thought; the *moral artefacts* view which states that all technological artefacts are, or could, function as moral agents and alternatively the *morally intelligent agents* view which states that certain highly evolved technological artefacts, namely those capable of autonomous behaviour and intelligent information processing, qualify as moral agents [*Brey*, 2012, p. 4]. According to the moral artefacts view, Latour argues that "humans and artefacts are programs of action that aim to enforce particular moral or social rules or configurations" (e.g., a police officer and a speed bump enforcing the rules of the road) [*Brey*, 2012, p. 4]. Both humans and artefacts are referred to as actors or actants and are called as such when the program of action inscribed in their technical content enforces a moral rule – meaning they are capable of steering moral behaviour in humans and/or influencing moral outcomes. In other words, artefacts are capable of altering practices through their presence; through their presence they impact the actions, reasoning and behaviours of the human actors in the network [*Verbeek*, 2008]. Essentially, moral agency is bestowed on a thing/object based on its interactions in context, it is something that happens (i.e., it is dynamic) rather than something that is (i.e., is static). Thus, a robot is a morally relevant entity meaning it has a moral impact; however, this impact is not intrinsic nor does it come from the robot's decision making capabilities (as in the organic view or the standard conception of the moral agent) but from its impact on the decision making capabilities, norms, customs, prioritization of values, shifts in roles and responsibilities, etc. of the human actors in the same network as the robot.

In line with the organic view, for Verbeek, moral agency rests on two fundamental components, that of intentionality and that of freedom. Both intentionality and freedom are explained as hybrid affairs between humans and their material environments (including the non-human agents/actants that comprise

that environment). Given that technologies, and robots, do not have intentions like humans, in that they cannot deliberately do something, they cannot be moral agents. Robots, however, do have a kind of intentionality in that they direct one's course. Thus, intentionality is a hybrid affair between a human and a robot. The same holds for the component of freedom – robots do not have freedom in the sense that humans do. While a robot may have a choice between one or more actions/outputs, the robot is not aware of the repercussions of one choice over another. Freedom in decision making for a human "is always bound to the specific situation in which decisions are made and the material infrastructure plays a key role here" [*Verbeek*, 2008, p. 98]. Thus, freedom is also a hybrid affair between a human actor and the non-human actors in the network.

Contrasting the *standard conception* of a moral agent (insofar as the moral agent is endowed with sophisticated reasoning and decision making capabilities) with the *artefacts view* of a moral actor (insofar as the moral actor is granted such a status based on its impact on the practice and how it changes), how do we conceive of robots in healthcare contexts, as moral agents or as moral actors? Let me take as an example the daVinci®surgical robot – it is an actor in the practice of Minimally Invasive Surgeries (MIS) in so far as it has changed the way the practice is fulfilled: the distribution of roles and responsibilities, the acquisition of new skills, the interpretation of a new skill set and the interpretation of a good/skilled surgeon, surgery and outcome. But can this robot be considered an agent in this practice? Its autonomous capabilities for scaling the movements of the surgeon are not enough to qualify it as an agent, rather, it is an actor in that it plays a role in the mediation of the practice of surgery (or a specific kind of surgery). If, however, the robot were capable of deciding whether the movements of the surgeon were accurate and could choose to accept or decline the input of the surgeon, one might claim that this kind of responsibility and decision making capability of the robot brings the robot's status closer to the definition of a moral agent. In this case, the robot's agency is attributed to its capacity for autonomous decision making .

Setting aside the criterion of empathic intelligence (or experience, sentience, consciousness), does the autonomy and sophisticated intelligence of the robot, which in the above example made it capable of overriding the decisions of the human surgeon, render it a moral agent? Such a question leads us to a discussion of *the morally intelligent* view of moral agency. Dominant proponents of this view include Luciano Floridi and Jeff Sanders who claim that artificial intelligence opens new avenues when speaking of moral agents. Specifically, that technologies with highly sophisticated mechanisms for reasoning, capable of interacting with their environment, acting in an autonomous fashion and

adapting to their environment ought to alter the discussion of moral agents. Their goal is to expand the category of moral agents such that it includes such sophisticated technical artefacts, rather than to alter the concept of morality such that artefacts and humans engage in a practice of hybrid morality. Within this conception, Floridi and Sanders aim to disentangle the relationship between moral agency, accountability and moral responsibility. They argue that moral accountability is a necessary but insufficient condition for moral responsibility. According to their view, a moral agent, and ultimately a robot, may be considered a moral agent insofar as it may be considered accountable for its actions (and thus subject to censure); however, it may not be held responsible for its actions given that it lacks the intentions guiding it to make said decisions [*Floridi and Sanders*, 2004].

If we refer back to the daVinci®example, we must ask whether or not we want the robot to be a moral agent if it is delegated these kinds of decisions without being responsible for the outcome. For now, the robot is still currently an actor in that it impacts the carrying-out of the practice but it is not responsible for high level decision-making – that remains in the domain of the human moral agent (the surgeon). But in the traditional form of Actor-Network Theory (ANT) both the human and non-human actors have the same status. Accordingly, the care robot would have the same moral status and associated responsibilities as the human actors. Can this be the case in care practices? What does it mean to be a moral agent in care? For care ethicist Joan Tronto, there are certain necessary and sufficient criteria which render a care-giver a moral agent as we are well aware of at this point (chapters 3 and 5). These criteria involve the aspect that care is only considered good care when the dimensions of caring about (the disposition to care) and caring for (the activity of care) are married. It is possible to claim that a care robot may be involved in the activity to care but what about the disposition to care? Is it possible that if the robot were endowed with capabilities to give the illusion that it had such a disposition then in fact the care robot meets the necessary requirements? This last point is quite interesting when we consider that not all care-givers will in fact have the same *feelings* for all patients but will still portray a caring disposition. Is this the same as the robot conveying or portraying a caring disposition?

What's more, Tronto's four moral elements, which act as the normative criteria for evaluating a care practice include the element of responsibility. Consequently, within the very framework that I am working in, the element of responsibility is placed at the fore. Responsibility refers to an actor taking the praise and or blame (being liable) for the outcome of events or behaviour elicited, based on the actor's action or decision. Within a care institution and the

therapeutic relationship, responsibility is a cornerstone for establishing trust and for ensuring that needs will in fact be met. If we adhere to the conception of a moral agent presented by Floridi and Sanders and the conception of a good care provider presented by Tronto (for which being responsible is necessary) then no robot should be delegated the role or responsibility of being the sole care-giver. This is so because the robot cannot be held responsible for its actions. In the same vein, if we adhere to a conception of moral agency based on the *standard conception* of a moral agent, again a robot cannot meet the criteria for moral agency. Thus, the robot cannot be labelled a moral agent regardless of the conception of moral agency one is working with. This claim has two outcomes. First, that the robot be designed intentionally such that it is not delegated a role for which a full moral agent would be delegated. Second, that the robot should be designed in a way that it enhances the capabilities of the human care-giver in so much as it helps to bring about the manifestation of the moral elements. Thus, the care robot may be designed in a way that it enhances the human care-giver's ability to take responsibility for an action by reminding them of important variables with respect to a particular patient. In the same case, if the care robot had the same capabilities for perceiving information about a patient but did not pass such information on to a human care provider, the responsibility for such information (and what to do with it) remains in the domain of the robot rather than the human. This is not the situation we want.

Aside from whether or not we might call the robot a moral agent, we must still acknowledge that the robot bears an impact, a moral impact, on the actors of the network. This impact is typified by the robot's ability to shift the decision making of the human actors in a network [*Brey*, 2012; *Verbeek*, 2006, 2008]. If we claim that the robot still bears a moral influence on the practices of the human actors in the network but that the robot and human do not maintain the same moral status, and further that the robot is not capable of higher moral deliberation, nor should it be (based on the normative aspects of care), then what do we call the robot? For this I turn to the structural ethics approach of Philip Brey; we call the robot a *moral factor*. Brey nicely bridges the standard conception of a moral agent (which claims that only humans are fully capable of being moral agents given that taking responsibility for moral actions is a necessary condition) with the belief that technological artefacts do in fact have moral influence on the behaviours and actions of human agents (the moral artefacts view). In this conception, the status of a moral factor mirrors that of a 'moral impact agent' according to James Moor [2006]. In short, the approach of structural ethics allows for the recognition that technological artefacts (what

he calls moral factors) bear a moral character but cannot be responsible, that responsibility always falls on the human agents in the same network. For Brey, "moral factors shape or influence moral actions and outcomes" [*Brey*, 2012, p. 10]. "Moral factors can be positive or negative, measured against a moral rule or principle. A positive moral factor is one that contributes positively to a moral principle being upheld, whereas a negative moral factor contributes negatively. In addition, factors can be accidental or intentional. An accidental moral factor is one that happened to contribute towards a moral outcome in a particular arrangement. An intentional moral factor is one that has been intended to contribute to an outcome in a particular way" [*Brey*, 2012, p. 10]. Moral factors can also be outcome-oriented or action-oriented. In short, the difference between the two has to do with whether or not a moral factor influences, positively or negatively, the behaviour of a human agent or a moral outcome (a moral outcome being the event or state-of-affairs) [*Brey*, 2012, p. 11]. Intuitively, one would conclude here that a care robot ought to be programmed according to a conception of it being an intentional positive moral factor.

But what does it mean to program a care robot according to a conception of it being a moral factor? If we look to the work of Wendell Wallach and Colin Allen we may begin a discussion along the lines of operational morality vs. functional morality[*Wallach et al.*, 2008; *Wallach and Allen*, 2010; *Wallach*, 2010]. The morality of a robot is measured according to two dimensions – sensitivity to values (along the x axis) and autonomy of the robot (along the y axis). These two axes are independent from one another. Wallach and Allen expand their idea of morality for robots by referring to the classification of James Moor: ethical impact agents, implicit ethical agents, explicit ethical agents and full ethical agents. An implicit ethical agent refers to a machine whose designers have attempted to decrease the negative ethical impact of the machine in terms of safety and reliability issues. Such a class of machines are what Wallach and Allen refer to as 'operationally moral' (or operational morality); the morality of the designers (values, norms, etc.) is embedded into the design of the system such that through the use of the system certain values are promoted (i.e., Value-Sensitive Design). In contrast, explicit ethical agents are those machines that can reason as part of their internal programming. This grouping of intelligence is what Wallach and Allen refer to as functional morality and there is quite the range of machines that may fall within this broad category. In contrast to operational morality/implicit ethical agents, the 'ethics' for explicit ethical agents (classified as functional morality) comes in as a capability of the machine rather than the exclusive programming of the designers. This then demands the question of when it will be possible to conclude that the machine has acted in

a way that was not the intentional programming of the designer, as the field of machine learning and autonomous systems is not wrestling with [*Wallach et al.*, 2008; *Wallach and Allen*, 2010; *Wallach*, 2010; *Moor*, 1995, 2006].

8.3 To Delegate or not to Delegate?

The concept of functional morality leads me to the crux of the question of a robot's moral status. Authors like James Moor claim that "explicit ethical agents should be the goal of the emerging field of machine ethics" [*Wallach and Allen*, 2010, p. 34]. Additionally authors such as Wallach and Allen have suggested that "artificial moral agents (AMAs) are necessary and inevitable". Such claims rest on the fact that robots will be used in morally charged contexts for tasks which will have moral consequences and thus the robot ought to be endowed with moral reasoning capabilities. An example of such thinking can be seen in the work of Susan and Michael Anderson (a philosopher and computer scientist respectively) seeking to incorporate the bioethics principlist [*Beauchamp and Childress*, 2001] approach to decision-making. The robot (a Nao platform commercially available) uses the bioethics principles of autonomy, non-maleficence and beneficence to arrive at the appropriate action when it interacts with patients to remind them to take their medication. In this case, decisions resulting in moral consequences are delegated to the robot [*Anderson and Anderson*, 2010a]. Consequently, in light of the role delegated to the robot, roboticists feel the need to endow the robot with certain moral deliberation capabilities.

In response, social scientists Batya Friedman and Peter Kahn warn of the danger such initiatives present in terms of humans relinquishing their own moral responsibility [*Wallach and Allen*, 2010, p. 40]. I would like to add to this warning. To do so, I challenge the assumption that robots, and care robots in particular, *ought* to be endowed with moral reasoning and/or decision-making capabilities based on the roles they will be assigned. My challenge has two parts. First, I claim that endowing the robot with such capabilities determines the roles and responsibilities delegated to the robot and not the other way around, and second, it wrongly presupposes a moral status of the robot. For the first point, I agree with the relationship between a robot's role and its capabilities presented in the initiative to create robots capable of moral deliberation; however the temporality of the argument is flawed. The design of the robot will determine its role. It is not the case that the robot will have a certain role already which in turn demands that it have certain moral reasoning capabilities. Thus, the role of the robot must be decided upon prior to an investigation of the capabilities

the robot ought to have in order to fulfil this role. Deciding on the roles of the robot must take into consideration the responsibilities associated with that role. As such, the role can then be adjusted if it appears as though too much responsibility is delegated to the robot. As we are now aware, the robot can be held accountable but not responsible and therefore this determines the kinds of roles the robot may be delegated especially in care contexts.

For the second point – delegating the robot a role which requires moral deliberation capabilities – presumes a specific moral status of the robot. In other words, if we delegate roles to robots for which a human moral agent has traditionally been delegated, the assumption is being made that the robot mirrors such a moral status and therefore may be delegated the same role. If this is not the case, if I am wrong in speculating this to be an assumption, then robots would not be programmed to have the capabilities thought necessary for these roles. Instead, the robot would fulfil a different portion of the role and a human would remain in command of the (moral) deliberative portion. To take an example, let us look to the 'reminder robot' developed by Susan and Michael Anderson. Traditionally a human nurse/care-giver would judge how often to remind a patient and when to inform the physician if the patient continually refused. With the inclusion of the robot, this is now the role and responsibility delegated exclusively to the robot. When understood in broader terms, as the Care Centered (CC) framework insists, the practice of medication reminders is much more complex. For starters, the patient is (hopefully) more inclined to trust the nurse and take their medication as a result of the establishment of the therapeutic relationship in other practices. In terms of attentiveness, the nurse has formed a rapport with the patient and is able to perceive the patient's moods and when to use more force if necessary (by force I do not mean physical force but rather the tone of voice, etc.). Based on these attributes, the nurse is capable of claiming responsibility for the patient taking their medication – the nurse is responsible for acquiring this kind of knowledge which permits (and demands) him/her to be responsible for using it in their daily care activities. The robot is not capable of such attributes and even if it were (which is the expected response by roboticists) it is not capable of accepting responsibility for the outcome of its actions. However, given that the robot has been programmed to fulfil this role, an assumption is being made that the robot can meet the requirements of the practice in the same way a human nurse does.

Without these attributes, how can we delegate responsibility to the robot when it is neither capable of perception along those lines nor is it capable of empathic reasoning – knowing that the patient is in a bad or testy mood because their spouse has just passed away or their son/daughter didn't visit them on

the day they were supposed to. If we recognize this and insist that the robot does not possess the same skills and attributes as the human nurse it therefore cannot be delegated the same responsibility. If the robot were not delegated the same role (with the associated responsibility) what would the practice look like? Instead of the robot making the decision of when to inform the physician as to the patient not taking their medication, the robot would inform the physician *each* time it reminded the patient as well as whether or not the patient took their medication so the physician could respond accordingly. As a result, the physician could then send a command to the robot that it needs to use more force or the physician could visit the patient and insist he/she takes their medication. Thus, the responsibility for medication taking remains in the domain of the human actors. The robot is as such a moral factor in that it impacts the decision-making capabilities of the human actors without making the final decision on its own based on internal programming.

Another way to look at this same scenario is to investigate if something were to go awry. Let us say that a patient refuses to take their blood thinning medication and the robot, aiming to support the patient's autonomy, in so doing fails to warn the physician and in consequence the patient suffers from a heart attack or stroke. We may suggest that the robot is accountable, meaning we could pin-point that the robot should have notified the physician earlier; however, the robot cannot be fired or sued as a repercussion for its action (or lack thereof). In this case, what happens to the profession of medicine and nursing when no one is held responsible for the life and death decisions made on a daily basis? The alternative picture of the robot being programmed according to the vision of a moral factor leaves the responsibility for such an outcome in the hands of the physician and would ultimately encourage the physician to remain engaged in the actions of both the patient and the robot. In this picture, one may then question whether or not the robot is providing a benefit at all. To this I would answer that if the physician were not geographically present in the hospital then the robot could be of assistance. Or, perhaps the robot could be used as an aid to the nurse rather than the physician. In this picture, the robot would provide information to the nurse as to when it reminded the patient and whether or not the patient took their medication. The nurse would then be responsible for deciding when to inform the physician of the patient's non-compliance.

Using the medication reminder robot as an example, it is possible to see the difference between a robot as a moral factor and a robot presumed to be a moral agent. Conceived of as a moral agent, the robot not only engages in the reflexive process but it also fulfills an action based directly on this reasoning

process. Thus, the robot is not a moral factor but is seemingly a moral agent presumed to be capable of taking responsibility for its actions. Hence, the robot is delegated a role to which traditionally a moral agent would be ascribed. With the recommendations I provided in terms of the robot's actions – that it be capable and required to transmit the necessary information to a human care-giver – the robot's status is conceived of as that of a moral factor. The responsibility for action lies in the hands of the human actor(s). What's more, the robot's actions and capabilities are adjusted in order to ensure the robot's status as that of a moral factor. As such, the robot may be considered an intentional positive moral factor. Thus, I return to the first portion of my argument against Moor (and others with the same claim as Moor) and the need to program robots with moral decision making capabilities, namely the temporality of the argument. Through an analysis of the capabilities of the robot, and the role to which the robot has been delegated via such capabilities, using the CC framework, one can decide what role the robot *ought* to be delegated. This is alternative to the current suggestion that given the roles these robots have and will have, they ought to have such capabilities. When we consider the prospective design of future care robots we must be just as critical and systematic in our analyses in deciding the role given to the robot.

8.4 Delegating Roles and Responsibilities to Care Robots

The classification of operational morality resembles that of a moral factor in that the robot as a moral factor will exhibit a moral impact; however, when the robot's design is intended to reflect a specific role and the robot will be acting according to this prescribed role, one might conclude that the internal programming of the robot arrives closer to the definition of functional morality. I must be very careful here though as I am arguing against programming the care robot with any ethical reasoning capabilities of its own. To be specific, the reasoning capabilities of the robot may not be used in place of a human's reasoning capabilities. This does not, however, demand that the robot cannot have certain capabilities for reasoning on its own. The robot may be there as an aid to the human care-giver or as a way of enhancing the manifestation of moral elements. Hence, the reasoning capabilities of the robot reflect this criterion. An example of the kinds of capabilities I am referring to are those investigated at the Technical University of Munich in the CoTeSys (Cognition for Technical

Systems) lab. If we recall, programming robots to be cognitive meant that semantic links were programmed into the robot such that it 'understood what it was doing and why'. If we consider a robot that is reminding a patient to take their medication, such a semantic link seems quite appropriate. The robot knows that if the patient doesn't take their medication at a given time (or with food if required), the patient has the potential to die or suffer from some other infarction. The robot may then be programmed to interpret this (in reinforcement learning terms) as negative feedback. Thus, the limits within which the care robot may be programmed as being either operationally moral or functionally moral are determined by the moral elements in care.

If we consider that the moral elements determine the limits of the capabilities of a care robot, and the moral elements dictate that a care-giver ought to be cognizant of the emotional state of the care-receiver, what impact does this have for the future capabilities of a care robot? Should a care robot be capable of perceiving the emotional state of a care-receiver. Further, should a care robot be capable of portraying emotions itself? Such a discussion brings me back to the question I posed in chapter 4, 'Care Robots and Robot Capabilities', asking whether a care robot, given its place in the therapeutic relationship engaged in care practices in the value-laden milieu of the care institution, demands that it be programmed with social capabilities. The question was posed to deliberate whether or not the care robot needed to have such capabilities given that care-givers were required to have both a care disposition in addition to fulfilling care actions. To reiterate, a care robot with social capabilities is distinguished from a social robot. The first is a robot intended to meet care needs and in so doing has an interface that allows it to communicate in a more human-like manner. The second is a robot intended to form a companionship with the human user as its end. The diet assist robot discussed in chapter 7, is an example of a care robot with social capabilities; its goal is weight loss and the manner in which it motivates the user to stick with their diet plan is through the formation of a bond. The question of interest in a discussion of what it means to delegate certain roles and responsibilities to the care robot demands we ask how far does one go in programming social capabilities into a care robot.

For Sherry Turkle, the question of social robots raises the most intriguing questions about human-human relationships, namely the origin and criteria of their meaningfulness. In Turkle's reflection on the use of social robots aimed at children and elderly users she states that:

> *children's evaluation of aliveness is less about cognition than about an object's seeming potential for mutual affection. If something asks*

> *for your care, you don't want to analyse it but take it 'at interface value'. It becomes alive enough for relationship. And with this, the heightened expectations begin. Now – for adults and children – robots are not seen as machine but as 'creatures' and then, for most people, the quotation marks are dropped. Curiosity gives way to a desire to care, to nurture. From there, we look toward companionship and more. So for example, when sociable robots are given to the elderly, it is with the suggestion that the robots will cure the troubles of their time of life. We go from curiosity to a search for communion. In the company of the robotic, people are alone, yet feel connected: in solitude, new intimacies.* [*Turkle*, 2011, p. 18].

The fear that Turkle points towards here is a de-valuing and/or a re-defining of intimacy and connectedness. Can we consider interactions between robots and humans to be meaningful and if so what is the impact this will have on traditional human-human relationships? From this one must ask what a meaningful interaction looks like. Perhaps we might suggest that a meaningful interaction is labelled as such when one or both of the humans form a bond through the interaction. This of course makes us question what a social relationship or bond is between humans. It is possible to claim that a social bond is only possible between one human and another, hence the word 'social' and requires at the very least reciprocity between both. Given this, what then do we call a bond between a human and an animal/pet? These types of bonds do not require reciprocity in the same way that a human-human interaction does. The interaction may still be considered authentic given that the human interacts in an authentic way and reciprocity remains an element; however, reciprocity differs in that the human and non-human do not interact using language as the medium of communication. Without language, the human is inclined to project meaning onto the non-human animal. This is still a bond, however, as the human is capable of taking something away from this, pleasure and/or pain, learning about oneself or meaning in some sense. For example, animal therapy has shown to have incredibly positive outcomes on individuals suffering from a range of disorders. People exhibit psychological, emotional, physical and neurological benefits. One would be hard pressed to conclude that no such bond was created in the interaction between these humans and the animals they were interacting with. Thus, reciprocity from a non-human is not a precondition for bonding.

If there is nothing morally problematic with such scenarios then how can we claim there to be something morally problematic with a human-robot bond? Is such a bond, because it is between a human and an artificial entity, an ex-

ploitation of the vulnerability of the human condition? Is it a de-valuing of an individual's dignity? Alternatively, could we not say that in many instances where a human has suffered from an unhealthy relationship that a care robot with social capabilities proves to be a more dignified means of interacting? Take for example the case of children who have suffered from abusive parents, a spouse who has suffered from an abusive partner, or, a high school student who is bullied at school. Could we not say that having a robot to interact with socially, to form a bond with and practice 'safe' interactions with, might actually mitigate some of the negative impacts of their previous detrimental interactions? One might suggest that this is where the *real* fear comes in; that the robot could in some way result in a labelling of human-robot interactions as being more dignified (in certain instances with specific users).

As in the case of care robots, we are once again faced with the dilemma of calibrating positive and negative aspects pertaining to care robots with social capabilities. On the one hand, their benefits are observable in service situations, for purposes of enjoyment, learning, therapy and for providing companionship to meet alternative end goals in care. But one must also consider whether such applications will be negative with pervasive use, even when the robots are used in more of a service application. Will human-human interactions be substituted with human-robot interactions in care contexts? Social robots are intended to meet a variety of needs of users in a highly successful manner and it follows that humans may prefer interacting with a robot. Such a trend is already visible with computer systems. What's more, given that the robot would be rational, predictable and controllable, this alone could lead to a preference for interacting with a robot. Along the same lines, will social interactions fundamentally change? If care-receivers and/or care-givers become accustomed to interacting with a robot, will their expectations for human-human interactions change? Meaning, will they expect humans to interact in the same rational, predictable and controllable way as a robot would.

The idea of substitution may be presented as the main fear with regards to care robots with social capabilities – that such robots will replace human care-givers. Such a fear becomes very real when considering that such robots may one day fulfil needs better than humans in certain instances or can change human-human social interactions. But this fear of replacement is not as much about the interactions humans will have with robots as it is about the larger context. Why will people want to interact with robots more so than humans? What is the current care situation and what is it lacking? What is their living situation like and what is it lacking? Would people actually prefer a care robot with such capabilities or is it a better alternative given no other option? The

question then is not whether we should use care robots with social capabilities but how they should be used, for what practices and in what contexts. When we endow care robots with social capabilities we are again making an assumption about the moral status of the robot. This is in line with the role the robot is delegated but is also a reflection of the sophisticated intelligence of the robot that renders it capable of interaction in a human-like manner. This means that the boundaries within which the robot can have social capabilities are again decided according to the manifestation of moral elements along with the role and responsibility of the robot understood.

In summary, the limits within which the care robot may be programmed are determined by the moral elements in care. If the robot, according to a conceptualization of it being functionally moral, nears a role or responsibility to that attributed to a full ethical agent, then the robot's capabilities must change. Alternatively, if the robot does nothing to enhance the care-provider's ability to meet the criteria of the moral elements, then once again the robot's capabilities must change. Thus, deciding on the robot's role and capabilities is a deliberative dynamic process rather than a static one. It follows that the care robot's capabilities are decided according to an understanding of the limits of its morality which are determined by the interpretation and understanding of the manifestation of moral elements based on a specific care practice in a given context. This question will greatly increase in importance in the coming years when more and more care robots will be designed with social capabilities. It is of the utmost importance to balance such social capabilities according to the limits established by the CC framework.

8.5 Conclusion

In this chapter I have argued that regardless of one's position concerning the robot as a moral agent, the robot cannot be responsible for its behaviour or an outcome resulting from its action(s). Thus, although I reject the claim by Floridi and Sanders that a robot can be a moral agent I adhere to their argument that the robot cannot be deemed responsible for a behaviour or outcome. The claim that the robot can be an agent according to the position laid out by Floridi and Sanders insists that the robot be accountable but not responsible. The underlying distinction between the two refers to the intentional state of the robot; that the robot would have to intentionally act in order to be responsible. Of course the robot could be goal-oriented; however, this is quite different from the robot having an intention to act based on beliefs and/or feelings. I argue

that at this time, in the development of care robots, it is not possible to claim that a robot has the in-depth capabilities for moral reasoning according to the organic and standard view of a moral agent. Thus, the robot is not considered a moral agent and as such cannot be held responsible. What's more, I argue that according to the requirements of the care ethics perspective for good care, a care-giver cannot be conceived of as being accountable and not responsible. If the requirements of care-givers were to exclude responsibility, and in its place demand accountability, one would seriously question whether or not they would have any trust in the system. Thus, the care robot must be delegated roles for which it assists and/or enhances the roles of care-givers rather than replacing care-givers.

Above all else, I challenge the claim made by roboticists and other robotics researchers that programming a care robot to be capable of moral deliberation is in fact desirable. Rather, I pay tribute here to the work of Turkle and claim that we ought to pay careful attention to the roles we delegate to the robot and the propensity this has to alter practices for good and/or for bad. Accordingly we ought to pay careful attention and design robots according to a conception of their role and responsibility as being that of a moral factor. The CC framework insists on understanding the distribution of roles and responsibilities within a care practice explicit prior to the introduction of the robot and to also make the re-distribution of roles and responsibilities explicit once the robot has been introduced into a care practice. As such, the role and responsibility of the care robot are made explicit retrospectively for evaluation. Based on the above discussion of how programming a care robot differs depending on one's starting point of the robot's moral status (as moral agent vs. moral factor) and the recognition that the robot cannot be held responsible for a behaviour or outcome, the robot must be programmed according to a conception of its status as a moral factor. Added to this is the recognition of the significance of responsibility in care, it must be possible to ascribe responsibility for a behaviour or outcome. This has far-reaching implications when it comes to the design of the robot. Consequently, the role delegated to the robot can be decided upon in a prospective manner in a dynamic process through which different capabilities are explored and their resulting role and responsibility ascription evaluated. Such a consideration is integrated into the Care-Centered Value-Sensitive Design Approach presented in the following chapter.

Chapter 9

Designing and Implementing Robots With Care

The best forms of institutional care will be those which are highly deliberate and explicit about how to best meet the needs of the people who they serve.
[*Tronto*, 2010, p. 169]

9.1 Introduction

In observing the current development of robotics technologies, Sherry Turkle claims that we are at a crucial moment, a moment in which we can decide to explicitly shape robots in a way that allows us to rethink the values and the general direction that robots are taking us in. Turkle presents her book "Alone Together" [2011] to mark the opportunity we have now to shape robotic technology in a way that protects those values which we hold in regard and want to safeguard. Most significantly that robots offer us the opportunity to rethink our conceptions of what a relationship is and what it means to be in relationship with someone else (or something else). I propose that 'now' marks the opportunity to shape the design and development of care robots in a way that safeguards the values which form the buttress of the care tradition. Moreover, that the entering robot allows us the opportunity to re-evaluate care at the institutional level and to design the robot in a way that reinforces the purpose of such institutions. This opportunity for rethinking and shaping requires a bridging of gaps so to speak, a coming together of disciplines in a way that allows

for an understanding on the part of engineers as to the impact this technology can (potentially) have from the ethicist's perspective, but also to encourage the translation of ethics into a tangible format for engineers to grasp. Like Turkle, I am encouraging a stewardship of values. Beyond Turkle's claim, however, I am presenting a format for accomplishing this feat. Consequently, this chapter pointedly and explicitly addresses my original research question: *how to design and implement care robots used in care practices, in a way that supports and promotes the fundamental values in care.*

To do this, I aim to show how robotic technology can be shaped in a way that safeguards the cornerstone values of the healthcare tradition. Creating technologies that act as a steward of care and care values, is accomplished through rigorous attention to the wider impacts of the technology – in care, the impact the technology has is not on the one care practice which the robot is used for but on the overall process of care that the robot is introduced into (care at an institutional level but also the process of care as it pertains to one patient – their overall care). Above and beyond using the framework in the prospective design of a care robot (shaping the design process according to the framework) I will argue for the use of the framework in the implementation of the care robot as well. Although there exists a wide variety of studies concerning the domestication of technologies (how technologies take on meanings through their pervasive use), there is little work done on the ethical implementation of technologies – the bridging of design studies with domestication studies. My goal is to show that given the normative force of the framework and the assumptions pertaining to how the robot ought to be used, these insights should be translated into both the policies governing the use of the robot as well as how the robot is first introduced into the network. For such a discussion, one that focuses on the widespread impact of the care robot, I refer to and expand on the ethical approach of structural ethics [*Brey*, 2012].

Accordingly, the following chapter is meant as the conclusion of this book bringing together the insights and conclusions of all the previous chapters. Accordingly, the chapter maps out the uses(s) I have in mind for the CCVSD approach; namely integrating ethics into the design process of future care robots and how to ethically implement care robots in context. For the first issue, I will show at what stage in the design process the framework is meant to be used (idea generation and onwards to implementation), how to proceed with the framework (its method for use) and finally how to interpret the reflections from using the framework (recommendations for design and/or implementation). Given that the prospective design of care robots according to the CCVSD approach begins at the stage of idea generation, I re-visit certain fears associated

with the initiative to create care robots and aim to illustrate how these fears can mark the starting point in the design of the care robot. The design process of future care robots ought to account for the fears associated with the use of care robots in order to mitigate such fears but also to highlight the significant values that must be protected. Throughout the methodology I also aim to show that robot's capabilities ought to reflect the role and responsibility attributed to the robot as a moral factor bearing a moral influence, as opposed to the care robot being conceived of as a moral agent (the conclusion of chapter 8). Shifting contexts from the lab to the hospital, I will show how the framework may also be used in the actual implementation of the care robot. To do all this I will present ideas for novel care robots generated from observations and interviews in a hospital and nursing home setting.

9.2 Designing Robots with Care: the Care-Centered Value-Sensitive Design Approach

The 'user manual' for the prospective use of the Care Centered (CC) framework, what I refer to as the Care-Centered Value-Sensitive Design (CCVSD) approach, differs from that used for the retrospective evaluations of current care robots (see Table 9.1 for a reminder of the CC framework). For starters, there is no current robot to evaluate. This means that the recommendations resulting from the analysis are directed towards designers engaged in an improved future design of the same robot or for policy makers creating guidelines delineating the proper usage of the robot. The CCVSD approach does not begin with a robot prototype, thus the initial value-based analysis results in recommendations for designers to begin the making of a care robot such that it embeds care values in the most promising way. With the initial prototype, the robot may then be re-evaluated (in context) again using the framework to arrive at additional recommendations for designers or alternatively recommendations for policy-makers.

Alongside the embedding of values into the design of the robot, prospective design analysis is also about addressing fears related to the use of a robot. Speculating the uses for care robots ultimately reduces the dialogue to the definitive question of which processes can be standardized. What ethics then asks is which processes SHOULD be standardized and according to what criteria. What is gained and what is lost through standardization. For example, if the goal of standardizing is to increase efficiency, this value is beneficial in care

scenarios but only if it can free up the nurse's time for addressing other needs of patients (other ADLs or social needs perhaps). For example, standardizing medication administration, feeding routines, check-ups on patients, etc. Alternatively, standardization presents itself as the main threat for a care ethicist as it promotes an objectification of persons. Thus, using the CCVSD approach prospectively means addressing the fears related to the very initiative to use robots in care contexts. By addressing the fears, I mean understanding what it is that these fears point towards – the values that are perceived as being in threat. Making these fears explicit allows one to uncover the values in direct risk and to use this as an additional tool to steer the design and design process of future care robots. For example, one might suggest that the value of efficiency comes with the threat of standardization. The CCVSD approach then asks, for a given practice how can efficiency be maintained without having to standardize entire practices or the overall care process of a patient? With the lifting robots, one might suggest that the human-operated robot allows for efficiency in terms of the mechanics of the practice of lifting while at the same time allowing for the dynamic completion of the practice. By dynamic completion I am referring to the care-giver's ability to "tinker" the practice of lifting if they see that the patient is in good/bad spirits, is improving/declining and thus their need for assistance changes, etc.

Table 9.1: The Care-Centered Framework

Context – hospital (and ward) vs. nursing home vs. home **Practice** – lifting, bathing, feeding, delivery of food and/or sheets and/or medications, **Actors involved** – nurse and patient and robot vs. patient and robot vs. nurse and robot **Type of robot and robot capabilities** – assistive vs. enabling vs. replacement **Manifestation of care values** – Attentiveness, responsibility, competence, responsiveness

The CCVSD approach thus begins with an account of the threats associated with the use of robots and continues with how to embed values into the architecture of the robot. The prospective methodology does not end, however, with the created robot. Not only is the design of the robot a crucial element but so too is how the robot is introduced into its context. In other words, how it is implemented. This is done using a dialogical approach. Discussions with the

direct and indirect users and stakeholders in the care institution in which the robot will be used aim at: mitigating unintended uses of the robot, empowering workers in the role as care providers, making clear that traditional roles and responsibilities are re-distributed but ultimate responsibility remains within their domain. The reason for addressing and incorporating the implementation of the robot in the prospective methodology has to do with the overall aim of the CCVSD approach and the lessons learned from the retrospective evaluations. Through the retrospective evaluations of care robots, the relationship between the robot's capabilities and the resulting expression of values was revealed. Once the robot became an actor in the network of a practice, the robot's potential to threaten the overall care process or the link between one practice and another was made clear. With this in mind the capabilities of the robot became a significant element for study. Additionally, the robot's presence as an actor was shown to play a role in shifting the delegation of roles and responsibilities within the network. It follows then that if the robot is created with a specific use in mind, in order to ensure the promotion of care values and a particular distribution of roles and responsibilities, then is becomes of paramount importance for the robot to be used in the intended way. Added to this is the belief that like many technologies, the robot presents the potential to alter existing norms and rituals and essentially takes on a meaning with pervasive use. One need only think of the cell phone, microwave, car or social networking like Facebook to understand such a phenomenon. Rules, norms, standards and social etiquette have altered with the pervasive use of these technologies. Both how the technology is used and the meaning it takes on have to do with the domestication of the robot. Once I have outlined potential future care robots I will address the aspect of implementation of the robots.

I begin with the CCVSD approach by re-visiting certain fears associated with care robots to pin point the underlying assumptions in order to account for these fears in the design process of future care robots.

9.3 Addressing Our Fears of Care Robots for their Future Design

I began this book by addressing the range of fears attributed to the initiative and use of care robots in care contexts. Illustrating this range of fears was meant to draw attention to the need for ethical analysis of these kinds of robots. I return to these fears, or ethical concerns, once again but with a different intention in

mind. This time I analyse certain fears to uncover the values thought to be overlooked or threatened and to incorporate these fears into the design process of the care robot. In particular, I address the fear of replacing care workers, the fear of a lack of human presence in the care of the elderly, and the fear of threatening the cultivation of goods internal to the care practice. All of these fears, as we will see, are linked with the general inquiry/fear of changing the standards and quality of care following the introduction of a care robot. If we consider what it means to have a fear or to be in fear of something, we may consider a fear as the anticipation that a value will not be met or promoted. A disvalue in its place will manifest itself.

Certain sceptics in the field of robotics warn of a technology push and ask whether the technical solution of a care robot is really the answer to a shortage in healthcare personnel [*van der Plas A. et al.*, 2010]. In Japan, where immigration laws prevent immigrants from coming in to act as care-givers, the technical solution appears quite appealing and necessary. In Western countries, however, without the same laws, one must wonder whether or not people can, and should, do the job of the robot. Many scholars believe that the initiative to use robots in care contexts is both an undervaluing of the care work that care-givers do, an undervaluing of care-givers and/or an undervaluing of care-receivers [*Sparrow and Sparrow*, 2006; *Turkle*, 2011; *Vallor*, 2011; *Wilson*, 2002]. This devaluation stems from the fact that designers fail to grasp the meaning, significance and particulars of care in context. Sparrow and Sparrow discuss this in terms of the meaning of human presence in the care of individuals and in the care of elderly persons in particular. They make the point that good care is dependent on the demographic that one is discussing and along those lines they articulate the significance of human presence in the care of elderly persons and the current lack in that area already [*Sparrow and Sparrow*, 2006]. From the standpoint that human presence is a value in the care of persons, the robot poses a threat to the promotion of this value. The script embedded in the robot is a devaluation of the value of human presence in care. One might conclude that this has to be with human presence alone or alternatively what human presence is required for, namely the establishment of the therapeutic relationship between care-giver and care-receiver. Is it possible that the design of the robot might mitigate this fear or is the presence of the robot enough to claim that the robot is promoting a disvalue? I claim that it is possible to design the robot in way that human presence is ensured for the practices within which it serves a direct role in the establishment of the therapeutic relationship. Even when a care robot is used, it may be designed in a way that demands the presence of the human care-giver while at the same time relieving them of a certain duty.

But the moral dilemma faced with robots does not stop at devaluation; it continues with a discussion of what the devaluation leads to. Care-givers often express their hesitation with care robots for fear that the robot will take their job [*van der Plas A. et al.*, 2010, p. 312]. This fear presupposes both that the robot can provide care similar to that of a human nurse and further that the nurse's role is replaceable. What's more, in order to make room for the robot a certain amount of standardization must take place. But, would standardizing care practices to allow for a robot change the actual manner in which care will take place in the future? Turkle speaks of this and warns that; "When we make a job rote we are more open to having machines do it. But even when people do it, they *and the people they serve* feel like machines" [*Turkle*, 2011, p. 146]. In Turkle's style of questioning, one might wonder whether the same trend will fall upon certain users of care robots, the nurses. Will care-givers begin to fulfil their roles in a machine-like manner to mirror that of the robot? Or perhaps more importantly, will care-givers lose sight of the social dimension in care if a robot can do it? I claim that in order to mitigate this last fear in particular, the care robot may be endowed with social capabilities but only to the extent that the capabilities enhance the human-robot interaction rather than meeting social and emotional needs of patients. This, remains the role and responsibility of the human care-giver.

This leads us to the question of whether or not certain goods internal to care practices will no longer have the (same) room for development if a robot takes over a social role or more generally if the robot is present at all. Although the moral elements are goods that are developed within and throughout the practice, Shannon Vallor speaks about the virtues of the care-giver as a good internal to the practice of care-giving. Vallor claims that the real threat inherent in care robots is their potential to threaten the cultivation of the care skills and virtues that are required both for the provision of good care but also in the development of the care-giver as a person capable of engaging in meaningful relationships and cultivating skills like empathic reasoning [*Vallor*, 2011]. In other words, that using the care robot takes something away from the care-giver in terms of their flourishing as an individual as well as in terms of their skills as a competent care-giver. Again we are reminded of the conceptualization of care as meaningful in the development of persons as persons [*Reich*, 1995; *Vanlaere and Gastmans*, 2011].

I do not deny this claim nor consider it irrelevant but I ask whether or not it is possible to design the robot according to this insight/fear/threat. If a virtue can be articulated and its relevance proven, there ought to be a way to facilitate the development of the virtue even while using the robot. This of

course demands an understanding of the virtues, or rather of the care virtues. Just like care values, however, there are a range of virtues that may be applied within the care domain. Care virtues are virtues directed at the overall aim of providing good care. Consequently, the root of the cultivation of said virtues has to do with another individual, the care-receiver. Alternative to the Aristotelian conception of virtues as integral components in the attainment of eudaimonia, happiness or flourishing of one's life, the virtues in care are directed at the flourishing of an individual in their role as care-giver. This is not to say that care for oneself is not important and in fact there is much to be said about care-givers not having the time or resources to care for themselves properly which bears an impact on their work. Rather, the point I am trying to make here has to do with the aim of cultivating care virtues as opposed to other (moral) virtues like courage, honestly, etc. Care virtues are cultivated in relation to others and so too does the criteria for their evaluation. With this in mind we may zoom in further on what care virtues might be. As I have claimed thus far, that the moral elements represent the range of care values. I add to this and claim that the moral elements also reflect the goods internal to care practices and further that the moral elements may also be labelled as the care virtues. Virtues in the sense that they are goods or skills cultivated on the part of the care-giver with the aim of providing good care for another. Consequently, by ensuring that the design of the robot accounts for the moral elements and acts to promote their presence (or the cultivation of the moral elements as skills of the care-giver) makes certain that the care-giver is afforded the opportunity to nurture the development of care virtues as represented by the moral elements. The robot may even present the opportunity to foster a greater development of the moral elements by drawing the attention of the care-giver towards important aspects of a patient's status or by reinforcing a role and responsibility of the care-giver.

I take the aforementioned warnings and fears seriously and I concur that to date, the development of care robots presents these fears as real threats. But I also write as an optimistic observer and believe there is a way to mitigate these fears and further to take these fears into consideration in the very design and design process of a care robot. To that end, I propose that the CCVSD approach allows for an incorporation of these warnings and offers a means for systematically designing the robot in a way that prevents these fears from becoming real. I claim that it is possible to design a robot that prevents the rote performance of care practices. For certain practices, ensuring the care-giver's role and responsibility remain in tact serves this goal. For others, designing the robot such that it actively engages the care-giver in a (novel) way acts to mit-

igate this fear. I claim it is also possible to design a robot that still manages to cultivate the skills and virtues of a good care-giver. In care practices, ensuring the presence of the care-giver along with the presence of the robot, and the co-operation between the two may fulfil this goal. For other care practices, it may not be desirable to have a care robot involved in the practice in any way if its presence alone will inhibit the development of a necessary skill/virtue. What's more, care robots may be designed for practices that do not actively promote the development of said virtues (as we will see later in this chapter).

Added to the fear relating to the cultivation of skills on the part of the care-giver, it is possible to align this discussion with one of the moral status of the robot. There exists a relationship between the moral status of the robot, the role ascribed to the robot, and what skills are cultivated through such a role. Given that I have already concluded that the robot's moral status be that of a moral factor (chapter 8) – meaning ultimately a human is responsible for an outcome or end effect – the robot cannot be delegated a role for which a full moral agent is required. A role in which empathic reasoning is required, a role which requires the actor to take responsibility for an outcome or end effect, a role which is essential for the cultivation of care virtues of the care-giver.

In short, from the above discussion we may conclude that a care robot must be designed in a way that addresses the threat of replacing care workers, of reducing human presence and of inhibiting the development of care skills and virtues.

9.4 Designing Robots with Care

The aim of re-visiting the fears identified by a variety of scholars to date was to bring attention to these fears as well as to argue that the design of the robot may in fact take these into consideration and incorporate them into the resulting design of the robot. Thus, the fear of replacing care workers is taken into consideration when brainstorming robot capabilities; how might the robot assist rather than replace the nurse? The fear that the use of the care robot will detract from the cultivation of the care-giver's skills and virtues is also one that may be taken into consideration by exploring what skills are cultivated through a given practice and the possibility to maintain a space/place for this within said practice or in another.

Thus, the first step in the CCVSD approach is to make explicit certain fears associated with the creation of care robots and to speculate the ways in which these may be overcome. Following this, still during the phase of idea

generation, the engineer and/or ethicist visits the hospital or nursing home to understand care in context; the number and variety of practices and their link with one another. In this way, the engineer/ethicist can also observe how values are translated and ranked in context which conforms with the suggestions of Nathan et al [2008] building on the traditional VSD methodology. Above all, however, the engineer/ethicist will understand where in fact there is a relevant need for a robot's assistance. I list both the engineer and ethicist here not to exclude one or the other from this role. It is important to note here that a division in moral labour is advisable. Rather, it is impossible for the designer to anticipate all the ways in which the robot will have an impact (and given this is not their training). As such, they will benefit from visiting the context to visualize care as a process. The ethicist must also visit the context to explain to the engineers the meaning associated with certain care practices and the relationship one practice has with another and with the overall process of care. Hence, the ethicist is trained for such work. Thus, the starting point differs not only in the timing of the robot's development but also the way in which needs are identified and targeted (by visiting the anticipated context and speaking with potential users, rather than speculating on the potential uses and needs of users from the lab). This process may also be enriched with focus groups to include multiple stakeholders but it is crucial that the ethicist and/or engineer be familiar with the context of use for themselves.

The next step is to describe the care practice for which a care robot idea is being developed. The practice is described in the same detail for the CCVSD approach as it was for the retrospective evaluations; how values are manifest through the actions and interactions of actors (human and non-human), how a particular practice is related to other practices and to the overall care of a patient and areas in which a robot may provide the possibility to re-introduce certain care values and elements that ought to remain intact if not strengthened through the use of the robot. With this, one begins to brainstorm the care robot in terms of its capabilities, features and functioning. During this phase of the design process, it is important to make explicit the relationship these capabilities will have in terms of the robot's responsibility to ensure the robot is conceived of as a moral factor and not a moral agent. As such, the ethical acceptability of the robot's capabilities (linked with their assigned roles) is studied on a case-by-case or design-by-design basis rather than making sweeping generalizations as to what any care robot ought to be capable of. Once again, the practice is described in the same terms as above, only this time with the addition of the care robot as a moral factor in the network. Continuing with the CCVSD approach I will know illustrate my intentions through the presentation of two

novel care robots "the wee-bot" and the "roaming toilet".

9.5 A Care Robot for Urine Testing – the "Wee-Bot"

9.5.1 The Practice

In the paediatric oncology ward in a hospital the nurse is responsible for a variety of activities; cleaning the patient, maintaining a sterile environment, keeping up to date with new research protocols and treatments as well as others. One responsibility of the nurse in this ward is to test the urine of the children undergoing chemotherapy for the presence of chemo toxins found in the urine. Essentially, the nurse is checking to ensure that the therapeutic intervention is present (the next step is to test if the intervention is working and whether or not higher or lower levels of the toxin are required). To do this the nurse puts on protective clothing as the chemicals are quite toxic and can cross the skin barrier, enters the bathroom, takes a sample of urine and tests it. With this information the nurse reports the findings back to the oncologist. For this practice, attentiveness is thought of in terms of the nurse's capability to know when to test the urine. This is not always up to the nurse, however, as it is also dependent on the nurse's schedule as well as the patient's ability to use the bathroom. Responsibility refers to the nurse's role in this process – this is an action that the nurse is responsible for fulfilling and for providing the results to the appropriate oncologist. Thus, this practice is inextricably linked with the overall care of the patient. Competence refers to the nurse being capable of retrieving the sample and accurately testing it. The urine sample acts as a vessel for reciprocity. The patient has not spoken, but the sample provides information pertaining to the patient's physiological status (the urine sample is an example of non-verbal communication).

This practice is linked with the overall care process in multiple ways. First, that the nurse is present to indicate to the patient what he/she is doing and in so doing reaffirms their role as care-giver within the therapeutic relationship. Second, the nurse continues to build the element of trust in the relationship – the patient trusts the presence of the familiar nurse (I have already shown how the element of trust is necessary in the ensuring the patient will comply with their care plan and do what the nurse says in other instances). I say familiar as this context (the paediatric ward) is one in which patients and nurses often have a long term relationship – the patient is under the care of the nurses for many

years in certain instances and builds a strong therapeutic relationship with both the patient and the patient's family/friends. This particular practice is not a moment in which the bond or relationship is strengthened but more so about the nurse "*caring for*" the patient through this action. Third, this practice is related to other care practices in that the results of the urine sample, coupled with results as to the efficacy of the treatment will determine the future of the treatment (whether higher or lower doses of toxins are needed and whether or not they are working at all). Thus, the nurse's role in this practice renders them responsible for a portion of the physical/therapeutic intervention of the oncologist.

Although the nurse is advised to completely cover their body to prevent exposure to the skin before entering the patient's bathroom, nurses often indicate that there is not enough time for this.

> *The well-being of the patient comes first and we don't have time to fully cover up so we just enter the bathroom covering our mouths and do the test* (personal communication).

The nurse, in this case, admits that due to time constraints she isn't able to cover up entirely but that due to the weight of the activity, the nurse adheres to the well-being of the patient and sacrifices his/her own safety to do so. Safety is interpreted in this practice as maintaining the care regime of the patient. Safety is also thought of in terms of the nurse's safety but the patient's safety trumps that of the nurse. One might question whether or not care is provided in a competent manner when the nurse is not protected; however, this is not something that can be blamed on the nurse (i.e., calling them incompetent) because they are fulfilling their role under time constraints in the only way they see possible.

What this current practice points towards is the same fear that feminist ethicists and care ethicists continue to point towards – a sacrificing of the well-being of the care-giver to meet the needs of care-receivers. In other words, a de-valuation of the care-giver when compared with the care-receiver. This undervaluing of the care-giver is a common and repetitive theme. The question then is whether or not it is possible to create an alternative means for this practice, one that prioritizes the well-being of the nurse higher in the ordering of values. To this end, I propose the use of a robot for the testing of urine.

9.5.2 The “Wee-Bot" robot

The robot, which I name "Wee-Bot", is a mobile robot that is human-operated. Thus, a human controls the movements of the robot. The robot is intended to drive into the bathroom of the patient and to collect and test a urine sample taken from the patient. To do this, the robot is driven by the nurse who is at close proximity. The robot is also capable of autonomous actions as well. Like the TUG platform, the robot is able to travel within hallways and elevators without the input of a human operator but this is only to travel from one floor and/or room to another. Once the robot is in place for its role, the nurse takes over and the robot is operated in real time using the input of the nurse for guidance. Once the robot has collected the specimen and tested it, the nurse gives a command and the information is sent directly to the appropriate oncologist (to their PDA for example).

By ensuring that the robot is human operated, the responsibility for the accurate completion of this practice still remains within the nurse's domain. The robot is recognized as a moral factor and as such the final responsibility remains in the hands of the nurse and not the robot. If the final responsibility were to be delegated to the robot, the robot would not require input from the nurse nor would it require the nurse to send the results of the test to the oncologist. With this in mind, the responsibility for testing is exclusively that of the robots. If something were to go wrong, however, the robot would be accountable but who would be to blame? Not the nurse as they were not required to be present and may not even be aware of when testing was happening, whether or not there were any problems with testing and whether or not the information was sent to and received by the appropriate oncologist. By ensuring that the nurse sends the information to the oncologist, again the nurse ultimately remains responsible for the acquisition and testing of the sample and the chain of responsibility is made clear – the nurse hands over responsibility for doing something with the sample once it has been sent to the oncologist. Once the robot has been included, attentiveness on the part of the nurse has to do with operating the robot to enter the bathroom instead of the nurse entering the bathroom. The nurse must be competent in this sense. Reciprocity still occurs in the same manner as prior to the robot's introduction – the urine sample is a form of reciprocity between the patient and nurse/oncologist. Thus, the addition of the robot with these particular capabilities ensures the promotion of the moral elements while at the same time relieving a burden of the nurse that threatens their own well-being. What's more, this practice is not identified as one in which the bond or relationship is established. It may be recognized as playing a role in

strengthening the relationship (or values like trust in the relationship) and thus ensuring the nurse's presence for this purpose may achieve the same end. It is necessary to ensure the nurse's presence when one considers that this moment is linked with the overall care of the patient in terms of the nurse's responsibility for the patient. Safeguarding the presence of the nurse to drive the robot and pass on the information guarantee's that the chain of responsibility is both made explicit and remains in tact regardless of the presence of the robot.

9.5.3 The Ethics of the "Wee-Bot"

As in the evaluations of current care robots, the very concept of this robot brings with it certain assumptions in the form of scripts. Using this robot for this practice ultimately aims at placing the value and safety of the nurse as a high priority, one which should not be disregarded for the well-being of the patient. The main assumption here has to do with the prioritization of the value of the care-giver; the script that the robot embodies is a conscious prioritization of their safety. In terms of the impact on the moral elements, using the robot does not detract from their manifestation and in fact the robot provides additional resources for ensuring a chain of responsibility. Given the value of human presence and the necessity for a human care-giver to bear the responsibility of such a crucial moment in the care of the patient, I would suggest further design considerations.

One might envision that the robot autonomously travels to the nurse's location; however, how does the nurse actually operate the robot? Is there a separate console or remote attached to the robot that the nurse picks up and uses to drive the robot? Will this not be too time consuming for the nurse or require a new skill set irrelevant to the other roles of the nurse? What's more, what if something is happening to the patient and the nurse's attention is directly focused on the robot and not on the patient? In light of these considerations, perhaps then we might suggest that the robot not only travel throughout the hospital autonomously but also travels inside the patient's room and collects the urine sample autonomously. With this suggestion one is left wondering where the nurse would be during all of this and further why the nurse's presence is required at all? This design recommendation may be thought of in terms of whether or not the human care-giver remains "in the loop" of both the practice and the decision-making of the practice; "whether a human being will still meaningfully be in the loop as robot care-givers emerge and become more pervasive is an overarching concern. For instance, will a person check on an elderly resident in a nursing home or monitor a robot's performance? Robots could work in

conjunction with human care-givers [*Decker*, 2008]. But, Sparrow and Sparrow suspect that this practice will not continue over time [2006, p. 150]. Endowing the robot with complete autonomous capabilities for makes it possible to suggest that in time a human being *in* the loop of decision making may shift to a human being *on* the loop, meaning, the human monitoring the robot's actions rather than making decisions for the robot.

To mitigate this concern, I would suggest that the nurse be responsible for identifying themselves to the robot prior to its entry into the hospital room (identification could be through voice commands, facial recognition, a finger swipe, etc.). This gives the robot permission to enter the room but also ensures that the nurse is present for the practice. The robot could be programmed with semantic links to endow the robot with the capacity of knowing 'why' it must ensure the presence of the nurse. The robot may also be designed such that when it leaves the hospital room it must also interact with the nurse prior to sending the information onto the oncologist. What's more, once the information has been sent to the oncologist the robot requires that the nurse 'sign-off' in a manner of speaking before the robot is able to leave the scene. The robot also keeps track of times, who the nurses are, when information has been sent and when information has been received. The nurse's role and responsibility for this practice remain in tact but their safety is prioritized in tandem with the efficiency of the practice. The nurse remains connected with the overall process of care and may in fact appreciate the initiative and presence of the robot in safeguarding their own interests. In other words, the robot provides the potential not only to benefit the care-giver in terms of the physical safety but also in terms of the satisfaction and worth they receive from the role as care-giver.

When we consider the skills cultivated on the part of the care-giver through this practice, I would suggest that the care-giver is faced with the existential vulnerability of the patient. He/she is acutely reminded of the delicacy of the patient which conjures the cultivation of empathy, compassion, a recognition for the integrity and dignity of the patient and the frailty of life. All may be considered virtues in care as well as virtues in the development of the care-giver as an empathic individual. The nurse is also reminded of the relational status of the patient given that the patient relies on the nurse's actions essential for their well-being – actions they cannot complete on their own. These are skills or virtues cultivated throughout multiple practices that the nurse engages in but it is important to recognize that this moment in the care process acts to reinforce such a recognition or cultivation. Thus, by ensuring the nurse's presence while the robot acquires the urine sample takes such considerations

into its very design.

9.6 A Care Robot for Waste Removal

9.6.1 The Practice

In the hospital and nursing home contexts, patients who are not able to get themselves to the bathroom have a bed pan and will excrete into the pan. At certain points in the day this pan is then changed. What also happens with these patients is a soiling of the sheets/linens which then have to be changed. When a nurse is aware of this, they will come in to change the sheets, and/or empty the bed pan, and will carry them to the cleaning station (where laundry is done or a drop off point on their particular ward). Alternatively, the nurse will empty the contents of the bed pan in the bathroom or will again bring the bed pan to a cleaning station. Carrying the sheets/bed pan through the halls, however, is both unappealing (and undignified) for the nurse and poses a sanitary risk in the institution – others are exposed to the bacteria in the fecal matter.

Attentiveness, in this practice, refers to the nurse's ability to observe the excretion patterns of the patient. The nurse must be attentive to the colour, size, smell, etc. of the feces, the differences between days and times in a day. There is much information to be gained through the observation of excretion (indicative of someone taking their medication, indicative to someone's eating patterns, indicative to someone's recovery post surgery). Although the end result of waste removal here is to literally remove the waste from the patient's immediate surroundings, it is also a moment in which the relationship between nurse and care-giver is fortified. I refer to a strength in the relationship because this is quite an intimate practice between nurse and patient; it is a moment in which the asymmetry of the relationship is quite visible as is the patient's dependence on the nurse. If we recall in chapter 3, I spoke of the therapeutic relationship and the responsibilities of the nurse within this relationship. Attention to the asymmetry in power is a responsibility of the nurse. The nurse must be attentive to the vulnerability of the patient and must ensure the dignity of the patient through his/her behaviour and response to both the excretion and the removal of excretion from the room. This practice is linked with the overall process of care for the patient given that the nurse is made aware of how fragile and/or vulnerable the patient is and may take this into consideration for other care practices, like bathing or lifting. Thus, by knowing the patient, the nurse is able

to tailor their behaviour for this particular practice but through this practice the nurse is also aware of how to treat the patient in other practices. If the nurse observes a high level of discomfort of the patient, the nurse may choose to be extra sensitive when the time for bathing comes. Consequently, attentiveness in this practice refers to the nurse's capacity to observe the excretion patterns of the patient but also to incorporate the patient's existential state, in terms of his/her vulnerability, in the treatment and removal of the excretion as well as in the treatment of the patient during other care practices.

Competence refers to the nurse's ability to accomplish the above level of attentiveness. What's more, competence here refers to the safety with which the nurse fulfills the practice. When seen in terms of the risk of infection, however, often times this practice is not fulfilled in the most competent manner – when the nurse carries soiled linens through the halls, other healthcare staff and visitors are potentially exposed to bacteria found in fecal matter, urine or vomit. Thus, the practice, as described here leaves room for improvement in terms of competence and ensuring a greater level of safety for both the nurse removing it as well as other hospital staff and visitors that may come into contact with the nurse while travelling throughout the hospital. Responsibility refers to attentiveness and competence in that the nurse is the care-giver who is delegated the role for notating the excretion patterns of the patient. This presupposes the nurse is present to observe the excretion of the patient. Responsibility in this practice is thus expressed through the presence of the nurse, not necessarily during the moment in which it occurs but during the moments following the act. This presence contributes to a trusting feeling of the patient for the nurse. Hence, responsibility, associated with human presence, is also important with respect to the overall care of the patient. Being present for and observing the excretion patterns of the patient intertwines the nurse's role with the physiological dimension of care for this patient. This is of course in addition to the nurse's attentiveness to the patient's status and tailoring other care practices accordingly. Once again, the relationship between attentiveness and competence is reinforced.

Reciprocity, similar to the urine testing practice, is not reduced to verbal communication between nurse and patient but also incorporates a corporeal communication in a manner of speaking – the patient communicates through the information provided through their body, their feces. In addition to this, however, the nurse and patient may also communicate throughout the practice as a way of fostering the relationship. This will be decided by the nurse depending on the comfort level of the patient – there will be patients that would prefer not to speak at this time.

This practice is an integral component of the care of a patient in terms of the physiological information gained by the nurse but also in terms of a strengthening in the relationship between nurse and patient. But as we can see through the description of the practice here that there are values that may have been neglected – the physiological safety of other healthcare workers and visitors to the hospital, the dignity of the patient and the dignity of the nurse. The moment the nurse carries the soiled linens through the hospital corridors is a moment in which their dignity, or comfort level is threatened. What's more, this uncomfortable moment also translates to the patient – their level of comfort and dignity is threatened by the discomfort of the nurse in this role and by the nature of the practice on its own. Of course I will not claim that this is so for all patients, but one may assume this is so for a large portion of patients. One must ask then whether or not there is an alternative that may still allow for attentiveness on the part of the nurse but that may relieve a certain portion of the discomfort experiences by nurses and patients in this practice.

9.6.2 The "Roaming Toilet" Robot

Here, I again pose the question whether or not a robot has the potential to re-integrate certain values that ought to be given higher priority – namely the safety or nurses and other hospital staff/visitors as well as the dignity of the nurse. This prioritization may or may not have been intentional, but rather a product of an efficient system where an alternative has not presented itself. To this end, I propose a waste removal robot for this purpose. I distinguish this robot from the "dustbot" created at the University of Pisa, Italy. The dustbot is intended as a garbage removal system without a particular context in mind. The robot uses a segway platform which does not provide the required stability when carrying the excretions of a patient. Instead, I propose a robot that uses the same, or a similar, platform as TUG with adjustments. The waste removal robot, which I call "the roaming toilet", should be capable of receiving a command and manoeuvring through hallways and elevators autonomously to meet the demand (i.e., to travel to the destination at which it is needed). The robot is equipped with a sterile compartment into which the excretions from the bed pan (urine and feces and throw up) may be placed. These are kept in a sterile compartment to ensure the particles of the excretion do not escape. The robot then brings the contents of the compartment to the linen washing area or garbage area in the case that there are no linens to be washed. Although the robot autonomously travels throughout the hospital, it requires human input at certain points. One of these, already referred to, is commands – the robot may

be programmed for routine visits to rooms but is also capable of responding to an urgent situation in which it is needed. The robot also responds to the nurse's commands at the site of its pick-up. It is the nurse who places the bedpan in the collecting bin of the robot with its contents (bodily excretions as well as soiled linens). The dirty bedpan is replaced with a clean bedpan, stored in another compartment on the same robot. Once the nurse has done this, he/she will give the robot the command to leave the room. It is the robot then that travels through the hospital corridors to bring the soiled linens and excrement to the cleaning station. At the cleaning station the robot's contents are emptied, bed pans are cleaned and the robot is loaded with a new collection of sterile bed pans. This may be done by the nursing staff or the cleaning staff already present in the hospital .

By ensuring that the robot requires the nurse to input the materials into the robot's sterile compartment, the robot ultimately ensures that the nurse still bears the responsibility for his/her role in observing the patterns of excretion of the patient. In other words, the nurse still complies with the requirements for attentiveness as described above and as such still cultivates the skills for observing and understanding the patterns of excretion of the individual patients. Ensuring the presence of the nurse also maintains a delegation of the responsibility for the practice in the hands of the nurse. This last point refers to the chain of responsibility, or the continuity of care. The nurse is still deemed responsible for the overall care of the patient, as expressed through their physical presence for the practice. What's more, the insurance of the nurse's presence also contributes to the element of reciprocity. If the robot were designed to remove excretions or soiled linens on its own without the nurse present, the nurse would not have the opportunity for observation. The moment in which the nurse observes the fecal patterns of the patient plays a key role in the reciprocity. Excretion patterns tell a lot about the patient in terms of whether or not they are taking their medications, are eating properly, or are recovering and these are all key factors in terms of reciprocity for this practice – there is much physiological information found in the excrement that the patient cannot express verbally. As such, excretion patterns become a vehicle for reciprocity in this practice .

In terms of the manifestation of competence in this practice, as mentioned the nurse is still afforded the opportunity for fostering his/her own competence, or attentiveness, of the individual patient, but the robot offers an opportunity to increase the level of competence when understood in terms of safety. By carrying patient excrement in a sterile compartment, the robot decreases the risk of transmitting potentially air-borne bacteria throughout the hospital. Thus,

when compared with the current practice, the robot re-prioritizes the values of safety defined by a reduction in exposure to potentially dangerous bacteria.

Although maintaining the nurse's presence plays an integral role in fostering attentiveness, competence, responsibility, reciprocity, and the continuity of care, it always plays a role in maintaining the relationship between nurse and patient. It affords an opportunity for communication between the two, whether this communication refers to the nurse verbally speaking to the patient, a pat on the arm or eye contact while engaging in the practice. Regardless of the manner of communication, the nurse's presence allows for this. In addition, the nurse is relieved of his/her role in the carrying the materials throughout hospital hallways and elevators and it is possible to suggest that the delegation of this role to the robot increases the comfort level and dignity of the nurse which in turn relieves a certain pressure of the patient – for a patient that feels vulnerable and uncomfortable with having the nurse in this role, the patient is appeased now. Given that the nurse will not need to leave the room immediately to remove the soiled items (the robot will instead do this), this nurse may have more time and may feel more inclined to sit and spend this time with the patient. This is of course speculation, but one may presume that when the nurse bears the responsibility for this particular burden, they may come to resent the patient, or may be less than friendly towards a patient. The assumption here then is that when the nurse is relieved of this burden the same opportunity for resentment does not exist. In this analysis of the robot, many assumptions pertaining to the script of the robot are being made. I turn now to investigate these assumptions and the overall script embedded in the design of the robot.

9.6.3 The Ethics of the "Roaming Toilet"

When we try to understand the inscribed script of this robot we are acutely aware of the robot's capacity to increase the comfort (and dignity) of the nurse and patient while at the same time ensuring the efficiency of the system. Added to this is the recognition of the significant role the nurse plays in this practice; the nurse is present to cultivate their own skills of attentiveness, competence and reciprocity. By skills of reciprocity I mean to refer to the nurse's ability to observe and 'read' excretion patterns, something that takes time to learn. What's more, the script adheres to the significance of having the nurse present with respect to the continuity of care; the nurse cultivates the skill of attentiveness which is applied in other care practices. Thus, the goods internal to this practice, the skills acquired throughout it, remain in tact when we ensure the presence of the nurse for the practice of waste removal. The design of the robot,

in terms of the role and responsibility the robot takes on, reflects a recognition of the importance of human presence in this practice as a means for fostering a trusting relationship. As such, the incorporation of the robot neither disengages the nurse from the overall care of the patient nor does it remove an opportunity for the nurse to learn more about the patient (in terms of their preferences and their state of well-being be it physiological or emotional). The role that the robot has taken on – namely a sterile carrier of waste – also inscribes a prioritization of the value of safety for the nurse and other hospital/nursing home visitors. The robot is thus responsible in this sense for promoting the value of safety as understood in this way. The robot's responsibility is limited in that it requires the nurse to load the bedpan on the robot's cart. Thus, the responsibility for the practice is shared between the two.

A potential problem that may occur would be that *any* human is capable of loading the sterile tray. There may be a time when the nurse does not reach the room quickly enough and perhaps there is a family member present. If the family member has been present for the patient's care on a daily basis, the patient and family may wish to allow the family member to load the tray, especially if this is more comfortable for them or if the nurse is not present immediately. This would defeat the purpose of ensuring the nurse's presence within the practice. Thus, a design recommendation to ensure the role and responsibility of the *nurse* (or another care-giver, for example a porter) in this instance may be to equip the robot with the feature of facial recognition or voice recognition. The robot is then capable of detecting the presence of particular persons/care-givers. In this way, the robot again maintains the chain of responsibility and acts to reinforce the role and the associated responsibility of the nurse. As in the case of the "wee-bot", this robot to may be programmed with semantic links to endow the robot with the capacity for knowing 'why' it must ensure the presence of the nurse. Thus, an additional redundancy is built-in to the system (i.e., an added mechanism for ensuring the presence of the nurse). What's more, to increase the dissemination and fluidity of information, the nurse may be required to verbally or manually input information pertaining to the patient at the moment during which time the waste is removed. The robot will of course keep track of the time of day and number of times it has visited a patient's room in a day/week/month and this information is stored to allow the other nurses and physicians access. Thus, the robot acts as a tool for linking healthcare professionals as to the status of the patient and any potential changes on a day-to-day basis.

This type of design recommendation naturally introduces questions of privacy; how long will such information be stored in the robot? Who will have access to that information, and so on? I would suggest here that only the care-

givers listed as being responsible have access to the information. If another physician wishes access they must first confirm permission with the physician and/or nurse in charge. In terms of how long that information is available for, I would suggest that the information be stored while the patient remains in the hospital. After the patient is discharged from the hospital the information may be kept in a database but the patient's identity remains anonymous (encoded through a patient number and not a name). If the patient re-enters the hospital, their information may be retrieved. If the patient is deceased, the information may be destroyed. If a healthcare professional wishes to use the information for future studies such permission must be obtained by the patient prior to its use.

Thus, the robot, designed as a tool that is added within the practice, is a tool that acts to promote certain values that may have been lost or overlooked. At the same time the robot, although only a tool, becomes an actor in the network in such a way that it increases the ethics of care-giving. What I mean by this has to do with feelings the patient may have pertaining to their satisfaction as a care-receiver and their vulnerability. To avoid the risk of stigmatization or undignified feelings of patients I suggest that the robot be used for all linen removal so that hospital staff and visitors do not attribute the use of the robot with an unpleasant event but rather with normal operating protocol. When compared with the use of TUG for linen removal one might question whether or not my description of a waste removal robot is all that different. Firstly, "the roaming toilet" is used not only for the removal of linens but also for the removal of bedpans. Second, this removal is done according to sanitary conditions. Thirdly, unlike the use of TUG strictly for everyday linen removal in which the role was previously delegated to a support staff and ultimately the nurse acquired an additional task, this use of the robot would not place an additional responsibility on the shoulders of the nurse but would relieve a responsibility of the nurse.

9.7 Implementing the Robots

Up to this point I have been engrossed in a discussion of the relationship between the design of a care robot and the resulting network in terms of the manifestation of values within said network as well as the re-distribution of roles and responsibilities among actors both human and non-human. With the inclusion of a care robot, the robot is then delegated a certain role and responsibility which influences the manifestation of values and it is through analysis using the CC framework that we come to understand how this happens and further what

meaning(s) might be attributed to the robot and the "new" practice. We have also come to understand the necessity of programming the robot – endowing the robot with specific capabilities – in order to maintain its moral impact, or moral status, as that of a moral factor. Thus, in line with the structural ethics approach, I claim that a human is always responsible in care but this does not take away from the fact that the robot is also recognized as having a moral impact. This impact, however, is in terms of its presence within a network and the consequences of this presence on the decision-making capabilities of the human actors along with a re-distribution in roles and responsibilities. My aim has been to show that the design process ought to be tailored to the moral status of the care robot, meaning how the artefact/robot will co-produce (new) distributions of roles, responsibilities, meanings and norms. Interestingly, it becomes evident that there is a specific normative force, derived from the care ethics tradition, embedded within the CC framework which is also ultimately embedded in the resulting design of the care robot. Despite this inherent force, is it possible to suggest that even with so much insight and thought going into the design (process) of the robot that it might still be used in an unintended and potentially negative way?

9.7.1 Domestication versus Implementation

The kind of question raised above refers to the domestication of the artefact, or robot. Domestication is a phenomenon in Science and Technology Studies to describe the enactment or performative character of artefacts [*Akrich*, 1992; *Latour*, 1992; *Silverstone et al.*, 1992; *Sorensen*, 2006; *Jelsma*, 2006]. The phenomenon of technology domestication explores how an artefact blends in with existing norms and meanings but also how the artefact co-produces new norms and meanings. Explaining and studying domestication rests on the semiotic approach of Akrich and Latour – script theory. Actor-Network theory comes into play here in that the visibility of the script, and the meaning of the artefact, is observed through the lens of the network, its actors and their interactions. Both the phenomenon and the study of domestication refer to something that is in the process of happening and the observance thereof, or to something that has happened and the denoting of. No normative approach to domestication exists to date. No studies testing whether or not domestication is something that can be steered. Of course for authors like Latour it is not possible to predict the impact of a technology until the technology is embedded and generates a sense of meaning; however I challenge this point and claim that through the CCVSD approach one is able to predict the distribution of roles and responsibilities with

respect to the addition of a care robot and further to enforce that their distribution occurs in the manner in which the designer/ethicist intended. With this, I make reference to the implementation of the robot.

Implementation is the carrying out, execution, or practice of a plan, a method, or any design for doing something. As such, implementation is the action that must follow any preliminary thinking in order for something to actually happen. In an information technology context, implementation encompasses all the processes involved in getting new software or hardware operating properly in its environment, including installation, running, testing, and making necessary changes. Implementation in a healthcare institution also refers to the policies and guidelines structured according to a technology; how it ought to be used and what constitutes misuse. Such is the work of Health Technology Assessment (HTA) [*Smits et al.*, 1995]. Recently, scholars in this field have begun to address how ethics might be integrated into HTA and whether or not such an inclusion is possible given the reflexive nature of ethics and the standardized nature of policies [*Hofmann*, 2008]. This of course also demands the question of how to know what ethical principles can be used to design and structure such guidelines and policies [*Moor*, 2006, 1995].

The benefits of HTA show us how to regulate and govern technologies prior to and once they have been implemented. But this kind of implementation – a list of how the technology is to be used, or a prescription as to what to avoid with the technology – does not reach the kinds of meanings the technology will take on in context. The benefits of domestication studies are the illustrations of how an artefact alters, reinforces and adjusts meanings, norms, values, roles and responsibilities. Without this knowledge one may not be aware of what happens with a technology once it enters its context of use. Domestication studies, however, fail to prescribe any manner in which a technology ought to be implemented and instead allow for a morality (norms, meanings, shifts in values) surrounding the technology to appear through time and use, and study how this came to be ad hoc. Design studies, like VSD, try to account for insights from domestication studies and to incorporate these insights into design. In particular, VSD takes as its starting point the very idea that the technology can and will promote or demote certain values and is thus designed accordingly. Added to this methodology is the insight that one using VSD must understand the values as they are expressed and understood in context which strengthens the intuition that the technology's implementation in context is of the utmost importance.

Other design studies emphasize context as the starting point. For example, contextual design or use-centered design. Focusing on context without appre-

ciating the network of human and non-human actors fails to grasp the moral quality the artefact may inherit. Alternatively, focusing on the usability of the artefact, while a value in itself, again fails to grasp the moral significance the robot has. Design studies like user-centered design focus on the users as the most important stakeholders in determining the direction the design will take [*Dubberly*]. As I have shown, such a focus fails to take the larger context and the interactions a user will have in a network, into consideration. What's more, focusing on a 'user', a 'use', or a 'context' holds the threat of embedding a kind of naturalistic fallacy – that what *is* becomes mistaken for what *ought* to be.

Accordingly, how does one incorporate the insights from design studies, domestication studies and implementation studies? I put forward that the CCVSD approach, when applied to implementation, does just that. The CCVSD approach manages to overcome certain limitations of domestication theories, design studies and HTA. I use the insights from both domestication studies as well as design studies but go further and claim that when one has the ideal intended use in mind based on a normative framework that this should be introduced along with the technology's introduction in context. The normativity of the CCVSD approach is grounded in the care ethics tradition. What's more, the approach embraces the care ethics tradition in its entirety and acts to incorporate its essence along with its components; an essence (or tradition) of deliberation and reflection, of allowing the moral question to reveal itself rather than on focusing on the traditional questions (pertaining to autonomy, justice, etc.). With this in mind then we don't want the care robot to enter its context and use and result in the formation of a novel, unanticipated or unintended morality. Rather, we want the robot to be used in the manner in which it was intentionally designed, a manner grounded in ethical consideration and foresight.

9.7.2 Ethical Implementation: The Marriage of Design Studies with Domestication Studies

With technologies like the mobile phone or the car, where it wasn't evident at the moment of design just how the technology would be used, the emergence of a morality pertaining to the use of the artefact following its implementation/domestication was observed (and in the case of the mobile phone or related technologies like the iPad, a formation of a related etiquette is still forming). But, on the other hand, with a technology like a care robot, for which there is an assumed use dictated by normative criteria, there is also an assumed morality of the artefact, a morality linked with its place in a network dictated by its use. According to script theory, Latour introduces the idea of an 'anti-program'. In

short, when people using a technology do so in a manner unintended by the designers. While it is true that the care robot may be used somewhat differently from its intended use, the morality of the network too may shift from the intended, anticipated one. The goal is not to place the robot in the network and allow for the emergence of a novel morality, one completely unintended and unexpected at the time of design. The goal is rather to encourage the use of the care robot in its intended way given that the design process and resulting design of the robot aim at the establishment of a morality in accordance with the care ethics tradition thereby protecting the provision of good care, the needs of care-receivers and care-givers and the formation of the therapeutic relationship.

So what happens when a criteria exists which might dictate the roles and responsibilities delegated to the robot and the human actors of the network? Shouldn't this be communicated with the micro and macro networks related to the robot's use? This does not assume that the robot will not be used in an alternative way but rather that its intended use, and the associated meaning ascribed to the robot, be made explicit and apparent. The implications for policy refer to the responsibility the nurse has to use it in a certain way. Is there something wrong with marrying the views of the engineer and the recommendations/policy for its implementation? Is this a good/desirable thing? To answer this I look to the work of Pak Hang-Wong who claims that engineers and philosophers shouldn't be so reluctant to give recommendations because of the built-in recommendation in the technology [*Wong*, 2011]. Consequently, I have already made normative recommendations as to how the technology in question (care robots) ought to be designed and I know make further recommendations as to how the technology ought to be implemented.

Of course my suggestions may draw critiques that I envision a utopian ideal, an ideal picture that cannot possibly be realized. But the question I would pose is why? If the response is for economic reasons, then why are we investing millions of dollars in robots for care in the first place? If the response is because the working environment doesn't allow for it in terms of time, then my retort would be to question how anyone will be able to use the system efficiently without first knowing how to. What's more, who is to be liable when the proper precautions for safe use haven't been made explicit. There are a number of reasons why someone might claim that my picture is too idealistic but to them I question their assumptions driving the use of care robots in the first place and their assumptions about best care practices. As Turkle enforces in her book, the interesting thing about robots is that they are now giving us (society) the opportunity to reflect on the values and principles that we as a society want to safeguard [*Turkle*, 2011]. Thus, in care, we have the opportunity

with the introduction of care robots, to tailor the implementation of the robot in a way that not only steers the domestication of the robot but ultimately shifts the way care is practised for the better. This shift comes in a reinforcement that although robotics technology is introduced, the core components in care (attentiveness, competence, responsibility, the therapeutic relationship, human contact, reciprocity, communication, touch, empathy, etc.) must be protected.

9.8 Ethical Implementation of a Care Robot – the Final Step of the CCVSD Approach

The robot's inclusion in the network already carries with it an assumed morality, one which has been reflected on and decided upon throughout its design. Thus, the introduction of the robot ought to be done in a way that ensures the manifestation of the presumed morality (distribution of roles and responsibilities, manifestation of care values, etc.). Such an idea conforms with the views found in the structural ethics approach, that the robot bears an impact not only on the micro network within which it is included but also on the macro network of the larger institution. Thus, by ensuring the manifestation of the morality which takes these aspects into consideration in design (and the design process), we are able to conform with the insights of the structural ethics approach.

The care ethics perspective is integral when outlining a care practice and the components of the care practice but also provides insights when it comes to the implementation of the robot – the care ethics tradition fosters a dialogical approach, one in which roles and responsibilities are made explicit through dialogue among actors. From this perspective, it is not only important to incorporate the components of a care practice into the design and development of the robot but the components are integral at the moment of implementation too. For this reason, I advocate in favour of using the framework as a guideline for the implementation of the care robot as well. But, what do I mean by this and further, how might this be achieved? To answer both of these, I offer a picture (or a scenario if you will) of a future in which the ethicist literally accompanies the technology into the context and facilitates a dialogue among healthcare workers.

> The ethicist, equipped with the CC framework in one hand and a technical knowledge of the robot in the other, enters the hospital context *prior* to the introduction of the robot. Over a short period of time (a couple of days), the ethicist observes the practices in the

> specific context (a nursing home, or a hospital, or a hospital ward) into which the robot will be added and the specific actors involved within that particular context to familiarize themselves with the contextual details. The staff members are then brought together to be introduced to the robot and explained about the robot's design and design process. The actors partaking in the discussions are those that will interact directly with the robot (the nurses using the "Wee-Bot" in paediatric oncology or the nurses using the "roaming toilet" along with the support staff who will receive the materials brought by the robot) as well as indirect users of the robot (the doctors who will receive information from the robot as to the concentration of chemotherapy drugs in the urine of the patient or the support staff who will no longer visit the rooms of patients to retrieve linens, respectively). In these meetings, the ethicist uses the components of the framework to structure a discussion among participants; the ethicist describes his/her interpretation of how the practice proceeded prior to the robot in terms of how and when values are manifest along with the distribution of roles and responsibilities of those involved. The ethicist then goes on to discuss what the robot is capable of doing, what role the robot will take on, what responsibility the robot is being delegated and how values might shift. The ethicist then encourages a discussion among care workers as to the robot's role once used pervasively in the context. This forum facilitates an opportunity for discussion of the potential unintended uses, the misuses of the robot, and/or the anticipated tensions in values and how they might be overcome. By engaging the healthcare workers in a dialogue about the robot, their roles and responsibilities are made explicit as well as where and how the robot fits in to their conception of roles and responsibilities. Once this has been accomplished, the robot is ready to enter the context and practice. The ethicist meets once again with staff after a short period of time to discuss the new pattern of the practice and to work out any unresolved intuitions or conflicts with the robot.

By applying the framework in the context within which the robot will be included, the goal is to shape the co-production of the social and technical characteristics of the network along with the material and semiotic characteristics of the network. Not only are things made explicit in terms of roles and responsibilities, the re-distribution thereof, and how values come into being but

the discussion session is also meant to trouble shoot potential conflicts. With the experts partaking in a discussion of what the robot will do, the expert's knowledge is tapped into to predict where a conflict might arise and how to pre-emptively deal with this conflict. Even when one incorporates the framework into the design process it is impossible to predict all the minor and major complications or disruptions. "One should not dismiss the fact that output is contingent on input decisions and design decisions about which information is relevant and how that information should be processed. Consequently this places a heavy burden on designers to predict the dynamics of socio-technical contexts within which a robot will be placed" [*Borenstein and Pearson*, 2011, p. 259]. With the CCVSD approach I am claiming that this isn't necessarily the case, that instead a division of moral labour may be introduced. Such a division would require that the ethicist be responsible for acquiring relevant contextual information and bringing it back to the designer and the relevant contextual information is dependent on the ethicist being present in the hospital or nursing home context to observe and conduct interviews. The action of visiting the context of use may also invite a kind of 'downstream flexibility' in the design process, which requires that systems are designed with a certain degree of "methodological insecurity" [*Kiran*, 2011] in order to allow 'room' for such flexibility. Prior to the robot's official implementation in its context of use, a kind of use- or user-centered design approach is adhered to whereby user needs are explicitly taken into account. This may be in terms of ergonomics or certain practical details of the context that had previously gone unnoticed.

Bringing together healthcare workers to discuss the introduction of the robot also affords a moment in which needs are discussed; what the predominant needs are, how they are currently met, what resources are still required, etc. While these discussions may not extend beyond the realm of the robot's inclusion it also gives the ethicist a plethora of information to bring back to the robotics company and the engineers they are working with. The discussion not only gives voice to the care experts but actively values their knowledge and opinions. This makes the implementation of the robot safer for all but it also empowers the healthcare workers to feel engaged with the robot, the practice, the process, the institution, the outcome and the people involved. These two points pay tribute to maintaining the care process and the linkage between care practices as well as ensuring the robot enters the context in its intended manner.

For the first point – the safety of the institution/context following the introduction of the robot – one need only look to the empirical evidence provided

by Salus Global[1]. Salus Global is a Canadian company dedicated to patient safety in the hospital, predominantly in obstetrics and gynaecology. The basic mantra of Salus is to improve patient safety by fostering communication among healthcare personnel that promotes the speedy, efficient and caring handling of anything going on in the ward. Here, safety is interpreted in terms of the morbidity and mortality rate of mothers and post natal infants. Safety is therefore prioritized very high on the hierarchy of values. Safety is achieved and maintained when each staff member has a clear idea of the role and responsibility delegated to themselves as well as the other care personnel which allows individuals to know when they need to act or who to ask when they are unsure of what to do. This also promotes an appreciation of each individual's role which in turn fosters a positive working environment. This, in turn makes the patient's experience less stressful and more enjoyable (for lack of a better word I used enjoyable although hospital visits are rarely depicted as enjoyable). Making the distribution of roles and responsibilities explicit contributes to creating a model or representation which is then used to reflect/evaluate/assess the provision of care. This model is not based solely on the current mode for performance but on a discussion of the current compared with the ideal. Actions become implicit through years of practice and thus by making the details of actions and their goals explicit the care-giver is afforded the opportunity to articulate it for them self as well as for others. Consequently, not only is patient safety maximized through the program but so too are the subjective experiences of the healthcare workers and patients. One might also conclude that the insights presented here pertaining to the implementation of a care robot might also be used for current care robots which are about to be introduced for the first time in certain hospitals (commercially available TUG, Paro, RP-7, daVinci®and/or Titan's Amadeus Composer^{TM}surgical robot among others).

9.8.1 The Role of Structural Ethics in the Implementation of the Care Robot

Using the CCVSD approach in the implementation of the robot we can see how domestication studies, while hugely insightful as to the mechanisms that drive the formation of a morality (through the pervasive use of a technology) are also limited in that they study the forming of a morality without steering any such formation. Counter to this, design studies which account for the domestication of an artefact, fail to implement the technology in a way that

[1] See salusgc.com

ensures the intended use. Added to this, technology assessments which aim to create guidelines and policies for the initial introduction and continued use of a technology fail to incorporate an adequate ethical analysis to guide such an introduction. If we take into account the claims made by Tronto regarding a good care institution along with the structural ethics approach, the robots role must be thought of in terms of its impact on the micro level (the care practice) and the macro level (the care institution). To balance all of this, using the CCVSD approach in the design, development and implementation of a care robot is shown to ensure an adequate ethical analysis of all related networks throughout the design process of the robot as well as steering the introduction of the care robot in a manner that reflects the intentions for the envisaged use of the robot. By making the initial roles and responsibilities of the human and non-human actors explicit as well as indicating how these roles and responsibilities will shift with the introduction of the robot, care-givers are given the chance to understand the initiatives and assumptions guiding the design and to understand how their role as care-giver remains in tact and respected. Thus, not only does the robot offer the opportunity for enhancing the provision of values through its design but through its implementation as well. The foundation of care rests on a valuation of both care-giver and care-receiver. By giving voice to the care-givers, their significant role is given value and meaning through this practice of empowerment. In other words, the kind of reflection insisted upon by Turkle and the kind of regulation insisted upon by robot scholar Peter Singer are achieved through the CCVSD approach.

The last point, referring to a targeting of the human actors at the micro level of the overall network, is the direct link with Trontos recommendations pertaining to care institutions [*Tronto*, 2010] and with the structural ethics approach [*Brey*, 2012]. The work of Tronto and Brey parallel each other. Brey claims that the overall network must be taken into consideration and Tronto specifies this in terms of care institutions. Namely, the purpose of the care institution is to be a guiding force for the multiple care practices within. Within the structural ethics approach, recognizing the network on a larger scale means the hospital for example is considered a network within which multiple micro networks exist. Or, perhaps the healthcare system in a province or state is considered a network within which multiple networks exist at a more micro level, and so on. Important to the approach of structural ethics is the recognition that all actors play a role in establishing a morality of the structure/network. Thus, the janitors or the kitchen cleaning staff are actors in the morality of the network as a whole. This may be seen in how they are themselves treated (and thus valued) and how they feel while complying with their roles and responsibilities.

What I mean here has to do with tasks like urine testing or waste removal and the potential undignified feelings such roles may elicit on the part of the nurse/care-giver. Alongside the consideration of human actors, the non-human technologies that exist within each network also contribute to the formation of a morality. The technologies, which are also considered moral factors (or ethical impact agents according to Moor) are known as such given that they bear a moral impact on the network according to the observation that they help to structure the actions, behaviours and outcomes of the network.

The relevance of structural ethics is seen on many layers when considering the CCVSD approach. For example, the design of the technology takes into account the actors, context and practice at stake and how institutional values might be preserved through the design of a care robot. Given that the design of the care robot must also reflect the understanding of the care robot as a moral factor and not a moral agent, the role and responsibility is dictated by the foundation of the structural ethics approach. Whats more, structural ethics approach rests on the belief that the morality of the macro network rests on, is supported by that of, the micro level and as such the micro level must adhere to the values structuring that of the macro level. Thus, ensuring a certain implementation of the care robot, one that empowers care workers and makes explicit how roles and responsibilities are reinforced or altered, works to insist on a preservation of the values which configure the morality of the network.

One might be asking at this point "what an interesting turn from the design of the robot to a focus on communication among a team of people who may or may not come into contact with the robot. This is for a number of reasons. The first reason has to do with a recognition once again of the care process – the vision of care being a holistic process, each action linked with the others. This is in contrast to the flat view of care actions as tasks – separate individual tasks serving only one purpose. When recognizing the holistic vision of care one is reminded of the interconnectedness of actions. Thus, if the robot performs one action, other actions and actors will be impacted. By discussing and preparing for this, the hope is to anticipate and troubleshoot but also to make explicit the shifts in roles and responsibilities. This of course can only be done when conventional roles and responsibilities have been made explicit. The second reason has to do with my claim that the robot be designed as a moral factor. The robot does not make ethical decisions itself but rather becomes part of the ethical dimension of a network when the decisions of human actors are impacted by the mere presence of the robot. In other words, the human actors are still making the ethical decisions even though there is a robot around and even if the robot "nudges" the actions of a human, the choice is still the humans to

make: the responsibility of such a choice is also still within the human's domain. The robot then is a tool within human-human care that enhances certain values whether they be efficiency, safety, dignity, human presence, touch, etc. I am claiming that essentially 'care' remains a human domain: a human activity to be fulfilled by a human, enhanced through the use of a care robot when designed and implemented according to the CCVSD approach.

9.9 Conclusions

There are many conclusions and benefits to be drawn from the work of this book. Firstly, this book presents a concrete methodology for the creation of future care robots, the Care-Centered Value-Sensitive Design (CCVSD) Approach. This approach explicitly integrates the care ethics perspective in the design process of care robots: in structuring the design process, by shaping the resulting design, and by steering the implementation of future care robots. Given the novelty of care robots, the lack of current standards guiding their design, and the morally charged contexts within which these robots will be placed, ethical guidelines pertaining to their design ad development are not only recommended but are crucial to merit the trust placed in both the designers and the resulting care robot.

Secondly, added to the utility of the CCVSD approach is its inherent interdisciplinary focus. Given that care robots draw on multiple disciplines in terms of both their design as well as their use, it is of paramount importance that multiple disciplines be involved in structuring their development and use. But such a task is not an easy feat. To accomplish this, I have translated ethical principles, values and norms into a tangible tool for engineers and robot designers. The CCVSD approach not only allows for interdisciplinary collaboration, but also provides a means to encourage and foster such collaborations. In so doing, the CCVSD approach provides a standardized vocabulary to be used across disciplines, a vocabulary that may be understood by all disciplines involved. Granted the ethicist will have a deeper understanding of ethical principles and the roboticist will have a deeper understanding of robot capabilities, but, all actors in the design process will now have access to a homogenized vocabulary. Not only that, but the CCVSD approach allows for the ethicist to enter the design process prior to a robot being embedded in practices. This is in line with a new wave of ethical technology assessment that seeks to be proactive given the current recognition that technologies are not mere instruments that we use to shape our practices, but rather that they are themselves shaped by society

and in turn shape practices and policies by their presence.

In conjunction with the interdisciplinary advantage, a benefit of the CCVSD approach is its commitment to a division of moral labour. The approach encourages an ethical reflection on the part of engineers and designers without demanding that they become specialists; the work of translating ethics has been done for them. Added to this, the approach demands that the ethicist have a technical knowledge of the robot, without demanding they become robot specialists. This is necessary in order to avoid a speculative ethical appraisal of future robots that may never come into existence. The focus is on the creation of robots that may be used currently, based on an understanding of what current robots are capable of.

The CCVSD approach also incorporates the element of implementation where there are additional advantages to draw upon. By proposing a means for the implementation of the care robot, my aim was to marry the intentions of design studies with the results of domestication studies. This would ensure that the built-in morality of the robot would be the same one produced through the use of the robot. By accompanying the technology into its context of use, and engaging in a dialogue with the direct and indirect users, the CCVSD approach once again makes real the care ethics tradition of giving voice to the under-represented (and often under-valued) groups in care; namely, the care-givers. By giving voice to the care-givers and explaining both the initiative and hopes for the care robot, the robot has a better chance of being used in the intended manner of it design. And the practice is a concrete valuation of the care-givers in their roles. Thus, the CCVSD approach not only bridges the gap between the disciplines involved in the design and the ethical evaluation of the care robot but it also bridges the gap between the intended and actual use of the care robot. One might also likewise conclude that the insights into the implementation portion might also be used for current care robots which are about to be introduced for the first time in certain hospitals (including surgical robots). In other words, that from this moment on, any robot introduced into a healthcare context is done with the assistance of an ethicist to make clear the re-distribution of roles and responsibilities that will result once the robot is integrated.

In terms of designing future care robots, the work of this book has provided an illustrative example of how to operationalize the Value-Sensitive Design (VSD) approach. Scholars have criticized VSD for its lack of aligning with any particular ethical theory and have also criticized VSD for being more of a theoretical approach rather than a practical one. To counter the first claim, in aligning VSD with the care ethics tradition this problem was overcome. Moreover, in using the VSD blueprint to create the Care-Centered (CC) framework, I was

able to provide a means for helping to clarify the steps of VSD. Added to this, the CCVSD approach strongly resembles the VSD approach in the creation of a future technology. In fact, it provides a pragmatic application of VSD in practice. As an alternative to traditional VSD approaches, I was able to provide a concrete methodology for the design of a future system, one that can be copied for the design of any future care robot rather than being applicable for the design of one specific computer system only.

An additional benefit of the work in this book is seen when we address the CC framework in isolation from the CCVSD approach. In short, the CC framework may be used in the evaluation and future design of other technologies, used in different contexts which are guided by different values. To do this, certain adjustments must be made. If one were to use the framework in the evaluation of an ICT system, one may wish to change the values of importance from the moral elements to values such as privacy, distributive justice, efficiency, or dignity. Accordingly, the practices for evaluation will also adapt to the technology in question. The values chosen for evaluation are dependent on the context of use, the goals of the institution within which it will be used and the end goals of the system. The methodology for describing such practices and the relationship between one practice and another remain the same. Although the values may change, their manner of interpretation (contextualized) does not, nor does the manner in which the practice is described with respect to the interactions between actors as the vehicle for such a manifestation (the methodologies for use).

Let us take social robots as example of how to alter the CC framework for evaluation. It should be clear at this point that I did not use the framework in the evaluation of social robots for a variety of reasons. Firstly, social robots do not fall within the realm of care robots given the definition and end goals of both (see section 4.5.1, "Care Robots, Social Robots and Companionship). Social robots aim at companionship, care robots aim at providing care. A relationship may be established in the case of care robots; however, this relationship is not one of companionship but of a therapeutic nature. The relationship is not an end in itself but is a means for the end of good care. I have previously stated that the kinds of practices social robots are engaged in are social practices. The evaluative criteria of the CC framework must reflect this to incorporate the ends which companionship serve as well as the subjective nature of such relations. In other words, added to the CC framework must be a tool for conceptualizing when the relationship between human and robot may be considered meaningful and according to what criteria. As such, the CC framework, the retrospective evaluations of care robots and the CCVSD approach may all be used in the

evaluation and design process of alternate technologies with adjustments.

The above discussion reinforces the tailoring of the CC framework and its methodologies for use to: the care ethics tradition, the concept of care, the institutional context of care, the significance of the therapeutic relationship in the provision of good care, and an understanding of care at a contextual level that illustrates the manner in which values come into being. I return to the quote given as the metaphor for this chapter; "The best forms of institutional care will be those which are highly deliberate and explicit about how to best meet the needs of the people who they serve" [*Tronto*, 1993, p. 169]. Tronto is mindful of the complexity of care and how demanding such claims are (to understand the care institution and care process in its totality); however, by enforcing her three central foci (politics, particularity and purposiveness) the ineffable dimensions of care are uncovered and are articulated. Just as Tronto argues in favour of understanding the care process as a whole in order to structure a good care institution. I too make such a claim with the creation of the CCVSD approach. Including both prospective design as well as implementation through the CCVSD approach I claim that the care process must be understood and critically examined as a whole prior to the design and introduction of a care robot. The manner in which care is understood takes into account the needs of care-receivers but places the needs of care-givers and the allocation of responsibility as central foci. This line of thinking falls under the umbrella of the structural ethics approach as well, that the care institution be understood in its totality. This demands that the relationship each care practice shares with another be made explicit and criticized. How the manifestation of a specific value in one practice impacts the manifestation, or lack thereof, of a specific value or the meeting of needs (of either or both the care-giver and care-receiver) in another practice. To that end, the strong recommendations of Tronto [1993; 2010] and of psycho-analyst Sherry Turkle [2011], coupled with the insights from the structural ethics approach [*Brey*, 2012], are made real through the CCVSD approach.

In short, the CCVSD approach reflects a commitment to understanding the starting point of care, being critical of care as it is currently practised, understanding the necessary elements of care and how they contribute to the provision of good care, and above all else how care might change (for better or for worse) with the introduction of a care robot. With these commitments in mind, the future of care, with or without the inclusion of a care robot is studied: the values in care are understood in-depth and preserved. The non-human actors are recognized for their role in forming a morality and meaning within a network. Most importantly, the human care-givers are recognized for the value of their

role, and the responsibility of their actions as the stewards of care. Added to this, the CCVSD approach also reflects a commitment to the design and implementation of future care robots that systematically accounts for the valuation of care practices and care workers. Seen through this lens, care robots hold the potential to provide a vital role in aiding the promotion of care values when designed and implemented according to the manner dictated by the CCVSD approach.

Bibliography

Akrich, M., The de-scription of technical objects, in *Shaping technologybuilding society*, edited by W. E. Bijker and J. Law, pp. 205 – 224, MIT Press, 1992.

Albrechtslund, A., Ethics and technology design, *Ethics and Information Technology*, *9*(1), 63–72, 2007.

Anderson, M., and S. Anderson, Machine ethics: Creating an ethical intelligent agent, *AI Magazine*, *28*(4), 2007.

Anderson, M., and S. Anderson, Robot be good: A call for ethical autonomous machines, *Scientific American*, *303*(4), 15 – 24, 2010a.

Anderson, M., and S. Anderson, Robot be good, *Scientific American*, *303*, 72 – 77, 2010b.

Arras, K., and D. Cerqui, Do we want to share our lives and bodies with robots? a 2000 people survey, 2005.

Asaro, P., What should we want from a robot ethic?, *International Review of Information Ethics*, *6*, 8 – 16, 2006.

Asaro, P., Modeling the moral user, *Technology and Society Magazine, IEEE*, *28*(1), 20 – 24, 2009.

Barras, C., Useful, loveable and unbelievably annoying, *The New Scientist*, pp. 22 – 23, 2009.

Beauchamp, T. L., and J. F. Childress, *Principles of biomedical ethics*, – pp., Oxford University Press, Oxford; New York, 2001.

Beetz, M., U. Klank, I. Kresse, A. Maldonado, L. Mosenlechner, D. Pangercic, T. Rühr, and M. Tenorth, Robotic roommates making pancakes, in *11th IEEE-RAS International Conference on Humanoid Robots*, Bled, Slovenia, 2011.

Bekey, G., Current trends in robotics: Technology and ethics, in *Robot Ethics: The Ethical and Social Implications of Robotics*, edited by P. Lin, K. Abney, and G. Bekey, Intelligent Robotics and Autonomous Agents, chap. 2, pp. 17 – 34, Mit Press, 2011.

Bensalem, S., M. Gallien, F. Ingrand, I. Kahloul, and N. Thanh-Hung, Designing autonomous robots, *IEEE Robotics and Automation Magazine*, *16*(1), 67 – 77, 2009.

Bicchi, A., M. Peshkin, and J. Colgate, Safety for physical human-robot interaction, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 1335 – 1348, Springer, Berlin, 2008.

Billard, A., S. Calinon, R. Dillmann, and S. Schaal, Robot programming by demonstration, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 1371 – 1394, Springer, Berlin, 2008.

Bischoff, R., and V. Graefe, Hermes - an intelligent humanoid robot designed and tested for dependability, *Springer Tracts in Advanced Robotics: Experimental Robotics VIII*, *5*, 64 – 74, 2003.

Borenstein, J., and Y. Pearson, Robot caregivers: Ethical issues across the human life, in *Robot Ethics: The Ethical and Social Implications of Robotics*, edited by P. Lin, K. Abney, and G. Bekey, Intelligent Robotics and Autonomous Agents, chap. 16, pp. 251 – 266, Mit Press, 2011.

Borgmann, A., *Technology and the Character of Contemporary Life: A Philosophical Inquiry*, University of Chicago Press, 1987.

Breazeal, C., and L. Aryananda, Recognition of affective communicative intent in robot-directed speech, *Autonomous Robots*, *12*(1), 83–104, 2002.

Breazeal, C., A. Takanishi, and T. Kobayashi, Social robots that interact with people, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 1349 – 1370, Springer, Berlin, 2008.

Breazeal, C. L., *Designing sociable robots*, MIT Press, Cambridge, Mass.; London, 2004.

Brey, P., Values in technology and disclosive computer ethics, in *The Cambridge Handbook of Information and Computer Ethics*, edited by L. Floridi, chap. 3, pp. 41 – 58, Cambridge University Press, 2010.

Brey, P., From moral agents to moral factors: The structural ethics approach, in *Moral Agency and Technical Artefacts*, edited by P. Kroes and P. Verbeek, Springer, 2012.

Brock, O., J. Kuffner, and J. Xiao, Motion for manipulation tasks, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 615 – 646, Springer, Berlin, 2008.

Buber, M., *I and Thou*, Scribner, New York, 1958.

Buss, M., and M. Beetz, Cotesys – cognition for technical systems, *Kunstliche Intelligenz*, 2010.

Butterfield, J., *Collins English Dictionary*, HarperCollins, 2003.

Callon, M., The sociology of an actor-network: The cae of the electric vehicle, in *Mapping the dynamics of science and technology : sociology of science in the real world*, edited by M. Callon, J. Law, and A. Rip, pp. 19 – 24, Macmillan, Basingstoke, 1986.

Campion, G., and W. Chung, Wheeled robots, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 391 – 410, Springer, Berlin, 2008.

Capek, K., and P. Selver, *R.U.R. (Rossum's Universal Robots) : a fantastic melodrama*, – pp., Doubleday, Page, Garden City, N.Y., 1923.

Capurro, R., Ethics and robotics, in *Ethics and robotics*, edited by R. Capurro and M. Nagenborg, pp. 117 – 123, AKA ; IOS Press, Heidelberg; [Amsterdam], 2009.

Chen, T., C.-H. King, A. Thomaz, and C. Kemp, *Touched by a robot*, Proceedings of the 6th international conference / Human-robot interaction (HRI '11), ACM, 2 Penn Plaza, Suite 701, New York, NY 10121-0701, USA, 2011.

Chung, W., L. Fu, and S. Hsu, Motion control, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 133 – 160, Springer, Berlin, 2008.

Coeckelbergh, M., Health care, capabilities, and ai assistive technologies, *Ethical Theory Moral Pract. Ethical Theory and Moral Practice*, *13*(2), 181–190, 2010.

Cooley, M., From judgment to calculation, *AI & Society*, *21*(4), 395 – 409, 2007.

Cooper, C., A. Selwood, M. Blanchard, Z. Walker, R. Blizard, and G. Livingston, Abuse of people with dementia by family carers: Representative cross sectional survey, *British Medical Journal*, *338*(7694), 583–585, 2009.

Correia, M., and D. Waitzberg, The impact of malnutrition on morbidity, mortality, length of hospital stay and costs evaluated through a multivariate model analysis., *Clinical nutrition (Edinburgh, Scotland)*, *22*(3), 235–9, 2003.

Cotin, S., H. Delingette, and N. Ayache, A hybrid elastic model for real-time cutting, deformations, and force feedback for surgery training and simulation, *VISUAL COMPUTER*, *16*, 437–452, 2000.

Cummings, M., Integrating ethics in design through the value-sensitive design approach, *Science and Engineering Ethics*, *12*, 701–715, 10.1007/s11948-006-0065-0, 2006.

Dautenhahn, K., Roles and functions of robots in human society - implications from research in autism therapy, *Robotica*, *21*, 443 – 452, 2003.

Dautenhahn, K., and I. Werry, Towards interactive robots in autism therapy: Background, motivation and challenges, *Pragmatics /& Cognition*, *12*(1), 1–35, doi:doi:10.1075/pc.12.1.03dau, 2004a.

Dautenhahn, K., and I. Werry, Towards interactive robots in autism therapy: Background, motivation and challenges, *Pragmatics & Cognition*, *12*(1), 1 – 35, 2004b.

Dautenhahn, K., S. Woods, C. Kaouri, M. L. Walters, K. L. Koay, and I. Werry, What is a robot companion - friend, assistant or butler?, *International Conference on Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ*, pp. 1192 – 1197, 2005.

Davies, A., and P. Snaith, Mealtime problems in a continuing-care hospital for the elderly., *Age and ageing*, *9*(2), 100–5, 1980.

Decker, M., Caregiving robots and ethical reflection: the perspective of interdisciplinary technology assessment, *AI & Society*, *22*(3), 315–330, 2008.

Dubberly, H., *How do you Design?*, Dubberly Design Process, 2501 Harrison St. 7, San Francisco, CA, 94110.

Engelberger, J. F., *Robotics in service*, MIT Press, Cambridge, Mass., 1989.

Feng, P., and A. Feenberg, Thinking about design: Critical theory of technology and the design process, in *Philosophy and design : from engineering to architecture*, edited by P. E. Vermaas, pp. 105 – 118, Springer, Dordrecht, 2008.

Feron, E., and E. Johnson, Aerial robotics, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 1009 – 1030, Springer, Berlin, 2008.

Floridi, L., and J. Sanders, On the morality of artificial agents, *Minds and Machines*, *14*(3), 349–379, 2004.

Fong, T., I. Nourbakhsh, and K. Dautenhahn, A survey of socially interactive robots, *Robotics and autonomous systems.*, *42*(3), 143, 2003.

Franklin, S., and A. Graesser, Is it an agent, or just a program?: A taxonomy for autonomous agents, *Lecture notes in computer science.*, (1193), 21, 1997.

Fraser, N., *Unruly practices : power, discourse, and gender in contemporary social theory*, University of Minnesota Press, Minneapolis, 1989.

Friedman, B., P. H. K. Jr., and H. Peter, The human-computer interaction handbook, chap. Human values, ethics, and design, pp. 1177–1201, L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 2003.

Friedman, B., P. H. K. Jr., and A. Borning, Human values, ethics, and design, in *Human-Computer Interaction and Management Information Systems: Foundations*, edited by P. Zhang and D. Galletta, Advances in Management Series, chap. 16, pp. 348 – 372, M.E. Sharpe, 2006.

Friedrich, H., S. Mnch, R. Dillmann, S. Bocionek, and M. Sassin, Robot programming by demonstration (rpd): Supporting the induction by human interaction, *Machine Learning*, *23*(2/3), 163 – 189, 1996.

Gadow, S. A., Nurse and patient: The caring relationship, in *Caring, Curing, Coping: Nurse, Physician, and Patient Relationships*, edited by A. Bishop and J. Scudder, chap. 2, pp. 31 – 43, University Of Alabama Press, 2002.

Gill, S., Socio-ethics of interaction with intelligent interactive technologies, *AI & Society*, *22*(3), 283–300, 2008.

Gilligan, C., *In a different voice : psychological theory and women's development*, Harvard University Press, Cambridge, Mass., 1982.

Goetz, J., and S. Kiesler, Cooperation with a robotic assistant, *CHI - CONFERENCE-*, pp. 578–579, 2002.

Hannaford, B., and A. Okamura, Haptics, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 719 – 758, Springer, Berlin, 2008.

Harbers, H., A. Mol, and A. Stollmeyer, Food matters: Arguments for an ethnography of daily care, *Theory, culture & society.*, *19*(5), 207, 2002.

Haselager, W. F. G., Robotics, philosophy and the problems of autonomy, *PRAGMATICS AND COGNITION*, *13*(3), 515–532, 2005.

Hayashi, T., Control method of robot suit hal working as operator's muscle using biological and dynamical information, in *International Conference on Intelligent Robots and Systems*, pp. 3063 – 3068, IEEE, 2005.

Heinzmann, J., and A. Zelinsky, 3-d facial pose and gaze point estimation using a robot real-time tracking paradigm, in *Automatic Face and Gesture Recognition IEEE*, pp. 142 – 147, 1998.

Hirose, S., and H. Yamada, Snake-like robots, *IEEE Robotics and Automation Magazine*, *16*(1), 88 – 98, 2009.

Hofmann, B., Why ethics should be part of health technology assessment., *International journal of technology assessment in health care*, *24*(4), 423–9, doi: 10.1017/S0266462308080550, 2008.

Hosoda, K., T. Takuma, A. Nakamoto, and S. Hayashi, Biped robot design powered by antagonistic pneumatic actuators for multi-modal locomotion, *Rob Autom Syst Robotics and Autonomous Systems*, *56*(1), 46–53, 2008.

Howcroft, D., N. Mitev, and M. Wilson, What we may learn from the social shaping of technology approach, in *Social Theory and Philosophy for Information Systems*, edited by J. Mingers and L. Willcocks, John Wiley Series in Information Systems, chap. 9, pp. 329 – 371, J. Wiley, 2004.

Introna, L., Disclosive ethics and information technology: Disclosing facial recognition systems, *Ethics and Information Technology*, *7*(2), 75–86, 2005.

J. Li, B. A. E., L. Wolf, Use of mechanical patient lifts decreased musculoskeletal symptoms and injuries among health care workers, *Injury Prevention*, *10*(4), 212 – 216, 2004.

Jain, A. K., and S. Z. Li, *Handbook of Face Recognition*, Springer Science+Business Media, Inc., [New York], 2005.

Jelsma, J., Designing 'moralized' products: Theory and practice, in *User Behavior and Technology Development: Shaping Sustainable Relations between Consumers and Technologies*, edited by P. Verbeek and A. Slob, pp. 221 – 231, Springer, 2006.

Kahn, R. E., M. J. Swain, P. N. Prokopowicz, and R. J. Firby, Gesture recognition using the perseus architecture, *Computer Vision and Pattern Recognition 1996 IEEE*, pp. 734 – 741, 1996.

Kajita, S., and B. Espiau, Legged robots, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 361 – 390, Springer, Berlin, 2008.

Kavraki, L., and S. LaValle, Motion planning, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 109 – 132, Springer, Berlin, 2008.

Kawamoto, H., and Y. Sankai, Power assist system hal-3 for gait disorder person, in *Proceedings of the 8th International Conference on Computers Helping People with Special Needs*, ICCHP '02, pp. 196–203, Springer-Verlag, London, UK, UK, 2002.

Kawasaki, H., T. Komatsu, K. Uchiyama, and T. Kurimoto, Dexterous anthropomorphic robot hand with distributed tactile sensor: Gifu hand ii, *IEEE INTERNATIONAL CONFERENCE ON SYSTEMS MAN AND CYBERNETICS*, *2*, II–782–II–787, 1999.

Kemp, C., P. Fitzpatrick, H. Hirukawa, K. Yokoi, K. Harada, and Y. Matsumoto, Humanoids, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 1307 – 1334, Springer, Berlin, 2008.

Kidd, C. D., and C. L. Breazeal, Designing a sociable robot system for weight maintenance, in *Consumer Communications Conference*, pp. 253 – 257, IEEE, 2006.

Kidd, C. D., M. I. of Technology. Dept. of Architecture. Program in Media Arts, and Sciences., Designing for long-term human-robot interaction and application to weight loss, Ph.D. thesis, 2008.

Kim, K. H., S. W. Bang, and S. R. Kim, Emotion recognition system using short-term monitoring of physiological signals, *Medical & biological engineering & computing.*, *42*, 419–427, 2004.

Kiran, A. H., Responsible design. a conceptual look at interdependent design-use dynamics, *Philosophy & Technology*, (1), 2011.

Koggel, C. M., *Perspectives on equality : constructing a relational theory*, Rowman & Littlefield Publishers, Lanham, MD, 1998.

Krapp, K., Activities of daily living evaluation, in *Encyclopedia of Nursing & Allied Health*, Gale Group, Inc., 2002.

Kubrick, S., et al., 2001, a space odyssey, 2001.

Kunze, L., T. Roehm, and M. Beetz, Towards semantic robot description languages, in *IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China, 2011.

Latour, B., Where are the missing masses? the sociology of a few mundate artifacts, in *Shaping Technology // Building Society: Studies in Sociotechnical Change*, edited by W. Bijker and J. Law, chap. 8, pp. 225 – 258, MIT Press, 1992.

Lauwers, T., A dynamically stable single-wheeled mobile robot with inverse mouse-ball drive, in *IEEE International Conference on Robotics and Automation*, pp. 2884 – 2889, 2006.

Le Dantec, C. A., E. S. Poole, and S. P. Wyche, Values as lived experience: evolving value sensitive design in support of value discovery, in *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, pp. 1141–1150, ACM, New York, NY, USA, doi: 10.1145/1518701.1518875, 2009.

Leininger, M., Leininger's theory of nursing: Culture care diversity and universality, *Nursing Science Quarterly*, *2*, 11 – 20, 1988.

Lin, P., K. Abney, and G. Bekey, *Robot Ethics: The Ethical and Social Implications of Robotics*, Intelligent Robotics and Autonomous Agents, Mit Press, 2011.

Little, M. O., Care: From theory to orientation and back, *The Journal of Medicine and Philosophy*, *23*(2), 190–209, 1998.

Lucivero, F., T. Swierstra, and M. Boenink, Assessing expectations: Towards a toolbox for an ethics of emerging technologies, *NanoEthics*, *5*(2), 129–141, 2011.

Lytle, M., Robot care bears for the elderly, 2002.

Manders-Huits, N., What values in design? the challenge of incorporating moral values into design, *Science and Engineering Ethics*, *17*(2), 271–287, 2011.

Maslow, A. H., *Motivation and personality*, Harper & Row, New York, 1970.

Maurer, M., Some ideas on ict as it influences the future, NEC Technology Forum, Tokyo, 2007.

Max-Neef, M., Economic growth and quality of life: a threshold hypothesis, *Ecological Economics*, *15*(2), 115–118, 1995.

Mayer, C., B. Radig, S. Sosnowski, and K. Kuhnlenz, Towards robotic facial mimicry: System development and evaluation, *Proc. IEEE Int. Workshop Robot Human Interact. Commun. Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, pp. 198–203, 2010.

Melchiorri, C., and M. Kaneko, Robot hands, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 345 – 360, Springer, Berlin, 2008.

Metzler, T., and L. Lewis, Ethical views, religious views, and acceptance of robotic applications: A pilot study, *Association for the Advancement of Aartificial Intelligence (www.aaai.org)*, pp. 15 – 22, 2008.

Minato, T., M. Shimada, H. Ishiguro, and S. Itakura, Development of an android robot for studying human-robot interaction, *Lecture notes in computer science.*, (3029), 424–434, 2004.

Minato, T., Y. Yoshikawa, T. Noda, S. Ikemoto, H. Ishiguro, and M. Asada, Cb2: A child robot with biomimetic body for cognitive developmental robotics, in *IEEE-RAS International Conference on Humanoid Robots*, pp. 557 – 562, 2007.

Minguez, J., F. Lamiraux, and J. Laumond, Motion planning and obstacle avoidance, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 827 – 852, Springer, Berlin, 2008.

Mitcham, C., *Encyclopedia of science, technology and ethics*, – pp., Macmillan Reference, Detroit, Mich., 2005.

Mitra, P., and G. Niemeyer, Model-mediated telemanipulation, *The International Journal of Robotics Research*, *27*(2), 253–262, doi: 10.1177/0278364907084590, 2008.

Mol, A., *Care and its Values*, pp. 215 – 234, Transcript Verlag, 2010.

Mol, A., I. Moser, and J. Pols, *Care in practice : on tinkering in clinics, homes and farms*, Transcript ; Distributed in North America by Transaction Publishers, Bielefeld; Piscataway, NJ, 2010.

Moor, J. H., Is ethics computable?, *Metaphilosophy.*, *26*(1-2), 1, 1995.

Moor, J. H., Machine ethics - the nature, importance, and difficulty of machine ethics, *IEEE intelligent systems.*, *21*(4), 18, 2006.

Moravec, H. P., *Robot : mere machine to transcendent mind*, Oxford University Press, New York, 1999.

Mori, M., Bukimi no tani: The uncanny valley, *Energy*, *7*(4), 33 – 35, 1970.

Morin, P., and C. Samson, Motion control of wheeled mobile robots, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 799 – 826, Springer, Berlin, 2008.

Morris, J., et al., Wall-e, 2008.

Mowshowitz, A., Technology as excuse for questionable ethics, *AI & Society*, *22*(3), 271–282, 2008.

Mutlu, B., J. Forlizzi, and P. of the 3rd ACM/IEEE international conference / Human robot interaction (HRI '08), *Robots in organizations*, – pp., ACM, ACM, 2 Penn Plaza, Suite 701, New York, NY 10121-0701, USA, 2008.

Nathan, L. P., B. Friedman, P. Klasnja, S. K. Kane, and J. K. Miller, Envisioning systemic effects on persons and society throughout interactive system design, in *Proceedings of the 7th ACM conference on Designing interactive systems*, DIS '08, pp. 1–10, ACM, New York, NY, USA, doi: 10.1145/1394445.1394446, 2008.

Nedelsky, J., Reconceiving rights and constitutionalism, *JOURNAL OF HUMAN RIGHTS -LONDON-*, *7*(2), 139–173, 2008.

Neven, L., 'but obviously not for me': Robots, laboratories and the defiant identity of elder test users, *Sociol. Health Illn. Sociology of Health and Illness*, *32*(2), 335–347, 2010.

Ng-Thow-Hing, V., K. thorisson, R. K. Sarvadevabhatla, J. Wormer, and T. List, Cognitive map architecture: Facilitation of human-robot interaction in humanoid robots, *IEEE Robotics and Automation Magazine*, *16*(1), 55 – 66, 2009.

Niemeyer, G., C. Preusche, and G. Hirzinger, Telerobotics, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 741 – 758, Springer, Berlin, 2008.

Nissenbaum, H., The cutting edge, *SIGCAS Comput. Soc.*, *28*(1), 38–39, doi: 10.1145/277351.277359, 1998.

Noddings, N., *Caring, a feminine approach to ethics & moral education*, University of California Press, Berkeley, 1984.

Noddings, N., *Starting at home : caring and social policy*, University of California Press, Berkeley, 2002.

Nordmann, A., and A. Rip, Mind the gap revisited, *Nat. Nanotechnol. Nature Nanotechnology*, *4*(5), 273–274, 2009.

Nurses of Ontario, C. o., *The Ethical Framework for Nurses in Ontario : Standards of Practice*, The College of Nurses of Ontario, 1999a.

Nurses of Ontario, C. o., *Standard for the Therapeutic Nurse-client Relationship : for Registered Nurses and Registered Practical Nurses in Ontario : Standards of Practice*, The College of Nurses of Ontario, 1999b.

Nussbaum, M. C., *Women and human development : the capabilities approach*, Cambridge University Press, Cambridge; New York, 2000.

Oh, J., D. hanson, W. Kim, Y. han, J. Kim, and I. Park, Design of android type humanoid robot albert hubo, in *IEEE - RSJ International Conference on Intelligent Robots and Systems*, pp. 1428 – 1433, 2006.

Oosterhof, N., Thinking machines that feel: The role of emotions in artificial intelligence research, Master's thesis, University of Twente, 2005.

Oosterlaken, I., Design for development: A capability approach, *Des. Issues Design Issues*, *25*(4), 91–102, 2009.

Orpwood, R., T. Adlam, N. Evans, and J. Chadd, Evaluation of an assisted-living smart home for someone with dementia, *Journal of Assistive Technologies*, *2*(2), 13 – 21, 2008.

Payne, B. K., and R. Cikovic, Empirical examination of the characteristics, consequences, and causes of elder abuse in nursing homes, *Journal of Elder Abuse & Neglect*, *7*(4), 61–74, 1995.

Pellegrino, E. D., The virtuous physician, and the ethics of medicince, in *Virtue and Medicine: Explorations in the Character of Medicine*, *Philosophy and Medicine*, vol. 1, edited by E. Shelp, D. Reidel Publishing Company, 1985.

Picard, R. W., *Affective computing*, – pp., MIT Press, Cambridge, Mass. [u.a.], 2000.

Pillemer, K., and D. W. Moore, Highlights from a study of abuse of patients in nursing homes, *Journal of Elder Abuse and Neglect*, *2*(1-2), 5–29, 1990.

Pineau, J., M. Montemerlo, M. Pollack, N. Roy, and S. Thrun, Towards robotic assistants in nursing homes: Challenges and results, *Robotics and autonomous systems.*, *42*(3), 271, 2003.

Podnieks, E., *National survey on abuse of the elderly in Canada*, Ryerson Polytechnical Institute, Toronto, Ont., 1990.

Pollack, M., et al., Pearl: A mobile robotic assistant for the elderly, in *AAAI Workshop on Automation as Caregiver: The Role of Intelligent Technology in Elder Care*, 2002.

Pols, A. J., *Good care : enacting a complex ideal in long-term psychiatry*, Trimbos-instituut, Utrecht, 2004.

Prattichizzo, D., and J. Trinkle, Grasping, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 671 – 700, Springer, Berlin, 2008.

Rayman, R., et al., Long-distance robotic telesurgery: a feasibility study for care in remote environments., *The international journal of medical robotics + computer assisted surgery : MRCAS*, *2*(3), 216–24, 2006.

Rayman, R., et al., Robotic telesurgery: a real-world comparison of ground- and satellite-based internet performance., *The international journal of medical robotics + computer assisted surgery : MRCAS*, *3*(2), 111–6, 2007.

Razavi, S., *The political and social economy of care in a development context : conceptual issues, research questions and policy options*, United Nations Research Institute for Social Development, Geneva, 2007.

Reich, W. T., History of the notion of care, in *Encyclopedia of bioethics*, edited by W. T. Reich, pp. 219 – 331, Macmillan Pub. Co. : Simon & Schuster Macmillan ; Prentice Hall International, New York; London, 1995.

Roach, M. S., *The human act of caring : a blueprint for the health professions*, Canadian Hospital Association Press, Ottawa, 1999.

Ruddick, S., *Maternal thinking : toward a politics of peace ; with a new preface*, Beacon Press, Boston, 1995.

Russell, S. J., and P. Norvig, *Artificial intelligence : a modern approach*, – pp., Prentice Hall, Englewood Cliffs, N.J., 1995.

Saenz, A., Incredible tug robots automate delivery in hospitals, *Singularity Hub*, 2010.

Sakamoto, D., T. Kanda, T. Ono, H. Ishiguro, and N. Hagita, Android as a telecommunication medium with a human-like presence, in *Proceedings of the ACM/IEEE international conference on Human-robot interaction*, HRI '07, pp. 193–200, ACM, New York, NY, USA, doi:10.1145/1228716.1228743, 2007.

Sandelowski, M., Exploring the gender-technology relation in nursing., *Nursing inquiry*, *4*(4), 219–28, 1997.

Santoro, M., D. Marino, and G. Tamburrini, Learning robots interacting with humans: from epistemic risk to responsibility, *AI & Society*, *22*(3), 301–314, 2008.

Satoh, H., T. Kawabata, and Y. Sankai, Bathing care assistance with robot suit hal, in *International Conference on Robotics and Biomimetics (ROBIO)*, pp. 498 – 503, IEEE, 2009.

Schoenhofer, S., A framework for caring in a technologically dependent nursing practice environment, in *Advancing Technology, Caring and Nursing*, rozzano loscin ed., pp. 3 – 11, Auburn House, 2001.

Schulz, D., W. Burgard, D. Fox, and A. B. Cremers, People tracking with mobile robots using sample-based joint probabilistic data association filters, *The International Journal of Robotics Research*, *22*(2), 99–116, 2003.

Sen, A., *Commodities and capabilities*, North-Holland ; Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co., Amsterdam; New York; New York, N.Y., U.S.A., 1985.

Sharkey, A., and N. Sharkey, Granny and the robots: ethical issues in robot care for the elderly, *Ethics and Information Technology*, *14*(1), 27–40, 2012.

Sharkey, N., and A. Sharkey, The rights and wrongs of robot care, in *Robot Ethics: The Ethical and Social Implications of Robotics*, edited by P. Lin, K. Abney, and G. Bekey, Intelligent Robotics and Autonomous Agents, chap. 17, pp. 267 – 282, Mit Press, 2011.

Shaw-Garlock, G., Looking forward to sociable robots, *International Journal of Social Robotics*, *1*(3), 249–260, 2009.

Shieh, M., C. Lu, C. Chen, C. Chuang, and Y. Lai, Design and implementation of an interactive nurse robot, in *Annual Conference on Instrumentation, Control, Information Technology and System Integration (SICE)*, pp. 2121 – 2125, 2007.

Siciliano, B., and O. Khatib, *Springer handbook of robotics*, – pp., Springer, Berlin, 2008.

Silverstone, R., E. Hirsch, and D. Morley, Information and communication technologies and the moral economy of the household, in *Consuming technologies: media and information in domestic spaces*, edited by R. Silverstone and E. Hirsch, pp. 15 – 31, Routledge, 1992.

Singer, P. W., *Wired for war : the robotics revolution and conflict in the twenty-first century*, – pp., Penguin Press, New York, 2009.

Smits, R., J. Leyten, and P. den Hertog, Technology assessment and technology policy in europe: New concepts, new goals, new infrastructures, *Policy Sciences*, *28*(3), 271–299, 1995.

Soraker, J., When does the mind matter? the strengths and limitations of the informational level of abstraction, *Ethics and Information Technology*, Forthcoming.

Sorensen, K., Domestication: The enactment of technology, in *Domestication of Media and Technology*, edited by T. B. et al., pp. 40 –61, Open University Press, 2006.

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

Stabell, A., H. Eide, G. A. Solheim, K. N. Solberg, and T. Rusten, Nursing older people: Nursing home residents’ dependence and independence, *Journal of Clinical Nursing*, *13* (6), 677–686, 2004.

Super, D. E., *Work values inventory*, Houghton Mifflin, Boston, 1968.

Swierstra, T., and A. Rip, Nano-ethics as nest-ethics: Patterns of moral argumentation about new and emerging science and technology, *NanoEthics*, *1* (1), 3–20, 2007.

Tamburrini, G., Robot ethics: A view from the philosophy of science, in *Ethics and robotics*, edited by R. Capurro and M. Nagenborg, pp. 11 – 22, AKA ; IOS Press, Heidelberg; [Amsterdam], 2009.

Tamura, T., Y. S, I. A, O. D, K. A, H. Y, F. T, and N. K, Is an entertainment robot useful in the care of elderly people with severe dementia?, *The journals of gerontology. Series A, Biological sciences and medical sciences*, *59* (1), 83–5, 2004.

Tenorth, M., Knowledge processing for autonomous robots, Ph.D. thesis, Technische Universität München, 2011.

Tenorth, M., and M. Beetz, Knowledge processing for autonomous robot control, in *AAAI Spring Symposium on Designing Intelligent Robots: Reintegrating AI*, Stanford, CA, USA, 2012.

Tenorth, M., D. Jain, and M. Beetz, Knowledge representation for cognitive robots, *Kunstliche Intelligenz*, *24* (3), 233–240, 2010.

Thaler, R. H., and C. R. Sunstein, *Nudge : improving decisions about health, wealth, and happiness*, Yale University Press, New Haven, Conn., 2008.

Thrun, S., Toward a framework for human-robot interaction, *Human-Computer Interaction*, *19* (1), 9–24, 2004.

Thrun, S., J. Schulte, and C. Rosenberg, Interaction with mobile robots in public places, *IEE Intelligent Systems*, pp. 7 – 11, 2000.

Torrance, S., Ethics and consciousness in artifical agents, *AI & Society*, *22*(4), 495 – 521, 2008.

Tronto, J., Creating caring institutions: Politics, plurality, and purpose, *ETHICS AND SOCIAL WELFARE*, *4*(2), 158–171, 2010.

Tronto, J. C., *Moral boundaries : a political argument for an ethic of care*, Routledge, New York, 1993.

Turing, A. M., Computing machinery and intelligence, *Mind : a quarterly review of psychology and philosophy.*, *LIX*(236), 433–, 1950.

Turkle, S., *Alone Together: Why We Expect More from Technology and Less from Each Other*, Basic Books, 2011.

Vallor, S., Carebots and caregivers: Sustaining the ethical ideal of care in the twenty-first century, *Philos. Technol. Philosophy and Technology*, *24*(3), 251–268, 2011.

van den Hoven, J., Ict and value sensitive design, *INTERNATIONAL FEDERATION FOR INFORMATION PROCESSING -PUBLICATIONS- IFIP*, (233), 67–72, 2007.

van der Loos, H., and D. Reinkensmeyer, Rehabilitation and health care robotics, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 1223 – 1252, Springer, Berlin, 2008.

van der Plas A., S. M., and W. C., Beyond speculative robot ethics: A vision assessment study on the future of the robotic caretaker, *Account. Res. Accountability in Research*, *17*(6), 299–315, 2010.

Van Gorp, A., and I. Van de Poel, Deciding on ethical issues in engineering design, in *Philosophy and design : from engineering to architecture*, edited by P. E. Vermaas, pp. 77 – 90, Springer, Dordrecht, 2008.

Van Koughnett, J., S. Jayaraman, R. Eagleson, D. Quan, A. van Wynsberghe, and C. Schlachta, Are there advantages to robotic-assisted surgery over laparoscopy from the surgeon's perspective?, *Journal of Robotic Surgery*, *3*, 79–82, 10.1007/s11701-009-0144-8, 2009.

van Wynsberghe, A., Designing robots for care: Care centered value-sensitive design, *Sci Eng Ethics Science and Engineering Ethics*, (4), 2012.

van Wynsberghe, A., and C. Gastmans, Telesurgery: An ethical appraisal, *Journal of Medical Ethics*, *34* (10), 2008.

van Wynsberghe A, and G. C, Telepsychiatry and the meaning of in-person contact: a preliminary ethical appraisal., *Medicine, health care, and philosophy*, *12* (4), 469–76, 2009.

Vanlaere, L., and C. Gastmans, A personalist approach to care ethics, *Nurs. Ethics Nursing Ethics*, *18* (2), 161–173, 2011.

Verbeek, P., Morality in design; design ethics and the morality of technological artifacts, in *Philosophy and design : from engineering to architecture*, edited by P. E. Vermaas, pp. 91 – 102, Springer, Dordrecht, 2008.

Verbeek, P.-P., *What things do : Philosophical reflections on technology, agency, and design*, Pennsylvania State University Press, University Park, Pa., 2005.

Verbeek, P.-P., Materializing morality, *Science, Technology, & Human Values*, *31* (3), 361–380, 2006.

Verbeek, P.-P., *Moralizing technology : understanding and designing the morality of things*, The University of Chicago Press, Chicago; London, 2011.

Verkerk, M., The care perspective and autonomy., *Medicine, health care, and philosophy*, *4* (3), 289–94, 2001.

Verkerk, M. A., J. J. V. Busschbach, and E. D. Karssing, Health-related quality of life research and the capability approach of amartya sen, *Quality of Life Research*, *10* (1), 49–55, 2001.

Veruggio, G., and F. Operto, Roboethics: a bottom-up interdisciplinary discourse in the field of applied ethics in robotics, *International Review of Information Ethics*, *6*, 3 – 8, 2006.

Veruggio, G., and F. Operto, Roboethics: Social and ethical implications of robotics, in *Springer handbook of robotics*, edited by B. Siciliano and O. Khatib, pp. 1499 – 1524, Springer, Berlin, 2008.

Vongsoasup, V., and M. Mataric, Path planning and navigation for geeves: A tour-guide/greeter robot.

Wada, K., T. Shibata, T. Saito, K. Sakamoto, and K. Tanie, Psychological and social effects of one year robot assisted activity on elderly people at a health service facility for the aged, in *IEEE International Conference on Robotics and Automation*, pp. 2785 – 2790, 2005.

Wallach, W., Robot minds and human ethics: The need for a comprehensive model of moral decision making, *Ethics Inf. Technol. Ethics and Information Technology*, *12*(3), 243–250, 2010.

Wallach, W., and C. Allen, *Moral machines : teaching robots right from wrong*, – pp., Oxford University Press, New York; Oxford, 2010.

Wallach, W., C. Allen, and I. Smit, Machine morality: bottom-up and top-down approaches for modelling human moral faculties, *AI & Society*, *22*(4), 565–582, 2008.

Wallach, W., S. Franklin, and C. Allen, A conceptual and computational model of moral decision making in human and artificial agents, *Top. Cognitive Sci. Topics in Cognitive Science*, *2*(3), 454–485, 2010.

Walters, M., K. Dautenhahn, S. Woods, and K. Koay, Robotic etiquette: Results from user studies involving a fetch and carry task, in *Proceedings from Human Robot Interaction Conference*, 2007.

Weaver, K., J. Morse, and C. Mitcham, Ethical sensitivity in professional practice: concept analysis, *Journal of Advanced Nursing*, *62*(5), 607–618, doi: 10.1111/j.1365-2648.2008.04625.x, 2008.

Widdershoven, G., Technology and care, from opposition to integration, in *Between technology and humanity : the impact of technology on health care ethics*, edited by C. Gastmans, pp. 35 – 48, Leuven University Press, Leuven, 2002.

Wilson, M., Making nursing visible? gender, technology and the care plan as script, *Information Technology & People*, *15*(2), 139–158, 2002.

Wong, P.-H., Technology, recommendation and design: On being a paternalistic philosopher, *Sci Eng Ethics Science and Engineering Ethics*, (5), –, 2011.

Wood, R., Fly, robot fly, *IEEE Spectrum IEEE Spectrum*, *45*(3), 25–29, 2008.

Wright, L., M. Hickson, and G. Frost, Eating together is important: using a dining room in an acute elderly medical ward increases energy intake, *Journal of Human Nutrition and Dietetics*, *19*(1), 23–26, 2006.

Yoshiro, U., O. Shinichi, T. Yosuke, F. Junichi, I. Tooru, N. Thoshihro, S. Tsuyoshi, and O. Junichi, Childcare robot papero is designed to play with and watch over children at nursery, kindergarten, school and at home, *Development of Childcare Robot PaPeRo*, pp. 1 – 11, 2005.

Zaeh, M. F., et al., Artificial cognition in production systems, *IEEE Transactions on Automation Science and Engineering*, *7*(3), 1–27, 2010.

Zhang, T., B. Zhu, L. Lee, and D. Kaber, Service robot anthropomorphism and interface design for emotion in human-robot interaction, in *International Conference on Automation Science and Engineering*, pp. 674 – 679, IEEE, 2008.

Summary

This book begins by highlighting the range of positive and negative aspects associated with the initiative to design and use care robots. The positive aspects vary from; the reality that care is required 24/7 and a human care-giver cannot possibly meet such a standard, the reality that patients are abused when in the care of others, the reality that current practices fail to meet the ideal standard of care and the foreseen lack of healthcare personnel and resources to meet ever increasing care needs of societies. When the actuality of a care robot is discussed, however, many relevant ethical concerns are introduced. What will the impact of the robot be on the provision of good care, on the manifestation of care values; will it change the standards of care and ultimately lower them? Will care robots displace and/or de-skill care workers? What will the existential impact be on the care-giver and the care-receiver; is the use of robots a de-valuing of either or both? Given the strength of both arguments, however, it is not possible to conclude firmly on either side of the debate. It is true that when we describe care practices and think of a robot being integrated, it looks better to have a human present. This does not, however, mean the human will provide care according to the ideal.

Many scholars to date have discussed the positive and negative ethical issues associated with the use of care robots and some have even suggested the need for a framework to evaluate robots in general [*Asaro*, 2006] and care robots in particular [*Sharkey and Sharkey*, 2012; *Turkle*, 2011]; however, none have presented such a framework. The work of this thesis is to take up the challenge of creating a framework not only for the ethical evaluation of current care robot prototypes but to create a framework that can be used to steer the design process of future care robots. As such, the research question guiding this work is as follows: how can care robots used in care practices be designed and implemented in a way that supports and promotes the fundamental values in care.

To answer this question, I create a framework to be integrated in the design

process of care robots. The framework can be used for retrospective evaluations of current care robots and can also be used to steer the entire prospective design process and implementation of future care robots. I label the framework, the "Care-Centered" (CC) framework given the pivotal role the care ethics tradition plays in both its creation and methodologies for use. The first seven chapters are intended to indicate: 1. how the CC framework is created, 2. the conceptual, theoretical and practical foundation for substantiating and justifying the framework, and 3. how the framework is intended to be used in retrospective evaluations of current care robots. To accomplish this, I fulfil the following steps: describe the manner in which the framework is created using the blueprint of Value-Sensitive Design (chapter 1); complete a conceptual investigation into what a value is and how it comes to be embedded in a technology (chapter 2); complete a conceptual investigation into the concept of care and the care ethics perspective (chapter 3); and, complete a technical investigation of current care robots and robot capabilities (chapter 4). Chapter 4 is also meant to illustrate the impossibility of translating care values into technical capabilities of a care robot independent of any consideration for context, practice and/or actors. Understanding the justification for the components of the framework and the manner in which it will be used (chapter 5) it is possible to label the framework accordingly; the Care-Centered framework (CC). With the framework proposed and all of its components justified, chapters 6 and 7 present two different care practices to show the manner in which the framework is used retrospectively but also as a way of highlighting once again the necessity of evaluating and designing robots on a design-by-design, or practice-by-practice basis. This is counter to the claim made by Asaro that robots should not be designed on this basis, that instead the design of a robot ought to reflect the general trends of human decision making/reasoning.

Following these retrospective evaluations, chapters 8 and 9 begin the prospective portion of this work. Chapter 8 explicitly addresses the moral status of the robot. Roboticists and robot ethicists are grappling with the question of how to program a kind of morality into the robot such that it is capable of moral decision-making and reasoning on its own. This is so given the morally charged contexts within which care robots will be placed and the roles which they will be assigned. However, what such initiatives fail to address is the assumption as to why we (society) would want to endow the robot with such capabilities, why robots ought to be delegated such roles and moreover, what such a delegation means in terms of the robots moral status. In revealing the relationship between the robots capabilities and the inscribed role based on such capabilities, it is also made clear that with this role comes an associated responsibility. Thus,

the design of the care robot understood as the capabilities of the care robot, dictate the re-distribution of roles within the network. But, as the discussion in chapter 8 reveals, the care robot is not capable of claiming responsibility related to an outcome or behaviour of its action(s). With this in mind, the care robot can never be delegated a role to which moral responsibility is assigned; a role traditionally delegated to a human moral agent. The robot, as a result, must be designed according to the assumption of its moral impact as that of a moral factor. Concluding that the robot ought never be delegated the role and responsibility attributed to a full moral agent, but rather to a moral factor (operational morality, an implicit moral agent), presents clear guidelines for the design of care robots in terms of their capabilities and the limits bounding their potential roles. These limits and boundaries are dictated according to the CC framework.

Chapter 9, the apex of this work, is intended to show how all of this information is amalgamated in order to be used prospectively in both the development (the design process and resulting design) and the implementation of future care robots. The user manual for the CCVSD approach mirrors that of the retrospective evaluations but takes additional criteria into consideration. To begin, the fears associated with the initiative, design and development of care robots is used to mark the starting point in the design process. Following this, based on the ethicists experience of observing care practices in context, novel robot uses are proposed. To prospectively analyse or evaluate robots for these uses the care practice is then described in detail by articulating the actors, the distribution of roles and responsibilities among actors, the manifestation of values throughout the practice and where certain values may have been overlooked or sacrificed. With this in mind, I then postulate a potential design infrastructure for the robot according to the mechanical description of the practice. Added to this, the proposed design is then criticized for its potential to impact the overall care process of a patient as well as the establishment and maintenance of the therapeutic relationship.

In consideration of the design requirements, special attention is paid to the moral status of the robot and how the delegation of one capability or another shifts the robots moral impact. The care robot ought always to be designed according to the assumption of its moral status as that of a moral factor (as was the conclusion of chapter 8). If the robot is granted the capability to monitor the performance of a human care-giver, such a role attributes the final responsibility to the robot. As we have seen, it is not possible for the robot to take responsibility when understood in terms of responsibility requiring that one be punished or praised for its actions. There are two ways in which this is problem-

atic. The first has to do with the distinction in responsibility and accountability – the care robot will of course be accountable for its actions but cannot take final responsibility (or be liable). If the care robot is delegated a role for which it bears final responsibility then where is responsibility, and as such liability, placed? The second problem has to do with the necessary requirements of good care and the attribute of responsible care providers. The professionalization of medicine and nursing relies on the basis that care providers claim responsibility for their roles. This translates into the demand that they must be skilful, and thus properly trained. They must engage in a professional/therapeutic relationship with patients which demands sensitivity to the asymmetry in power within this relationship . Consequently, it is neither possible nor advisable to delegate the care robot a role for which it bears final responsibility, but rather the robot is delegated a role in which it bears a minimum responsibility. Such a minimum means the robots role is to enhance the manifestation of moral elements within a care practice. Whats more, the robots capabilities might also act as a platform for re-integrating certain care values (within the conceptualization of the moral elements) that may have been lost in previous introductions of healthcare technologies (the lifting robots for example). The robot thus displays a commitment to maintaining the stewardship of the moral elements. The urine testing robot ("the wee-bot) and the waste removal robot (the "roaming toilet), are illustrations of what the design process of a care robot will look like when using the CCVSD approach as the foundation for the design process.

Central to the creation and methodology of the CCVSD approach as well as the resulting evaluations using the framework, is actor-network theory (ANT) and its associated theory of scripts and domestication studies. Script theory is used as a means of uncovering potential hidden assumptions that find their way in the design of current care robots. Assumptions pertaining to what an ideal care practice looks like. ANT is used as a means for describing the manifestation of values as the result of the actions and interactions among and between actors in a network. Whats more, that actors in a network referred to both the human and the non-human actors (technological artefacts as well as the material environment) and played a role not only in the manifestation of values but of norms and rules as well. Thus, all actors in the network of a practice contribute to the forming of a morality of the practice. With this crucial insight in mind, we must then consider the implementation of the robot as it will now be understood to contribute to the re-establishing of a morality or the establishment of a new morality (in terms of new values and/or norms created). Integrating the framework in the implementation of the robot presupposes that the network of actors, in a specific context, into which the robot is stepping play an integral

role in the construction of patterns of use and the co-production of meaning and identity in relation to the robot. One might also conclude that the insights into the implementation portion might also be used for current care robots which are about to be introduced for the first time in certain hospitals (including surgical robots).

Using the CCVSD approach in the implementation of the robot we can see how domestication studies, while hugely insightful as to the mechanisms that drive the formation of a morality (through the pervasive use of a technology) are also limited in that they study the forming of a morality without steering any such formation. Counter to this, design studies which account for the domestication of an artefact, fail to implement the technology in a way that ensures the intended use. Added to this, technology assessments which aim to create guidelines and policies for the initial introduction and continued use of a technology fail to incorporate an adequate ethical analysis to guide such an introduction. If we take into account the claims made by Tronto regarding a good care institution along with the structural ethics approach, the robots role must be thought of in terms of its impact on the micro level (the care practice) and the macro level (the care institution). To balance all of this, using the CCVSD approach in the design, development and implementation of a care robot is shown to ensure an adequate ethical analysis throughout the design process of the robot as well as steering the introduction of the care robot in a manner that reflects the intentions for the envisaged use of the robot. By making the initial roles and responsibilities of the human and non-human actors explicit as well as indicating how these roles and responsibilities will shift with the introduction of the robot, care-givers are given the chance to understand the initiatives and assumptions guiding the design and to understand how their role as care-giver remains in tact and respected. Thus, not only does the robot offer the opportunity for enhancing the provision of values through its design but through its implementation as well. The foundation of care rests on a valuation of both care-giver and care-receiver. By giving voice to the care-givers, their significant role is given value and meaning through this practice of empowerment. In other words, the kind of reflection insisted upon by Turkle and the kind of regulation insisted upon by robot scholar Peter Singer are achieved through the CCVSD approach.

In short, the CCVSD approach reflects a commitment to understanding the starting point of care, being critical of care as it is currently practised, understanding the necessary elements of care and how they contribute to the provision of good care, and above all else how care might change (for better or for worse) with the introduction of a care robot. With these commitments in mind, the

future of care with or without the inclusion of a care robot is investigated: the values in care are understood in-depth and preserved. The non-human actors are recognized for their role in forming a morality and meaning within a network. Most importantly, the human care-givers are recognized for the value of their role, and the responsibility of their actions as the stewards of care. Added to this, the CCVSD approach also reflects a commitment to the design and implementation of future care robots that systematically accounts for the valuation of care practices and care workers. Seen through this lens, care robots hold the potential to provide a vital role in aiding the promotion of care values when designed and implemented according to the manner dictated by the CCVSD approach.

Samenvatting

Dit boek begint met het duiden van een reeks van positieve en negatieve aspecten die gepaard gaan met het ontwerpen en het gebruik van zorgrobots. De positieve aspecten variëren van het feit dat zorg 24/7 nodig is en dat een menselijke zorgverlener onmogelijk kan voldoen aan een dergelijke standaard, het feit dat patiënten soms worden mishandeld door hun zorgverleners, het feit dat de zorg die nu geboden wordt niet voldoet aan de ideale standaard van zorg en het verwachte gebrek aan medisch personeel en de middelen die nodig zijn om tegemoet te komen aan de steeds groter wordende behoefte aan zorg in de samenleving. De suggestie dat zorgrobots een goede oplossing zouden bieden voor voornoemde problemen, stuit echter op tal van relevante ethische bezwaren. Hoe zal de robot het verlenen van goede zorg en de morele waarden die daarmee samenhangen beïnvloeden; zal die de zorgstandaarden veranderen en uiteindelijk verlagen? Zullen zorgrobots hulpverleners gaan vervangen en/of een verlies van deskundigheid veroorzaken? Wat zal de existentiële invloed op de zorgverlener en de zorgontvanger zijn; laat het gebruik van robots hen in hun waarde? Gezien de kracht van de argumenten aan beider zijden is het niet mogelijk om tot een eenduidige conclusie te komen ten voordele van een specifieke positie in deze discussie. Wanneer we nadenken over het integreren van robots in zorgpraktijken, lijkt het beter om daar ook mensen bij te betrekken. Dit betekent echter niet dat dat de menselijke hulpverlener altijd ideale zorg biedt.

De positieve en negatieve ethische kwesties die gepaard gaan met het gebruik van zorgrobots zijn veelvuldig besproken in de wetenschappelijke literatuur. Sommige wetenschappers suggereren dat er behoefte bestaat aan een kader waarbinnen het gebruik van robots in het algemeen beoordeeld kan worden [Asaro, 2006] en het gebruik van zorgrobots in het bijzonder [Sharkey en Sharkey, 2012; Turkle, 2011], maar tot dusver is zon kader nog niet opgesteld. Dit proefschrift heeft tot doel om een kader op te stellen dat niet alleen geschikt is voor de ethische beoordeling van de huidige generatie zorgrobot-prototypes, maar

dat ook inzetbaar is om het ontwerp van toekomstige zorgrobots te sturen. De onderzoeksvraag die ten grondslag ligt aan dit proefschrift is aldus de volgende: hoe kunnen zorgrobots die worden ingezet in zorgpraktijken worden ontworpen en gebouwd op een wijze die de fundamentele waarden die samenhangen met zorg ondersteunt en bevordert.

Om deze vraag te beantwoorden, creëer ik een kader dat bedoeld is om gebruikt te worden in het ontwerpproces van zorgrobots. Dit kader kan worden gebruikt voor de retrospectieve beoordeling van de huidige generatie zorgrobots, maar het kan ook worden ingezet om het gehele ontwerp- en implementatie proces van toekomstige zorgrobots te sturen. Ik noem dit kader het Care-Centered (lett. zorg-gecentreerd) oftewel CC kader, vanwege de prominente positie die de zorgethiek inneemt in zowel het opstellen van het kader als in de methoden voor het gebruik van het kader. De eerste zeven hoofdstukken hebben tot doel aan te geven: 1. hoe het CC kader is opgebouwd, 2. wat de conceptuele, theoretische en praktische fundamenten van het kader zijn, en 3. hoe dit kader gebruikt moet worden in de retrospectieve beoordeling van de huidige generatie zorgrobots. Om dit te bewerkstelligen, neem ik de volgende stappen: een beschrijving van de manier waarop het kader is opgebouwd met behulp van het concept van Value-Sensitive Design (lett. waarde-gevoelig ontwerp) oftewel VSD (hoofdstuk 1); het vervolledigen van een conceptueel onderzoek naar wat waarden zijn en hoe deze kunnen worden geïntegreerd in een technologie (hoofdstuk 2); het vervolledigen van een conceptueel onderzoek naar het concept van zorg en het zorgethisch perspectief (hoofdstuk 3); en een technisch onderzoek naar de huidige generatie zorgrobots en hun capaciteiten (hoofdstuk 4). Hoofdstuk 4 heeft ook tot doel te laten zien dat het onmogelijk is om waarden die samenhangen met zorg te vertalen in technische capaciteiten van een zorgrobot als niet wordt gekeken naar de context en praktijk waarbinnen die wordt ingezet en/of de actoren die daarbij betrokken zijn. Wanneer duidelijk is gemaakt uit welke componenten het kader bestaat en hoe het gebruikt kan worden (hoofdstuk 5), is het mogelijk het kader de naam Care-Centered (CC) te geven. De hoofdstukken 6 en 7 laten vervolgens aan de hand van twee voorbeelden zien hoe het kader retrospectief gebruikt kan worden en benadrukken dat robots per specifieke situatie beoordeeld en ontworpen moeten worden. Dit gaat in tegen het idee van Asaro dat robots zo ontworpen zouden moeten worden dat ze de algemene kenmerken van het menselijk redeneren en menselijke beslissingen reflecteren.

Na deze retrospectieve beschouwingen vormen de hoofdstukken 8 en 9 het prospectieve gedeelte van dit werk. Hoofdstuk 8 gaat specifiek over de morele status van de robot. Technici en ethici worstelen met de vraag hoe een ro-

bot zo geprogrammeerd kan worden dat deze in staat is zelfstandig moreel te redeneren en morele beslissingen te nemen. Deze vraag is van belang, omdat de omgeving waarin de robot zal functioneren en de taken die hij uit zal voeren moreel geladen zijn. Men zou zich echter ook af moeten vragen waarom wij als maatschappij zouden willen dat robots hiertoe in staat zijn, waarom wij robots zulke taken willen geven en wat het toekennen van die taken betekent voor hun morele status. Duidelijk wordt dat er een verband bestaat tussen de capaciteiten van de robot en de taken die deze uitvoert. Ook wordt duidelijk dat de robot met de taak tevens een zekere verantwoordelijkheid krijgt. Dus indien een robot ontworpen wordt die bepaalde capaciteiten heeft, leidt dit ook tot een herverdeling van taken. Echter, in dit hoofdstuk wordt onthuld dat de robot geen verantwoordelijkheid kan nemen voor diens gedragingen of de gevolgen daarvan. De robot mag daarom nooit taken toebedeeld krijgen die een morele verantwoordelijkheid met zich meebrengen; taken die van oudsher aan mensen als morele actoren zijn overgelaten. De robot moet dus ontworpen worden met het idee in het achterhoofd dat het een morele factor (een impliciete morele actor) en geen (expliciete) morele actor is. Dit is een duidelijke richtlijn die voortvloeit uit het CC kader.

Het laatste hoofdstuk, hoofdstuk 9, heeft tot doel te laten zien hoe al deze informatie in de toekomst gebruikt kan worden in het ontwerpproces en het gebruik van zorgrobots. In dit hoofdstuk wordt een prospectief CC kader geboden dat ik het CC VSD kader zal noemen. Het CCVSD kader lijkt op het CC kader, maar er is een aantal aanvullende criteria. Om mee te beginnen, worden de angsten die het ontwerp en gebruik van zorgrobots oproepen als uitgangspunt genomen in het ontwerpproces. Vervolgens wordt vastgesteld voor welke nieuwe taken zorgrobots zouden kunnen worden ingezet. Deze vaststellingen zijn gebaseerd op de waarnemingen van ethici die zorgpraktijken in hun context hebben bestudeerd. Om vervolgens de geschiktheid van robots voor deze taken te kunnen analyseren en beoordelen, worden de zorgpraktijken waarbinnen deze zullen moeten worden uitgevoerd tot in detail beschreven. Het accent ligt daarbij op de actoren, de waarden die met de praktijk gemoeid zijn en waar bepaalde waarden eventueel over het hoofd zijn gezien of opgeofferd. Ik stel dan een potentiële ontwerp infrastructuur voor, die is gebaseerd op de gedetailleerde beschrijving van de praktijk. Daarbij wordt het voorgestelde ontwerp kritisch beoordeeld op zijn potentiële impact op het zorgproces van de patiënt en op het ontstaan en verloop van de therapeutische relatie.

Bij het bepalen van de ontwerpeisen wordt specifiek gekeken naar de morele status van de robot en hoe de capaciteiten die de robot heeft diens morele impact beïnvloeden. De zorgrobot moet altijd zo ontworpen worden dat deze een mo-

rele factor is en geen morele actor (dit was de conclusie uit hoofdstuk 8). Indien de robot in staat zou worden gesteld het functioneren van de menselijke zorgverlener te controleren, zou de eindverantwoordelijkheid bij de robot komen te liggen. Het is niet mogelijk een robot deze verantwoordelijkheid te geven, omdat een robot niet beloond of gestraft kan worden voor diens daden en dat is vereist voor verantwoordelijkheid. Dit problematisch vanwege twee redenen. Allereerst bestaat er een verschil tussen verantwoordelijkheid en aansprakelijkheid de zorgrobot zal aansprakelijk zijn voor diens daden, maar kan daarvoor niet verantwoordelijk gehouden worden, ook niet in rechte. Indien een robot een taak krijgt toebedeeld waarvoor deze eindverantwoordelijkheid draagt, rijst dus de vraag wie (in rechte) verantwoordelijk gehouden kan worden. Het tweede probleem heeft te maken met de noodzakelijke vereisten voor goede zorg en het kenmerk van verantwoordelijke zorgverleners. De professionalisering van de genees- en verpleegkunde berust op het idee dat zorgverleners verantwoordelijk zijn voor de taken die zij uitvoeren. Dit idee vertaalt zich in de vereisten dat zij bekwaam en goed geschoold moeten zijn. Zij moeten in staat zijn een professionele / therapeutische relatie met hun patiënten te onderhouden en zich bewust zijn van de machtsongelijkheid in deze relatie. Daarom is het onmogelijk en onwenselijk om de zorgrobot taken te geven waarvoor deze eindverantwoordelijkheid draagt en is het beter de robot taken te geven waarvoor deze minimale verantwoordelijkheid draagt. Minimale verantwoordelijkheid betekent dat de robot morele elementen in een zorgpraktijk naar voren moet kunnen laten komen. Daarbij kan het zo zijn dat de capaciteiten van de robot bepaalde morele waarden kunnen herintroduceren die van belang zijn voor zorg, maar verloren waren gegaan door de introductie van nieuwe technologieën (zoals de tillift). De robot moet dus in staat zijn de morele elementen van zorgpraktijken te bewaken. De voorbeelden van de urinetest robot (de wee-bot) en de robot die afval opruimt (de roaming toilet) laten zien hoe het ontwerpproces van zorgrobots eruit ziet wanneer de CCVSD-benadering gebruikt wordt als uitgangspunt.

De actor-netwerktheorie (ANT) en de daarmee samenhangde theorie van script- en domesticatiestudies staan centraal in de opzet en methodologie van de CCVSD-benadering en de daaruit voortvloeiende beoordelingen. Scripttheorie wordt gebruikt als een manier om potentiële verborgen aannames te ontdekken die terecht zijn gekomen in het ontwerp van hedendaagse zorgrobots. Het gaat om aannames die verband houden met hoe de ideale zorgpraktijk er uitziet. ANT wordt gebruikt als een manier om waarden te beschrijven die ontstaan als resultaat van acties en interacties tussen actoren in een netwerk. Bovendien verwezen deze actoren in een netwerk naar zowel de menselijke als de niet-menselijke actoren (technologische artefacten en de materiële omgeving) en speelden niet

alleen een rol in het ontstaan van waarden, maar ook van normen en regels. Op die manier dragen alle actoren in het netwerk van een praktijk bij aan de vorming van een praktijkmoraal. Dat cruciale inzicht indachtig moeten we nadenken over de invoering van de robot, omdat we nu weten dat deze bijdraagt aan het herinvoeren van een moraal of het tot stand brengen van een nieuwe moraal (in de betekenis van nieuwe waarden en/of ontstane normen). Het integreren van het kader in de ingebruikneming van de robot vooronderstelt dat het netwerk van actoren, in de specifieke context waarin de robot staat, een belangrijke rol speelt bij het maken van gebruikspatronen en het geven van betekenis en identiteit aan de robot. Je zou ook kunnen concluderen dat de inzichten uit het implementatiedeel tevens gebruikt zouden kunnen worden bij hedendaagse zorgrobots die op het punt staan voor het eerst in bepaalde ziekenhuizen gebruikt te worden (incluis operatierobots).

Door de CCVSD-benadering te gebruiken bij het invoeren van de robot zien we dat domestiecatiestudies, hoewel ze nuttig zijn om inzicht te verkrijgen in de mechanismen die de vorming van een moraal sturen (door het volop gebruiken van een technologie), beperkt zijn, omdat daarmee de vorming van een moraal bestudeerd wordt zonder deze te sturen. Ontwerpstudies die de domesticatie van een artefact verklaren zijn, daarentegen, niet in staat een technologie te implementeren op een manier die het bedoelde gebruik ervan waarborgt. Daarbij komt nog eens bij dat het niet lukt om met technologie-beoordelingen, die richtlijnen en beleid pogen te maken, een toereikende ethische analyse te ontwerpen om de introductie van robots te sturen. Als we kijken naar de beweringen van Tronto over een goede zorginstellingen en rekening houden met de structureelethische methode dan moet de rol van robots gezien worden in termen van hun invloed op microniveau (de zorgpraktijk) en het macroniveau (de zorginstellingen). Om dat allemaal in balans te brengen, blijkt de CCVSD-benadering bij het ontwerpen, ontwikkelen en ingebruiknemen van een zorgrobot een toereikende ehtische analyse te garanderen gedurende het ontwerpproces van de robot en ook de introductie van de zorgrobot te sturen in de richting van het gebruik van de robot op de vooraf bedoelde manier. Door duidelijk te maken wat de eerste verantwoordelijkheden van zowel menselijke als niet-menselijke actoren zijn en hoe die verantwoordelijkheden zouden kunnen verschuiven door de introductie van de robot, worden de intitiatieven en aannames die het onwerpproces sturen voor de zorgverleners inzichtelijk gemaakt en wordt tevens inzichtelijk gemaakt hoe hun rol als zorgverlener intact en gerespecteerd blijft. Zodoende biedt niet alleen het ontwerp, maar ook de implementatie van de robot de kans om waarden te versterken. De basis van zorg is de waardering van zowel de zorgverlener al de zorgontvanger. Door zorgverleners een stem te geven, krijgt hun belangrijke rol

waarde en betekenis. Met andere woorden, met de CCVSD-methode bereiken we het soort van reflectie waarop Turkle hamert en het soort van regulering die robotonderzoeker Peter Singer benadrukt.

In het kort weerspiegelt de CCVSD-benadering een toewijding aan het begrijpen van het beginpunt van zorg, is het kritisch op de manier waarop de zorg op dit moment uitgevoerd wordt, begrijpt het de noodzakelijke elementen van zorg en hoe die bijdragen aan het verbeteren van goede zorg en bovenal hoe zorg zou kunnen veranderen (ten goede of ten kwade) met de introductie van een zorgrobot. Deze toewijdingen indachtig wordt de toekomst van zorg met of zonder zorgrobot onderzocht: de waarden in de zorg worden diepgaand begrepen en vastgelegd. Niet-menselijke actoren krijgen erkenning voor hun rol in het vormen van een moraal en betekenis in een netwerk. Het belangrijkste is dat menselijke zorgverleners erkenning krijgen voor de waarde van hun werk en de verantwoordelijkheid van hun handelingen als hoeders van de zorg. Daabij is de CCVSD-benadering ook een uitdrukking van het ontwerp en de ingebruikneming van toekomstige zorgrobots die systematisch rekening houden met de waardering van zorgpraktijken en zorgverleners. Zo bezien hebben zorgrobots de potentie om een cruciale rol te vervullen bij het uitdragen van zorgwaarden als ze ontworpen en geïmplemeneerd zijn op de manier die de CCVSD-benadering voorschrijft.