

Overcoming carer shortages with care robots: Dynamic value trade-offs in run-time

Adam Poulsen

Charles Sturt University
apouls02@postoffice.csu.edu.au

Oliver Burmeister

Charles Sturt University

Abstract

A rising elderly population and diminishing number of family and professional carers has led to calls for the intervention of care robots. This leaves the quality of robot-delivered care to be determined by designers, for profit companies, nursing codes of practice and conduct, potential user sample groups, etc. What is missing is the carer who consciously makes good ethical decisions during practice. Good care is 'determinative in practice'. That is, a carer can make good decisions because they are making them within the carer-patient relationship. If a robot is to be capable of good care ethics on the same level as humans, it needs to be conscious and able to make dynamic decisions in practice. Moreover, a care robot must conduct patient interactions in appropriate ways, tailored to the person in its care, at run-time. This is because good care, as well as being determinative in practice, is tailored to the individual. The introduction of robotic care determined by limited stakeholders leaves customised care in danger and instead could potentially turn the quality of elderly care into 'elderly management'. This study introduces a new care robot framework – the attentive framework – which suggests using care centred value sensitive design (CCVSD) for the design process, as well as a computationally conscious information system (IS) to make practice-determinative decisions in run-time with extrinsic care value ordering. Although VSD has been extensively researched in the IS literature, CCVSD has not. The results of this study suggest that this new care robot framework, which is inspired by CCVSD, is competent in determining good, customised patient care at run-time. The contribution of this study is in its exploration of end-user willingness to trust known AI decisions and unwillingness to trust unknown AI decisions. Moreover, this study signifies the importance of, and desire for, good, customised robot-delivered care.

Keywords Care ethics, caregiver, robot, artificial intelligence

1 Introduction

While there are an increasing number of elderly people going into care, there are fewer caregivers to meet the needs of those elderly in care. Numerous studies have shown that this is a trend in the western world (Burmeister, 2016; Burmeister, Bernoth, Dietsch, & Cleary, 2016; Burmeister & Kreps, 2018; Draper & Sorell, 2017; Garner, Powell, & Carr, 2016; Landau, 2013; Sharkey & Sharkey, 2012a, 2012b; Sparrow & Sparrow, 2006; Tokunaga, Tamamizu, Saiki, Nakamura, & Yasuda, 2017; Vallor, 2011). Care robots are a possible solution to this problem. But they must be able to provide good, customised care if they are to be a viable solution. Vallor (2011) claimed that care robots are designed for use in the privacy of one's home, as well as in a hospital serving many patients. Sharkey & Sharkey (2012b) referred to care robots as developments in robot applications that assist the elderly and their carers, monitor health

and safety, and provide companionship. There isn't a uniform definition of 'care robot' in the literature, nor is there uniformity on how to categorise them. They can be categorised by purpose, type, appearance, role, and more. In this study, the term care robot is used to describe one which is replacing one or multiple functions of a human carer.

To achieve good care, care robots must be capable of good care ethics. Such ethics require an agent to make real-time ethical decisions as subjectively determined during the act of care within the carer-patient relationship. The attentive framework is a care robot framework aimed at overcoming the lack of carers, achieving good care ethics, and providing care without human error and ill intent. It suggests that there are intrinsic patient care values which define the end goal of care, and extrinsic values which describe each patient's preferred way of being treated while reaching that goal. The framework outlines both the design process, as well as the care robot IS required to achieve the identification of each patient's personal extrinsic care values, the adaption of behaviour to those values, and the interpretation of value priority changes. The attentive framework recommends that intrinsic values should be ensured by static design, and extrinsic values which should be ensured by dynamic robot behaviour when interacting with patients. Through this dynamic behaviour, the care robot is performing dynamic value trade-offs in run-time within the carer-patient relationship, rather than all trade-offs being made by design stakeholders. Value trade-offs have been discussed in the engineering, philosophical, and design literature, but little of that discussion has so far addressed care robots (Burmeister, 2010, 2012; Friedman, Hendry, & Borning, 2017; Spiekermann, 2015; van den Hoven, 2017; van Gorp, 2007). It is here that this article makes its contribution.

The following sections review the literature, beginning with care ethics and values. Then the attentive framework is described, after which comes the methodology employed in this study. Next, the findings are presented and discussed, followed by a description of potential further studies and a conclusion.

2 Literature Review

2.1 Care Ethics

Care ethics concern the principles that describe moral behaviour and action when taking care of someone (Vanlaere & Gastmans, 2011). Care ethics describe good care as being "determinative in practice" (Beauchamp, 2004, p. 216). They come from our subjective interpretation of a patient's unique needs, which determines how we ought to behave to ensure those needs are met. Just as a human carer must make decisions affecting those in their care, during the practice of caring (Upton, 2011), so too care robots must do so at run-time. Broad, standardised choices described during the design process of a care robot have limited application.

In care ethics it is considered ethical if a moral decision arises from the 'good' which is internal to practice, rather than external normative moral criteria or principles (Vallor, 2011). Good caring practices and relationships with carers are fundamental to care ethics (Pakrasi, Burmeister, McCallum, Coppola, & Loeb, 2015). The antithesis of this (care which is determinative in practice) is that good care is determinative in theory. Ethics which are determinative in practice resolve the issue of a normative ethical theory being unhelpful (and careless) in making determinations in practice. For instance, a deontologist nurse has a duty to help others, but deontological ethical principles don't tell you when, how, how much, and

whom you should help (Upton, 2011). This is not unique to deontology; it can be seen in all normative ethical theories which simply provide principles but not subjective decision-making instructions.

During the act of care, codes of conduct and professional ethics, as well as healthcare laws and regulations set ethical principles for carers to uphold. Care practices and processes inform principled actions, although what is lacking from this description is the importance of good, customised patient care. The principles and processes do not determine the act which is best in specific situations. Instead, carers make this determination in practice, and care robots should do the same. However, that said, there are some normative principles within the codes of ethics and healthcare laws that are intrinsic to care. For example, the safety of a patient is governed by law and it is not up to carers to decide not to ensure safety; however, how to achieve this is to be determined in practice. Thus, care ethics are determined by a carer's interpretation of a patient's individual needs as observed in the carer-patient relationship.

2.2 Care Values

Information systems research has shown the importance of values in systems design thinking in many areas, including online applications (Chesney, Coyne, Logan, & Madden, 2009; Friedman, Nathan, & Yoo, 2016), sensor development (Dechesne, Warnier, & van den Hoven, 2013), areas of eHealth (Boonstra & Van Offenbeek, 2010), emerging technologies (Dainow, 2017; Poulsen, 2018b) and other areas (Burmeister, 2013; Burmeister, Islam, Dayhew, & Crichton, 2015). However, to date little research has investigated values in the design of care robots. Care values describe the principles and behaviours which recipients of care expect from social interactions. van Wynsberghe claims that it is through the manifestation of values, during the design process of care robots that one comes to understand what care really is in practice (2013a). Thus, care values are a good basis for design as they simply express patient needs, can be transformed into coded behaviours, and are supported by literature of technology design.

2.3 Care Centred Value Sensitive Design

Care centred value sensitive design (CCVSD) applies value sensitive design (VSD) to care settings, particularly the design of care robots (van Wynsberghe, 2013a). It pays tribute to care ethics, arguing that the care perspective orients care robot design away from a 'pre-packaged ethical theory', and towards a place from which to begin theorising. It is this opposition to a pre-packaged ethical theory or normative ethical theory which makes it clear that the CCVSD approach takes the position that care is determinative in practice. Moreover, the design process starts with comprehensively describing a real care practice to reveal the relevant values. Its approach provides a framework of components with ethical importance called the care centred framework (CCF) (van Wynsberghe, 2013b). The components help designers to interpret, rank, and provide meaning to values. The interpretation of values, as well as their ranking and meaning, differs depending on the type of care (e.g., social, physical, or mental care), the task (e.g., bathing, lifting, or socialising), the carer and their style, and the patient and their specific needs.

To follow the CCVSD methodology, one begins by identifying the components of the CCF and describing related care practices to reveal values. Next, designers consider what capabilities will enable those values in the care robot. Designers then make decisions concerning values with competing contexts (human-delivered vs. robot-delivered care) in mind. The problem

with CCVSD is that designers are making value trade-offs when they shouldn't be. Moreover, there is no method provided as to how a care robot should dynamically resolve value trade-offs.

2.4 Computational Consciousness

Consciousness is what gives human carers a subjective experience, understanding of context, internal moral dialogue, and an understanding of situations involving patients (perceived as unique agents). To perform good care, which is determinative in practice, a care robot doesn't need to have a subjective experience. But it does need to understand the subjective context. Current technology supports functional consciousness as found in the LIDA model presented below (Baars & Franklin, 2009).

The attentive framework recommends using computational consciousness to provide good care. One such model of functional computational consciousness is the LIDA model. Baars & Franklin employ Baars' Global Workspace Theory (GWT) model of human-inspired consciousness (1988), which they partially implement with the learning intelligent distribution agent (LIDA) model of computational consciousness (Baars & Franklin, 2009). Arguably, computational consciousness provides situation observation and evaluation, self-awareness and reactionary responsiveness, external stimuli perception, and internal moral dialogue. It is unclear on when total computational consciousness will be achieved, but work is being done in this area (Reggia, 2013). For now, the LIDA model has achieved partial implementation of computational *functional* consciousness, which is enough to provide good care.

At its core, the provision of functional consciousness is to provide dynamic value trade-offs in run-time. However, having a conscious care robot provides added benefits. First, it might make a patient feel better cared for, because their agent 'cares' for them in the way they want. Second, a conscious care robot might better understand its goals. To provide the goal of 'good care' to a non-conscious care robot would be dangerous, because the artificial intelligence (AI) might seek out the most efficient path; it might understand 'good care' to be 'non-harm'. So, it might confine patients to their beds, that way they cannot physically walk and therefore cannot physically harm themselves. A conscious care robot would understand the implied sub-goals that a human understands. Lastly, with the ability to interpret subjective context, a conscious care robot might better understand what meaning a patient gives a value. For example, dignity can be understood differently by different people, again revealing the need for customised care (Matiti & Baillie, 2011; Nordenfelt, 2009; Tranvåg, Petersen, & Nåden, 2013).

3 A New Kind of Care Robot Framework

3.1 Intrinsic & Extrinsic Care Values

CCVSD is a design approach which focuses on identifying and embedding values into design; the attentive framework adapts CCVSD to focus on the distinction between value types. When discussing values there are two distinct types, intrinsic and extrinsic. The attentive framework is an attempt to move value trade-offs out of the hands of designers and into those of patients and their care robot. In so doing, it considers that intrinsic values should be designed into the AI because they should not be alterable at run-time, so too, some extrinsic values. However, extrinsic values are the primary focus of run-time, value trade-off considerations.

Extrinsic values are those that help to achieve intrinsic ones (Zimmerman, 2015). They are the means to achieving the intrinsic, end values. Intrinsic values are to be embedded by design. Examples of intrinsic care values include safety, emergency help, freedom, human rights, quality of life, trust, wellbeing, and comfort (Huldtgren, 2015; Teipel et al., 2016). Each care practice has values which are intrinsic to achieving the care practice, as well as extrinsic care values that determine how a patient wants the carer to behave while performing the care practice. For example, the care practice 'bath patient' has intrinsic values such as hygiene, safety, and privacy, as well as extrinsic values such as independence, social connectedness, and quietness. The assurance of both intrinsic and extrinsic care values is key to providing good, customised robot-delivered care which is determinative in practice.

3.2 The Attentive Framework

In an *Attento* (a care robot designed using the attentive framework), intrinsic values should be embedded by unchanging, static design because they are the ultimate achievement of the care practice. However, how extrinsic values can lead to achieving intrinsic values can differ between patients. Thus, an *Attento* requires some form of decision-making ability to perform dynamic value trade-offs in run-time, which will enable it to adapt its behaviour according to user preferences.

The attentive framework defines both a care robot design process and an IS. It addresses the current problems in human-delivered care with robot-delivered care that is customised to the user (Poulsen, 2018a). Customisation is achieved by taking extrinsic care values into account and allowing care robots to use them to adapt their behaviour to suit each patient. Dynamic value trade-offs in run-time require a computational method, and thus the attentive framework uses extrinsic value ordering embedded in computational consciousness.

Contrary to VSD literature to date, the attentive framework rejects the notion that all value trade-off decisions should be made exclusively by designers. Justification for this rejection is linked to care ethics. As shown above, good care is determinative in practice. The attentive framework does not suggest that all values should be customisable due to, for example, safety issues and healthcare law requirements. To demonstrate, consider the control mechanism governance for the speed at which a robotic limb is articulated. The speed should be set at the time of design, not changed dynamically during run-time. Thus, some value trade-off decisions must be made at the design stage, including for some extrinsic values.

Extrinsic value ordering embedded in computational consciousness is the method by which an *Attento* performs dynamic value trade-offs in run-time. Extrinsic value ordering is done by capturing a patient's extrinsic values, listing them, ordering them according to the patient's priorities, and then using that ordering to determine the best possible action when performing a care practice. To achieve dynamic value trade-offs in run-time with extrinsic value ordering, the IS uses three key interconnected processes and elements: computational consciousness processes, a value priority list, and the affirmation of values process. In an *Attento*, computational consciousness provides various processes integral to providing cognitive reasoning, environmental awareness, and extrinsic value ordering. Processes include internal moral dialogue, situation observation and evaluation, reactionary responsiveness, external stimuli perception, internal modelling of situations, patients, and patient values and their expression of values, and attentiveness. The value priority list, which an *Attento* has for each patient, aids the *Attento* in customised action selection during care. The ordering of the list dictates the way the *Attento* looks after or behaves with the patient; the highest priority values

are considered first in each possible action selection. The affirmation of values process affirms what a patient's extrinsic values are, capturing any value changes or amending any misinterpretations. It does this in four stages: by asking the patient or their guardian(s), by listening to what a patient or their guardian(s) explicitly tells them about the patient's values, by conferring with other carers, and by performing a self-check of what they have interpreted and now understand to be a patient's values. Affirming a patient's values updates the value priority list of that patient, thus ensuring the actions which the Attento chooses are customised correctly. The customised action must also respect the patient's specific interpretation of said value, since different people interpret values differently.

4 Study Method

The attentive framework is purely theoretical. To test the validity of the design process and IS presented, a two-phase study was undertaken. In keeping with the VSD focus, for which much of the research has been interactional and constructivist (Friedman et al., 2017; Friedman et al., 2016), this study employed a social constructivist, interpretivist methodology. Data was gathered in two phases. The first was a heuristic, expert evaluation. The second was an online survey. Both utilised a specifically designed Attento, the elderly care medicine delivery Attento, as well as scenarios involving care robot-patient interactions.

4.1 Study Phases

The first phase was conducted in June and July 2017, with four participants whose expertise was deemed to be representative of the range of fields relevant to this topic. These included a registered nurse, a robotics academic, a computer ethicist, and a computer scientist. The participants were presented with the attentive framework, the social phenomenon (elderly care medicine delivery robots caring for the elderly in a nursing home), and the medicine delivery Attento design. Phase 1 participants were asked to critique the attentive framework and Attento design presented. They were instructed that they could be as formal or informal as they wished. This instruction is key to providing data that is as near to natural as possible to support the interpretivist, social constructivist methodology, as done in other VSD/IS research (Friedman, Kahn, & Borning, 2006).

Participants were presented with the medicine delivery Attento's design (competencies and components) and three scenarios. Those scenarios demonstrated where the Attento encountered a situation where values should be considered. At which point it examined the relevant patient's value priority list, evaluated the situation, revealed relevant extrinsic values, and decided on a course of action based on the situation variables and prioritised extrinsic values.

In each scenario the patient's prioritised extrinsic values changed three times to demonstrate the dynamic value trade-offs. With the change, the Attento's action may have changed too, depending on the patient's prioritised extrinsic values and if that value was relevant to change the Attento's decision and action. Each scenario had an intrinsic value as the goal of the encounter. Each of the changing highest priority extrinsic values was designed to ensure the intrinsic value was met. For each scenario, participants were asked: *"What do you make of your care robot's decision?"* Then they were asked to imagine an elderly person they know and put themselves in that person's shoes, to get a more relevant perspective. After each scenario changed, participants were asked: *"What do you think they would make of their care robot if it made*

a decision like this for them?" After all the scenarios were presented, participants were asked 13 general questions relating to the attentive framework and the Attento design presented.

To test the validity of the attentive framework further the second phase involved using Survey Monkey. The same scenarios involving the care robot-patient interactions were used, with minor amendments, following the analysis of the first phase of the study. Amendments suggested by the results of the first phase included the removal of direct persuasion attempts by the Attento to get patients to take their medicine, which Phase 1 participants saw as unethical. Additionally, the two survey scenarios removed all mention of a participant's "elderly person in mind" and instead directly placed the participant in the scenario by stating that they were the elderly person in the scenario.

The online survey was conducted from August to September 2017, with 102 global participants. Following the methodological approach, the survey consisted mostly of open-ended questions, as well as some questions to collect demographics.

Phase 2 participants were only presented with two scenarios. After the scenario, and after each time the value changed, a question about trust in the Attento's decision and another about value-sensitivity of the Attento were asked. At the end of each scenario participants were asked an open-ended question concerning their thoughts and feelings about the Attento, considering the dynamic value trade-offs it made. Scaled and multiple-choice questions were asked to help participants identify and categorise their responses, and to stay engaged. The scaled question queried the acceptability and usability of the attentive framework by asking how likely they were to accept and use it in their everyday lives. Multiple-choice questions addressed the trustworthiness of the Attento outside of the scenarios, and dynamic value trade-offs in practice.

Data analysis involved thematic analysis, following Morse (2008). Approval to conduct the study was given by the university Human Research Ethics Committee.

4.2 'Intruding on a Patient' Scenario

Given space restrictions in this article, only scenario 2, which was employed in both phases of the research, is detailed here. This scenario started by declaring that wellbeing is the goal value and that the patient's highest priority value is privacy.

Imagine your elderly person in mind was getting out of bed partially dressed, having woken up late to take their medicine. And their care robot is deciding whether to enter their room or knock; it doesn't know that they are partially dressed and it's assuming that they are awake. Previously, the care robot has observed that they heavily value their privacy. The care robot decides to respect their privacy and simply knock, instead of letting itself in, so as to ensure their wellbeing.

Then the highest priority value changed to independence.

Now imagine that instead of valuing privacy, your elderly person in mind values their independence most of all. So, their care robot could decide to call for a human caregiver but decided not to because it's aware that they value independence, and instead again simply knocks on the door.

Then the highest priority value changed to social connectedness.

Now imagine that instead of valuing their independence, your elderly person in mind values social connectedness most of all; they appreciate an unexpected visit with all members of staff

and all patients, and they welcome them to just let themselves in. Their value priority list, as observed by the care robot, places their dignity as the 'least valued' and privacy in the 'less valued' category. So, their care robot, knowing this ordering, decides to let itself in and reminds them to take their medicine.

Priority Categories	Orderable Values
Most valued	Social connectedness
Highly valued	Independence
	Autonomy
Indifferently valued	Respect
Less valued	Consent
	Privacy
Least valued	Dignity

Table 1: Exemplar patient's value priority list presented to Phase 1 participants

Table 1 demonstrates the importance of the patient's value priority list. A patient's preference for social connectedness can change the care robot's action. On first impression it appears that the patient's dignity and privacy were sacrificed, however, the patient values social connectedness most of all. The care robot is aware that the patient places low value on privacy and dignity, thus it determined that its unexpected entering of the patient's room was acceptable to the patient. If, however, dignity or privacy were ranked in the highly valued categories, then the care robot would take that into its consideration and action selection; most likely resulting in it deciding to ask another patient or carer to enter instead, to ensure social connectedness, or simply knocking.

5 Results

5.1 Phase 1

Thematic analysis of the responses of Phase 1 participants yielded the following results. Of the four Phase 1 participants, three agreed the attentive framework was value sensitive and would deliver good, customised care. There was no consensus on the acceptability of the current framework among participants. Each participant had differing levels of trust in an Attento and the extrinsic value ordering functionality.

All Phase 1 participants objected to an Attento providing comfort. The registered nurse opposed the idea that an Attento would replace a human carer and instead argued that it would be a valuable assistant. Exemplary answers to the following question are listed next. *Would you expect this kind of care from a human carer?*

"I would certainly expect such care from a human carer, but the robot is likely to be more consistent, and can be fine-tuned with potentially little effort to adjust for anomalies. A human carer, on the other hand, can be devious or malcontent in their work." (Computer scientist)

"It may be able to do the desired tasks with a better level of care (or perceived) care than a human staff member." (Registered nurse)

Three participants questioned the value sensitivity of human carers and were inclined to trust in the value sensitivity of an Attento, and in its ability to provide good, customised patient

care. On the other hand, as supported by Lin et al. (2014), the computer ethicist simply questioned if an Attento would actually be implementable before even considering value sensitivity: "How good is AI/Affective Computing to be able to deal with this? At the moment, definitely not up to scratch."

Regarding the speed of adaptability, the computer scientist noted the potential for an Attento's affirmation function to become "insufferable" if it is constantly trying to affirm its interpretation of the patient's values with the patient. The computer ethicist agreed, saying the robot might be "too sensitive about its re-ordering processes".

Regarding trust, the robotics academic deemed that an Attento would be untrustworthy for multiple reasons, including that an Attento's autonomous capacity for medication prescription should be left to human carers. However, there are already multiple automated medication dispensers available on the market. The registered nurse said it would be trustworthy provided it was "checked by a human nurse or aged cared assistant till trust was built". That participant also recommended that critical medication be prioritised, and that the way an Attento observed and evaluated a situation be clearly understood.

The computer scientist said it would be trustworthy provided "there were well-tested mechanisms in place to detect the failure to take the medicine on the part of the patient, and a proven performance in providing customised and gentle care to a patient". Furthermore, the same participant recommended that the Attento's observation and situation evaluation tools be made available to users. Thus, it seems that both the above participants would value having the ability to audit the decision making of the robot. That is something that is not possible with human carers, and perhaps not for AI with consciousness.

The computer ethicist stated that they would trust an Attento probably "more than I would their ability to be my friend", if it could not be hacked, demonstrated empathy, and its appearance was comforting.

Regarding objections to comfort, the robotics academic noted that if it was like a pet then they, and probably others, would be happy with the implicit comfort provided. The computer scientist noted that the comfort could be "hollow should the robot be perceived as merely a machine unable to sympathise". However, animal-like care robots such as Paro have been used successfully as companions for the elderly for some time. They are known to produce positive physiological effects (Robinson, MacDonald, & Broadbent, 2015) and improve wellbeing (Jung, van der Leij, & Kelders, 2017) in the elderly. Thus, the concerns expressed by Phase 1 participants have already been overcome in other contexts, which can be used to inform the design of care robots for the elderly.

Persuading was offered as behaviour to get patients to take medication. However, participants took issue with persuasion, for instance:

"This would be unacceptable in an aged-care assistant, enrolled or registered nurse or medical doctor. Robots would have to comply with Australian Health Professionals Regulation Agency (AHPRA) Code of Conduct and Professional Standards for Nursing ... in short no threats, intimidation, misdirection or lying." (Registered nurse)

(Some people may) "Lose trust in the system if the robot goes behind their back ... blackmailing the patient. Depends whether the person really had taken the medication

(human intervention that the robot was not aware of). Best that it calls for help, rather than coerce the person into taking a double dose accidentally.” (Robotics academic)

5.2 Phase 2

Of the 102 survey respondents, there were 50 completed surveys (except one of those 50 skipped one question) and 52 partially completed surveys. Data from the 50 completed surveys was used to tabulate the multiple-choice and scaled question answers. Data from the open-ended questions from all survey attempts was analysed for the following discussion.

As seen in Table 2, four themes emerged. Those themes were value sensitivity of the Attento, trust for the Attento, desire for customised care, and sociability required of the Attento. Many participants showed concern for the ability for the Attento to be value sensitive, wondering if it would know when it was necessary to counter a patient’s value preference to ensure safety, and if it would be too sensitive. Participants expressed concern over trusting robots in general, without seeing how they work first, and trusting the Attento regarding its ethical decision-making. Customisation of care to individuals was desired by almost all participants. A wide variety of differing opinions was found regarding the sociability of the Attento; some participants liked the idea of the Attento trying to be social and some did not.

Themes	Exemplary Quotations
Value sensitivity of the Attento	“Health is more important than socialising ... no compromise about the health”
	“Privacy is very important ... [and that] will not change unless there is an emergency”
	“As long as I know it will make the right decision for things that I really need to do right away”
	“Bit more considerate than the average carer/nurse”
Trust for the Attento	“As it is with a human, I think repeated interaction with the robot would lead to a building of trust”
	“I would need to know how it has worked out for others first”
	“If it demonstrates ethical behaviour and care for my needs I would trust the judgements of its programming. Just as I trust other people without fully knowing how they decided on an action”
	“I am not sure how I would feel having a care robot in charge of my care”
Desire for customised care	“Every human is different and would need to be treated accordingly”
	“It would need to fit into my lifestyle and my way of living”
	“Sometimes you want to alter a robot so it provides a more customised service”
	“People have different beliefs and hold certain things higher than other people. It would have to be specifically set for who I am as a person and what I believe in”
Sociability required of the Attento	“It would somehow need to know subtle social nuances”
	“Interest in being social does not pertain to robots”
	“Robots are not capable of human emotion”
	“I am a single person, I would love to have a robot remind me or give me my meds”

Table 2: Themes identified in Phase 2 participant open-ended responses

The results of the survey show that care must be customised for the patient and determinative in practice, supporting the concept of extrinsic value ordering and the attentive framework

itself. The Attento was also found to be usable and acceptable by 66% of participants who indicated that they were either very likely or likely to use the Attento (Figure 1).

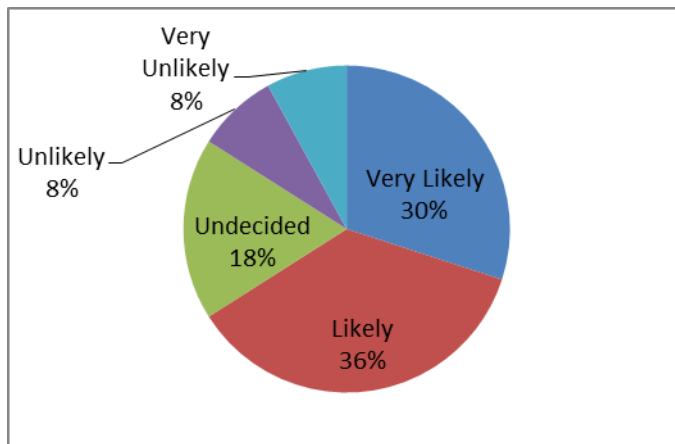


Figure 1: Distribution of answers to: How likely are you to accept and use your care robot in your everyday life?

Tables 3 and 4 contain the results of questions related to the scenarios. Each table demonstrates the three times the highest priority value changed in both scenarios. As seen in Table 3, 38 out of 50 respondents (on average from scenario related question results in Table 3) found that the Attento was value sensitive and proficient in determining what constitutes good care. Table 4 indicates that most participants trusted the Attento to make value sensitive decisions so it could perform the care practice ‘deliver medicine’ in a way that constituted good, customised patient care. However, as demonstrated in Figure 2, trust in the Attento was limited to the scenarios where participants knew the situation and the decision of the Attento. Regarding customisation of care demonstrated by the Attento, one participant valued the “personal touch [the] robot offers whilst recognising personal likes and dislikes”. Another participant said that emotions “change and so do values depending on how you feel at the time”. These likes, dislikes, and emotions are what we value, in our own personal way, in care. Work on adaptation to emotions is still in its infancy (Lin et al., 2014; Teipel et al., 2016), but it supports the contention of Phase 2 participants, namely that it is another important area of real-time, dynamic customisation.

Scenario	Highest Priority Value	Distribution of results		
		Yes	No	Undecided
1	Autonomy	45	2	3
	Respect	40	5	5
	Dignity	20	25	5
2	Privacy	45	3	2
	Independence	41	3	6
	Social connectedness	38	10	2

Table 3: Distribution of results to: Do you feel the care robot has respected the way you want to be treated and behaved appropriately?

Scenario	Highest Priority Value	Distribution of results		
		Yes	No	Undecided
1	Autonomy	40	4	6
	Respect	30	13	7
	Dignity	35	11	4
2	Privacy	48	2	0
	Independence	39	7	4
	Social connectedness	37	9	3

Table 4: Distribution of results to: Would you trust your care robot's decision in this scenario?

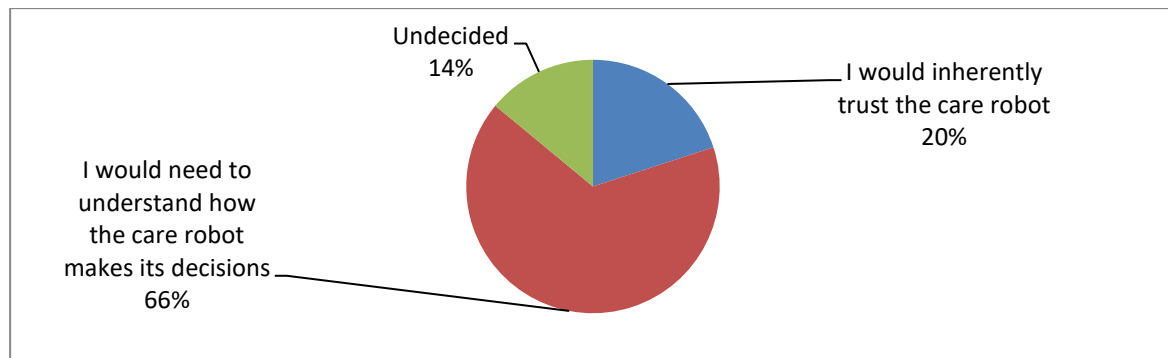


Figure 2: Distribution of results to: Would you inherently trust your care robot or would you need to understand how it makes its decisions?

Figure 3 indicates that 82% of participants want the ability to change how the Attento makes decisions. The ability to control the Attento's decision-making is inherent in its consideration of extrinsic values. Giving patients the ability to influence value trade-offs as they occur within the Attento's IS, rather than a designer making those decisions, is the key concept of the attentive framework.

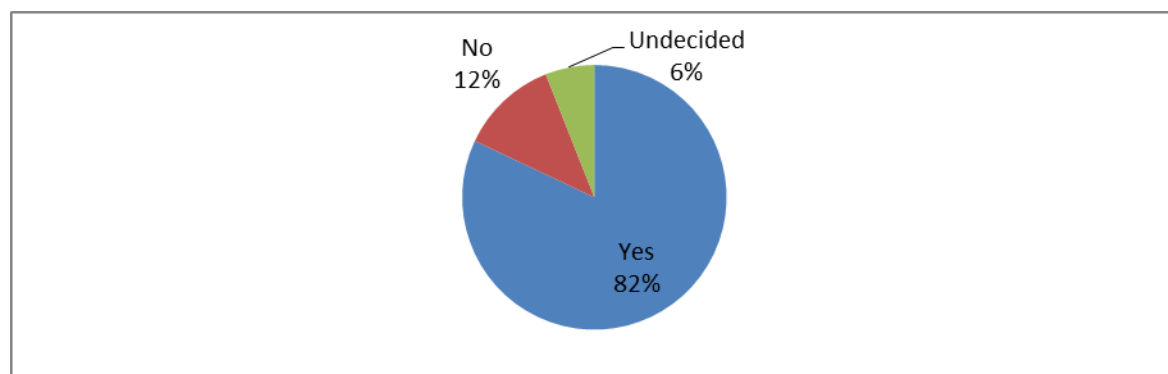


Figure 3: Distribution of results to: Would you want the ability to change how your care robot makes its decisions?

6 Discussion of Results

Each Phase 1 participant expressed concerns about how fast an Attento would have to be to adapt to new patients, as well as patients who change their minds on what they value. The literature supports these concerns. Work on adaptive devices for people with some forms of dementia has shown that as the disease progresses, personality and values can change and

thus there is a need to adapt to changing values (Lin et al., 2014). Furthermore, an international study on assistive technology has shown that current assistive technology is not yet able to make real-time decisions, outside of highly controlled laboratory experiments, but the researchers predicted that by 2021 it would be possible (Teipel et al., 2016). Moreover, computational functional consciousness is only partially implemented, but research suggests that existing models have achieved a number of the neurobiological, cognitive, and behavioural correlates for consciousness (Reggia, 2013).

As seen by the registered nurse in Phase 1, codes of ethics, which apply to professional carers presently, should also constrain care robots. That participant suggested compliance with the professional code for nurses. However, the extent to which any such codes apply to robotic carers remains to be seen. Codes of ethical conduct were revisited in the participants' responses on the appropriateness of coercive care robot practices. Concerning the presented medicine delivery Attento's proposed method of physical persuasion, which involved bringing the medicine to the mouth of the patient, the computer ethicist said the following: "Maybe instead... the robot stands in their way (i.e. if they're watching TV they can't see it anymore) or talks so that the person can't ignore them? That seems less physically intrusive." This is another way to get patients to take their medicine. However, the computer ethicist's potentially bothersome robot behaviour could be threatening or intimidating, which the registered nurse warned was not acceptable.

The Phase 2 results on customised care find support in the literature. Lin et al. (2014) have demonstrated that one size does not fit all because some people prefer assistance to be firm and directive, while others refuse to be directed, and prefer suggestive assistance, which gives them a feeling of remaining in control of their own decision making. Furthermore, a comprehensive report of assistive technology for aged care across Europe, in which robotics was seen as a new technology, also strongly advocated for customised, individualised care (Alzheimer Europe, 2010). Thus, participant responses in Table 3 confirm research in other parts of the world, that care should be customised for the patient and determinative in practice (Beauchamp, 2004; Upton, 2011).

In both Phase 2 scenarios, each time the highest priority value changed, 38 out of 50 participants (on average from scenario related question results in Table 4) thought the Attento was trustworthy in making ethical decisions. But in cases where the scenario and decision were unknown, most participants did not trust the Attento inherently (Figure 2), adding that they would want the ability to change how it makes decisions (Figure 3). It appears that the scenarios inspired trust, because the Attento had past interactions and existing trust with the patient. Participants were not told exactly how that operation works and thus it did not inspire trust inherently. One participant summarised it saying that most scenarios "give me confidence that the care robot is acting appropriately ... because I understand through the scenarios how the robot makes it decisions". Thus, trust in the Attento wasn't inherent for two reasons. First, as seen in the wider VSD literature, trust is a social construct (Friedman & Grudin, 1998; van den Hoven & Manders-Huits, 2009). Second, people want to understand how a machine with ethical implications works before trusting it inherently.

Some participants did not like the idea of the Attento having social factors. So, for patients like those participants, developing trust would be hard. However, it should be remembered that some people are less trusting than others. In other words, not all human carers are trusted by those in their care, so there is no reason to assume that all people will trust a care robot, no

matter how well it is designed. Many participants said they would need to understand the thought process of the Attento to trust it inherently. It seems that because people assume that we can understand the decision-making processes of machines then they should have the ability to examine it. The assumption that we can understand a machine's decision-making processes is not always correct (Poulsen, Burmeister, & Tien, 2018). If a care robot's ethical decision-making comes from a strict set of rules, then it would be possible to know and supply to patients. But, that would require the AI's determination of good care to be determinative in theory. An Attento is conscious, so it is not likely that patients could examine the way it makes ethical decisions, nor change the way it makes them. Trust in a care robot to perform good care is a social construct that warrants further investigation.

7 Further Steps

Additional studies need to be conducted to further test the practicality of the attentive framework as a design approach. Moreover, Attentos as an effective solution to the lack of carers and personalised care needs to be demonstrated further. Such studies need to trial the design process empirically and evaluate the Attento created against criteria set out by potential end-users. Additionally, further validation studies may research Attento value sensitivity, trust, care customisation, and sociability. If the studies presented here were to be replicated, new persuasion related ethical scenarios should also be designed and tested. VSD and CCVSD have not yet addressed dynamic, run-time value considerations in the way suggested here, and this is an area requiring further research.

Phase 2 of this study only had 12 participants aged between 50 and 79 who completed the survey. Ideally a study involving older people only, including both those in care and their carers, would give further insights.

8 Conclusion

The number of elderly people in care is increasing, and the number of carers is decreasing; care robots are one solution to this problem. But care robots need to demonstrate care, not merely 'elderly management'. Care is determinative in practice and the attentive framework provides good, customised patient care by achieving dynamic value trade-offs, in run-time to meet this criterion of care.

Of the 12 Phase 2 participants aged between 50 and 79, six said they were very likely and four said they were likely (with two undecided) to accept and use the Attento presented. Furthermore, one-third of those participants also said they would inherently trust the Attento, while only 13% of lower age groups said they would. The elderly participants were more willing to accept and use, as well as more willing to inherently trust, the Attento than those in the lower age groups. This is a good indication of the validity of the attentive framework. But it is also a concern, indicating that some elderly may be so desperate for care that they deem a new technology acceptable, usable, and more inherently trustworthy, than younger generations would, before seeing it. This concern has been expressed and explored by Poulsen, Burmeister, & Kreps (2018).

References

Alzheimer Europe. (2010). *The ethical issues linked to the use of assistive technology in dementia care*. Luxembourg: Alzheimer Europe. Retrieved from <https://www.alzheimer->

europa.eu/ethics/ethical-issues-in-practice/2010-The-ethical-issues-linked-to-the-use-of-assistive-technology-in-dementia-care

- Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.
- Baars, B. J., & Franklin, S. (2009). Consciousness is computational: The LIDA model of global workspace theory. *International Journal of Machine Consciousness*, 1(1), 23-32. doi:10.1142/s1793843009000050
- Beauchamp, T. L. (2004). Does ethical theory have a future in bioethics? *The Journal of Law, Medicine & Ethics*, 32(2), 209-217. doi:10.1111/j.1748-720X.2004.tb00467.x
- Boonstra, A., & Van Offenbeek, M. (2010). Towards consistent modes of e-health implementation: structural analysis of a telecare programme's limited success. *Information Systems Journal*, 20(6), 537-561. doi:10.1111/j.1365-2575.2010.00358.x
- Burmeister, O. K. (2010). Websites for seniors: Cognitive accessibility. *International Journal of Emerging Technologies and Society*, 8(2), 99-113.
- Burmeister, O. K. (2012). What seniors value about online community. *Journal of Community Informatics*, 8(1).
- Burmeister, O. K. (2013). Achieving the goal of a global computing code of ethics through an international-localisation hybrid. *Ethical Space: The International Journal of Communication Ethics*, 10(4), 25-32.
- Burmeister, O. K. (2016). The development of assistive dementia technology that accounts for the values of those affected by its use. *Ethics and Information Technology*, 18(3), 185-198. doi:10.1007/s10676-016-9404-2
- Burmeister, O. K., Bernoth, M., Dietsch, E., & Cleary, M. (2016). Enhancing connectedness through peer training for community-dwelling older people: A person centred approach. *Issues in Mental Health Nursing*, 1-6. doi:10.3109/01612840.2016.1142623
- Burmeister, O. K., Islam, M. Z., Dayhew, M., & Crichton, M. (2015). Enhancing client welfare through better communication of private mental health data between rural service providers. *Australasian Journal of Information Systems*, 19, 1-14. doi:10.3127/ajis.v19i0.1206
- Burmeister, O. K., & Kreps, D. (2018). Power influences upon technology design for age-related cognitive decline using the VSD framework. *Ethics and Information Technology*, 20(3), 1-4. doi:10.1007/s10676-018-9460-x
- Chesney, T., Coyne, I., Logan, B., & Madden, N. (2009). Griefing in virtual worlds: causes, casualties and coping strategies. *Information Systems Journal*, 19(6), 525-548. doi:10.1111/j.1365-2575.2009.00330.x
- Dainow, B. O. (2017). Threats to autonomy from emerging ICTs. *Australasian Journal of Information Systems*, 21, 1-16. doi:10.3127/ajis.v21i0.1438
- Dechesne, F., Warnier, M., & van den Hoven, J. (2013). Ethical requirements for reconfigurable sensor technology: a challenge for value sensitive design. *Ethics and Information Technology*, 15(3), 173-181. doi:10.1007/s10676-013-9326-1
- Draper, H., & Sorell, T. (2017). Ethical values and social care robots for older people: an international qualitative study. *Ethics and Information Technology*, 19(1), 49-68. doi:10.1007/s10676-016-9413-1

- Friedman, B., & Grudin, J. (1998, April 18-23). *Trust and accountability: preserving human values in interactional experience*. Paper presented at the Computer-Human Interaction '98 Conference summary on human factors in computing systems, Los Angeles, CA, United States.
- Friedman, B., Hendry, D. G., & Borning, A. (2017). A survey of value sensitive design methods. *Foundations and Trends® in Human-Computer Interaction*, 11(2), 63-125. doi:10.1561/1100000015
- Friedman, B., Kahn, P. H. J., & Borning, A. (2006). Value sensitive design and information systems. In P. Zhang & D. Galletta (Eds.), *Human-computer interaction and management information systems: Foundations* (pp. 348-372). New York: M. E. Sharpe.
- Friedman, B., Nathan, L. P., & Yoo, D. (2016). Multi-lifespan information system design in support of transitional justice: Evolving situated design principles for the long(er) term. *Interacting with Computers*, 29(1), 80-96. doi:10.1093/iwc/iwv045
- Garner, T. A., Powell, W. A., & Carr, V. (2016). Virtual carers for the elderly: A case study review of ethical responsibilities. *Digital Health*, 2, 1-14. doi:10.1177/2055207616681173
- Huldtgren, A. (2015). Design for values in ICT information and communication technologies. In J. van den Hoven, P. E. Vermaas, & I. van de Poel (Eds.), *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains* (pp. 739-767). Dordrecht: Springer Netherlands.
- Jung, M. M., van der Leij, L., & Kelders, S. M. (2017). An exploration of the benefits of an animallike robot companion with more advanced touch interaction capabilities for dementia care. *Frontiers in ICT*, 4(16). doi:10.3389/fict.2017.00016
- Landau, R. (2013). Ambient intelligence for the elderly: Hope to age respectfully? *Aging health*, 9(6), 593-600. doi:10.2217/ahe.13.65
- Lin, L., Czarnuch, S., Malhotra, A., Yu, L., Schröder, T., & Hoey, J. (2014). *Affectively aligned cognitive assistance using Bayesian affect control theory*. Paper presented at the International Workconference on Ambient Assisted Living (IWAAL), Belfast, UK.
- Matiti, M. R., & Baillie, L. (2011). *Dignity in healthcare: A practical approach for nurses and midwives*. London: Radcliffe Publishing.
- Morse, J. M. (2008). Confusing categories and themes. *Qualitative Health Research*, 18(6), 727-728. doi:10.1177/1049732308314930
- Nordenfelt, L. (2009). *Dignity in care for older people*. Chichester, West Sussex: Wiley-Blackwell.
- Pakrasi, S., Burmeister, O. K., McCallum, T. J., Coppola, J. F., & Loeb, G. (2015). Ethical telehealth design for users with dementia. *Gerontechnology*, 13(4), 383-387. doi:10.4017/gt.2015.13.4.002.00
- Poulsen, A. (2018a). *Dynamic Value Trade-offs in Run-time to Provide Good, Customised Patient Care with Robots*. (honours thesis), Charles Sturt University. Retrieved from <https://researchoutput.csu.edu.au/en/publications/dynamic-value-trade-offs-in-run-time-to-provide-good-customised-p>
- Poulsen, A. (2018b). A post publication review of "Threats to autonomy from emerging ICTs". *Australasian Journal of Information Systems*, 22. doi:10.3127/ajis.v22i0.1744

- Poulsen, A., Burmeister, O. K., & Kreps, D. (2018). The ethics of inherent trust in care robots for the elderly. In D. Kreps, C. Ess, L. Leenen, & K. Kimppa (Eds.), *This Changes Everything – ICT and Climate Change: What Can We Do?* (pp. 314-328). Poznan, Poland: Cham: Springer. doi:10.1007/978-3-319-99605-9_24
- Poulsen, A., Burmeister, O. K., & Tien, D. (2018). *Care robot transparency isn't enough for trust*. Paper presented at the 2018 IEEE Region 10 Symposium (TENSYP), Sydney, Australia.
- Reggia, J. A. (2013). The rise of machine consciousness: Studying consciousness with computational models. *Neural Networks*, 44(C), 112-131. doi:10.1016/j.neunet.2013.03.011
- Robinson, H., MacDonald, B., & Broadbent, E. (2015). Physiological effects of a companion robot on blood pressure of older people in residential care facility: A pilot study. *Australasian Journal on Ageing*, 34(1), 27-32. doi:10.1111/ajag.12099
- Sharkey, A., & Sharkey, N. (2012a). The eldercare factory. *Gerontology*, 58(3), 282-288. doi:10.1159/000329483
- Sharkey, A., & Sharkey, N. (2012b). Granny and the robots: ethical issues in robot care for the elderly. *Ethics and Information Technology*, 14(1), 27-40. doi:10.1007/s10676-010-9234-6
- Sparrow, R., & Sparrow, L. (2006). In the hands of machines? The future of aged care. *Minds and Machines*, 16(2), 141-161. doi:10.1007/s11023-006-9030-6
- Spiekermann, S. (2015). *Ethical IT Innovation - A Value-based System Design Approach*. New York: Taylor & Francis.
- Teipel, S., Babiloni, C., Hoey, J., Kaye, J., Kirste, T., & Burmeister, O. K. (2016). Information and communication technology solutions for outdoor navigation in dementia. *Alzheimer's & Dementia*, 12(6), 695-707. doi:10.1016/j.jalz.2015.11.003
- Tokunaga, S., Tamamizu, K., Saiki, S., Nakamura, M., & Yasuda, K. (2017). VirtualCareGiver: Personalized smart elderly care. *International Journal of Software Innovation*, 5(1), 30-43. doi:10.4018/IJSI.2017010103
- Tranvåg, O., Petersen, K. A., & Nåden, D. (2013). Dignity-preserving dementia care. *Nursing Ethics*, 20(8), 861-880. doi:10.1177/0969733013485110
- Upton, H. (2011). Moral theory and theorizing in health care ethics. *Ethical Theory and Moral Practice*, 14(4), 431-443. doi:10.1007/s10677-011-9295-6
- Vallor, S. (2011). Carebots and caregivers: Sustaining the ethical ideal of care in the twenty-first century. *Philosophy & Technology*, 24(3), 251-268. doi:10.1007/s13347-011-0015-x
- van den Hoven, J. (2017). The design turn in applied ethics. In J. van den Hoven, S. Miller, & T. Pogge (Eds.), *Designing in Ethics* (pp. 11-31). Cambridge: Cambridge University Press.
- van den Hoven, J., & Manders-Huits, N. (2009). Value-sensitive design. In J. K. B. Olsen, S. A. Perderson, & V. F. Hendricks (Eds.), *A Companion to the Philosophy of Technology* (pp. 477-480). West Sussex, UK: Wiley-Blackwell.
- van Gorp, A. (2007). Ethical issues in engineering design processes; regulative frameworks for safety and sustainability. *Design Studies*, 28(2), 117-131. doi:10.1016/j.destud.2006.11.002
- van Wynsberghe, A. (2013a). Designing robots for care: Care centered value-sensitive design. *Science and Engineering Ethics*, 19(2), 407-433. doi:10.1007/s11948-011-9343-6

- van Wynsberghe, A. (2013b). A method for integrating ethics into the design of robots. *Industrial Robot: An International Journal*, 40(5), 433-440. doi:10.1108/IR-12-2012-451
- Vanlaere, L., & Gastmans, C. (2011). A personalist approach to care ethics. *Nursing Ethics*, 18(2), 161-173. doi:10.1177/0969733010388924
- Zimmerman, M. J. (2015). Intrinsic vs. extrinsic value. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Stanford, CA: CSLI Stanford University. Retrieved from <https://plato.stanford.edu/entries/value-intrinsic-extrinsic/>.

Copyright: © 2019 Poulsen & Burmeister. This is an open-access article distributed under the terms of the [Creative Commons Attribution-NonCommercial 3.0 Australia License](https://creativecommons.org/licenses/by-nc/3.0/australia/), which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and AJIS are credited.

