Radboud Honours Academy

Robots against Anorexia Nervosa

An interdisciplinary assessment of the possible use of Socially Assistive Robots in the treatment of Anorexia Nervosa

Think tank *Thinking Machines* 2016 – 2017

Marco Dessi
Dylan Henssen
Sophie Horsman
Rutger Meijers
Annika Schiefner
Fenja Schlag
Luuk Schmitz
Vika Shimanskaya
Ricarda Weiland

Radboud University

Interdisciplinary Honours Programme for Master's Students

The Radboud Honours Academy offers talented and motivated students at Radboud University the opportunity to take an additional, challenging study programme. Students are selected based on their ambition, potential and study results.

Participants of the Interdisciplinary Honours Programme for Master's Students are stimulated to look beyond the borders of their own field of study. They work together in interdisciplinary think tanks and conduct research on a socially-relevant issue. Their final report is addressed to an external organisation. The extra knowledge and skills that students acquire during the programme are of great value to their personal and academic development, and for their further career. The extra study load is equivalent to 15 ECs.

© Authors and Radboud Honours Academy, 2017
www.ru.nl/honoursacademy

Acknowledgements

We would like to thank Radboud Honours Academy for their efforts and for providing us with necessary structures and for realizing our conference visits to “Get Into The Future 2016” in Amsterdam, to the “HRI 2017” in Vienna and to the “ERF 2017” in Edinburgh.

Special thanks go out to Noortje ter Berg, for her supervision and relentless efforts in bringing out the best in each one of us.

Furthermore, we would like to thank our supervisor, Lucien Engelen. He taught us to open our minds to new (technological) possibilities and encouraged us to be bold and explore new grounds.

Another important person in this project was Lorraine Faulds. She tirelessly provided us with valuable feedback on our writing style and helped us improving this final version of the report.

Many thanks also go out to various people that provided us with their expertise: Prof. Jan Buitelaar, who gave us valuable insights into treatments of anorexia nervosa and robots in autism therapies; Prof. Peter Hagoort who provided us with useful feedback on our research question and interview design; and Robin Hooijer, who helped us to think outside the box.

Finally, we would like to thank all the experts and stakeholders who agreed to give us an interview. Without their input this report would have lacked an essential part.

"I make mistakes growing up. I'm not perfect; I'm not a robot." Justin Bieber

Executive Summary

Within the next decades, technology is going to play an increasingly important part in our lives. Machines are progressively becoming more complex, so that they can even give the impression to be 'thinking'. These developments render the distinction between human and machine less clear. In particular, robots seem to embody the very idea of "thinking machines". However smart such artificial intelligence might be or become, its creator should be even smarter in steering its application in a desirable way. Under careful guidance, the development of robotics could be geared towards the preservation and enhancement of the quality of human life. For such an endeavor, recently robotic assistance has been put at service of healthcare thus engendering the promising field of Socially Assistive Robots (SARs). Whether or not SARs would eventually change healthcare practices for the best is a crucial question for humanity and a well-informed approach is needed.

This entails the following research question:

- *How can we enhance the likelihood of a desirable outcome for the proliferation of robots in our society?*

The many exciting opportunities that Socially Assistive Robots provides are met with potential concerns. For instance, who is responsible when a 'thinking machine' makes a mistake, and how could privacy be reconciled with the introduction of more personalized robots? These questions call for a more fundamental discussion on how exactly the proliferation of robots in our society should be shaped, thus far not found in the discourse around SARs. In order to shed light on these comprehensive questions, we decided to focus on the possible design and implementation of SARs for a very sensitive case of mental disorder, namely Anorexia Nervosa (AN).

The choice for AN is informed by:

- Previous research on SARs showing promising results in amongst others the treatment of elderly with dementia and children with Autism Spectrum Disorder (ASD);
- The successful introduction of SARs as coaches for children with diabetes and people with obesitas;
- The potential to combine the results of previous research with the specific desiderata of AN patients;
- The potential to generalize the results of this study to other mental disabilities.

In the present report, we tackle those issues in a twofold approach:

- A theoretical background combining the scientific literature on SARs and on AN treatments in a ethical, legal, and social framework;

- An empirical study exploring how patients, therapists, and experts potentially involved in the design of an SAR for AN patients conceive of such a development.

The main results from the theoretical chapter were:

- Ethical, legal, and social concerns should be comprehensively taken into account when developing SARs for healthcare. The framework found in the first chapter of our study provides the basis for doing so.
- AN is a complex mental disorder with multifactorial origin, unfavourable prognosis and strongly affected quality of life.
- SARs could improve patients' cognitive abilities, social interaction competencies, coping strategies and quality of life.

The main results of the empirical chapter were:

- Stakeholders hold differing opinions on how the robot should function, what the role of the robot should be, and how interaction between robot and patient should function.
- a companion-type robot is most suited for chronic AN patients, while the coach-type robot is suited more for non-chronic AN patients.
- A confirmation of the important role that ethical, legal, and social considerations play throughout the development and implementation of SARs

Conclusion

- The theoretical part revealed a potential for introducing SARs to enhance current treatment practices, but only as long as ethical, legal, and social concerns are taken into account.
- The empirical part shows that the needs of patients in therapy are extremely complex, necessitating advanced communication skills and complex social behaviours on the side of the SAR.

Recommendations

The following list of recommendations have been proposed. Further elaboration on these recommendations can be found in the report.

- The introduction of Socially Assistive Robotics (SARs) into both existing and new fields of healthcare requires an approach that centres stakeholders' needs, while remaining sensitive to ethical, legal, and social concerns.
- With regard to the specific case of Anorexia Nervosa, a differentiation should be made between adolescent and adult patients.
- Development of SARs in healthcare always necessitate a personalized approach.
- Future studies assessing the introduction of SARs in healthcare should seek to conduct focus group discussions with stakeholders to further clarify their needs.
- Scientific research on SARs should seek to deploy controlled trials and good experimental designs to enhance their explanatory power and generalisability.
- Researchers must avoid approaching the topic only in a problem-solving manner, and also dare to ask more fundamental questions.
- An interdisciplinary approach is the way to go forward for enhancing the likelihood of a desirable outcome for introducing more complex and capable robots in society.

Acronyms

AN	Anorexia Nervosa
SAR	Socially Assistive Robot
ELS	Ethical, Legal, Societal
TPP	Therapeutic Play Partner
CCVSD	Care-Centered Value-Sensitive Design
VSD	Value-Sensitive Design
SBTC	Skills-Based Technological Change
RBTC	Routine-Biased Technological Change
DSM-5	Diagnostic and Statistical Manual of Mental Disorders
BN	Bulimia Nervosa
BED	Binge-Eating Disorder
BMI	Body Mass Index
CSF	Cerebrospinal Fluid
CBT	Cognitive Behavioural Therapy
FBT	Family-Based Therapy
EDE-Q	Eating Disorder Examination Questionnaire
SCOFF	Questionnaire to assess an eating disorder
HRQoL	Health-Related Quality of Life
GRP	Guideline Relapse Prevention Anorexia Nervosa
GRADE	Grading of Recommendations, Assessment, Development and Evaluations
ASD	Autism Spectrum Disorder
HRI	Human-Robot Interaction
AI	Artificial Intelligence

Table of Contents

Introduction

The last decades have brought forward a remarkable degree of technological change. Technology is changing the face of everyday life, and there is no sign that the relentless march of technological progress will slow down. Technology is becoming increasingly smart and is moving towards a point where conventional notions on who or what is 'thinking' are challenged. Many machines nowadays appear to be 'thinking'. A particularly interesting group of 'thinking machines' are robots. Robots combine the best of both mechanical engineering and artificial intelligence (AI) programming. This results in machines that have both a physical embodiment that is able to interact with the world, and an internal component that allows them to do so in a thoughtful manner. This powerful combination is leading to a fast growth of increasingly far reaching appliances that robots take over. This naturally raises many concerns as well, for example regarding the responsibility in the case of adverse events caused by robots. Therefore, a crucial question for the current generation is:

- *How can we enhance the likelihood of a desirable outcome for the proliferation of robots in our society?*

Perhaps the most interesting and challenging introduction of robots lies in social domains where robots interact and communicate with human users. An example of an area in which these social robots are already implemented, is the field of healthcare. Recently, there have been advancements in the introduction of so-called Socially Assistive Robots (SARs) in healthcare. SARs provide assistance to human users through social interaction. Depending on the type of end user and their prescribed therapeutic treatment, SARs are typically designed to fulfill a specific role in the treatment, for example as a companion, play partner, or coach (Rabbitt, Kazdin, & Scassellati, 2015).

As additions to classical treatments, SARs have various advantages. First of all, their appearance and their functionalities can be customized to fulfill the needs of a specific target group. Their ability to engage with people in both a social and emotional way is used to target both physical and psychological needs of patients. Furthermore, SARs are thought to improve the quality and accessibility of mental health care (Rabbitt et al., 2015). An increasing share of patients gets treated in an ambulant setting. SARs could therefore be proved to be an effective and cost-efficient addition to existing treatments.

SARs have already shown promising results in treating geriatric patients with dementia, children with autism spectrum disorder (ASD) and patients that suffer from depression. Several studies showed that SARs could help people to improve their cognitive abilities, social interaction skills and coping strategies. Moreover, they can reduce feelings of loneliness and improve the quality of life in certain groups of patients (Gustafsson, Svanberg, & Müllersdorf, 2015; Moyle et al., 2014).

Whereas many studies investigated the effect of SAR on children and the elderly, little is known about the effect on adolescents and adults. An example of a mental disorder that affects adolescents and young adults is anorexia nervosa (AN). AN is a serious eating disorder that affects both physical and psychological health. It is characterized by the inability to maintain a body weight at or above a minimally normal weight, an intense fear of becoming overweight and a distorted self-image (American Psychiatric Association, 2000). AN has a prevalence of 0.4% and mostly affects young women between the age of 15-19 years old, yet 10-25% is male. Current treatment consists of medical, nutritional and psychological interventions, is not always effective and usually takes years, with approximately 20% of the patients develop a chronic course, accompanied with a low quality of life. It is conceivable that some tasks, such as the monitoring of weight, food intake and vital signs, as well as more social parts of the treatment could be performed by an SAR.

AN can be viewed as a particularly challenging case for social robots designed to deal with mental conditions due to its severity and the complexity of the treatment. In this sense, AN can be considered a crucial case: design choices that can hold here are more likely to be generalizable for other psychiatric conditions.

The challenges related to developing a possible SAR for anorexic patients, combined with the challenges and opportunities that technological progress brings are the motivation behind conducting this study. We argue that previous studies and existing SAR applications have thus far lacked the broad scope that is necessary to tackle challenges that go beyond the scope of a single field such as robotics. Existing research on SARs has only covered children and aging populations thus far in healthcare. We therefore aim to expand possible target groups. Moreover, SARs are often designed and implemented from the top down, rather than in a more bottom-up, stakeholder-driven approach. This leads to a mismatch between what patients need and how these needs can be met by the SAR. Additionally, the advance of robots brings forward ethical concerns, legal concerns over problematic current notions of product liability, and finally distributary concerns over how robots and the surplus value they generate will be divided among society. In contrast to existing studies, we would also like to address the question of whether SARs in healthcare are desirable at all. This leads to the following guiding question:

How could and should a socially assistive robot be implemented in the treatment of anorexia nervosa?

Aim of this study

This study provides first insights into the factors that can contribute to the functioning of SARs in the treatment of AN. In doing so, we hope to establish the most optimal design for social robots for

treating patients with AN. Results of this study will provide practical and normative guidelines that facilitate successful implementation of SARs in the treatment of AN.

Approach

This report takes an interdisciplinary approach, which is a "widespread mantra" (Klein, 2007: 117) for conducting academic research where various perspectives are combined in the study of a broad topic. Members from our think tank have backgrounds in medical sciences (Dylan, Henssen, Rutger Meijers, Vika Shimanskaya), medical biology (Fenja Schlag), philosophy (Marco Dessi, Sophie Horsman), political science (Luuk Schmitz), artificial intelligence (Sophie Horsman), psychology (Annika Schiefner, Ricarda Weiland), psycholinguistics (Annika Schiefner), and neuroscience (Fenja Schlag, Ricarda Weiland). Although combinations of different backgrounds are thought to enhance problem-solving capacities of research it might come at the cost of the possible detriment of exploring more fundamental questions (Klink and Takema, 2012: 12). Another fundamental trade-off in interdisciplinary research exists between the possibility to tackle broad topics on the one hand, and the potential to lose the depth of intradisciplinary discussions on the other hand (Kanakia, 2007). Moreover, when the perspectives are not properly integrated, a study runs the risk of different perspectives talking past each other (Kanakia, 2007).

These issues have been taken into account in this research project by following a two-step approach. In the first step, the perspectives of each discipline are explored in depth: At first, we provide extensive in-depth discussions about three areas crucial to our research question: (i) ethical, legal and social implications of SARs in healthcare in general, (ii) AN and its current treatment options and (iii) existing literature of SARs with other target groups. These three in-depth analyses, will then be integrated to discuss the role that an SAR could and should take in the therapy for AN patients. Secondly, we conducted interviews and discussions with relevant stakeholders in this matter, including AN patients, healthcare professionals, policy makers, ethicists and experts on AI. Results from these interviews were analyzed and integrated with the findings from the first part. This allows us to combine the advantages of in-depth intradisciplinary analyses and broad interdisciplinary discussions.

This report is addressed to our client SingularityU The Netherlands, a think-tank that is concerned with maximizing the potential of technology to have a positive impact on society. SingularityU The Netherlands seeks to achieve this by raising awareness of the impact and opportunities that technological change will bring, and by functioning as a nexus for dialogue between citizens, corporations, and the government. Ultimately, we hope to provide a number of insights and recommendations that help our client in furthering that goal.

1. Theoretical Background

a) Ethical, Legal, and Social Framework in the Design of Socially Assistive Robots in Healthcare

The advance of technology goes hand in hand with exciting opportunities on the one hand, and concerns and questions from those involved and by external observers on the other hand. No emerging field of technology presents as many opportunities and potential concerns as robotics. Thinking about possible future roles for robots seems to reveal deep-rooted and thought-provoking concerns about how robots could change our society. Of course, this is not to downplay the many exciting opportunities that robots will provide. Robots could perform or assist with tasks previously impossible or very dangerous. They can make our lives more efficient and provide us with more time for leisure and recreation. However, the prospect of more capable robots poses questions of crucial ethical, legal, and social relevance. For instance, when we assume increasingly capable and responsible robots, how much autonomy[1] should we grant them? Should a set of values or even morality be programmed into robots to inform their decision-making processes, and would we even want robots to have such advanced moral responsibility? These questions are becoming increasingly salient and should be well thought-through before Socially Assistive Robots (SARs) take on a more serious role in society. We can go one step further and argue that perhaps questions of morality, autonomy, and responsibility ought to precede those of design and implementation of SARs. After all, deep-rooted issues relating to how robots could challenge our own sense of humanity seem to suggest the need for a more comprehensive and structured investigation of these concerns.

It is indeed in this evaluative spirit that the guiding question of our project can be addressed. The healthcare domain is a prime example of a field in which robots are on the verge of being introduced on a large scale (Kachouie, Sedighadeli, Khosla, & Chu, 2014). In particular, the focus on SARs gives rise to some fundamental ethical, legal, and social concerns. The most widely debated ethical issues concerning care robots are the fear of human replacement and that care will be centred around efficiency-maximization at the expense of the needs of the person behind the patient (Royakkers & van Est, 2015; Stahl & Coeckelbergh, 2016). From a legal point of view, the question of who is responsible for the robot's actions is a poignant issue that needs careful examination. Finally, the distribution and accessibility of SARs on the societal level should also be taken into account. We believe that these concerns should be addressed by starting at the design process, since important choices will already have to be made at this early stage.

[1] Differing definitions of autonomy exist. An in-depth discussion of the concept follows later in the chapter. For now, it is sufficient to know that we approach autonomy from a robotics point-of-view, i.e., autonomy is seen as the capacity and extent to which robots can perform unsupervised actions (Haselager, 2005).

The aim of this chapter is to hold a literature-based discussion to provide a framework that stimulates a design process for SARs in healthcare where ethical, legal, and social (ELS) concerns are taken into account. It is important to note that such a framework is not meant to provide any definitive answers to the problems at stake. Rather, it is meant as a tool for systematically asking the relevant questions when dealing with SARs in healthcare. The issues in the ELS framework will be addressed in said order, thus starting with a discussion on ethical issues.

Care-Centered Value-Sensitive Design

Ethicist and robot specialist Aimee van Wynsberghe proposes a framework called Care-Centered Value-Sensitive Design (CCVSD), in which she advocates that ethics should play an important role in the design of robots in healthcare (van Wynsberghe, 2013; van Wynsberghe, 2016)[2]. Such a framework is intended to be applicable not only in retrospect, but also at the beginning of and during the design process. The basic idea behind this framework is that technology can never be free of values. Values are here defined as something desirable, something one wants to have or happen (van Wynsberghe, 2013).

The starting assumption is that technologies embody values, which means technologies are not neutral and thus not only dependent on how the user employs them. Rather, most technologies inherently have tendencies that promote or demote certain values. This could be the result of either thought-through or negligent design choices. An example of an intended effect is when the company Silent Circle designed their product in a way which prevents the tracking or tracing of phone conversations to promote the value of privacy (van Wynsberghe, 2013). Thus, the imposition of constraints on the technology and/or the concession of allowances for the technology can result in the promotion or demotion of ethical values. Therefore, many researchers have concluded that one should design technology in such a way that facilitates the choice and thus the realization of values of ethical importance (Friedman, Kahn, & Borning, 2006). This is the main idea behind Value-Sensitive Design (VSD) approach.

Before designing desirable technologies for the use in healthcare, one should know the values of ethical importance in this domain. In order to identify the morally relevant values, Wynsberghe discusses influential works from the history of care ethics. This provides a better understanding of care in general and the meaningful interactions between caregivers and caretakers. Within the care ethics tradition, care practices play a central role: such practices are the combination of attitudes, actions and interactions between actors in a care context that work together in a way that realises

[2] Previous work of ethicists has mainly addressed ethical concerns after robots have already been introduced (van Wynsberghe, 2016). We believe that incorporating ethics into the design process helps to find the right balance between the beneficial potential of robots in healthcare and taking the related ethical concerns seriously.

care values (van Wynsberghe, 2016). Before planning to introduce a robot into a certain care practice, we should first understand the current practice and how morally relevant issues are currently tackled within this practice.

Crucially, care ethics holds that a holistic perspective on care is paramount. Care should never be viewed as an isolated product to meet standardized needs. In this sense, good care needs to be viewed as a full package that in its entirety meets the needs of the caretaker. Accordingly, a robot should never be designed for the sole purpose of fulfilling a certain task, without taking into account how its task-relevant functions relate to the overall care practice.

According to Wynsberghe (2016), care itself can be seen as a value, since it seems meaningful to recognize the dignity and needs of one another. Furthermore, care encompasses many other values. In order to identify the values of ethical importance in institutional care, she adopts a top-down approach. In particular, she examines the abstract values of the World Health Organisation and how they relate to the more concrete institutional values listed in hospital policies and guidelines. Subsequently, she argues that it is hard to choose amongst the many different interpretations of values that all those institutions provide. Furthermore, there are many (possibly rather obvious) values usually not listed at the established institutions. Drawing on the influential work of Tronto (1993), this is why Wynsberghe (2016) suggests basing the importance of moral values on the multi-layered needs of the patients.

According to Tronto, there are four moral elements that need to be all integrated in any good care practice. These four elements are attentiveness, responsibility, competence and responsiveness. Attentiveness refers to the caregiver's ability to see the changing, unique needs of the patient. Responsibility means that an individual or institution is responsible in replying to the needs of the patient. Competence regards a skilled caregiver who is capable of performing the required tasks and it refers not only to the content of the actions carried out by the caregiver but also to their form. Finally, responsiveness refers to the willing attitude and engagement from the patient's side. According to Wynsberghe, these four elements are, on the one hand, the criteria for the ethical evaluation of a good caregiver and, on the other hand, the starting point for the evaluation of the appropriate use of robots in healthcare.

The interpretation of these core moral elements could vary between different contexts and care practices. For instance, the meaning of competence changes based upon the type of care practice: while in the practice of lifting a patient, it could mean being strong enough to carry out the action, in the practice of prescribing medicine it refers to the knowledge of the type and amount of medication appropriate for the patient. Similar examples of shift in meaning can be easily made for the other core values. Therefore, we should examine every care practice in its specific context and

how the different roles and responsibilities amongst the involved actors are divided within the practice.

This provides us with the following ethical framework that one can use to assess current care practices in a given context (see Table 1). By understanding the different roles and responsibilities of the involved actors within these care practices, we can assess how the core moral values are currently dealt with. When a robot is introduced in a certain care setting, we should evaluate what type of robot would be most appropriate for that context. Finally, one can evaluate the effect of introducing the robot in the given context by assessing whether all the moral elements are still in place and none of them is promoted at the expense of another. In sum, the integration of care ethics and VSD provides the necessary ethical elements to assess the desirability of technology in healthcare.

Table 1

Wynsberghe's ethical framework for the design of robots in healthcare

Context	Hospital vs. nursing home vs. home setting …
Practice	Lifting vs. bathing vs. feeding vs. delivery of food and/or sheets, playing games …
Actors involved	Human (e.g. nurse, patient, cleaning staff, other personnel) and nonhuman (e.g. care room, mechanical bed, wheelchair, mechanical lift, robot …)
Type of robot	Assistive vs. enabling vs. replacement
Manifestation of moral elements	The core moral values should not be demoted by introducing a robot into the care practice. It should not be the case that one of the values gets promoted at the expense of another. <u>Attentiveness</u> The capability of recognizing the changing and dynamic needs of the patient. <u>Responsibility</u> The capability of an individual or institution of being responsible for the needs of the patients. It requires the identification of the appropriate responses to the needs and the delegation to meet them.

	Competence The capability to of executing means/action to fulfil the identified needs in a skilled manner. Responsiveness The capability to engage with the care-receiver regarding the meeting of their needs.

The framework is taken over from Wynsberghe (2016), modified to include what the author meant by the manifestation of moral elements. The rows 'context', 'practice', 'actors involved' and 'type of robot' are intentionally left unchanged to show the reader this framework could in principle be applied to any care setting. In Chapter 4, we will use take the abstract framework and make it more tailored to the treatment of AN patients in a home setting.

Ethical machines

Thus far, we have discussed how one can use the framework as a starting point for designing healthcare robots in an ethical manner. In light of the VSD approach, robots are considered as any other piece of technology with regards to the ethical challenges they pose. However, given the highly interactive nature of SAR systems, it is conceivable that the expected increase in their capacities for interacting with end-users might correlate with their involvement in situations of substantial moral relevance (Sullins, 2006). This means that the more these robots enter dynamic and unpredictable situations, the higher the likelihood that they will be forced to make decisions that go beyond their pre-programmed explicit set of rules. Therefore, it is crucial to specifically problematize the role and possible responsibilities that the robot could have once placed in a care setting. Accordingly, one might ask whether it is preferable not only to consider how we can ethically design machines but also to examine whether we can and to what extent we should design ethical machines (Malle, 2016).

In this regard, considerable research has been conducted on whether it is possible to build machines capable to decide what is right and wrong (i.e. "ethical machines") (Anderson & Anderson, 2007; Powers, 2006; Wallach & Allen, 2008). Although conceiving ethical machines does not seem to lead to any logical contradiction, questions about their realisability are rather difficult to answer. For instance, how can an ethical model be implemented in a robot and who decides which ethical system is to be preferred? An extensive debate exists on these questions, from which no satisfying conclusion can easily be drawn. Ethicists such as Wynsberghe argue that we should avoid such questions altogether and conclude that robots cannot and should not be seen as moral agents (van

Wynsberghe, 2016). Instead, we should rather seek to constrain the robot's decision-making processes (ibid.). Perhaps one solution to the moral agency problem would be to have an entirely "reactive" robot, namely an embodied system which can only respond in a predictable manner to a set of predetermined cues.

However, the very nature of Socially Assistive Robots seems to elude such a sheer reactive design. In fact, since one of the primary functions of SAR systems is to socially interact with their users, the idiosyncratic character of their social partner's behaviour suggests the unsuitability of adopting an entirely reactive paradigm for their design. Therefore, in the context of SARs, we believe it is far too complex to determine explicit constraints that will prevent the robot from making any morally relevant decision. This does not seem to be merely a problem of complexity from a design perspective: on the one side, we want SAR systems to be able to socially interact with its user in a fluid (perhaps human-like) manner and, on the other side, we do not want to grant them any freedom to choose to behave in a way that we do not approve.

Autonomy and human control

The preceding discussion has pointed to the need of carefully examining the nature of interactions between robots and their environment. In this sense, a reflection on the possible normative dimension of robotic actions and interactions seems to require a proper consideration of the concept of autonomy. This is motivated by the fact that the moral ability to distinguish between appropriate and inappropriate behaviour logically necessitates the capacity to make choices.

In philosophy, autonomy entails the capacity to choose goals for oneself (Haselager, 2005). In light of this definition, robots have been traditionally considered to lack any kind of autonomy since they do not have the capacity to choose and act upon their "own" goals. However, this conclusion has been highly disputed by questioning human autonomy itself and has ultimately lead to the problem of free will (Haselager, 2005). In particular, the question of whether human actions are the product of autonomous or free agency or are merely predetermined by genetics and/or nurture seem to elude any straightforward answer. Thus, for reasons of space, it might be more fruitful for our purposes to steer the discussion on a deflationary and more technical notion of autonomy which originated in the field of artificial intelligence (AI).

In AI, autonomy refers to "the capacity to operate under all reasonable conditions without recourse to an outside designer, operator or controller while handling unpredictable events in an environment or niche"[3] (Haselager, 2005). Since this definition does not draw on the notion of goal

[3] In Artificial Intelligence there is a research paradigm entirely focused on the study of autonomous agents. Autonomous agents can be both software agents (such as online chatbots) and hardware robots.

ownership, which can be differently interpreted on the basis of one's own intuitions, it seems to be more workable for our case. Essentially, this interpretation of autonomy is centred on the question of how much human intervention is needed for a robot to be functional in a dynamic environment. Accordingly, one could assess the autonomy of robots on a continuous scale from not autonomous to fully autonomous.

For descriptive purposes, it might be fruitful to come up with an ideal typical distinction between several types of robots based on the AI definition of autonomy. On the bases of these ideal types, we shall later assess how the problem of moral accountability and its related legal implications might be different in each case.

1. The Inflexible Robot

This type of robot is completely pre-programmed and the programmer determines beforehand how the robot should respond to certain fixed cues in its environment. Since this type of robot cannot deal with unpredictable, dynamic situations, it lacks autonomy. Thus, it is "inflexible" simply because it cannot go beyond its pre-programmed set of behavioural rules.

2. The Marionette-like Robot

The behavioural potential of this robot can be understood in an analogous way to that of a marionette. Similar to a person controlling a marionette, the makeup of this robot is not completely fixed in the sense that another actor besides the programmer is involved to steer the behaviour of the robot. In this analogy, a therapist could play the role of the marionette player and thus influence the behavioural pattern of the robot. This could either be directly in a Wizard-of-Oz[4] setting, or indirectly by having impact on changes in the software during the therapy trajectory. What really matters to make a marionette-like robot as such, is that someone (e.g. a therapist) is involved either in the control or the modification of the clinically relevant behavioural pattern of the robot. With the guidance of the therapist, the robot might be modified to become more suited for particular clinical needs of her/his patient. In this case, the robot is not completely autonomous since it needs human intervention to deal with certain dynamic and unpredictable scenarios.

3. The Flexible Robot

This type of robot can be described as "flexible" since it attempts to fulfil its goals without following a strict set of predetermined rules. Although it is the programmer who decides which core goals shall inform the behaviour of the robot, the robot itself is enabled to reason how to best achieve these goals. Therefore, it is not only reactive towards its environment but also pro-active in the sense that it will flexibly work towards the realization of its goals. To do so, the robot will need to make online

[4] A Wizard-of-Oz setting refers to a commonly employed technique in Human Robot Interaction research in which a person remotely operates a robot, controlling one or multiple aspects of the robot's behaviour, such as its movement, navigation, speech, gestures etc. (Riek, 2012).

decisions[5]. For instance, the robot should autonomously decide whether to initiate a conversation, bring coffee to a thirsty patient, or play a cheerful song to affect the patient's mood. This robot can be thought of as autonomous in the sense given its ability to act in unpredictable environments without relying on the control of a human operator.

Obviously, real-world implementations of SAR systems can fall somewhere in between these three ideal types of robots. To a certain extent, autonomy seems to be necessary since the social nature of their interactions requires these robots to cope with unpredictable situations. At the same time, it also seems legitimate to include some normative constraint on the behavioural potential of the robot so that it cannot perform actions that we deem to be undesirable. Although reality is always more complicated, we believe that these three types of robot could help to shed some light on the concept of autonomy and to assess the moral and legal responsibilities of the actors involved.

Distribution of responsibility

As previously noted, ethical discussions on autonomy also problematize the notion of responsibility. Questions such as who exactly and in what circumstances should be responsible for the robot are widely debated not only for their obvious ethical implications but also for their legal relevance. The discussion of the three aforementioned ideal types of robots points to a complex set of actors responsible for the functioning of the robot. Since unexpected behaviour in real-world scenarios is all-but inevitable[6], we should therefore have a clear conception of who and in what circumstance is responsible for the behaviour of the robot.

The way in which responsibility is distributed is tangent to the type of robot employed. As a rule of thumb, the more it is capable of unsupervized action in a dynamic environment, the more complex the problem of distributing responsibility becomes. In relation to our framework, which type of robot would be most desirable for a healthcare related SAR? A trade-off seems to exist between maintaining control over the robot and enacting satisfactory social interaction. Although the high behavioural predictability of the 'Inflexible Robot' seems to greatly simplify the question of who is responsible when it malfunctions, due to its pre-programmed nature it is unlikely to achieve much social interaction at all (Sullins, 2006), thus limiting its potential as an SAR On the contrary, the 'Flexible Robot' is fully autonomous in the sense that it will always perform its actions unsupervized. This seems both unlikely to work for and undesirable to have in SARs in a healthcare context. A reasonable motivation for this is the intuition that the vulnerability and the complex needs of patients might be better treated by a human doctor than by a robotic platform. Moreover, the

[5] Online decisions are decisions on the course of actions at the runtime of the program that cannot be completely foreseen by merely viewing the software code.

[6] Even fully pre-programmed robots will exhibit unexpected behaviour in the form of software bugs.

ethical part of our framework requires the robot to be an addition rather than a substitution for existing health care practices.

These considerations push the discussion more into the direction of a 'Marionette-like Robot', whose actions can for example be controlled in a Wizard-of-Oz style, creating the impression of autonomous behaviour whilst actually being controlled by an unseen human. When the robot is not supervized, it could fall back to a simpler, though still somewhat dynamic, set of actions. Such an adaptive implementation of autonomy seems to form the best middle-ground in managing responsibility and achieving social interaction between patient and robot[7]. This middle-ground seems to go a long way in avoiding the robot from being a moral agent, preventing complex philosophical and legal discussions about who is responsible when the robot makes a mistake[8]. Nevertheless, the exact distribution of responsibility remains complex (Malle, 2016). As long as the robot is limited in its ability for unsupervized and dynamic action, however, it falls within normal product liability law. This means that only those involved in designing, building, shipping, and selling the robot can be held liable in case the robot malfunctions. When the robot becomes more adaptive to the needs and preferences of the end user, responsibility will also be shared with end-user(s) (ibid.).

In conclusion, an SAR in healthcare contexts can best strike a middle-ground between functionality for responding to unexpected cues and control by humans to prevent it from acting in an undesirable or even immoral manner. Table 2 summarizes the points of this section.

Table 2

Legal issues in the ELS Framework

Moral Agency and autonomy	A balance should be struck between the autonomy necessary for socially interactive robots and the control that human agents have over the actions of the robot.
Distribution of Responsibility	Responsibility should be conferred to the actor(s) or institution(s) that have caused the misuse or mistakes of the robot.

[7] This is not to say that a marionette-like robot will never make online-decisions. On the contrary, such online decisions might be necessary to enact or strengthen the element of social interaction.

[8] The threshold for a legally responsible robot is its capacity to be a moral agent (Asaro, 2007). This is arguably not met by a marionette-like robot with limited online decision making capabilities, preventing complex philosophical and legal discussions on conferring responsibility to the robot.

Socio-economic issues

The framework that has been developed thus far offers a firm basis to tackle ethical and legal issues on the micro-level (i.e. the actors directly involved with the SAR). However, in assessing what a desirable outcome of the robotic revolution is, one also needs to go beyond the micro-level, and consider the introduction and proliferation of SARs at the societal level. Here, there are several clear-cut challenges that the introduction of SARs would present. Firstly, such a robot would require adjustment costs for those working in an environment with SARs since affected employees will have to attain new skills and work routines. Moreover, such a robot would ostensibly be an expensive device, meaning that not everyone will have access to the robot without some form of compensation. Finally, the way in which additional surplus value generated by SARs is distributed also requires careful consideration. The aim of this section is to discuss these issues to inform the socio-economic elements of the framework.

Issues of employment

Ever since the Luddites destroyed weaving machinery as a form of protest against the socio-economic consequences of the First Industrial Revolution, the relationship between technological change and conditions of employment has been hotly debated in macro-economics. In this sense, the anxiety that robots might turn some of us from unemployed to unemployable is nothing new. However, the topic is no less relevant, exemplified by a recent eurobarometer poll showing that 73 percent of Europeans worry that robots might steal their jobs (European Commission, 2015). This section serves to provide a concise overview of the debate on the effects of robotics on employment, separating facts from fiction.

The debate on robotics should be contextualized in the broader debate about the effects of technological change on employment. The crucial question here is whether this time differs from previous cycles of technological innovation. Since the 1990s, consequences of computer-based technological changes have become a topic of interest for economists (Levy & Murnane, 2003). Back then, the Skill-Biased Technological Change (SBTC) hypothesis was developed to explain the recent shift that has favoured high-skilled jobs over low-skilled jobs. Prior to the computer revolution, technological change was seen as factor-neutral, meaning that the effects of a new technology were expected to apply equally to the factor employment (L. F. Katz, 1999). However, what happened in the 1980s and 1990s directly contradicted this expectation: technological changes induced a bias in favour of high-skilled labour (Berman & Machin, 2000). The logic behind this is that productivity for high-skilled labour is more positively affected by recent technological changes than low-skilled labour. This in turn increases productivity for high-skilled jobs, and can come at the expense of work previously done by low-skilled jobs (L. Katz, Autor, Ashenfelter, & Card, 1999).

Although the SBTC hypothesis has been successful in explaining the first wave of the computer revolution, the mid 2000s marked the arrival of an issue which seems to elude the SBTC: job polarization. Job polarization is a phenomenon where middle-skilled jobs are displaced by both high-skilled and low-skilled labour (Goos, Manning, & Salomons, 2014). Concurrent with this development is the increasing gap in wages between low- and high-skilled labour (Abel & Deitz, 2012). This shift in the labour market development requires a new understanding of these recent developments. Goos et al. (2014) provide such an insight with their Routine-Biased Technological Change (RBTC) hypothesis. The authors argue that recent technological changes have skewed towards displacing labour with a high intensity of routine-based tasks (e.g., accountancy, financial analysis). The consequences are twofold: On the one hand, high-skilled labour benefits from this development in a similar way to SBTC: not only the technology displacing middle-skilled jobs makes their jobs more efficient, it also increases productivity and thus demand for high-skilled labour. Additionally, falling costs of technology also make it more attractive to invest in technology that replaces routine-based labour (Abel & Deitz, 2012). On the other hand, low-skilled service sector labour such as waiters and healthcare aides are protected from most of these technological changes, since physical proximity and face-to-face contact continues to matter for these types of professions (Abel & Deitz, 2012). The RBTC hypothesis also marks the moment that robots come into the equation. Contemporary robots and AI-systems are at their best in replacing routine-based jobs (ibid.). This makes robots and AI-systems a major driver of RBTC and, as a consequence, job polarization.

Based on this discussion, it now becomes possible to address the fear of 73% of European citizens. A review of recent literature on the topic yields mixed results. Muller et al. (2017) argue that the long-term economic effects of robotization will not differ from the effects of technological change since the First Industrial Revolution. In other words, they argue that some jobs will eventually disappear, but this displacement will create additional surplus value and demand for new jobs. However, not all economists agree. The most extreme case is exemplified by Frey and Osbourne's (2013) estimate that in the next two decades, 47% of jobs in advanced economies will be at risk of being automated. Similar claims have been made by Bowles, 2014; Brzeski & Burk, 2015; and Pajarinen & Rouvinen, 2014.

Nevertheless, Muller et al. (2017) and Bonin et al. (2015) dispute these claims by arguing that proponents of the job-losses expectation overestimate the automation potential of most types of work by deploying a flawed methodology[9].

[9] Specifically, it is argued that even jobs at high risk of automation will likely only be partially automated, and hence not completely disappear (Bonin et al, 2015). Additionally, the expected capabilities of robots are based on subjective assessments of experts, who tend to overstate the potential future capabilities of technology

Where does this discussion leave the case of socially assistive robotics? Surprisingly, there are no studies that specifically examine the relationship between SARs and effects on employment. However, a number of reasonable inferences can be drawn from the previous discussion. Firstly, for the considerable future, socially assistive robotics will form a complement rather than substitution to existing labour (Dahl & Boulos, 2013), limiting the potential for human displacement. However, as future robots will become more capable, they might become viable replacements for humans in healthcare, at least from a competence point of view. Consistent with our ethical framework, however, we suggest that robots should be an addition, not a substitute, regardless of its competences. In spite of the limited expected effects on employment, the introduction of SARs will require adjustment costs for those whose day-to-day work will change, since the development of new skills to properly handle SAR systems are to be expected. We believe that such adjustment costs should be shared according to the principle of solidarity, since this is arguably the fairest way to mitigate these costs[10].

Issues of (re-)distribution

Besides the expected adjustment costs that SARs will bring to the labour market, another broad socio-economic issue should be discussed, namely the issue of (re-)distribution. Widespread introduction of socially assistive robotics in healthcare has distributary implications in at least two ways. Firstly, there is the issue of distributing the robots amongst potential benefactors. Since this technology will be arguably very expensive, not all potential users will be able to afford one. In this respect, we believe that when a clear added benefit to the treatment of a patient exists, a difference in socio-economic status should not form any form of obstruction towards accessing the SAR.

Besides the clear-cut re-distributary effect that accessibility of the SAR could have, a more complex distributary issue should also be taken into account: the distribution of surplus value generated by the SAR. As is argued by Muller et al (2017), the biggest economic question regarding robotization is how to handle the surplus value that the new technology will generate. Boosts to productivity, and in our case: healthier patients, will generate additional economic growth. We now find ourselves at an important crossroad between status quo and a fairer distribution of added wealth. Ceteris paribus, surplus value generated by selling (socially assistive) robotics would flow

(Autor, 2014). Finally, additional surplus value extracted from an increase of productivity has the potential to increase demand for new types of labour (Arntz, Gregory, & Zierahn, 2016).

[10] Practically speaking, an example could be that the adjustment costs of introducing an SAR for the treatment of a specific condition could be shared by the therapist and supporting personnel in the hospital or clinic, the patient, the insurance company, and finally the hospital itself. Contributions could be made not only in monetary form, but also in the form of time to acquire the required new skills.

directly to corporations responsible for designing and selling them, increasing their profits, and adding to the recent trend where in spite of having more profitable corporations, economic inequality actually increases (Stiglitz, 2012).

Fortunately, a more attractive alternative is available. Several authors (e.g. Smith, 2017; Varoufakis, 2017) have discussed the possibility of instigating a fund where (parts of) the surplus value generated by robots and AI systems can be used to provide financing schemes for the costs of the robotic revolution. Such a fund could be created by imposing a tax on the profits generated by robotics. The fund could be used to share liability, pay for adjustment costs, and to more evenly distribute the benefits of SARs among society. However, such a fund comes with methodological difficulties[11], and, therefore, it seems more reasonable and prudent to conclude the discussion by simply calling for a comprehensive discussion on how robots, and technological innovations in general could eventually benefit society at large.

Conclusions

This chapter was an attempt to explore the most important ethical, legal, and social aspects of the robotic revolution in the healthcare domain. Based on this discussion, normative choices can be made to shape the direction of this revolution in a way that we consider desirable. Table 3 contains a summary of the most important conclusions, recommendations, and points of discussion within our framework. Topics are grouped based on the sequence of discussion throughout the chapter. The contents in the framework should be seen as something in between values that can be maximized and recommendations for careful consideration of certain issues. We believe that this framework can provide the basis for steering the robotic revolution in such a way that the likelihood of a desirable outcome could be enhanced.

Table 3:
Complete ELS Framework for the design of SARs in Healthcare

Context	Hospital vs. nursing home vs. home setting …
Practice	Lifting vs. bathing vs. feeding vs. delivery of food and/or sheets, playing games …
Actors involved	Human (e.g. nurse, patient, cleaning staff, other personnel) and nonhuman (e.g. care room, mechanical bed, wheelchair, mechanical left, robot …)
Type of robot	Socially Assistive Robot (SAR)
Manifestation of	The core moral values should not be demoted by introducing a robot into the

[11] For instance, it might be difficult to demarcate between a robot and robotic technology included in conventional technology.

moral elements	care practice. It should not be the case that one of the values gets promoted at the expense of another. Attentiveness The capability of recognizing the changing and dynamic needs of the patient. Responsibility The capability of an individual or institution of being responsible for the needs of the patients. It requires the identification of the appropriate responses to the needs and the delegation to meet them. Competence The capability to of executing means/action to fulfil the identified needs in a skilled manner. Responsiveness The capability to engage with the care-receiver regarding the meeting of their needs.
Moral agency and autonomy	A balance should be struck between the autonomy necessary for socially interactive robots and the control that human agents have over the actions of the robot.
Distribution of Responsibility	Responsibility should be conferred to the actor(s) or institution(s) that have caused the misuse or mistakes of the robot.
Socio-economic consequences	Issues of adjustment Adjustment costs for those employed in professions disrupted by SAR should be proportionally shared by all benefacting parties. Issues of distribution All patients, regardless of socio-economic status, should have equal access to the SAR when it is envisioned to have an added benefit to their treatment process. Surplus value generated by the robot should benefit society as a whole.

b) Anorexia Nervosa: a Medical and Psychological Overview

Before an assessment can be made about the development and implementation of SARs in the treatment of AN, further knowledge about the disorder, its current treatment and patient's features is essential. This chapter will give an overview of AN. Firstly, the definition and characteristics of AN are discussed, followed by the medical complications that can occur. Next, the psychological and physiological factors that may contribute to the etiology of the disorder are explained. Consequently, there is a description of the treatment options and the prognosis. At last, the economic impact on

society is mentioned. In a following chapter, this information is used to discuss which implications this may have for an SAR as a future part of the treatment of AN.

AN is an eating disorder with severe impact on both a physiological and psychological level. It is characterized by the inability to maintain a body weight at or above a minimally normal weight, an intense fear of becoming overweight and a distorted self-image as described in The Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2000). The failure of maintaining a healthy weight can be due to a restricted calorie intake, an elevated level of exercise, or purging by self-induced vomiting, misuse of laxatives or diuretics. Eventually, also binge eating occurs. When this behaviour evolves in extreme starvation, its effects on the bodily state can reach life threatening dimensions. Moreover, patients often have psychiatric comorbidities and suicide reveals an increased prevalence in AN patient populations with 1,5%, which is a relative risk of 35 compared to the healthy population (Preti et al, 2011). With a mortality rate of 5-10%, AN is classified as one of the most fatal mental disorders (Arcelus et al, 2011;(American Psychiatric Association, 2000; Hoek, 2006; Steinhausen, 2002). The incidence of AN is approximately 8 per 100 000 people per year in general population, though the highest incidence is observed in women between 15 and 19 years which is 270 per 100.000 per year. The lifetime prevalence of AN is 2,2%. Although most of the patients are female, 10-25% of the patients are male (Hoek, 2006; Keski-Rahkonen et al., 2007; Smink, van Hoeken, & Hoek, 2012) Hudson et al, 2007,

DSM 5

Besides anorexia nervosa, the DSM-5 classifies two other categories of eating disorders, which are bulimia nervosa (BN) and a residual category including binge-eating disorder, pica, rumination disorder and avoidant/restrictive food intake disorder.

DSM-5 Diagnostic criteria for Anorexia Nervosa are:

A. Restriction of energy intake relative to requirements, leading to a significantly low body weight in the context of age, sex, developmental trajectory, and physical health. Significantly low weight is defined as a weight that is less than minimally normal or, for children and adolescents, less than that minimally expected.

B. Intense fear of gaining weight or of becoming fat, or persistent behaviour that interferes with weight gain, even though at a significantly low weight.

C. Disturbance in the way in which one's body weight or shape is experienced, undue influence of body weight or shape on self-evaluation, or persistent lack of recognition of the seriousness of the current low body weight.

Specify whether:

Restricting type: During the last 3 months, the individual has not engaged in recurrent episodes of binge eating or purging behaviour (i.e., self-induced vomiting or the misuse of laxatives, diuretics, or enemas). This subtype describes presentations in which weight loss is accomplished primarily through dieting, fasting, and/or excessive exercise.

Binge-eating/purging type: During the last 3 months, the individual has engaged in recurrent episodes of binge eating or purging behaviour (i.e., self-induced vomiting or the misuse of laxatives, diuretics, or enemas).

Specify if:

In partial remission: After full criteria for anorexia nervosa were previously met, Criterion A (low body weight) has not been met for a sustained period, but either Criterion B (intense fear of gaining weight or becoming fat or behaviour that interferes with weight gain) or Criterion C (disturbances in self-perception of weight and shape) is still met.

In full remission: After full criteria for anorexia nervosa were previously met, none of the criteria have been met for a sustained period of time.

Specify current severity:

The minimum level of severity is based, for adults, on current body mass index (BMI) (see below) or, for children and adolescents, on BMI percentile. The ranges below are derived from World Health Organization categories for thinness in adults; for children and adolescents, corresponding BMI percentiles should be used. The level of severity may be increased to reflect clinical symptoms, the degree of functional disability, and the need for supervision.

Mild: BMI ≥ 17 kg/m2

Moderate: BMI 16–16.99 kg/m2

Severe: BMI 15–15.99 kg/m2

Extreme: BMI < 15 kg/m2

Whereas AN is a relatively rare disorder, BN affects 2-3% of females in the United States (Harrington et al, 2015). The age and gender distribution of AN and BN appears to be similar.

Medical complications

As a result of severe long term malnutrition several medical complications can occur, which can be observed in various organ systems and lead to a generally bad medical condition of patients with AN. The number of affected organ systems and the severity of complications are correlated with the degree of weight loss. Furthermore, vitamin deficiencies and metabolic disturbances due to the reintroduction of nutrition in starved patients are commonly observed in AN. Due to the severe

malnutrition, protein and fat catabolism is induced that leads to loss of cellular volume and atrophy which eventually leads to deterioration of organ functioning. AN patients with extreme loss of body mass and fat tissue lose strength and endurance and move more slowly (Harrington et al, 2015). Cardiovascular complications, including myocardial atrophy, heart failure and arrhythmias can be potentially fatal. Gastrointestinal complications include epigastric pain, bloating sensation and haemorrhoids and rectal prolapse as a result of laxative abuse. Severe hypoglycemia may lead to epileptic seizures. Bone marrow changes and cytopenia, including anemia, leukopenia and thrombocytopenia, are frequently observed in AN and are reversible with weight restoration and nutritional rehabilitation. AN is associated with multiple neuroendocrine abnormalities, including hypothalamic-pituitary axis dysfunction. This results in amenorrhea, osteoporosis and hypo- or hypernatremia. Decrease in bone mineral density leads to overuse injuries and stress fractures.

More than half of all deaths in patients with AN is due to these complications. Most of the complications are treatable with weight gain, though some of them may not be completely reversible, such as osteoporosis. The complications in male and female patients are similar with the exception that males start with a lower reserve percentage of body fat and a higher lean muscle mass, allowing them less weight loss before the onset of ketosis and protein breakdown (Mehler & Brown, 2015).

Pathophysiology

The pathophysiology of AN is multifactorial and still poorly understood. Risk factors for the development of the disorder are gender, cultural factors idealizing an ultra-thin body image as feminine beauty, family prevalence of the disorder and a predisposition to personality traits such as perfectionism, obsessionality and anxiety (Garner and Keiper, 2010). The experience of adverse life events as abuse, neglect, sexual abuse and experiences as bullying, criticism or teasing often contribute the onset of the disorder.

A complex construct of psychological aspects underlies the development of behavioural symptoms of anorexia nervosa. Generally, patients with AN tend to have negative core beliefs about themselves that come along with personal withdrawal and self-preoccupation (Garner and Bemis 1982). Further, they may experience a sense of depression, helplessness, and a loss of control (Garner and Keiper, 2010).

In initial stages of AN, patients often discover that dieting and the body figure are one of the things they can exercise control on. From these experiences, they are able to derive a gratifying sense of power, which might be further strengthened by the support from the environment giving compliments on losing weight (Garner and Bemis 1982). Soon, a phobic orientation toward food and weight gain is developed giving thinness a predominant importance. Constructed on reasoning errors

and disturbed information processing, anorexic believes develop to hyperactive cognitive sets that become functionally autonomous from outside stimuli (Beck,1970). Both positive and negative reinforcement contingencies help to maintain the state. Positive reinforcement contingencies are feelings of accomplishment, pleasure, power, and pride that makes patients cling to an anorexic identity (Garner and Keiper, 2010). Considered as negative reinforcement contingencies are roles of the disorder in avoiding unpleasant feelings as negative emotions, conflict, and anxiety. Eventually, the avoidance of a mature body shape might also negatively reinforce the disorder as this might be associated with fears towards developmental expectations for patients at the verge of puberty (Garner and Keiper, 2010).

Several behavioural and social traits have been put forward as premorbid vulnerabilities that influence the severity of the disorder (Treasure and Schmidt 2013, Garner and Keiper, 2010). These predisposing factors are obsessive compulsive features, anxiety, social communication impairments, and emotional disturbance. Obsessive compulsive features can be characterized by a preoccupation with details, orderliness, perfectionism, mental and interpersonal control and a high behavioural rigidity. This results in cognitive processing styles with poor set shifting and a weak central coherence (Treasure and Schmidt 2013). Poor set shifting manifests itself by the inability to alter perspectives in thinking and behavioural routines. This is most likely influenced by a weak central coherence, which leads to patients being unable to consider a more general picture of their situation (Treasure and Schmidt 2013). In AN, persistence in rules and rituals related to details in weight control and inflexibility lead to the success in achieving the goal of persistently losing weight. More than that, these characteristics might even be necessary to develop and maintain the anorexic state. As Bruch stated: “Developing the anorexic state [...] demands active and alert attention from its victim. It is not just a habit they cannot break; to maintain it, it requires suffering and continuous hard work (Bruch, 2001).”

Several anxiety disorders as simple phobia, social phobia, and panic disorder are associated with AN (Halmi, Eckert et al. 1991, Perdereau, Faucher et al. 2008). The anxious avoidance of emotions, especially in social settings, is another frequently observed predisposing factor often manifesting itself in the social isolation of the patient (Treasure and Schmidt 2013, Garner and Keiper, 2010). Closely connected to the social isolation and the anxious avoidance are impairments in social communication of patients with AN, revealing itself in difficulties in reading intentions, mental states and emotions of both others and oneself. The disturbance in the perception of emotions becomes apparent in difficulties in labelling and expressing the legitimacy, desirability and acceptability of inner experiences. This also includes the mal-interpretation of sensations perceiving sentiments as pleasure, relaxation, or sexual drives as being wrong, frivolous or threatening (Garner and Keiper, 2010).

Consequences of starvation play a pivotal role in the psychological state of anorexia patients. Many features of the disorder as emotional distress, social withdrawal, anxiety, depression, rigidity, poor concentration and concrete thinking can also be understood as consequences of semi-starvation (Treasure and Schmidt 2013), which could also be observed in normal volunteer groups exposed to periods of sustained weight loss (Garner, 1997). As a result, starvation and malnutrition may lead to a vicious circle further enhancing the psychological features crucial for the maintenance of the anorexic state.

AN has traditionally been considered a disorder with social and cultural etiology, rather than a developmental disorder (Nasser, Katzman, & Gordon, 2000). Starting from 1990's, new evidence from twin studies emerged showing a more complex genetic and neurobiological component(Collier & Treasure, 2004).

Current literature describes various structural and functional changes in neurobiology in the AN course. Numerous structural brain differences have been reported, as for example the enlargement of cortical sulci and ventricles, enlargement of the inter-hemispheric fissure and reduced grey and white matter (Phillipou et al. 2014). The findings about the reversibility of this changes following weight recovery remain inconsistent in literature. Reduced sizes of the pituitary gland (Doraiswamy et al., 1991), areas of the limbic system including the amygdala, hippocampus and cingulate cortex (LeDoux et al. 2000) and the putamen were also reported (Packard et al. 2002, Philipou et al. 2014).

The serotonin system (5-HT) has been of particular interest in relation to AN. The main metabolite of 5-HT, 5-hydroxyindoleacetic acid (5-HIAA) and the 5-HT receptors have been the main focus for research (Philipou et al., 2014). Decreased levels of 5-HIAA were found in cerebrospinal fluid (CSF) in individuals suffering from AN in comparison to recovered AN patients (Kaye, Ebert, Raleigh, & Lake, 1984; Kaye, Gwirtsman, George, & Ebert, 1991). Altered binding potential of serotonin receptors in different brain regions has been investigated in numerous neuroimaging studies (Audenaert et al., 2003; Bailer et al., 2017; Frank et al., 2002; Galusca et al., 2008). The 5-HT system plays an important role in numerous symptoms and behaviours, including obsessional behaviour, anxiety, impulse control and mood (Philipou et al., 2014).

The dopaminergic system has also been considered relevant for the the course of AN as it plays an important role in eating behaviours, motivation and reward (Philipou et al., 2014). Altered levels of homovanilic acid, the metabolite of dopamine, have been found in CSF in individuals with AN (Kaye et al., 1984, Phillipou et al., 2014). In addition, altered dopamine receptor sensitivity and concentration is described (Bailer et al., 2012; Frank et al., 2005). It is still unclear whether the changes in the 5-HT and the dopaminergic system contribute to the development of AN or occur due to starvation.

Regarding genetics, the field of AN and other eating disorders is still at an early stage (Brandys, de Kovel, Kas, van Elburg, & Adan, 2015). The early results of genome-wide association studies have not yet improved the understanding of the AN. However, the optimal sample size has not been reached yet. The majority of the candidate-gene approach studies failed to provide the confirmation of the findings in the replication phase (Brandys et al., 2015).

All eating disorders share a complex multifactorial aetiology. In general, eating disorders are associated with concurrent psychiatric disorders. More than half of patients meet criteria for a current or past episode of major depression (Yager & Andersen, 2005) (Harrington et al, 2015. Other disorders associated with eating disorders include obsessive-compulsive disorder, social phobia, anxiety disorders, substance use disorders, and personality disorders (Harrington et al, 2015).

Treatment of anorexia nervosa

Initiation of treatment is often hard. AN patients generally view their symptoms as functional, necessary, or even desirable, and may be ambivalent about seeking treatment. Furthermore, patients are considered as being resistant, stubborn and defiant towards their condition and are often unmotivated to change.

Therefore, the treatment of AN is a complex process, which is usually organized from a multidisciplinary and multidimensional approach. In most treatments, psychiatrists, psychologists, dieticians, general practitioners and social workers are involved. Medical care, if necessary, is provided by paediatricians, physicians and nurses. With this diverse team of caretakers, strong cooperation is highly required.

At the beginning of the treatment, extensive physical, psychological and psychosocial assessments are done. Moreover, a therapist explores the personal development and family histories. Subsequently, the team develops an individualized treatment plan, which depends on the wishes and motivation of the patient, the severity of the disease, the comorbidities of the patient and his or her social support (American Psychiatric Association, 2000, 2006, Hey et al., 2014, NICE clinical guidelines, 2004).

The treatment of AN targets multiple areas. The first goal is the restoration of the patient's weight within a normal range for their age, height and sex and the treatment of medical symptoms. Moreover, the treatment focuses on modifying the behaviour and cognitions that led to the eating disorder and the distorted self-image. Furthermore, it addresses the identification and resolution of personal and family problems. Psychiatric comorbidities are treated along with that. The entire treatment course focuses on all the aforementioned aims concurrently, although the weight restoration and medical complications have greater priority during the initiation phase. Throughout

the treatment, the priority shifts to the psychological conditions that cause and maintain the eating disorder.

In the last decades, there has been a shift from long-lasting hospitalization or inpatient treatment to outpatient care (NICE clinical guidelines, 2004). Inpatient treatment is an intensive type of therapy where patients are residents in a facility with the possibility of 24 hours of care a day, usually for multiple weeks. In contrary, outpatient treatment is a part-time program where the patient still lives at home and is able to attend school or work. Currently, hospitalization is limited to brief acute weight restoration and refeeding in the case of serious medical symptoms. Furthermore, failure of outpatient treatment or severe psychiatric conditions, such as suicide ideations, are reasons for inpatient treatment (Campbell & Peebles, 2014; Rosen, 2010).

Multiple therapies can be considered and combined. The three main groups are medical, nutritional and psychological therapy. Medicine and nutritional advice are necessary to restore the low weight and to treat and prevent medical symptoms. However, these two cannot cure the patient. Psychological therapy is crucial to ameliorate the eating pattern and the dysfunctional behaviour and thoughts.

At present, there is still no effective pharmaceutical therapy (Pederson, Roerig, & Mitchell, 2003; Zhu & Walsh, 2002). Tricyclic antidepressants , selective serotonin re-uptake inhibitors and antipsychotics are not effective in restoring weight and do not improve the eating pattern, despite the similarities between AN and depression, anxiety and obsessive compulsive disorder, which are psychiatric diseases that can be treated with these types of medication (Pederson et al., 2003; Zhu & Walsh, 2002). However, medication can play a small role in the treatment of these psychiatric comorbidities. The development of new pharmaceutical therapies is complicated by the fact that the mechanisms that contribute to the aetiology are still poorly understood, as mentioned above.

For the restoration of the weight, the calorie intake has to exceed the energy consumption. This means that not only the food intake has to increase, but also that the amount of exercise should be restricted. A dietician can help creating a healthy food intake with a sufficient number of calories and nutrients. It is important to obtain a daily schedule of at least three full meals and some snacks. If possible, this intake is monitored by parents and/ or school.

A frequently used first target for weight restoration is 90% of the average expected weight for the patient's height, age and sex (Golden et al., 1997; Golden, Jacobson, Sterling, & Hertz, 2008). However, in adolescents, the premorbid weight, the pubertal stage and the preceding growth record should be also taken into account6. Another indicator of a healthy weight is the resumption of the menstrual cycle, as this indicates a restoration of the hypothalamic-pituitary-ovarian axis. The menstrual cycle returns within 3-6 months in the large majority of female patients who achieve a healthy BMI (Golden et al., 1997).

Psychotherapy is the cornerstone in the curative treatment of AN. At present, there is no conclusive evidence that any therapy is superior to others in the treatment of adults (Hay, Claudino, Touyz, & Abd Elbaky, 2015). There are only a few randomized controlled studies, all with small sample sizes and remarkable biases and limitations. Therefore, the evidence of these studies is of very low quality (Hay et al., 2015). Furthermore, this lack of evidence may be partly explained by the low prevalence and high morbidity, which makes it ethically difficult to conduct randomized controlled trials with novel treatment modalities.

There are several therapies used in the treatment of AN, of which the most recognized are Cognitive Behavioural Therapy (CBT), and Family-Based Therapy (FBT).

For adolescents, FBT has been investigated the most and seems to be an effective therapy, on the long-term possibly even superior to individual therapy (Couturier, Kimber, & Szatmari, 2013). FBT is also known as Maudsley family therapy, as it was developed in the Maudsley hospital in London. It is an intensive outpatient therapy that is most effective in adolescents, since they are often still integrated in a family system.

In FBT, the parents are considered to be essential for the successful treatment of AN and they should have an active role. Both patient and immediate family members are intensively guided by a single therapist to gradually gain control over their own food intake. This is achieved by having family mealtimes at home, which the therapist will observe and try to steer in a healthy direction. On top of that, the therapy focuses on forming a healthy adolescent identity, maintaining family relationships and helping with individual psychological issues. Moreover, a physician is involved to check and improve the physical status (Lock & Le Grange, 2015). The treatment takes about 9-12 months (Lock & Le Grange, 2015).

CBT is the most investigated form of individual therapy in adults with AN (Wilson, Grilo, & Vitousek, 2007). It concentrates on the development of personal coping strategies and changing distorted thoughts and maladaptive behaviour that plays a role in the maintenance of the disorder. There are three phases in the manualized version for AN patients. In the first phase, there is the development of a positive therapeutic relationship. Moreover, the therapist explores key symptoms, provides psycho-education and nutritional counselling. In the second phase, dysfunctional thoughts and behaviours are identified and reframed, resulting in a healthier eating pattern and weight gain. In addition, interpersonal problems can be addressed. During the third phase, the patient is prepared for termination of the treatment and for prevention of a potential relapse.

At the beginning of the CBT, there are usually two sessions a week. This frequency will decrease during the course of the therapy, to approximately one session every two weeks in the last phase. The therapy mostly lasts for 1-2 years (Wilson et al., 2007).

An essential element of the treatment is a therapeutic alliance between patient and caretaker (la Rie, Noordenbos, Donker, & Van Furth, 2008; Pereira, Lock, & Oggins, 2006; Zeeck & Hartmann, 2005). This is complicated by the fact that most patients experience difficulties in forming a relationship with the therapist. According to patients, a good therapist should be trustworthy, take the patient seriously, and be able to talk about feelings, thought and eating behaviours (la Rie et al., 2008). Moreover, the therapist should be focused on the individual patient and work on the patient's self-esteem.

Prognosis

The recovery of AN is not only a complex, but also a lengthy process (Herzog et al., 1999; Strober, Freeman, & Morrell, 1997). The mean duration of the eating disorder is over 12 months after start of treatment, with some patients even recovering after eight years (Bergh et al., 2013; Strober et al., 1997; Wentz, Gillberg, Anckarsäter, Gillberg, & Råstam, 2009). The prognosis of AN is unfavourable, with complete recovery rates of only 50-70% ten years after the first onset (Steinhausen, 2002, 2009). Of the remaining group, 50% show partial recovery, and the other 50% have a chronic course of AN. Patients who get AN when they are adolescents seem to recover faster than patients with a later onset. Moreover, patients with a shorter duration of AN have a better prognosis (Herzog et al., 1999). Consequently, an early detection and thus treatment of AN by screening people at risk may be beneficial. The Eating Disorder Examination Questionnaire (EDE-Q) and SCOFF questionnaire have been validated in adults for screening purposes (Mond, Hay, Rodgers, Owen, & Beumont, 2004; Morgan, Reid, & Lacey, 1999). For example, many general practitioners fail to consider eating disorders as a possible diagnosis in children presenting with typical eating disorder features (Bryant-Waugh et al. 1992). In The Netherlands, on average, only 40% of the AN cases are detected by general practitioners and 79% of these patients are referred on for mental health care (Hoek, 2003).

Generally, AN is associated with significantly reduced Health-Related Quality of Life (HRQoL) compared to healthy individuals and patients with other psychiatric conditions (Agh et al., 2016; Sy, Ponton, De Marco, Pi, & IsHak, 2013). In addition, concomitant depression and anxiety are associated with further deterioration of HRQoL.

The long-term prognosis of patients with AN that fully recover seems promising (Mustelin et al., 2015; Strober et al., 1997). Strober et al. showed in 1997 that approximately 90% of this group reports a good outcome in the domains of work status, interpersonal relating and overall life satisfaction 10 years after the onset of the disease16. In contrast, the quality of life of the chronically ill patients remains low, with only 6% reporting a high life satisfaction 10 years after the onset. Moreover, only a small minority of the chronic patients had a good work status or experience difficulties in social functioning.

It appears that patients with AN that are fully recovered, no longer have an elevated risk of additional psychiatric comorbidities or personality disorders, whereas patients with a chronic course of AN are more likely to still suffer from psychiatric comorbidities as depression and anxiety disorders, as well as from personality disorders (Herpertz-Dahlmann et al., 2001). Treatment of chronic patients focuses on elevating the quality of life and maintaining a safe weight rather than recovery (NICE clinical guidelines, 2004).

Relapse is a common problem in recovered AN patients, with a relapse rate estimates ranging from 10 to 41% (Berends et al., 2016; Bergh et al., 2013; Carter, Blackmore, Sutandar-Pinnock, & Woodside, 2004; McFarlane, Olmsted, & Trottier, 2008). Possible predictors for a relapse are a long duration of the illness, a higher age at presentation and high-level of exercise (Berends et al., 2016; Deter & Herzog, 1994; Strober et al., 1997). Moreover, patients who receive both in- and outpatient care show a higher relapse rate than patients who just receive outpatient care (Berends et al., 2016). As a method for relapse prevention, the Guideline Relapse Prevention Anorexia Nervosa (GRP) was developed and implemented in specialized therapy in the Netherlands. Due to this guideline, relapses are early recognized, which ideally results in early interventions to prevent a further development in a full relapse. The highest rate of relapse is between 6-17 months after the termination of the treatment (Carter et al., 2004). Therefore, it is recommended to periodically monitor discharged patients for at least 18 months (Berends et al., 2016).

Economic consequences

AN has a considerable financial effect on society. Eating disorders in general are associated with higher healthcare costs and increased healthcare utilization. Furthermore, a substantial part of AN patients is not able to contribute to the economy, resulting from disruption to education, employment and professional development. In addition, their carers may experience loss of earnings, since the disease has impact on entire families. The exact costs of a disease are hard to assess and, moreover, differ between countries. The estimated total annual treatment costs in literature vary between €4,900 and €5,952 for every AN patient in Germany (Krauth et al., 2002; Haas et al., 2012), are up to €69,776 in the UK (Byford et al., 2007) and between €3,221 and €55,270 in the USA (Crow & Nyman, 2004; Lock, Couturier, & Agras, 2008; Mitchell et al., 2009; O Brien & Ward, 2003). Different components were included in the calculation for each study, which partly explained the large differences. Part of the high costs may be explained by the expensive inpatient treatment. Research showed that the average daily costs for inpatient treatment of AN patients are €256 in Germany, €512 in the UK and up to €1,790 In the US per patient (Toulany et al, 2015, Krauth et al, 2002, Kalisvaart et al, 2007).

In the most countries, AN is associated with higher healthcare costs in comparison to other eating disorders. However, the total economic burden of AN is comparable with other major psychiatric conditions (Agh et al., 2016).

Conclusion

AN is a relatively rare but severe eating disorder with multiple medical and psychiatric comorbidities and complications. The aetiology is multifactorial and still poorly understood. Patients are characterized by disrupted core beliefs about themselves which is combined with morbid anorexic reasoning. Positive and negative reinforcement contingencies of the disorder further enhance the patient's attachment to the disorder. Personality traits as perfectionism, obsessive compulsive features, social phobia and disturbances in both emotional perception and social communication are considered predisposing factors which might wind up in a viscous circle of enhancing effectiveness in starvation, which further leads to the deterioration of the predisposing traits. The initiation of treatment is often resisted due to strong denial of the condition and lack of motivation to change. The approach to treatment of AN is multidisciplinary with focus on nutritional, medical and psychological aspects. The evidence for the effectiveness of psychotherapy is still inconclusive. Family-based therapy, an intensive treatment in the home setting in which both patient and parents are guided by a single therapist, is proved to be effective in adolescents with AN. The treatment usually takes years to achieve stable recovery. The prognosis for patients who fully recover is on the long term comparable to healthy population in terms of quality of life. However, approximately twenty percent of patients become chronically ill with high medical and psychiatric comorbidity rates and low quality of life. The economic burden in AN is high which is explained by the long duration of the disease, expensive inpatient treatment and psychiatric comorbidities that disrupt education, employment and personal development.

c) Socially Assistive Robots in Mental Healthcare

Robots are advancing fast in our society, in factories and homes (e.g. "Roomba"). Yet, their capabilities are rather limited so far. Doubtlessly, robots excel at simple tasks such as repetitive production line tasks in a factory or vacuum-cleaning, but their usage for therapeutic treatment is still somewhat underdeveloped. Nonetheless, many advances have been made in designing robots capable of basic social interaction. For instance, the robot "Pepper" (*SoftBank Robotics America, California USA)* is one of the most successful, commercially available robots which is able to perform spoken conversation and gestures. In a more assistance-oriented line of research, robots are used to physically aid patients, for example, by physically guiding them through exercises (Prange, Jannink, Groothuis-Oudshoorn, Hermens, & Ijzerman, 2006).

Recently, these two branches have been combined in socially assistive robots (SARs). These are robots that have a purpose of aiding people or patients through social interaction (Rabbitt et al., 2015). Given that the goal of SARs is to best assist a user population, the various functionalities of the robots should be tailored to the specific needs of the patient and the existing therapy regimen. Depending on the type of end user and their prescribed therapeutic treatment, SARs are typically designed to fulfil a specific role such as companion, therapeutic play partner, therapist, learning aid, coach or instructor (Rabbitt et al., 2015).

The goal of this chapter is to provide an overview over existing experimental literature of SARs in mental healthcare. Based on a systematic review, we will then suggest some recommendations for future research. This chapter is structured as follows. Section 1 describes the details of the systematic review on SAR in mental healthcare. In Section 2, we outline the results on the basis of the target patient population of SARs. Section 3 examines the current state of the research on SAR by discussing which types of robotic platforms have been recently applied to mental healthcare. In Section 4, we conclude by presenting some directions for further research.

Methods

Protocol

Prisma guidelines (Moher et al., 2015) were followed during to carry out this systematic review. An extensive literature search was conducted on Pubmed, MEDLINE and Embase by three researchers. The consulted databases were selected in order to retrieve articles that were published in the field of medical sciences. The search string included "*Artificial Intelligence; Social robots; Mental Illness; Mental Care; Robot*; Delivery of Health Care; Autistic Disorder; Autism; Autism Spectrum Disorder; Dementia"*. Medical Subject Headings terms were used with major subheadings. The reference lists of retrieved articles were reviewed for missing references. Based on title and abstract, articles were independently selected for full article review. In case of disagreement, a fourth author was consulted. Inclusion criteria were the use of English, Dutch or German language; the inclusion of at least two patients; the use of social assistive robots; and the implementation of SARs in mental healthcare. Furthermore, the articles had to report which robot was used, what the role of the robot was, the diagnosis, age and sex of the patients and the cognitive, social and emotional challenges in the interaction of SARs and patients. Reviews and articles that were animal based or those that investigated the neurophysiological or –psychological basis for SARs in mental healthcare were excluded. The quality of the studies and the risk of bias were determined according to the GRADE (Grading of Recommendations, Assessment, Development and Evaluations) approach. GRADE is a systematic and explicit approach to allow judgements about quality of evidence and strength of

recommendations. It was developed by the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) Working Group, and it is now seen an effective method of linking evidence-quality evaluations to clinical recommendations in medical sciences (G. Guyatt et al., 2011; G. H. Guyatt, Oxman, Kunz, Atkins, Brozek, Vist, Alderson, Glasziou, Falck-Ytter, & Schünemann, 2011; G. H. Guyatt, Oxman, Kunz, Brozek, Alonso-Coello, Rind, Devereaux, Montori, Freyschuss, & Vist, 2011; G. H. Guyatt, Oxman, Kunz, Woodcock, Brozek, Helfand, Alonso-Coello, Falck-Ytter, et al., 2011; G. H. Guyatt, Oxman, Kunz, Woodcock, Brozek, Helfand, Alonso-Coello, Glasziou, et al., 2011; G. H. Guyatt, Oxman, Montori, Vist, Kunz, Brozek, Alonso-Coello, Djulbegovic, Atkins, et al., 2011; G. H. Guyatt, Oxman, Sultan, Glasziou, Akl, Alonso-Coello, Atkins, Kunz, Brozek, et al., 2011; G. H. Guyatt, Oxman, Vist, Kunz, Brozek, Alonso-Coello, Montori, Akl, Djulbegovic, et al., 2011).

Results

156 articles were retrieved from the searches that were conducted, 55 of which were excluded based on title and abstract. 101 articles were thus included for full text analysis. After the removal of double articles, the full text of 59 articles were analyzed. After reading the full text, 16 articles were excluded because they were reviews and 21 more articles needed to be excluded due to inaccessibility (*n*=1), the fact that the article described a study protocol and did not analyse data (=12) or the use of an excluded language (*n*=8). The remaining 22 articles were included in the final analysis. The quality of the included articles according to the GRADE classification can be found in Table 4.

Table 4

GRADE scores per included article

Authors	Year	Reference number	GRADE-score
Libin, A. and J. Cohen-Mansfield	2004	13	2
Odetti, L., et al.	2007	14	0
Begum, M., et al.	2013	15	0
Robinson, H., et al.	2013	16	0
Sabanovic, S., et al.	2013	17	0
Moyle, W., et al.	2014	18	0
Valenti Soler, M., et al.	2015	19	3
Gustafsson, C., C. Svanberg, and M. Mullersdorf	2015	20	0
Bemelmans, R., et al.	2015	21	1
Joranson, N., et al.	2015	22	3

Kim, G.H., et al.	2015	23	3
Wu, Y.H., et al.,	2016	24	0
Kanamori, M., et al.	2003	25	0
Wada, K., et al.	2005	26	0
Wada, K. and T. Shibata	2006	27	0
Wada, K., et al.	2004	28	0
Banks, M.R., L.M. Willoughby, and W.A. Banks	2008	29	0
Costa, S., et al.	2009	30	0
Kim, E.S., et al.	2013	31	3
Costescu, C.A., B. Vanderborght, and D.O. David	2015	32	1
Srinivasan, S.M., et al.	2016	33	2
Simut, R.E., et al.	2016	34	1

SARs for elderly with dementia-related cognitive impairments

In 2004, an animaloid SAR was introduced into the daily lives of 9 participants (age range 83-98 years; *Median*= 90 years) with dementia. The level of intensity of manipulation of the SAR and the amount of attention paid to the SAR were strongly associated with the level of cognitive deterioration. Therefore, the SAR seemed suitable for agitated behaviour and was able to keep the participants' interest (Libin & Cohen-Mansfield, 2004).

Odetti et al. (2007) described 24 participants (Median age 76.6 years) suffering from dementia and illustrated that different attitudes to technology (rejection vs. a generically non-negative attitude) do not affect the affective components of interaction between an SAR and older people significantly. Women seemed to be slightly more likely to form an attachment with electronic goods and technology. The study used an animaloid robot with a companionship role. The role of an animaloid companion robot embedded in an ubiquitous robotic system seemed to vary greatly, depending on the user's functional and cognitive profile (Odetti et al., 2007).

In 2013, five participants (age range 59-88 years; mean age *Median* = 77.2 years) with dementia were studied to investigate how they interacted with humanoid SARs in daily live activities. The preliminary analysis of the study data showed that SARs have an enormous potential in forming a social interaction with elderly patients with dementia (Begum, Wang, Huq, & Mihailidis, 2013).

Robinson and colleagues (Robinson, MacDonald, Kerse, & Broadbent, 2013) argued that SARs in healthcare should be simple and easy to use, as well as being stimulating and entertaining. In their study, they employed two SARs; an animal-like robot and a robot with a technical appearance and a touch screen. It was concluded that the sounds of the animal-like SAR should be modified to be more

acceptable to the elderly patients. Moreover, the more technical SAR could have more simplified software in order to optimize the interaction with the target population (Robinson, MacDonald, Kerse, & Broadbent, 2013).

Another study including 10 participants with dementia found that most participants needed the therapist's assistance to communicate with others. The participants discussed how the animaloid SAR worked and what it was, and also reminisced about their family members, previous pets, and other life experiences as they were interacting with the robot. In total, the SAR increased the patients' communicative behaviour and improved the activity level of the participants. Participants showed higher levels of engagement with their environment and other people, becoming increasingly attentive and interactive as the study progressed. Even participants who were not directly interacting with the SAR displayed increasingly attentive behaviour toward the SAR and others in their environment through the duration of the study (Sabanovic, Bennett, Chang, & Huber, 2013).

Another study (*N* = 18, *Median* = 85.3 years) showed that animaloid companion SARs may be an alternative psychosocial intervention for older adults with dementia providing improved quality of life and improved pleasure (Moyle et al., 2014). Valenti and colleagues described 211 participants (*M* = 84,7 years) with dementia who were introduced to a humanoid robot, an animal-shaped robot, or conventional therapy (Soler et al., 2015). The robots seemed to trigger a decrease in apathy and a decrease in mini-mental state examination scores. Humanoid compared to animal-shaped SARs did not show statistical significant differences (Soler et al., 2015).

Gustafsson discussed four participants (age range 82-90 years) with dementia who were confronted with a cat-like SAR. The authors saw more opportunities for connecting with others although some patients had an aversion to cats (and other animals). The results showed reduced feelings of loneliness and moreover, the robot provided comfort and peace. However, the authors mention that a possible Hawthorne effect might have confounded the results as the patients received more attention during the experimental period (Gustafsson et al., 2015).

Another study, published by Bemelmans, concluded that 71 participants with dementia treated with an animaloid SAR showed an improvement in their self-reported mood. The authors concluded that the SAR should be seen as a tool for the staff and not as a replacement (Bemelmans, Gelderblom, Jonker, & de Witte, 2015). However, the same SAR was presented in another study including 60 participants (age range 62-95 years; *Mean* = 84 years) with dementia. This study showed that not all participants show interest in the same animal-like companionship robot (Mustelin et al. 2915) .

One study investigated the cortical thickness in patients with dementia that were exposed to SARs compared to demented patients that were not. In this study, 85 participants (*Median* = 67

years) with dementia were included. Two types of SARs were used in the robot-assisted cognitive training. Results suggest that robot-assisted cognitive training can help reduce the cognitive disabilities associated with age-related cortical thinning compared to the traditional treatment (G. H. Kim et al., 2015).

In 2016, Wu et al. studied 11 participants (ranging in age from 75-85 years; *Mean* = 79.3 years) with mild cognitive impairments. SARs were readily accepted by the elderly as long as they were introduced by family members or caretakers. The robot had a humanoid appearance and served as a companion. The authors recommended that robot functionalities should include cognitive stimulation and object-finding systems. They mentioned that some issues needed to be carefully addressed, including the maintenance of autonomy and self-image. Furthermore, they concluded that although distinct advantages could be provided by technologies, older adults did not always seem to be ready to embrace them (Wu et al., 2016).

SARs in depression

In 2003, an animaloid robot was introduced into the lives of elderly patients in nursing homes or their homes for seven weeks. Quality of life was measured using self-assessment questionnaires and biochemical markers in saliva (salivary chromogranin A). After activities with the SAR, reported loneliness and salivary chromogranin A decreased significantly, indicating an improvement in mood and quality of life in general (Kanamori et al., 2003). In 2005, an animaloid SAR was implemented in a study with 13 elderly patients suffering from dementia and depression. Mood and severity of depression were self-reported using the geriatric depression scale and face scales, respectively. After one year of exposure to the SAR, the feelings of the elderly patients improved, which was thought to be achieved through interaction with the animaloid SAR (Wada, Shibata, Saito, Sakamoto, & Tanie, 2005). Another study illustrated that patients also spend increased time in public areas and around other patients and staff members after the SARs were introduced (Wada & Shibata, 2006).

SARs in stress-related conditions

In 2004, Wada and colleagues investigated the modifications in reaction to stress in elderly patients and the nursing staff. In the elderly, stress levels were monitored using urinary tests, whereas in the nursing staff, the mental state was evaluated using a "burnout scale." The day service center was provided with animaloid SARs for five weeks. As a result, the feelings of elderly people seemed to improve. Urinary tests showed that the ability to overcome stress was improved as well. Furthermore, the stress levels of the nursing staff decreased. It was hypothesized that this was due to the elderly people requiring less supervision when interacting with the SARs. Consequently, the used SARs were judged to be useful at institutions for the elderly, such as the day service center (Wada, Shibata, Saito, & Tanie, 2004).

SARs in loneliness

In 2007, Banks and colleagues presented that loneliness in long term care facility residents can be tackled using alive animals or using animaloid SARs. They found that attachment was not the mechanism by which animal-assisted treatment decreases loneliness and conclude that interactive robotic dogs can reduce loneliness as well. Furthermore, it was shown that residents become attached to the SARs (Banks, Willoughby, & Banks, 2008).

SARs in autism spectrum disorder (ASD)

In 2009, Costa and colleagues presented 2 participants with ASD and mild cognitive impairments that showed playing with a SAR can be experienced as very entertaining. The SAR was made of Lego, had animal-like features, and had a companionship role (Costa et al., 2009). A couple of years later, in 2013, 24 children with ASD (age range 4.6-12.8 years; *Mean* = 9.4 years) were found to interact better with an animaloid, companionship SAR than with an adult or a novel touch screen computer game. The SAR accompanied ASD children in their cognitive therapy by moving (e.g. walking, jumping) and socially interacting through vocalizations or behaviour (e.g. moving its tail) (E. S. Kim et al., 2013). In 2014, a larger sample of 81 patients (age range 4-13 years; *Mean* = 6.95 years) with ASD showed that SARs increased the cognitive performances of the participants. In this study, a humanoid SAR was introduced in a teaching environment. Patients seemed to enjoy their tasks more when interacting with the SAR compared to interacting with an adult. The cognitive flexibility of the patients was not different in the SAR-setting compared to the adult setting except for the learning phase were the SAR can interfere with performance. It was suggested that the appearance of the SAR played an important role in these settings (Costescu, Vanderborght, & David, 2015).

In 2016, Srinivasan and colleagues investigated three groups of ASD patients, totalling 36 participants (age range 5-12 years; *Mean* = 7.63 years) with ASD. They were divided in a rhythm-, SAR-, and control group. They implemented a humanoid robot with a companionship role. The results of this study show that after controlling for baseline levels of social verbalization the rhythm- and robot groups increased social verbalization levels across sessions. Activities that focused on social interactions and/or singing without concurrent motor demands generated increased levels of verbalization in all groups. Across sessions, the robot group paid less attention to the robot and more attention elsewhere. Apparently, the subjects that were interacting with the SAR still showed interest in the surroundings, not only to the robot, which is a key finding that supports further investigation concerning SARs in ASD (Srinivasan, Eigsti, Gifford, & Bhat, 2016). Another study discusses 30 patients (age range 5-7 years; *Mean* = 6.7 years) with ASD and showed that SAR tasks were more engaging and motivating, and elicited more eye-contact. They used an animaloid SAR. The SAR however did not function as a social mediator (Simut, Vanderfaeillie, Peca, de Perre, & Vanderborght, 2016).

Discussion

Given that the goal of SAR is to best assist a user population, the various functionalities of the robots can be grouped together according to the role assigned to a robot. In particular, three different classes of SARs emerged in the literature review: companion, therapeutic play partner, and coach/instructor.

Companion SARs have been extensively explored in their application to mental health care. Analogously to trained therapy animals (e.g. dogs), contemporary companion SARs such as the animaloid "Paro" should be thought of as embodied forms of technology which can improve the mood of the patient through basic forms of social interactions (Yu et al., 2015). Although research on the value of pet therapy has been largely investigated (Nimer & Lundahl, 2007), studies on the clinical relevance of socially assistive robots have mainly focused on elderly patients suffering mostly of either dementia-related cognitive deficits (Shibata & Wada, 2011) or symptoms of depression (Banks et al., 2008). Notably, Wada et al. reported clinically relevant improvements in elderly living in care facilities after a long-term application of SAR systems (Wada et al., 2005).

The second group includes SAR systems serving as therapeutic play partner. The "playful" component of these robots is meant to capture the motivational goal to engage the end user (i.e. almost exclusively children) with cognitive tasks consistent with her therapeutic needs. In this context, the biggest patient group is composed of children with ASD where the SARs help to develop clinically relevant skills (Dickstein-Fischer & Fischer, 2014; Simut et al., 2016; Ueyama, 2015). For instance, in Kim et al. an SAR looking like a baby dinosaur, "Pleo", accompanied ASD children in their cognitive therapy by moving (e.g. walking, jumping) and socially interacting through vocalizations or behaviour (e.g. moving its tail) (E. S. Kim et al., 2013).

Finally, the third group of SARs is represented by SARs performing in the role of a coach or instructor. The main idea behind this application in robotics is that SAR systems can describe tasks, like physical exercises, to patients and monitor their performance, provide corrective feedback, and encourage them in their therapeutic activities (Rabbitt et al., 2015). In this sense, these SARs resemble human coaches and instructors insofar as their main goal is to motivate and guide the user in her exercises by offering constructive feedback. Notably, whereas companion and therapeutic play robots have been almost exclusively with specific patient groups (i.e. elderly and children, respectively), this type of coaching or instructing SAR have been tested both with young and adult users for miscellaneous clinical needs. In particular, post-stroke rehabilitation (Matarić, Eriksson, Feil-Seifer, & Winstein, 2007), weight loss coach (Kidd & Breazeal, 2008), motivational support to maintain a diary for diabetic children (Van Der Drift, Beun, Looije, Blanson Henkemans, & Neerincx, 2014), support for attention and memory task in old adults with dementia (Tapus, Ţăpuş, & Matarić, 2009) are some interesting examples of how SAR has been variously tested as an motivational means

for patients to comply with their therapeutic activities both inside and outside of their treatment sessions.

Another application of SAR in mental health care does not seem to belong to any of the three cluster applications mentioned above; namely, social robot-assisted therapy (SART). This application was developed by Alemi et al. who programmed a robot to serve as a humanoid assistant in psychotherapy sessions of young oncologic patients. The aim of the robot was to create a friendship bond with a group of children diagnosed with cancer and to alleviate their negative feelings (e.g. pain, anger, distress, anxiety). Through pre-programmed dialogs, the NAO robot attempted to inform the young patients about their condition and its related effects via human-like vocalizations, gestures, music, and dancing. SARTs serve the therapist as a tool during the therapy session in an entertaining and educative way. They can be thought of as integrating both the playful and instructive aspects belonging to the second and third group of robotic applications to mental health care (Alemi, Ghanbarzadeh, Meghdari, & Moghadam, 2016).

Perhaps due to the infancy of SAR research in mental healthcare, very few studies have investigated long-term consequences of SAR applications. Most studies only included a very limited amount of sessions. Especially studies using a playful companion SAR often only have one session where the patients interact with the SAR (for example Costescu et al., 2015; Libin & Cohen-Mansfield, 2004; Simut et al., 2016). For companion SARs, one study investigated the effects of an SAR over a span of multiple years (Wada, Shibata, & Kawaguchi, 2009). However, many therapies for mental disorders usually take more than a couple of sessions. The impact of SARs in long-term therapy sessions is therefore a much-needed investigation (see also Rabbitt et al., 2015).

Limitations can be found in the relatively low quality of the articles that were found (**Table 1**) according to the GRADE scores which made it complicated to draw conclusions. A striking result of the systematic review of the literature is the fact that just a limited number of articles have been published on this field and that approximately 30% of the body of the existing literature are reviews.

Another limitation of the studies so far is the lack of SAR-implementations at homes. When SARs are used as a tool for therapy (mostly playful companion SARs), they are usually implemented during a therapy session and are controlled by the therapist or confederate (Scassellati, 2007). An SAR in a home setting would have the considerable advantage of being available 24-7, as opposed to therapists or even family members. For that to be successful however, SARs must be developed that can operate more independently. Strengths of this literature review are the systematic approach of the search, the combining of various disciplines and the up-to-date overview of literature on SARs in mental healthcare that is provided.

Future research in SARs in mental healthcare should start by focusing on the expectations and desires of stakeholders within the field of interest. The outcomes of these studies should then

guide engineers in developing SARs tailored to their target groups. When the tailored robot has been developed, it should then be tested in clinical practice in randomized controlled trials with generalisable outcome measures, with validated questionnaires and combined with qualitative methods.

Conclusion

SARs could be a valuable addition to the therapeutic interventions for people with mental disorders. Many studies seem to indicate the potential of SARs in various groups, such as elderly people with or without dementia, and children with ASD. Most SARs can be classified according to their relationship with the patient as being either a companion, playful partner or coach. Another consideration is that for practical reasons and perhaps a lack of communication across fields, the standard procedure seems to involve the use and, at best, the modification of pre-existing SARs to test their clinical efficacy on a case by case basis. However, more preliminary need-driven research may help roboticists to improve SARs by tailoring them to specific target groups. More research is also needed with regards to long-term success and the implementation of SARs in home settings.

d) Problematizing the Use of SARs for the Treatment of AN: A Synthesis of Different Perspectives

The last three chapters had the function to provide the relevant theoretical background for discussing whether and how SARs could benefit AN patients. Perspectives from various disciplines such as ethics, law, economics, medicine, and robotics will now be combined in this section to identify all the relevant problems underlying such a multifaceted question. In particular, the goal of this chapter is to problematize whether and how SARs could be used as a therapeutic aid for AN patients in a home setting. This explorative attempt will be conducted under the guidance of the Ethical, Legal, Social (ELS) framework as described in Chapter 1. The chapter is structured as follows. First, the general motivation behind the possible introduction of a robotic platform for AN treatment in a home setting will be outlined. In the second part, the current treatment of AN will be evaluated based on the ethical considerations of the framework. In light of this, in the third part of this chapter we question whether the introduction of SAR could qualitatively improve the treatment of AN. The fourth part of this chapter represents a discussion of the possible functions of an SAR system for an AN user in a home setting. The chapter will be concluded by examining various issues related to the introduction of the envisioned robot, such as autonomy, legal accountability, and the socio-economic implications both at a micro- and macro-level.

Possible uses of SARs for a home-based treatment of AN

Since sophistication does not seem to necessarily entail qualification, one might wonder why a robotic platform should be employed as a clinical tool and what, in particular, motivated the choice to focus on mental healthcare. Such a choice presents serious challenges. First of all, robots are expensive pieces of technology and perhaps less sophisticated machineries could provide a cheaper but still qualified assistance. For instance, one could claim that the disembodied nature of smartphone applications could, in principle, empower the patient in circumstances that are not limited to the clinical and/or home setting. After all, smartphones have become everyday commodities, and it is not straightforward to assume that the same could and should happen with SARs. More importantly, previous chapters have outlined the ethical, societal, legal, and practical challenges concerning SARs and the challenges involved in the treatment of AN. Therefore, the choice to focus on SARs in particular calls from some solid justifications.

Before explaining the choices of our project, it is essential to emphasize that we remained neutral regarding the possibility and desirability of non-robotic technologies for treating AN. Our investigation is meant to be exploratory and we do not reject the possibility that other types of technologies could be more desirable and/or efficacious for this patient group. Nevertheless, it is important to emphasize that we agree in principle on the starting assumption motivating our research to investigate the implementation of SAR for AN treatment. We believe that the embodied and social nature of SAR systems could offer an important motivational support for the patient's compliance of the treatment. Recent evidence suggesting a high degree of acceptance of SAR from various patient populations seems to be consistent with this intuition (Rabbitt et al. 2015).

Although AN patients admittedly represent a very complex case study, there are several reasons for biting the bullet. First and foremost, the high mortality and the long duration of AN calls for further investigation aiming at improving the current treatment and increasing the quality of life of this target population. Second, in light of previous experimentations on SARs, robots have shown to contribute to cognitive and emotional improvements in other patient groups. Specifically, as research on SARs in the treatment of ASD has highlighted (e.g. Dickstein-Fischer & Fischer 2014; Ueyama 2015; Simut et al. 2016), robots can be of particular additional value because of their strong persistence, rigidity, and unlimited repetitiveness. Therefore, we believe that SARs could assist anorexic patients in practicing therapeutic relevant exercises on a more frequent and continuous basis. Finally, we believe that in case our investigation will provide positive results for such a complicated study case, it might not only encourage further research on the topic of SAR for AN, but also for other prima facie difficult patient populations.

Having clarified our basic assumptions, it is now necessary to address the choice to focus our case study on a home setting. As Chapter 3 showed, there are two target settings in which SARs are

generally applied, namely the clinical setting and the home- (or residency) setting. To a certain extent, these two different contexts of application map into the distinction between short-term and long-term Human-Robot Interaction (HRI). Admittedly, the fact that both home-based SAR applications and long-term HRI have been less thoroughly studied makes its investigation somewhat more difficult (see Chapter 3 for a more detailed discussion). Despite the fact that it would require more challenges than the more explored long-term HRI in a clinic setting, we believe that it is a worthwhile approach given the opportunity it would offer to reach areas of everyday life where most of the times the practice of care is hard to administer in a way that would be highly appreciated.

With regards to AN patients, given the recent tendency of conducting the therapy closer to the patient's everyday life (see Chapter 2), the use of SARs in a home setting seems consistent with the more common outpatient treatment of AN[12]. In the remaining part of the chapter we will elaborate further on the potential added value of the use of SARs in the home-based treatment of AN.

Ethical evaluation of current home-based AN therapy

As a prerequisite to consider SARs as an additional tool in the treatment of AN, the need for qualitative improvement of the current therapeutic setting should be discussed. In this regard, the ELS-framework offers a systematic approach to assess the current care setting, roles, and responsibilities of the actors involved. From an ethical perspective, two points are crucial to consider. First of all, care should never be seen as an isolated product to meet standardized needs (Wynsberghe, 2016). Therefore, a potential new application for AN patients should not take over any substantial role without reflecting on how it relates to other parts of the treatment. Secondly, the ELS-framework prescribes four core values that need to be maximized in any care setting: attentiveness, responsibility, competence and responsiveness. It states explicitly that none of the moral values should be demoted by introducing an SAR (see Chapter 1). This means that if the SAR is able to promote the value of competence (for example by making parts of the treatment more efficient), this should not come at the expense of another core value (for example by forgetting to see the changing needs of the patient).

In order to evaluate how the needs of AN patients are currently dealt with, we should first quickly recapitulate the nature of the condition. Briefly, AN is a mental disorder identified by an unhealthy low weight due to food restriction, purging and/or excessive exercise, a fear of becoming overweight and a distorted self-image (DSM-5).This can lead to severe, even life-threatening, medical

[12] Inpatient treatment is an intensive type of therapy where patients are residents in a facility with the possibility of 24 hours of care a day, usually for multiple weeks. In contrary, outpatient treatment is a part-time program where the patient still lives at home and is able to attend school or work.

complications and strongly affects the daily life of the patients (Harrington et al, 2015). Moreover, patients regularly have psychiatric comorbidities such as depression and anxiety disorder (Herpertz-Dahlmann, 2001). Therefore, the treatment requires complex and time consuming supervision in order to help the patients to cope with their disorder within everyday life (Hay et al. 2014).

With regard to the moral values in the ELS-framework, the first moral value is attentiveness, which refers to the capability of recognizing the (changing and dynamic) needs of the patient. Young adolescent AN patients, who still live at their family's house, will ideally be surrounded by their family members who can keep track of their changing needs (Lock and Le Grange, 2015). For example, in case of deterioration in the patient's situation, the family members can intervene and thus try to find solutions for urgent changes in their needs. For chronic patients who (most likely) live by themselves the situation is utterly different. Although the condition of these patients might be considered more static, it remains a challenge for the therapist to dynamically respond to their needs when sessions are not on a frequent basis.

The second moral value is responsibility, which refers to the capability of the involved actors' to take care of the needs of the patient (Wynsberghe, 2016). Similarly, in case the patients live with their family, the responsibility of having a successful treatment is most likely distributed between the therapist, family members and the patient (Lock and Le Grange, 2015). Especially for young patients, parents play a crucial role in pursuing goals that might even conflict with the interests of their own child, thus taking over an important part of the patient's own responsibility.

The third moral value is competence, which refers to the capability of executing the required actions (Wynsberghe, 2016). For the therapist the treatment of patients with AN is highly demanding, since it requires expertise in several domains: nutrition, psycho-education, psychotherapy and medical supervision. Finally, the last moral value is responsiveness, which refers to the willingness of the patient to cooperate to the treatment (Wynsberghe, 2016). When we translate this to AN, this is probably the hardest part, since patients often deny their own condition or even see it as desirable (Bruch, 1978). In particular, adolescents require family members to encourage initiation of and compliance with the treatment.

Potential of improving the current treatment of AN

AN patients often do not easily comply with the therapeutic treatment. Most likely, this is due to the strong phobic orientation towards changes in behavioural patterns (Treasure and Schmidt, 2013). In this respect, both family members and the therapist have a highly demanding role in facilitating the compliance of the prescribed therapy. Given the therapist's restricted schedule, the enhancement of outpatient care is of crucial importance. That is, the use of SARs could pave the way for extending some aspects of the therapy in the home setting. By means of a regular and reliable

supervision of an SAR, AN patients could better cope with their everyday challenges. It is worth adding that the robot could be designed to execute tasks (e.g. reminding of mealtimes) which, if performed by family members, might negatively affect the internal dynamics of the family.

As it has already been discussed, from a care ethics perspective, it is important that potential new elements in the treatment would not replace current well-functioning elements, especially those that address individual human needs rather than only therapeutic needs. In order to achieve this, the SAR should only be seen as an addition to the current treatment. This does not neglect the need of the expert supervision of a human therapist. Rather, such an SAR could potentially help to reduce the workload and the responsibilities associated with the assistance of both the therapist and family members. In the remainder of this section, we will suggest some possible roles and functionalities for an SAR designed to assist the various needs of AN patients in a home setting.

Possible functions of an SAR for an AN patient in a home setting

Since one of the goals of SARs is to provide assistance to a given patient population, the role and functions of SARs should comply with their specific needs. In line with this idea, we shall suggest here some possible home-based prototypes of SARs for both chronic and non-chronic AN patients. It is worth emphasizing that although our suggestions are informed by careful literature reviews on AN and SAR in mental healthcare, other possible prototypes are conceivable. Notably, the AN population is a heterogenous group and their needs may vary depending on factors such as age (adolescents, adults), chronicity or non-chronicity of their disorder, social situation (living alone or with a family, extension of social network, employed/unemployed), and personality traits (Touyz et al, 2016, Treasure and Schmidt, 2013). Based on our literature study, we recommend a specific type of robot according to the age-group (i.e. adolescents, adults) and chronicity or non-chronicity of the condition of a given AN patient.

As reviewed in Chapter 3, there are three main types of SARs which differ on the basis of their assigned role, that is, therapeutic play partner, companion, and coach/instructor. It has been noted that age-group of the patient population seems to play a major role in determining the role of the SAR. In particular, therapeutic play partners and robot companions have been almost exclusively tested with either children or elderly, respectively. Notably, coaching or instructing SARs have been tested both with young and adult users for miscellaneous clinical needs (e.g. post-stroke rehabilitation, weight loss coach, motivational support to maintain a diary for diabetic children, execution of cognitive exercises) (Mataric et al. 2007; Kidd & Breazeal, 2008; Drift, 2013; Tapus, Tapus, & Mataric, 2009).

In addition to these observations, it is paramount to highlight that our suggestions will be based on the particular goals of the treatment of chronic and non-chronic patients. This is motivated

by the fact that the treatment of these two patient populations is geared towards different ends: the enhancement of the quality of life for the former and recovery for the latter (NICE clinical guidelines, 2004). Accordingly, we hold that on the one hand the role of a companion might be appropriate for chronic AN patients and, on the other hand, a coach might be more suited for non-chronic ones. While a 'coaching SAR' might mainly function as a motivational tool to facilitate patients to comply with their therapeutic program, an SAR serving as a companion might be especially suitable for improving the overall mood. Accordingly, we hold that the coaching function might be a useful tool for the overall adult population, whereas the companion might be the most appropriate for chronic AN patients[13].

We aim to suggest some general functions that may be appropriate for any SAR regardless of what patient group is targeted. Since AN is an eating disorder, the SAR could help monitoring both the weight and the dietary trajectory of the patient[14]. Not only the end user could directly check this on a user interface, but the SAR itself could remind the patient about daily tasks from the dietary program (either via the user interface or vocalization). A similar reminding function can be performed with the cognitive exercises prescribed by the therapist. In doing so, the robot might help reduce aspects of thoughts and behaviour detrimental to the quality of life such as phobic orientations towards food. Importantly, the inexhaustible persistence of the robot might help counteract the rigid mindset and behavioural habits of the patient as prescribed by any treatment of AN. Moreover, the SAR could have the educational purpose of instructing the user about dieting, as well as the relationship between the consequences of starvation and the psychopathological and medical condition. Ideally, monitored vital parameters would be integrated to provide the educational information as feedback to the current medical condition. This could help in stimulating the treatment compliance of the patient. By labelling and designing the role of the robots differently, the form of the interactions realized by these various functions might vary accordingly. The same might be applicable for the physical appearance of the SAR.

In conclusion, it is worth emphasizing some additional remarks regarding the diversity between the “companion SAR” and the “coaching SAR”. As explained earlier, given that the goal of their addition to the therapy differs, their manner in conveying information also differs. For instance, in performing the delicate task of encouraging the patient to eat during mealtimes, the language of the SAR together with its rewarding tone should match the assigned role. Obviously, the linguistic complexity of these dialogues will depend on the sophistication of the prototypes. However, some

[13] Importantly, this does not exclude the possibility that the young age or infantile personality of a given patient might rather call for a therapeutic play partner.

[14] In this regard, an additional emergency function might be desirable. For instance, analogously to a home security system, the SAR might signal to a relative or the therapist, etc. when the patient’s vital signs are below or in the vicinity of the critical level.

functional limitations, either intentional or due to technological constraint, might even be advantageous since they might encourage a more proactive attitude in the patient. Ideally, this might contribute to a more interactive relationship between the SAR and the end user.

Issues of autonomy, responsibility and socio-economic consequences

After discussing possible roles and functionalities for SARs for AN patients as the ELS-framework prescribes, the broader picture of introducing a robot in the treatment of AN should be assessed. First of all, it is central to understand how the autonomy and the role of the robot are related[15]. In Chapter 1, we have concluded that an SAR should not be seen as a moral agent and that it therefore cannot bear any (moral) responsibility. In this regard, it is crucial to question to what extent the robot should be able to autonomously behave in unpredictable situations or whether we should limit the functionality of the robot towards completely pre-programmed scenarios. In light of these questions, we shall later assess who should be held responsible for potential misuses and undesirable behaviour of the envisioned SAR.

First of all, the extent of autonomy and distribution of responsibility depends on the exact type and functions of the given robot. In this regard, some examples might help to understand how the ascription of autonomy could vary in different realizations of the same type of SAR. On the one hand, one could think about a companion toy-like robot which is limited in its ability and mostly designed to help improve the overall mood of the patient. On the other hand, the realization of a companion robot which could verbally and/or nonverbally communicate with the patient would require much more flexibility and thus autonomy. The same applies to an SAR in the role of a coach. Since this type of robot is geared towards more specific therapeutic objectives, it could be programmed to perform various clinically relevant scenarios which could be updated in accordance with the progress of the patient. The programmer and therapist could thus work together to implement desirable responses of the robot in the envisioned scenarios. By doing so, the flexibility of the robot itself is constrained.

In case the robot malfunctions, moral and legal accountability would depend on the contributions of those involved in the design, implementation and employment. For a highly predictable companion, such as the toy-like robot, the question of who is responsible is rather simple. Since all possible errors can be traced back to either mechanical or software flaws, such a technological device will fall under normal product liability law (see Chapter 1). That means only people involved in designing, building, shipping and selling the robot can be held responsible for its

[15] Recall that In A.I., autonomy refers to "the capacity to operate under all reasonable conditions without recourse to an outside designer, operator or controller while handling unpredictable events in an environment or niche" (Haselager, 2005: 5)

malfunctioning. Conversely, the higher complexity of a more advanced companion SAR entails a larger risk for undesirable behaviour, since more autonomy is granted to the robot in choosing appropriate responses to unpredictable cues. Thus, the potential risks should be well thought-through by all the involved actors before determining the degree to which the robot is autonomous. In particular, if the therapist informs the range of the behavioural potential of the SAR (in the design and/or update phase) in accordance with the needs of the patient, she or he might be partially accountable for the undesirable behaviour of the robot. In this case, it is important to reflect on the cause or the misuse or misbehaviour in order to decide how to distribute the responsibility amongst the involved actors.

Finally, we would like to consider the possible economical and societal effects of introducing SARs in the treatment of AN. As discussed in Chapter 2, AN has considerable financial effects on society, due to the frequent inability of adult AN patients to participate in the labour market on the one hand, and the high costs of AN treatment options on the other hand (Agh et al. 2016). Therefore, a robot that would aid in achieving faster progress in the patient's recovery would be desirable in terms of AN-related costs. However, this conflicts with some of the roles that SARs could take in the home setting. Companion robots will not serve as tool to advance treatment, but rather to relieve some of the emotional burden of the patient. Only the coach robot potentially advances treatment of AN patients. Finally, since the robot will likely be very expensive, it is hard to balance its costs against the potential benefits that might be brought to the treatment of AN patients. In the case of the companion robot, such costs could never be balanced, since there is likely no economic benefit. However, approaching the decision whether or not to have the robot strictly on the related financial costs has a dehumanizing effect and treats the patient into an object of only economic value.

In terms of adjustment costs, the companion robot offers little to adjust for, since it is merely an addition in a field where no previous therapeutic practices existed to disrupt, and very little human intervention is required for making the robot function. On the other hand, the coach role does require adjustment, since both therapist, patient, and family will have to adjust to a robot that serves a role in the therapy process. This would entail training for all the aforementioned stakeholders on how to handle such a robot.

The different roles the robot can take may determine the outcome in terms of accessibility and distribution. The crucial difference lies in whether such a robot is implemented in the treatment of newly-diagnosed patients, or an addition to the treatment of chronic patients. The former is most conceivable with the coaching robot, as such a robot will become part of a therapy. Therefore, it will ostensibly be more easily covered by insurance schemes. This makes the desire to provide potential benefactors equal access more feasible. On the other hand, the companion robot would primarily be used to improve the quality of life which implies a different outlook, and is thus less likely to attract

financial backing from insurance companies. That induces the potential for this robot to increase inequality, since not all patients will then be able to afford the robot. Therefore, a more proactive role for governments in assuring equal access would be necessary here.

Conclusion

For legitimate practical reasons, it seems to have become common practice to test the clinical efficacy of SARs with pre-existing robots. In the previous chapter, we concluded that the ad-hoc modification of pre-existing robots for specific therapeutic purposes is a pragmatic, but perhaps suboptimal choice. Instead, we suggest that robotic applications could benefit from a more preliminary user-oriented research. We believe a more stakeholder-driven approach is needed, in which the involved actors could inform the design and implementation of the SAR with their opinions and expectations. In this regard, we propose that a combination of a literature-based approach and a stakeholders-driven approach could help design and implement ethical and clinically effective robots. Only by bringing together perspectives from the various stakeholders, one could reduce the risk of developing undesired (and perhaps even inadequate) assistive technology from the outset.

In this chapter, we endeavor to establish an initial basis for this approach through a literature-based discussion. By integrating perspectives from ethics, law, economics, medicine and robotics, we aimed to investigate the potentials of expanding the current treatment of AN patients both on a micro- and a macro-level.

2. Interviews

a) Motivation and Approach

The report has thus far explored which theoretical considerations are relevant to the deployment of a robot in the treatment of AN in a home setting. We have considered the particularities of the condition and the treatment settings, how SARs are used in mental health care to help a variety of patient groups, and which social, legal, and ethical issues are worth considering when designing a robot for therapeutic purposes. Finally, we have integrated these different perspectives into a unified framework, thus illustrating how past research can inform the design process of an SAR for an entirely new field.

At the same time, we also seek to explore our topic from the bottom-up, i.e., the issues and preferences that stakeholders in relevant areas have with regards to our project. Stakeholders were identified on the basis of the literature review, and the same broad interdisciplinary scope applies to which persons or groups were eventually included in the project. Conducting semi-structured interviews was identified as the best research strategy to gain an in-depth understanding of the issues and challenges that stakeholders would encounter.

Interviews can be used as a qualitative research method to establish the motivations and preferences that drive someone's actions and decisions. Separating thought processes from the actions they inform is of great importance in the creation of falsifiable theoretical arguments (Rathbun, 2008, p.7). While interviewing data is often criticized for being subjective and non-generalisable, this subjectivity of opinion is a valuable source of information for our research. The direct and targeted methods of interviewing allow us to explore the individual accounts of personal perceived realities of all our interviewees. We can use it to explore the unique worlds of individuals with different backgrounds and thereby investigate how expertise and experience shape perceived reality. Differences and overlap between these perspectives will allow us to map potential problems and solutions to these problems that will satisfy a variety of stakeholders. While we have no way to verify the subjective truth of our participants' accounts, it seems reasonable to assume that participants who volunteer to be interviewed in face-to-face are interested in presenting us with their true opinions. The status of the interview data should still be treated with caution, taking responses as a reflection of a subjective, rather than the objective truth.

To asses a desirable outcome of the use of SAR, the preliminary practical goal of our project is to identify functions and a setting for SAR in which not only the enhancement of the current therapy is pursued, but rather the diversification in therapy potentials through which the SAR would offer an added value. By these means, we aim to circumvent the replacement of already existent crucial and well-functioning elements of therapy and care. Through the direct integration of

stakeholders, we attempt to maximize the likelihood that the different arguments and interests related to an implementation of SARs are well represented and that suggested approaches are in line with the integration of relevant agents and already well-functioning therapy elements.

The suggestions and conclusions that we draw from a literature based attempt are helpful in assessing potentials and concerns based on previously positively assessed applications. However, this might be insufficient and misleading as we want to explore unexpected potentials and concerns when studying a new patient-group that has yet to be investigated within the context of SAR applications. By conducting semi-structured interviews with important stakeholders, we hope to derive a good estimate of the potential and concerns related to the implementation of SAR applications in the home care setting of AN. The stakeholders' perspectives should indicate how the setting of the robot implementation might change in relation to our previous assumptions and whether there are additional concerns that need to be taken into consideration.

We would like to use the results of the conducted interview to put the elements of the ELS-framework into perspective. By approaching the research question from different viewpoints, we would like to analyse potential changes between the interpretation and weighing of the moral values as attentiveness, responsibility, competence and responsiveness and the terms of accountability in case of misuse or mistakes within the SAR care setting, the autonomy of the robot, and the socio-economic consequences of the SAR implementation. Finally, we would like to integrate the distinct perspectives within the literature based assumptions on the SAR implementation in AN patient care to come up with a suggestion on how the design of such robots can be further approached.

b) Experts and Actors

For the development and implementation of an SAR as an addition to the treatment of AN, the following stakeholders were considered as essential: First, patients with AN, since the SAR would be developed for their treatment, and they would have to encounter the robot frequently. If the patients do not accept the robot, an implementation is not possible. Therefore, thoughts and preferences of the patients are required. Second, psychiatrists specialized in AN, because they are in charge of the treatment and, in addition, they are aware of the strengths and limitations of the therapies that are used nowadays. Furthermore, they are familiar with the patients and their behaviour towards social contact and therapy. For identifying future possibilities of SARs, both the perspectives from a specialist in AI and an engineer are crucial. The AI specialist has an understanding about the current and future possibilities in the field of AI, whereas the engineer can further elaborate on technical challenges and opportunities within the field of robotics. Furthermore, a policy maker is necessary for assessing legal and practical challenges in the implementation of

SARs. Because of the considerable effect of SARs on society, an ethicist specialized in the field of AI is also indispensable.

Patients who were diagnosed with AN and who were member of the Dutch patient society for eating disorders *Weet*, were contacted by a message on the forum about AN on the website of *Weet*. Other stakeholders were contacted due to their expertise on each individual area. For each of the aforementioned stakeholders, one expert within this field was contacted.

We interviewed a female patient, 33 years old, who suffers from AN for more than ten years. She considers herself as almost fully recovered, as she has a stable and healthy weight but still experiences anorexic thoughts.

A psychiatrist specialized in the outpatient treatment of adolescents and young adults with eating disorders was consulted. In the last part of the interview, a psychiatrist specialized in patients with a chronic course of AN joined the discussion. We interviewed an AI specialist who has major expertise in the social dimension of human-robot interaction. The engineer that was interviewed is an expert in the field of robotics and mechanical technology. A policymaker who acts in the field of the innovation of technologies in healthcare was interviewed. Moreover, he is familiar with health insurance companies that are relevant in the funding of therapies. We discussed the effect of SARs on society with an ethicist who is an expert in the field of human-robot interaction and the ethics of robot care.

c) Interview Guide Design

The interviews were conducted as semi-structured interviews on the basis of interview guides which contained open-ended questions to the participants. These were designed to allow us to explore the respective topics by the interviewee, resulting in answers that would reflect their thought processes. The interviews used direct questions about the interviewees' own experiences and professional opinions.

Wherever possible, face-to-face interviews were conducted with two researchers present. Locations were chosen based on the preferences of the interviewees. The interviews lasted between 30 and 90 minutes. All interviews were recorded and transcribed verbatim afterwards. Informed consent for participation in and recording of the interview were obtained from all participants. Each participant was interviewed once. Interviewees were chosen based on their expertise with regards to a specific aspect of the field of research and their availability.

The interviews were conducted by six teams of researchers. The setting of those interviews was informal and they were conducted as conversations. Participants were encouraged to express their own views freely and were regularly asked for clarification to ensure they were being

understood correctly. When required to facilitate expansion of the issues, sub-questions were incorporated into the interview. Prior to the interview, participants received an email with a short introduction to the project and the specific goals of the interview. All interviews were opened by repeating the brief introduction to our project and the specific goal of the respective interview. Participants were informed about the general premises and background information on SAR and AN. At the beginning of each interview, participants were asked for their first impression and opinion on the research question.

Agents in therapy setting

The goal of the interviews with actors in the therapy setting is the analysis of the potential setting in which an SAR could be employed for the treatment of AN. The interviews with these actors were based on a number of premises. To stimulate creative and visionary discussions, the actors were told to assume an unlimited potential for the robots, in terms of design, capabilities, financial aspects, and technical feasibility. For maximum openness towards the topic, ethically controversial issues identified in the theoretical framework (see Chapter 1) were tackled by assuming the following aims: The robot should prevent isolation and be deployed only as a temporary addition to current therapeutic approaches. It would not be intended as a replacement for any actor involved in the therapy, but rather broadening the potential in the therapy of AN.

In preparation, the interviews were structured to cover two main topics. The first part of the interview was designed to explore the perception of the agent's own role in the therapy and any unrealized potentials and limitations this role encompasses. In the second part, the agent's perception of the roles of other actors in the therapeutic setting was explored, specifically with regards to the interaction between themselves and any other parties involved. Additionally, the specific role of the robot as a non-human actor was investigated, by tackling questions concerning autonomy and control of the robot, and data collection and management.

Psychiatrist

The interview with the psychiatrist was designed to first explore the current therapy situation based on their own expertise and perception. The interview should elicit a description of current therapeutic approaches and their limitations, as well as suggestions on how these could be overcome given unlimited resources. Furthermore, potential cognitive exercises to be done at home should be identified. This information should be used as a basis to consider the potential for the deployment of a robot in this setting.

The interview guide covered control and autonomy of the robot as a second topic. It covered both the exploration of a scenario in which the therapist would be in control of the robot and one

with only the patient interacting with the robot. Additionally, the questions of access to and use of data from the robot, and the effects of such issues on the therapy were included in this part.

Since the goal of the project is to explore the deployment of an SAR in a home setting, the family setting was covered as a third broad topic. This topic included the consequences of AN on family dynamics, tested by contrasting the roles and tasks of parents of healthy children to those of parents of children with AN. By identifying the differences between these situations, the potential role of the SARto compensate for the discrepancies can be investigated.

The final topic covered in the interview guide was the role of the patient in the treatment process. We targeted the patient's role and tasks in the family dynamics as opposed to the role and tasks of a healthy child in such a setting, in order to identify potential support mechanisms that could be implemented in an SAR.

Patient

The main focus in the interview with the patient was their personal experience with the disorder and the treatment. In particular, their opinion on therapeutic approaches and the role they played for their own recovery was centered to create a clear picture of which aspects of therapy leave room for improvement. Based on this, the potential of a robot to improve therapy could be explored by the patient. This could be exercises or other elements improving therapy, but also general well-being within the home situation.

The patient's perception of the impact of their own disorder on the family situation was targeted in a second set of questions. As in the interview with the psychiatrist, roles and tasks with and without the presence of the disorder were targeted to gain insights into possible conflicts of interests that a SAR could help solve.

The final topic was the patient's creative approach to the design of the SAR. This includes ideas on the SAR's role in the treatment of AN, it's functionalities, and any ideas on the appearance of the SAR. Possible functions of the SAR were inquired by also naming suggestions such as cognitive exercises, social interaction, monitoring, coaching and educating on aspects of diet, daily activities, and the perception of the disorder. This topic also included issues of control and autonomy of the robot in relation to the patient and the therapist.

Other experts

The expert interviews targeted issues of feasibility and potentials and limitations of the setting, as well as technical, ethical, and financial issues. The topics for the expert interviews were different in every interview, allowing us to cover a wide range of topics across the interviews.

The topics covered in the interview with the engineer included the feasibility of introducing an SAR into a home setting, necessary functionalities for an SAR in the treatment of AN, technical

challenges, and cost factors in the design and deployment of such an SAR. All topics were addressed in an applied manner, resulting in questions about specific aspects of appearance, mobility, necessary hardware, sensors, usage interfaces, and data processing and access. Additionally, the interview guide included questions on the current state of the art of SARs used in mental healthcare in a home setting, as well as required steps to reach the goals laid out in the interview. Overall, the interview was structured in such a way to let the engineer suggest robotic solutions to improve care for patients with AN in a home setting and to present a time scale on which such a solution could be realized.

The interview guide for the AI specialist included the same topics as that for the interview with the engineer. A special focus was put on the questions on interaction between the SAR and the patient, and on the robot's ability to initiate interaction. The interaction of modalities and the robot's use of information across modalities was also explored. Overall, the focus was on software aspects of robot design.

In the interview with the ethicist, issues of autonomy and control of the SAR, responsibility, human dignity, and the patient's self-determination. The questions elicited a distinction between what a robot can potentially do and what it ought to be able to do. This provides us with an idea of how ethical concerns can be taken into account in the design phase as well as opening up a discussion about the regulation of potential discrepancies.

The interview with the policy maker focused on theoretical and practical issues related to the introduction of SARs in mental health care as part of a therapy for AN on the level of political regulations and legal implications. The interview guide covered regulations at a national and EU level and their relevance in policy making. Furthermore, the legal and financial consequences of AI technologies within the healthcare system, issues of responsibility, and data availability were included. The questions focused on processes in policy making and the roles of different stakeholders. Additionally, the specific policy-related challenges posed by the introduction of SAR in the care for patients with AN in a home setting were addressed.

Each interview centered the interviewee's expertise while staying relevant to our project. That way, we were able to use the direct, targeted method of interviewing to enrich the information from our theoretical background. This implies that every interview had a slightly different focus, concentrating on the unique perspective that each participant could add to the discussion. The interview data can thus be used to both explore the opinions of a variety of experts (Chapter 2F), and to falsify the theoretically driven claims of the ELS framework (Chapter 3).

d) Data Analysis Procedure

The data collected in the interviews was analysed in an inductive process. The interviews were individually summarized, extracting topics and key statements covered in each interview. This list was then compared to a topics list that was generated from the interview guides. In a next step, the topics list was validated by combining and grouping statements in each interview under the predetermined topics. Wherever new topics emerged in the interview, they were added to the topic list. The results from this data extraction are described in Section 2E of this report. The final list of codes can be found in Table 5.

Table 5:

Final list of codes

Code	Explanation
Role of the SAR	Description of the role of the SAR for the patient or other stakeholders, including issues of relationship, hierarchy, and framing of the SAR
Human values	Consequences that robots might have for human values such as privacy and problems related to the SAR deceiving patients into believing a mutual social bond exists
Non-compliance / manipulativeness of patient	Avoidance of and dealing with issues of non-compliance or manipulativeness of the patient when interacting with the SAR
Active / passive role of the SAR	Description of the SAR's role towards the patient with respect to its ability to initiate contact and its behaviour in interaction
Appearance of the SAR	Descriptions of the outward appearance and hardware of the SAR, including shape, colour, humanoid/animaloid/object characteristics, perceived character or personality, hardware
Technical possibilities	Assessment of technical possibilities, limitations, and issues of feasibility, includes concrete information about software, programming, modeling
Functionality of the SAR	Description of the functions of the SAR and how they are implemented
SAR vs. app	Characteristics of SARs and apps and their differential potential in

	therapy
Data	Issues of data collection, storage, protection, analysis, and use
Privacy	Threats to privacy violations and potential solutions
Legal issues	Legal issues that would accompany the use of SARs in therapy settings in general and specifically in our scenario
Financial aspects of the SAR	Financial aspects regarding development, production and maintenance of the SAR
Others	Any other topics introduced by the interviewee

e) Results of the Individual Interviews

This chapter will provide an overview of the results of the stakeholder-interviews. It will do so by adhering to the final list of codes, which in itself is a result of a meta-analysis conducted on the codes of the individual interviews.

Patient

The patient with AN declared that her first thoughts about SAR mainly concerned curiousness with regard to the possibilities. She patient saw potential in implementing SARs in a home setting. She particularly stressed the companionship role of the SARs. The SARs could provide a discussion partner, someone to talk to when feeling alone or someone to talk to when feeling desperate. The most important thing that was mentioned concerned the fact that the SAR would not be able to judge the patient. However, the patient also reported to see the future SAR as a form of a conscience; someone who could help remind her what was agreed upon with the therapist. However, the SAR was not seen as a therapist or a therapeutic device. The SAR was instead perceived as a “buddy”.

The interviewee mentioned that the monitoring of vital parameters would be a possible feature, but should not form the main focus of the SAR. With regard to retrieved data, the patient claimed ownership of the data. The patient underlined that the data was only to be shared if they wished it to be shared. The patient also claimed responsibility for the treatment and of recovery. The SAR was not held responsible for this. Another possible role for the SAR in younger patients would be as a coach for parents and other family members. The patient stressed that the SAR could never

replace the human interaction completely, but it could be a major improvement to patients' quality of life.

A physical SAR was found to be superior to an app on a tablet or smartphone. The interviewee mentioned that it was comparable to getting a hug from a fur character at a major amusement park: everybody knows it is fake but it still feels great and can provide you with a comfortable feeling. The outer appearance of the SAR was also discussed. The interviewed patient stressed the importance of a humanoid design. An animaloid SAR would look more like a toy and would make it feel less serious. The gender of the humanoid SAR must be optional as both genders might be needed in different situations. The gender of the SAR was discussed as more than just a male or female voice. The stereotypical behaviour of men and women was included in this discussion. Patients reported to have been in need of both types of companionship and feedback at different moments in their life. For example, complimenting, motivating and informing patients are different types of behaviour that were mentioned as equal important but were gender-dependent.

Psychiatrist

In the interview with the psychiatrist, she mentions that her first thoughts about SARs in taking care of AN patients concerned companionship of older AN patients with feelings of loneliness. The companionship role of the SARs was underlined several times, but a coaching role was also discussed. Furthermore, the psychiatrist thought the robot would be a more neutral companion, someone who you can tell anything without judgment. The psychiatrists suggested that the SAR could coach the patient at home by monitoring the daily calorie intake and ensuring that the patients will follow through on agreements made with the psychiatrist. The observing of vital parameters was also mentioned to be innovative and helpful in the treatment of AN. Mainly, the SAR was thought to improve quality of life by treating feelings of loneliness.

The psychiatrist stated that she had had positive experiences with primitive robotics in taking care of AN patients. A few years ago, a computer device could monitor the intake of food per day by weighing each dinner before starting to eat. Although easily manipulated, the psychiatrists noted that this was not usually the case. AN patients were actually content with this primitive device. Therefore, the psychiatrist thought that the non-compliance of patients with the SAR would not be a problem.

With regard to the appearance of the robot, an embodied SAR was preferred. The psychiatrist stated that the requirements of the patients and the exact functions are essential in order to design a successful SAR. Her own suggestion was a plant-like or butterfly-like SAR with soft colours.

Ethicist

The interview with the ethicist yielded several insights regarding the role of the robot. According to the ethicist, the robot should be deployed as a tool for monitoring the patient's dietary habits. Social interaction might function as a way to engage the patient in cooperating in this regard. However, the ethicist stressed the importance of taking moral values such as human dignity into account. According to the interviewee, the dignity of the patient could be threatened if the implementation of a robot created the suggestion of a social bond between patient and robot. The monitoring purpose of the robot might also jeopardize the patient's privacy, another moral value. On the other hand, the robot could enhance the patient's dignity if it really benefitted the recovery.

With regards to the patient's role towards the robot, the interviewee argued that the manipulative behaviour and non-compliance inherent to anorexia could be avoided by necessitating prior informed consent as a requirement for receiving the robot. At the same time, such requirements should not be too stringent, since this might discourage patients.

The ethicist brought forward several concerns regarding the possibility of having a pro-active robot. Firstly, she cast doubt upon the possibility of ever having a fully autonomous robot, as being autonomous for her implies being sentient. This hampers its ability to make moral decisions and has major consequences for what we should and should not want the robot to do. In the end, the interviewee argued that one should be very cautious about transferring responsibilities to the robot. The robot should remain predictable in what it will do when not supervized by a human. The needs of the patient do not necessarily have to be met with an embodied robot. Instead, the ethicist also mentioned the possibility of having an app instead. However, an embodied robot might be more effective in achieving social interaction.

The interviewee problematized the issue of data availability. According to the interviewee, the account for who has access to the information at what time should be clear. Moreover, the therapist should have access to all the data the SAR generates, and this should be clear to patient prior to getting access to the SAR. This mirrors the position on privacy that the interviewee took.

According to the ethicist, a robot that is not autonomous should also not be held responsible for its actions. Instead, responsibility should be distributed among those who build the robot and those who put it into practice. These actors can be both natural and legal persons. In practice, this means that actors such as programmers, corporations who design and sell robots, and therapists who steer the robot can be held responsible. Who is liable at what time should be traced back from the nature of the accident, meaning that different actor(s) can be responsible in different contexts.

Finally, the interviewee stressed the necessity of including ethical concerns both in the design process and during the deployment of the robot.

Policymaker

The policymaker argued that the role of the robot should be to stimulate healthy behaviours in the patient. Regarding the human values that the robot should adhere to, the policymaker argued that it should not exhibit unethical behaviour, but that it is difficult to define what unethical behaviour is and how to measure whether the robot is exhibiting the moral behaviour we expect.

Regarding technical possibilities, the interviewee stressed the importance of programming a script that the robot would operate on, which also implies that the robot acts upon a strictly pre-programmed script.

Data and privacy are two important areas for the policymaker. Determining who has access to the data, how it should be saved, and how it can be protected. Ownership is also problematic, but currently not tackled by the government. The privacy-problem can be partially solved by advances in encryption technology, which prevent unwanted access. Moreover, the privacy-problem can hypothetically be solved by a new legal governance framework in the future.

Regarding certification, so long as SARs do not include any diagnostic functions or perform restricted actions, they are treated as tools or toys and are only loosely regulated. Moreover, legal regulations are only established after implementation of new technology. In terms of liability, the policymaker argued that robots should be faultless and must not interfere with laws and regulations. Regarding decision-making processes of the robot, no clear regulations exist for self-learning robots. According to the policymaker, this is because we cannot predict behaviour of self-learning robots. Regarding the possibility of introducing a legal/electronic personhood for robots in the future, the policymaker did not provide a clear answer. However, he argued that having a code of conduct for robot designers would be necessary in the future.

The costs of the robot could be sponsored by introducing commercial capabilities (e.g., advertising by the robot) in SAR, though the policymaker personally objected to this possibility. Insurance companies were expected to play a major role in allowing patients to use SARs. The government also plays a role through funds for developing new therapies. These could also be used to stimulate or finance the development of robots for therapeutic purposes. The policymaker further stressed that economic consequences such as the cost of robots would require public-private partnerships between corporations building the robot, the government, and healthcare service providers.

With regard to consequences for employment, the policymaker believed that robots might displace workers in healthcare. This might hinder the acceptance and implementation of robots in healthcare.

Finally, the policymaker generally stressed the role that the government has in stimulating technological innovations. The government was described as a hub for bringing together

stakeholders responsible for developing new technologies. The government also signs 'health deals', where stakeholders are brought together in a more binding way.

Engineer

The engineer envisioned the SAR to be a friendly companion to the patient. It should be touchable and cosy, and create a trusted relationship with the patient. One way to achieve such a relationship would be a robot which is designed to "play dumb", i.e. voice or show its inability to do certain tasks. This may help the patient in increasing their self-esteem, by making them realize that they are more capable than the robot because they are able to help the robot out, or by making them feel more valuable.

In order to tackle a patient's potential non-compliance or even manipulativeness, the robot needs to achieve a balance between creating a trusting relationship with the patient and using the data from different sensors in combination to spot attempts at manipulation. However, the engineer also pointed out that a robot might be more difficult to manipulate than a human because it reacts to clearly defined external stimuli and is not biased by emotions and experiences.

Interactions between the patient and the robot can happen through the modalities of speech processing, the robots looking behaviour, and touch. According to the engineer, both the patient and the robot should be able to initiate such interaction, with the robot acting on scenarios such as the patient moving towards it, or wait intervals in which the patient isn't active.

The engineer was mainly focusing on the idea of a robot that has the shape of a chair with an integrated weighing system in the sitting surface and cameras as eyes in the back. Additionally, he mentioned that the robot could also be a stuffed toy, but he didn't specify what the functions or outer appearance of such a robot would be.

The main technical challenge in designing an SAR according to the engineer is to model the progress of the disease and therapy. In order to program the SARs behaviour towards the patient, the psychological knowledge of how the disease progresses and when to take which steps in therapy needs to be translated into mathematical models. Once such models are created, they can be stored in modules the SAR system moves through as the therapy progresses. Self-learning was proposed by the engineer as a tool to optimize and personalize behaviour on top of this baseline. Having an SAR learn optimal strategies by learning on-site would be both inefficient and unethical, as such a system needs large amounts of data for training, defining success criteria may be extremely difficult, and the SAR would be inefficient during a long training phase. The SAR would receive its input from a pressure sensor in the sitting surface of the chair, ultrasound, cameras, and a microphone.

Using these sensors, the SAR can collect weight data and monitor changes in weight, measure the proximity of the patient through ultrasound in order to start interaction when the

patient is close, and interact with the patient through speech processing and looking behaviours. The SAR is not mobile, as the engineer did not see any added value in mobility.

Using an SAR has a number of advantages over an app, according to the engineer. These advantages all result from the embodied nature of the robot, which allows the patient to establish a trusting relationship and perceive the robot as a touchable agent. This supports rapport building between the patient and the SAR.

The sensors will yield data about when someone was close to the SAR, weight data, and the data from speech processing and interaction. The SAR could store this data in the cloud or on an external computer and it should codify and store the progress of the therapy in order to make the data meaningful. Asked who should have access to and control over the data, the engineer initially thought the therapist should be in charge of the data. However, after reconsidering the question, he changed his mind in favour of a solution in which the patient has access to and control over all the data and chooses which data to share with the therapist. According to him, such an approach might strengthen the trust between SAR and patient and alleviate the violation of privacy entailed in bringing a robot into the home of a patient.

The costs of such an SAR will depend on the features and hardware, but also on the market size and specificity of the design. A non-mobile SAR with only a limited number of sensors, as proposed by the engineer, would be easier to build and cheaper than a multipurpose robot that can move around the house. Additionally, a smart weighing device with a sitting surface may also be interesting to other populations such as the elderly or people with a mobility impairment for whom standing on a scale may pose problems. This would increase the market size and thereby reduce production costs. The affordability for the individual patient depends on both the cost of building the robot and financing schemes put into place. The engineer proposed that insurances might pay for it, or that schemes may be put into place in which clinics, insurances, or private companies rent out the robots to patients. These proposals are not mutually exclusive.

Additionally to the topics identified in the topic list, the engineer also saw a potential role for a smart system in the prevention of AN. He thought that a smart weighing system could nudge at-risk teenagers towards healthy eating behaviours or alert them to the risk of developing an eating disorder. This could potentially help in prevention of the disorder or lowering the threshold of seeking therapeutic help in a very early phase of the disorder.

AI expert

The AI expert differentiated between an interactional and a functional role for the robot. In terms of interaction, he saw the SAR as a coach or peer to the patient, with questions of hierarchy between

patient and robot playing a central role in the design of the interaction. The functional role would entail informational and behavioural features, which are discussed at a later point of this section.

An SAR could make use of the auditory and visual modality, as well as of touch and smell. The AI expert found that touch can be a potent tool for initiating contact in human-robot interaction. This implies a robot that is able to initiate interaction and can therefore have an active role in such exchanges. Speech-processing based interaction is currently only possible in a question-answer format, where any variation in the speech input from the patient as well as in the answers of the robot is computationally very complex.

The software of SARs could be personalized by borrowing from smartphone applications. One example would be learning to disambiguate references, such as identifying where "home" is for an individual patient. Speaker identification is currently not reliable, making personalization within one household, e.g. between patients, parents and siblings impossible. The AI expert stressed that speech-based interaction is still in its infancy and may be very difficult to implement.

As mentioned above, the robot has two functional roles. In its information-related function, it is gathering, storing, and analysing objective data from its sensors. In its behavioural function, it is influencing the patient's behaviour following a therapeutic plan.

The AI expert problematized the use of a robot in a home setting, stressing the lower costs and higher accessibility and acceptability of apps. On the other hand, he saw some potential advantages for robots in their being touchable and embedding the system in a character.

The data gathered by the robot in the home should be stored and controlled by a computer in the home, which is external to the robot. This would provide better data security and privacy protection than uploading the data to the cloud. The patient would thus have access to the data on the computer at home and could choose which data to share with the therapist by bringing it to sessions.

With regards to legal regulations, the AI expert discussed the distinction between coaching and clinical tools. While the former are easier to get approved and may be cheaper in terms of insurance, the latter may have a higher acceptance by patients, parents, and clinicians. Stakeholders may also hold higher expectations towards clinical tools.

f) Analysis and Discussion

The interviews with the actors and experts provided us with an overview about the different perspectives that inform the possibility space for the design of an SAR to support the treatment of AN in a home setting. In this discussion, we will go through the issues discussed in the interviews and contrast the opinions brought forward by the different interviewees. This approach clarifies the

issues at stake, illustrates diverging approaches, and highlights possible solutions to bring the different perspectives together.

Role of the SAR

The possible roles of the SAR were discussed with all the interviewed experts. The psychiatrist, patient, ethicist, AI specialist and policymaker agreed on the role of a coach. The psychiatrist proposed the SAR to be used in order to regenerate social contacts and activate the patients to leave the house. SARs could also help with the cognitive remediation therapy and behavioural experiments in the home setting. The psychiatrist was reserved about the added value of a robot. The patient, ethicist and the policymaker discussed the possibility to monitor eating habits and intake and instigate healthy behaviours. The AI specialist did not further specify the tasks of SARs. The patient also discussed the role of SARs in family setting: it could help the parents deal with the child with AN and it could help adult patients with nutritional advices and monitoring.

Both the patient and the engineer proposed the SAR as a companion. However, their approach to the role differed. The engineer discussed a less intelligent SAR in order to increase the patient's feeling of value and usefulness. The patient, in contrast, valued a companion to share experiences and emotions.

The interviews revealed that the key roles an SAR may take on in the treatment of AN are the role of a coach, or the role of a companion. Social interaction seemed to be of particular value, though the SAR may also use such skills to affect the patient's behaviours and thoughts.

Human values

The topic of human values in SAR was discussed with all the experts. The psychiatrist could not provide the answer to the question whether a robot could be seen as more trustworthy compared to human.

The hierarchy in the human robot relationship was addressed by the AI specialist, the patient and the engineer. Both the patient and the engineer stated that a human is placed above a robot in a relationship. The AI specialist could not provide the definite answer on this question.

The ethicist also warned for a possible threat to the dignity for the patient if the implementation of a robot created the suggestion of a social bond between patient and robot. The patient's dignity could enhance if it really benefitted the recovery, though.

Overall, the interviewees did show some concerns regarding the protection of human values when introducing an SAR into the home setting. However, only the ethicist conceived SARs as potential threats for human dignity, while the other interviewees focused on aspects of a human-robot relationship.

Non-compliance/manipulativeness of the patient

The issues of non-compliance and manipulativeness were considered to not be specific to the introduction of an SAR into the treatment plan. The ethicist and patient agreed that the patient is ultimately responsible for their recovery and that an SAR can only be effective if the patient accepts it. The psychiatrist considered non-compliance a potential problem but did not have any concrete proposals on how to solve this. The ethicist proposed to make compliance a requirement for an SAR approach to overcome this problem.

Both the ethicist and the engineer centered the patient's willingness to interact as a central requirement for overcoming non-compliance. However, their approach to realizing this differed. While the ethicist thought that compliance from the patient could be achieved by putting the patient in control of when to start interaction, the engineer thought that a solution would be to put the patient in charge of the data collected by the SAR. They would thus get to choose which data to share with the therapist, increasing their perceived control and potentially willingness to accept this interference with their private sphere.

Additionally, the engineer pointed out that any SAR will react to concrete external stimuli and are thus more difficult to manipulate than humans. The system could also use a combination of data from different sensors to spot attempts of manipulation.

Manipulativeness was not seen as an important issue by the experts and actors. However, non-compliance was seen as a potential problem. The interviewees agreed that making interactions attractive and voluntary for the patient may help. By giving the patients control over interactions and data generation, they might be more willing to make use of the SAR.

Interaction between SAR and patient

Only three experts directly addressed the interaction between SAR and patient. The patient opinionated that interactions should happen on the same hierarchical level. The engineer and the AI specialist were more concerned with the form and modality of interaction.

Modalities deemed available for SARs were the auditory, visual, tactile, and potentially olfactory modality. However, neither expert could think of a useful application for smell in the described scenario. While the engineer saw no fundamental issues with implementing speech-processing for interaction, the AI specialist thought that interaction would remain confined to a question and answer format for any short or medium term developments. He also pointed out that any variation on the part of the robot or the patient would be computationally highly complex.

The experts who addressed interaction specifically showed a clear preference for interactions that mirror the modalities of human-human interaction. It is important to keep in mind that the setting and the aims of the interviews created some disagreement about the feasibility of such

human-like interaction. One of the goals of the interviews with the AI specialist and the engineer was to gain more knowledge on current developments in the human-robot interaction and communication means. The patient, on the contrary, was told to imagine limitless possibilities during the interview as we aimed to focus on the needs of this patient group.

Active/passive role of the SAR

The ethicist problematized the question whether an SAR can be considered autonomous and what this entails. She sees a programmed entity as incapable of autonomy in a human sense, and therefore as incapable of theory of mind and moral decision making. Based on this, she finds it important to limit the behavioural options of the SAR, by allowing it to perform only a limited set of predetermined tasks. She thinks that any complex interactions should be reserved for human actors.

The psychiatrist, patient, engineer, and AI expert talked about autonomy in the sense of initiation of interactions. According to the psychiatrist, only the patient should be able to initiate interactions through unambiguous behaviours. The patient, engineer, and AI expert are more open to initiation on the part of the SAR and the patient. However, the patient stressed the importance of retaining the patient's autonomy by allowing them to deny interaction or refuse to follow advice. The engineer specified that the SAR should never disrupt the patient's activities, but only initiate interaction if the patient is not doing anything.

The interviews revealed that autonomy in a human sense is impossible and not desirable for an SAR. However, autonomy in a robotic sense was perceived as useful by most experts. Such autonomy should be limited by the patient's willingness to interact with the robot and should not overrule their decisions.

Appearance of the SAR

The appearance of the SAR was touched upon by the patient and the psychiatrist, as well as by the engineer. The psychiatrist focused a lot on the aesthetic value of the SAR. She argued that the patients place a great importance on aesthetics and that the SAR should therefore be very girly, softly coloured and lovely to look at. She also thought that it should be small, so that they can also take it outside the home. The patient's ideas are contrary to some of those claims. She thought that the SAR should be humanoid and soft, with the feel and looks of a human and of substantial size. She said that it should have no toy-ish appearance or sound and speak with an adult voice. She also thought an option to choose a male or female voice is very important. Finally, she found intonation and emotional expression to be desirable.

The engineers imagined SAR diverged very much from those ideas. He came up with two distinct models, one which looks like a chair with arms, a sitting surface, and eyes in the back, and

with a stuffed toy with unspecified characteristics. The chair would be relatively simple, but could vary in colour and similar external design features.

The interviews revealed that the experts and actors had no clear idea about what an SAR could look like and which characteristics of the patient or SAR should drive design choices. While the engineer based his choices on functional considerations, the patient and psychiatrist found it extremely difficult to make such choices. The appearances they described seemed to be based on personal preference, past experiences with AN patients and the envisioned role of the SAR in therapy.

Technical possibilities

Various opinions were encountered on the topic of technical possibilities. The ethicist proposed to limit the learning abilities of SARs in order to prevent it from taking over bad habits. Similarly, the policymaker and the engineer discussed pre-programmed behavioural scenarios for SARs. The policymaker mentioned the importance of a strict script upon which the robot could act. Based on technical, rather than ethical reasoning, the engineer proposed a mathematical model to monitor the phases of the disease and to adjust the behaviour of the robot upon it. However, this would require very profound knowledge of the disease which might not be available yet. Ideally this model could be optimized and personalized through self-learning on-spot. The AI specialist also elaborated on the importance of personalization of the applied SAR with approaches borrowed from smartphone applications.

The patient stated the importance of good conversational skills and person recognition. However, the AI specialist noted that the speaker identification is not reliable yet and the speech interaction is technically difficult. Overall, the questions about technical possibilities revealed a large gap between the hopes and ideas of the patient on what the SAR could do, and what is technically feasible at this point.

Functionality of the SAR

The topic of the functionality of the SAR was discussed with the psychiatrist, the AI specialist, the patient and the engineer. Vital signs monitoring was considered optional by the patient, the engineer and the psychiatrist. The psychiatrist argued that this kind of monitoring would not have any direct therapeutic consequences. The engineer also discussed the possibility to monitor weight.

Both the AI specialist and the psychiatrist elaborated on collection of objective data in home setting. The psychiatrist specified that the data about the behavioural interventions at home could be collected and discussed during therapy.

Communication was considered important by the experts. The engineer proposed the use of pressure sensors, ultrasound, microphone and camera in order to improve it.

The proposed functionalities can be categorized in three categories, namely behavioural monitoring, data collection, and communication. There was general agreement on the desirability of such functions, but some disagreement about their feasibility.

SAR vs. apps

The topic of comparing SAR and apps was brought up by the ethicist, the AI specialist, and the engineer. All three experts perceived apps as cheaper, more accessible, and more accepted. The main advantage of using an embodied SAR was perceived to be their touchability and the possibility to see the SAR as a character facilitating rapport building. The engineer also suggested that the embodied nature of an SAR might support the building of a trusting relationship.

Data and privacy

The matters of privacy, data ownership and access were addressed by all the experts. The psychiatrist proposed to use the collected data in therapy. Furthermore, the patient and the psychiatrist could program the SAR together during the therapy in order to improve the use. The patient also discussed the possibility of the use of the data in therapy, though it was optional as the data belongs to the patient. The ethicist agreed with the patient to have control over the data and stressed the importance of clear account for all the involved parties of who has the access to the data.

Data storage protection were addressed by the AI specialist, the engineer and the policymaker. The AI specialist proposed to connect the SAR to a computer at home and to store all the data there. The engineer advised to codify and store all the data in the cloud or on the external computer. The policymaker stressed that the data access, storage and protection is an issue that is being considered by the government now. Due to the advances in encryption technology the privacy problem can be solved with current regulations. Though a good governance legal framework will be needed in the future.

Issues of data and privacy were important to all interviewees. The question of access to the data was of particular interest, revealing disagreement about whether or not the therapist should have direct access to the data, or whether all access should be monitored by the patient. These differences could be explained by the conflict of interests between the stakeholders.

Legal issues

Legal issues were discussed by the ethicist, AI specialist and the policymaker. The policymaker and the AI specialist mentioned that the robots are regulated loosely as long as they serve as a coaching tool and do not include diagnostic functions or perform restricted actions. The policymaker discussed that there are no clear regulations for self-learning robots at this moment as there is no way to predict their behaviour. More conduct will be necessary in the future.

The ethicist proposed to distribute the responsibility amongst the involved actors as the robots cannot be held responsible. In that case, different actors would be responsible for different types of mistakes.

The interviews showed that concrete legal frameworks that take into account the specific characteristics and challenges of SARs are currently missing. Legal solutions seem to depend on framing and definitions, rather than specific features and capabilities.

Financial aspects of the SAR

The policymaker and the engineer discussed the financial aspects of the SAR. The policymaker stressed the importance of cooperations between companies and health service providers. The engineer elaborated that and proposed a system where a robot can be rented from a company, a clinic or an insurance company. The policymaker also mentioned the funds for the development of new therapies.

According to the policymaker the robots can make healthcare cheaper and more efficient, eventually. However, the long term effects should be taken in consideration as increased use of robots may alter employment and profit. This may lead to a reduced acceptance of the new technology, which can hinder the implementation.

Financial aspects seemed to be dependent on the legal status, the technical complexity, and the market size. Several financing schemes were proposed for the deployment of the SAR, some of which again depend on the legal status. If interpreted as therapeutic tools, insurances would be responsible for the payment, while a coaching tool may depend on individual financing schemes.

Others

Only very few topics were brought up by the interviewees themselves, despite the open question about further issues at the end of every interview. The policymaker stressed the importance of governmental agencies in innovation process as they initiate the contact between the stakeholders. The engineer elaborated on the topic of disease prevention by proposing smart weighing systems which can be used by at risk teenagers as a tool to promote healthy behaviours.

Strengths and limitations

Strengths of this study can be found in the multidisciplinary view of all the actors, the broadness of topics discussed and the relatively new approach of research within this field. A limitation is formed by the structured interviews, which complicates gaining in depth knowledge. With regard to qualitative investigation, the rule of saturation must be followed. By interviewing only a limited amount of stakeholders of each group, saturation cannot have been reached on all the topics discussed in this study. The little amount of time in which this research has been conducted, formed a final limitation.

Future investigations should implement a qualitative research methodology with semi- or unstructured interviews, which enables the gaining of in-depth knowledge in this field. In addition, the use of a focus group discussion with all the involved stakeholders may provide a solution for the encountered disagreements in feasibility. The proof of concept presented in this study, combined with the aforementioned limitations, can be used as a blueprint for further research, not only in AN but also in other fields in which the way that robots could be beneficial to society is investigated.

3. Evaluation

The guiding question motivating our project was: How can we enhance the likelihood of a desirable outcome for the proliferation of robots in our society? This question was combined with a more specific research question, namely: How could and should a socially assistive robot be implemented in the treatment of anorexia nervosa? The two-step design of this study, including an extensive literature review and interviews with experts and stakeholders in various fields, provided a broad insight into the matter and possible challenges associated with the topic.

To succinctly address the both questions, both steps of the study provided possibilities for the introduction of an SAR for the treatment of AN. Depending on which theoretical perspective and which stakeholder is consulted, details about how such a robot can be realized might be different. The integrated theoretical perspective sees potential for different robots based on the type of AN patient. For instance, a companion-type SAR might be most appropriate for chronic patients, whereas a coach-type SAR is more suited for non-chronic patients. Interviews with the patient and therapist gave reasons to support this expectation. However, in some cases a large gap exists between expectations of various stakeholders on the one hand, and theory on the other hand. For instance, the patient preferred a sophisticated companion robot that is able to interact on a complex social level, whereas the engineer saw more potential in a simple companion robot. Moreover, the patient wished to have complete control over the robot, whereas this might pose a problem when the purpose of the SAR is to be a valuable addition to the treatment.

These contradictions complicate and decentralize the answer to the more focused research question. The results show that no clear-cut, unambiguous answer exists for how an SAR for the treatment of AN should and could look like. The most desirable robot for each expert and stakeholder differs in terms of appearance, functionality, role, and whether the robot should be proactive in its interaction and how that interaction should take shape. Moreover, the many forms of AN treatments make it difficult to determine what the needs of the individual patients are. Therefore, if any such robot would ever be developed, it would always require a high degree of personalization towards the needs and preferences of both the patients and therapist.

Beyond the initial positive answer to the guiding question, it is important to stress that the positive answer is conditional to addressing several points of concern. The results from theory and practice are hard to reconcile on multiple points. The first of those is what we define as human values. Human values are ideas that relate to the patient's dignity and the human element in the care that they receive. On the one hand, the very fact that our target group has a severe condition such as AN could affect their dignity. The emotional burden on patients, family, and the difficulty of the therapist in treating AN patients could be relieved by allowing an SAR to play a role in the

therapy process. However, such robots could also complicate roles and responsibility of stakeholders, or even effectively arrogate certain tasks to themselves that in reality should be conducted by human actors, according to the ethical and legal parts of the ELS framework. In general, theory and practice collide when more complicated robots are introduced. Whereas the patient preferred a complicated companion-like robot, this is currently not technologically possible. However, technology develops fast and it is imaginable that such a complicated robot will be feasible within the next decades. Yet, in the case of a complicated SAR, the ELS framework would see major problems with regard to (moral) responsibility. The sheer complexity of such a robot would complicate questions of who is responsible, while the robot would also become more of an agent in between patient, family, and therapist, thus hampering the ability to effectively monitor the patient and keep track of the needs and progresses of the therapy.

Another concern in introducing SARs to AN treatment are questions addressing data sharing and privacy. From an AI perspective, more data is always better. For an appropriate and personalized functioning of the robot, it needs a large database to learn its behaviour from. It is therefore required that the SAR collects and stores data about the patient. This however, raises concerns about the ownership and the sharing of the data. Different stakeholders have colliding notions about this: To optimize the use of SAR for therapy success, the therapist would prefer to have access to as much data as possible. Per contra, the patients might want to protect their very personal data. Additionally, knowing that data will be accessible by the therapist could have a negative impact on the patient-SAR relation, in the sense that the patient does not feel comfortable around it and has the feeling of being watched. Another institution that might have an interest in accessing the data are companies developing SARs since an improvement and optimization of SARs is only possible with feedback from actual users. These different interests are difficult to reconcile. In the end, this conflict can be defined as a trade-off between patients' privacy (i.e. not sharing data) and optimization of (future) treatment (i.e. sharing data). This might be best solved on an individual basis, letting the patient, or in the case of underaged patients their parents or legal guardians, decide.

During the interviews, several practical aspects of SAR implementation emerged that were not entirely covered by the background literature review. One point that stood out was the financial aspect. Although it was postulated that every potential benefactor should have equal opportunity to access novel treatment possibilities such as SARs, the realization of such an ideal situation is often difficult. Especially in the initial stages, the development and the production of the SAR will be very costly. These costs, however could be lowered by producing a large number of robots once they have been developed. Additionally, costs for technologies in general are thought to decrease continually in the future. Some possible solutions for the immediate situation emerged particularly from the policy maker's interview. He proposed that either insurances could finance a personal SAR, and that

hospitals or comparable healthcare facilities rent out SARs. In either case, it should be made sure that patients are not disadvantaged because of financial or accessibility reasons.

The final concern that should be kept in mind is the current technological limitation of SARs. Working with AN patients requires high level of intelligence to recognize the mood, emotions and behaviour which has not been reached by AI yet. Furthermore, various functions like mobility, conversational abilities, speech recognition and differentiation between persons are still limited.

Outlook: robots in our society

This project attempted to investigate the rather specific case of the possible use of SARs in the care and treatment of AN patients. However, the specific considerations could also have more wide-reaching implications. In this part, it will be evaluated how far our guiding question can be answered based on results of this project.

So far, SARs have only been implemented in a somewhat restricted set of populations, both in terms of age (children and elderly people) and needs (ASD, dementia, depression/ loneliness). This case presents an initial investigation to significantly expand that set. On the one hand, AN patients fill the age gap between children and elderly since most cases of AN occur during adolescence and early adulthood. On the other hand, AN patients introduce a new set of both physiological and psychological needs that can be targeted by SARs. Therefore, this study presents an ideal context for expanding appliances of SARs to new areas of mental healthcare.

Additionally, AN has proved difficult for already existing treatment options. Despite a large body of research on AN treatment, AN patients have an unfavourable recovery prognosis, a relatively large risk for developing a chronic course of the disorder, and high suicide rates (see Chapter 1a). A successful implementation of SARs in this context could present a crucial case to further develop SAR-based treatments for other (mental) disorders.

An even wider-reaching implication are the inferences this research allows us to make for introducing SARs to the general population. With our interdisciplinary literature-based approach and stakeholder-driven empirical investigation, considerations from a vast range of perspectives were included. This led to the discovery of several new insights on if and how SARs should be implemented. Many of those considerations also hold in a non-clinical population, such as issues regarding the safety of the robot and who is responsible in the case of mistake or misuse, the privacy of the patient, and finally human values such as the dignity of the patient when a robot is involved in the treatment process. Crucial points in how a desirable outcome can be achieved have been summarized in Table 3, Chapter 1.

Another general issue identified in the course of this project concerns the approach of developing an SAR. For our study, we decided to follow a two-step approach, the first step being an

in-depth review of existing literature, and the second step being interviews with experts and potentially affected persons. Every part contributed in a significant way, and we therefore validated the original decision to take a broad outlook when addressing the needs and challenges in designing a new SAR. Additionally, the contradictions between what different stakeholders found desirable show how crucial it is to first assess existing needs in a target population and to subsequently analyse how these needs could be met and improved upon. In a final step, it can be assessed whether and how SARs can help to improve the situation. By using this need-driven approach, an implementation of SARs that is more sensitive towards ethical, legal, and social concerns can be achieved.

This in turn reflects the need for an interdisciplinary exchange. Many research topics can be addressed by different disciplines but nowhere is the need to do so as big as in the development of SARs. Only patients themselves can communicate their needs, only therapists can point to gaps in current treatments, and only engineers and AI experts can inform about state-of-the-art technical possibilities and limitations. Additionally, now that we are in the initial phases of introducing SARs into society, it is crucial to include policy makers in the discussion since they can point out legal and practical factors of financing and introducing a robot. Finally, it is desirable for a good outcome to discuss possible long-term consequences with experts in ethics to make sure no human values are violated in the process of introducing and proliferating SARs. Research teams that fail to cover the whole spectrum of perspectives run the risk of neglecting crucial points. Possible consequences range from an SAR that presents no valuable addition to the target group (e.g. when not including a therapist's perspective) or are inappropriate for daily use (e.g. when neglecting the patient's perspective) to unrealistic expectations for an SAR (e.g. without the input from AI experts and engineers) or an SAR that violates principles of human values (e.g. without consulting an ethicist).

Another point of further discussion that was encountered during the research was the difficulty SARs provide to our current notions of responsibility and product liability. Flexible robots that make online decisions autonomously from human control are a likely next step in how SARs will evolve. This, combined with the increasingly blurry line between human and robot, requires a new legal framework to cope with. Possible solutions such as an electronic personhood are far-reaching, but in anticipation of the future we should not eschew from this discussion.

Reflection on the approach and design of the research

In the introduction, the interdisciplinary nature of this study was discussed. To briefly recapitulate, the key trade-off in conducting this type of research is the possibility to tackle broad topics on the one hand, and the potential to lose the depth of intradisciplinary discussions on the other hand. Additionally, interdisciplinary research could benefit the problem-solving capacities of a research project, to the possible disadvantage in exploring more fundamental questions. Moreover, without

careful integration, different perspectives run the risk of talking past each other. This section seeks to reflect on the challenges, strengths, and weaknesses regarding the interdisciplinary approach that were encountered during the project, and more generally the strengths and limitations of the chosen research design.

A major strength of the report is the broad outlook that the combination of different perspectives allows for. Approaching the topic of SARs from perspectives that are not directly connected to their instrumental implementation allows to more fundamentally assess the desirability of implementing them in the first place. The intradisciplinary depth that is possibly at risk in combining the different perspectives has been retained by first allowing distinct theoretical perspectives to evolve around the guiding question and research question. Finally, the bifurcation of guiding and research question has allowed for both fundamental and problem-solving questions to be tackled.

The biggest challenge regarding the interdisciplinary approach taken in this study has proved to be the integration of different theoretical perspectives. This challenge recurred during all phases of the project. Firstly, when the project was being conceptualized, the difficulty lied in defining the approach and scope since such a wide variety of perspectives convened. Some perspectives naturally require the researcher to investigate clearly defined, narrowly focused explananda, whereas other perspectives tend to approach the same issues more broadly. Moreover, several of the perspectives normally operate on a micro-level (i.e., the level of individual patients), whereas others operate on the macro-level (i.e., when the population of individual patients is aggregated). The differences in approaches were eventually mitigated by posing both a broad guiding question and a more narrowly focused research question.

Secondly, during the research phase difficulties were encountered with regard to conducting the literature study on the current usage of SARs in healthcare settings. On the one hand, part of our group preferred a systematic approach where literature from medical trials would be studied. Other members preferred a more inductive approach that included literature from robotics, in order to specify the roles that Socially Assistive Robotics could take in a therapy setting. In the end, this conflict was resolved by allowing the latter group to introduce the chapter and by writing the discussion where the roles were inferred from the results of the systematic approach.

Finally, it proved difficult to integrate the various perspectives during the writing phase. Even though the ELS framework was envisioned to aid in structuring the various findings that the theoretical perspectives produced, actual integration remained difficult due to differing starting points that members of the think tank involved in the chapter had. Whereas some group members approached the integration of perspectives with the therapeutic setting and patients in mind, others sought an approach that was inspired by the elements of the ELS framework and by the possible

roles for the SAR based on the literature. Thorough discussions and intensive collaboration was necessary to overcome the colliding approaches.

Regarding the chosen research method of semi-structured interviews, we argue that a sufficient in-depth understanding of stakeholders' motivations and preferences has been attained. At the same time, it has proven especially difficult to engage current and former AN patients. This might be due to the exploratory nature of our study, but possibly also due to the severity of the patients' condition, often combined with social anxiety problems and the reluctance of AN patients to acknowledge their condition in the first place.

Limitations

One serious limitation in our approach is that the limited number of interviewed experts reduce the generalisability of our findings. Practical reasons required us to only consult one person per expertise. This does not allow us to make inferences on which parts of the answers are based on personal opinion or taste and which parts are based on a more objective notion, especially when stakeholders also advocate personal interests.

Another limitation of this project is its exploratory approach. In this case, it was impossible to conduct confirmatory, hypothesis-driven research, since this is the first time that SARs are investigated in the care of AN patients. Also, past research investigating SARs in other areas of mental health has to be treated with caution. Most studies have only relatively low sample sizes and barely a suitable control group (see Chapter 1c, and Rabbitt et al., 2015). This is also reflected in the low GRADE scores (see Table 4). The theoretical background that this report is based upon, therefore lacks substance. However, we believe that this gives even more reason to add to the research field.

Finally, this project does not allow inferences on economic and financial consequences. Ideally, an SAR would eventually lower costs, both for the therapy by providing a more effective and more successful treatment and for society by making it possible for (adult) patients to re-enter the job market. This is speculative however, and should be investigated in further research.

4. Conclusion

This report sought to assess how a Socially Assistive Robot (SAR) could enhance the treatment of patients suffering from Anorexia Nervosa, while seeking ways to mitigate potential concerns and obstacles. The guiding motivation behind this effort was the assumption that robots will likely evolve into more competent and complex machines, introducing potential to enhance our lives, while also raising concerns about the increasingly blurred distinctions between human and machine. A literature-based study of various perspectives allowed for the formulation of a framework that assessed the needs of patients and potential obstacles in meeting those needs. Empirical research was conducted to see whether the stakeholders involved saw potential for such a robot and how they envisioned it.

The theoretical part revealed a potential for introducing SARs to enhance current treatment practices, but only as long as ethical, legal, and social concerns are taken into account. It also revealed a strong methodological discrepancy between the different fields working on using SAR in therapeutic settings. Additionally, a balance needs to be struck between the needs of patients, families, and therapists, and the technical possibilities of creating the envisioned SAR.

The empirical part shows that the needs of patients in therapy are extremely complex, necessitating advanced communication skills and complex social behaviours on the side of the SAR. At the same time, human-robot interaction is currently limited to simple question-answer formats and prespecified scenarios. These technical limitations are mirrored in the fields of mobility, the ability of robots to perform a variety of tasks, and their capability for autonomous behaviour. All of which are technically complex to realize and therefore expensive. Finally, the empirical study confirmed the important role that ethical, legal, and social considerations play throughout the development and implementation of SARs.

In terms of future outlook, we hope that the findings of this study will be used to strengthen explorations in introducing robots to new fields, by allowing for a more holistic and integrated approach that is both need-driven and sensitive to ethical, legal, and social concerns.

5. Recommendations

- The introduction of Socially Assistive Robotics (SARs) into both existing and new fields of healthcare requires an approach that centres stakeholders' needs, while remaining sensitive to ethical, legal, and social concerns. This can be achieved by:
    - Using the Ethical, Legal, and Social (ELS) framework to establish the role of the robot, integrate concerns regarding the human values essential to good care, distribute responsibility, and ensure equal access of the robot to all potential benefactors;
    - Using insights from relevant fields to understand the specifics of the medical condition and the technical requirements for designing the SAR;
    - Inductively assessing the preferences of stakeholders in the healthcare domain regarding the functioning, appearance, and role that the robot should take;
    - Striking a clear balance between preferences of patient and therapist regarding the degree of control over the functioning of the robot and management of the collected data.
- With regard to the specific case of Anorexia Nervosa, a differentiation should be made between adolescent and adult patients.
    - For adolescents, the robot should take the role of a coach, to provide the patient with dietary advice and motivate compliance for the therapeutic program.
    - For adults, we see more potential for a companion robot to alleviate the patients' social isolation.
- Development of SARs in healthcare always necessitate a personalized approach. This can be achieved by:
    - Consulting the patient and other stakeholders to learn their preferences;
    - Capitalizing on the wide variety of options available in hardware and software.
- Future studies assessing the introduction of SARs in healthcare should seek to conduct focus group discussions with stakeholders to further clarify their needs.
- Scientific research on SARs should seek to deploy controlled trials and good experimental designs to enhance their explanatory power and generalisability.
- Researchers must avoid approaching the topic only in a problem-solving manner, and also dare to ask more fundamental questions.
- An interdisciplinary approach is the way to go forward for enhancing the likelihood of a desirable outcome for introducing more complex and capable robots in society.

6. References

Abel, J. R., & Deitz, R. (2012). Job polarization and rising inequality in the nation and the New York-northern New Jersey region.

Agh, T., Kovács, G., Supina, D., Pawaskar, M., Herman, B. K., Vokó, Z., & Sheehan, D. V. (2016). A systematic review of the health-related quality of life and economic burdens of anorexia nervosa, bulimia nervosa, and binge eating disorder. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, *21*(3), 353-364-4909.

Alemi, M., Ghanbarzadeh, A., Meghdari, A., & Moghadam, L. J. (2016). Clinical application of a humanoid robot in pediatric cancer interventions. *International Journal of Social Robotics*, *8*(5), 743-759-4791.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. American Psychiatric Pub.

Anderson, M., & Anderson, S. L. (2007). Machine ethics: Creating an ethical intelligent agent. *AI Magazine*, *28*(4), 15 %@ 0738-4602.

Arcelus, J., Mitchell, A., Wales, J., & Nielsen, S. (2011). Mortality Rates in Patients With Anorexia Nervosa and Other Eating Disorders A Meta-analysis of 36 Studies. *Arch Gen Psychiatry*, *68*(7), 724–731. https://doi.org/10.1001/archgenpsychiatry.2011.74

Arntz, M., Gregory, T., & Zierahn, U. (2016). The risk of automation for jobs in OECD countries: A comparative analysis. *OECD Social, Employment, and Migration Working Papers*, (189), 0_1 %@ 1815-199X.

Asaro, P. M. (2007). Robots and responsibility from a legal perspective. *Proceedings of the IEEE*, 20–24.

Audenaert, K., Van Laere, K., Dumont, F., Vervaet, M., Goethals, I., Slegers, G., … Dierckx, R. A. (2003). Decreased 5-HT2a Receptor Binding in Patients with Anorexia Nervosa. *Journal of Nuclear Medicine*, *44*(2), 163–169. Retrieved from http://jnm.snmjournals.org/content/44/2/163.abstract

Bailer, U. F., Frank, G. K., Henry, S. E., Price, J. C., Meltzer, C. C., Mathis, C. A., … Kaye, W. H. (2017). Exaggerated 5-HT1A but Normal 5-HT2A Receptor Activity in Individuals Ill with Anorexia Nervosa. *Biological Psychiatry*, *61*(9), 1090–1099. https://doi.org/10.1016/j.biopsych.2006.07.018

Bailer, U. F., Narendran, R., Frankle, W. G., Himes, M. L., Duvvuri, V., Mathis, C. A., & Kaye, W. H. (2012). Amphetamine induced dopamine release increases anxiety in individuals recovered from anorexia nervosa. *The International Journal of Eating Disorders*, *45*(2), 263–271. https://doi.org/10.1002/eat.20937

Banks, M. R., Willoughby, L. M., & Banks, W. A. (2008). Animal-assisted therapy and loneliness in nursing homes: use of robotic versus living dogs. *Journal of the American Medical Directors Association*, *9*(3), 173-177-8610.

Beck, A. (1970). Cognitive Therapy: nature and relation to behaviour therapy. *Behaviour Therapy*, *1*, 184–200.

Begum, M., Wang, R., Huq, R., & Mihailidis, A. (2013). Performance of daily activities by older adults with dementia: The role of an assistive robot. In *IEEE International Conference on Rehabilitation Robotics*. https://doi.org/10.1109/ICORR.2013.6650405

Bemelmans, R., Gelderblom, G. J., Jonker, P., & de Witte, L. (2015). Effectiveness of robot paro in intramural psychogeriatric care: A Multicenter Quasi-Experimental Study. *Journal of the American Medical Directors Association*, *16*(11), 946-950-8610.

Berends, T., van Meijel, B., Nugteren, W., Deen, M., Danner, U. N., Hoek, H. W., & van Elburg, A. A. (2016). Rate, timing and predictors of relapse in patients with anorexia nervosa following a relapse prevention program: a cohort study. *BMC Psychiatry*, *16*(1), 316 %@ 1471-244X.

Bergh, C., Callmar, M., Danemar, S., Hölcke, M., Isberg, S., Leon, M., … Olofsson, B. (2013). Effective treatment of eating disorders: Results at multiple sites. *Behavioral Neuroscience*, *127*(6), 878 %@ 1939-0084.

Berman, E., & Machin, S. (2000). Skill-biased technology transfer around the world. *Oxford Review of Economic Policy*, *16*(3), 12–22 %@ 0266–903X.

Bowles, J. (2014). The computerisation of European jobs--who will win and who will lose from the impact of new technology onto old areas of employment. *Bruegel Blog*, *17*.

Brandys, M. K., de Kovel, C. G. F., Kas, M. J., van Elburg, A. A., & Adan, R. A. H. (2015). Overview of genetic research in anorexia nervosa: The past, the present and the future. *International Journal of Eating Disorders*, *48*(7), 814–825. https://doi.org/10.1002/eat.22400

Bryant-Waugh, R. J., Lask, B. D., Shafran, R. L., & Fosson, A. R. (1992). Do doctors recognise eating disorders in children? *Archives of Disease in Childhood*, *67*(1), 103–105. https://doi.org/10.1136/adc.67.1.103

Brzeski, C., & Burk, I. (2015). Die Roboter kommen. Folgen der Automatisierung f{ü}r den deutschen Arbeitsmarkt. *Economic Research*, *30*.

Byford, S., Barrett, B., Roberts, C., Clark, A., Edwards, V., Smethurst, N., & Gowers, S. G. (2007). Economic evaluation of a randomised controlled trial for anorexia nervosa in adolescents. *The British Journal of Psychiatry*, *191*(5), 436-440-1250.

Campbell, K., & Peebles, R. (2014). Eating disorders in children and adolescents: state of the art review. *Pediatrics*, *134*(3), 582-592-4005.

Carter, J. C., Blackmore, E., Sutandar-Pinnock, K., & Woodside, D. B. (2004). Relapse in anorexia

nervosa: a survival analysis. *Psychological Medicine*, *34*(4), 671-679-8978.

Collier, D. A., & Treasure, J. L. (2004). The aetiology of eating disorders. *The British Journal of Psychiatry*, *185*(5), 363. Retrieved from http://bjp.rcpsych.org/content/185/5/363.abstract

Costa, S., Resende, J., Soares, F. O., Ferreira, M. J., Santos, C. P., & Moreira, F. (2009). *Applications of simple robots to encourage social receptiveness of adolescents with autism*. *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*. IEEE.

Costescu, C. A., Vanderborght, B., & David, D. O. (2015). Reversal learning task in children with autism spectrum disorder: a robot-based approach. *Journal of Autism and Developmental Disorders*, *45*(11), 3715-3725–3257.

Couturier, J., Kimber, M., & Szatmari, P. (2013). Efficacy of family-based treatment for adolescents with eating disorders: A systematic review and meta-analysis. *International Journal of Eating Disorders*, *46*(1), 3–11 %@ 1098–108X.

Crow, S. J., & Nyman, J. A. (2004). The cost-effectiveness of anorexia nervosa treatment. *International Journal of Eating Disorders*, *35*(2), 155–160 %@ 1098–108X.

Dahl, T. S., & Boulos, M. N. K. (2013). Robots in health and social care: A complementary technology to home care and telehealthcare? *Robotics*, *3*(1), 1–21.

Deter, H.-C., & Herzog, W. (1994). Anorexia nervosa in a long-term perspective: results of the Heidelberg-Mannheim Study. *Psychosomatic Medicine*, *56*(1), 20-27-3174.

Dickstein-Fischer, L., & Fischer, G. S. (2014). *Combining psychological and engineering approaches to utilizing social robots with children with Autism*. *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. IEEE.

Doraiswamy, P. M., Krishnan, K. R., Boyko, O. B., Husain, M. H., Figiel, G. S., Palese, V. J., … Ellinwood, E. H. (1991). Pituitary abnormalities in eating disorders: Further evidence from MRI studies. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *15*(3), 351–356. https://doi.org/10.1016/0278-5846(91)90066-A

European Commission. (2015). Autonomous Systems - Report. *Special Eurobarometer*, *427*. Retrieved from http://ec.europa.eu/public_opinion/archives/ebs/ebs_427_en.pdf

Frank, G. K., Bailer, U. F., Henry, S. E., Drevets, W., Meltzer, C. C., Price, J. C., … Kaye, W. H. (2005). Increased Dopamine D2/D3 Receptor Binding After Recovery from Anorexia Nervosa Measured by Positron Emission Tomography and [11C]Raclopride. *Biological Psychiatry*, *58*(11), 908–912. https://doi.org/10.1016/j.biopsych.2005.05.003

Frank, G. K., Kaye, W. H., Meltzer, C. C., Price, J. C., Greer, P., McConaha, C., & Skovira, K. (2002). Reduced 5-HT2A receptor binding after recovery from anorexia nervosa. *Biological Psychiatry*, *52*(9), 896–906. https://doi.org/10.1016/S0006-3223(02)01378-1

Friedman, B., Kahn, P. H., & Borning, A. (2006). Value-Sensitive Design and Information Systems. In P. Zhang & D. Galletta (Eds.), *Human-Computer Interaction and Management Information Systems: Foundations*. M.E. Sharpe.

Galusca, B., Costes, N., Zito, N. G., Peyron, R., Bossu, C., Lang, F., … Estour, B. (2008). *Organic Background of Restrictive-Type Anorexia Nervosa Suggested by Increased Serotonin1A Receptor Binding in Right Frontotemporal Cortex of Both Lean and Recovered Patients: [18F]MPPF PET Scan Study*. *Biological Psychiatry* (Vol. 64). https://doi.org/10.1016/j.biopsych.2008.06.006

Garner, J. (1997). Dementia: An intimate death. *British Journal of Medical Psychology*, *70*(2), 177–184. https://doi.org/10.1111/j.2044-8341.1997.tb01897.x

Garner, D. M., & Bemis, K. M. (1982). A cognitive-behavioral approach to anorexia nervosa. *Cognitive Therapy and Research*, *6*(2), 123–150. https://doi.org/10.1007/BF01183887

Garner, D. M., & Keiper, C. D. (2010). Anorexia and Bulimia. In *Handbook of Clinical Psychology Competencies* (pp. 1429–1460). New York, NY: Springer New York. https://doi.org/10.1007/978-0-387-09757-2_52

Golden, N. H., Jacobson, M. S., Schebendach, J., Solanto, M. V, Hertz, S. M., & Shenker, I. R. (1997). Resumption of menses in anorexia nervosa. *Archives of Pediatrics & Adolescent Medicine*, *151*(1), 16-21-4710.

Golden, N. H., Jacobson, M. S., Sterling, W. M., & Hertz, S. (2008). Treatment goal weight in adolescents with anorexia nervosa: use of BMI percentiles. *International Journal of Eating Disorders*, *41*(4), 301–306 %@ 1098–108X.

Goos, M., Manning, A., & Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *The American Economic Review*, *104*(8), 2509–8282.

Gustafsson, C., Svanberg, C., & Müllersdorf, M. (2015). Using a robotic cat in dementia care: a pilot study. *Journal of Gerontological Nursing*, *41*(10), 46-56-9134.

Guyatt, G. H., Oxman, A. D., Kunz, R., Atkins, D., Brozek, J., Vist, G., … Schünemann, H. J. (2011). GRADE guidelines: 2. Framing the question and deciding on important outcomes. *Journal of Clinical Epidemiology*, *64*(4), 395-400-4356.

Guyatt, G. H., Oxman, A. D., Kunz, R., Brozek, J., Alonso-Coello, P., Rind, D., … Vist, G. (2011). GRADE guidelines: 6. Rating the quality of evidence—imprecision. *Journal of Clinical Epidemiology*, *64*(12), 1283-1293-4356.

Guyatt, G. H., Oxman, A. D., Kunz, R., Woodcock, J., Brozek, J., Helfand, M., … Akl, E. A. (2011). GRADE guidelines: 7. Rating the quality of evidence—inconsistency. *Journal of Clinical Epidemiology*, *64*(12), 1294-1302-4356.

Guyatt, G. H., Oxman, A. D., Kunz, R., Woodcock, J., Brozek, J., Helfand, M., … Vist, G. (2011). GRADE guidelines: 8. Rating the quality of evidence—indirectness. *Journal of Clinical Epidemiology*,

64(12), 1303-1310-4356.

Guyatt, G. H., Oxman, A. D., Montori, V., Vist, G., Kunz, R., Brozek, J., … Falck-Ytter, Y. (2011). GRADE guidelines: 5. Rating the quality of evidence—publication bias. *Journal of Clinical Epidemiology*, *64*(12), 1277-1282-4356.

Guyatt, G. H., Oxman, A. D., Sultan, S., Glasziou, P., Akl, E. A., Alonso-Coello, P., … Montori, V. (2011). GRADE guidelines: 9. Rating up the quality of evidence. *Journal of Clinical Epidemiology*, *64*(12), 1311-1316-4356.

Guyatt, G. H., Oxman, A. D., Vist, G., Kunz, R., Brozek, J., Alonso-Coello, P., … Falck-Ytter, Y. (2011). GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). *Journal of Clinical Epidemiology*, *64*(4), 407-415-4356.

Guyatt, G., Oxman, A. D., Akl, E. A., Kunz, R., Vist, G., Brozek, J., … Jaeschke, R. (2011). GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology*, *64*(4), 383-394-4356.

Haas, L., Stargardt, T., Schreyoegg, J., Schlösser, R., Danzer, G., & Klapp, B. F. (2012). Inpatient costs and predictors of costs in the psychosomatic treatment of anorexia nervosa. *International Journal of Eating Disorders*, *45*(2), 214–221 %@ 1098–108X.

Halmi, K. a, Eckert, E., Marchi, P., Sampugnaro, V., Apple, R., & Cohen, J. (1991). Comorbidity of psychiatric diagnoses in anorexia nervosa. *Archives of General Psychiatry*. https://doi.org/10.1001/archpsyc.1991.01810320036006

Harrington, B. C., Haxton, C., & Jimerson, D. C. (2015). Initial Evaluation, Diagnosis, and Treatment of Anorexia Nervosa and Bulimia Nervosa. *Am Fam Physician*, *91*(1), 46–51.

Haselager, W. F. G. (2005). Robotics, philosophy and the problems of autonomy. *Pragmatics & Cognition*, *13*(3), 515–9943.

Hay, P. J., Claudino, A. M., Touyz, S., & Abd Elbaky, G. (2015). Individual psychological therapy in the outpatient treatment of adults with anorexia nervosa. *The Cochrane Library %@ 1465-1858*.

Herpertz-Dahlmann, B., Müller, B., Herpertz, S., Heussen, N., Hebebrand, J., & Remschmidt, H. (2001). Prospective 10-year follow-up in adolescent anorexia nervosa—course, outcome, psychiatric comorbidity, and psychosocial adaptation. *Journal of Child Psychology and Psychiatry*, *42*(5), 603-612-7610.

Herzog, D. B., Dorer, D. J., Keel, P. K., Selwyn, S. E., Ekeblad, E. R., Flores, A. T., … Keller, M. B. (1999). Recovery and relapse in anorexia and bulimia nervosa: a 7.5-year follow-up study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *38*(7), 829-837-8567.

Hoek. (2006). Incidence, prevalence and mortality of anorexia nervosa and other eating disorders. *Current Opinion in Psychiatry*, *19*(4), 389–394.

Hoek, H. W., & van Hoeken, D. (2003). Review of the prevalence and incidence of eating disorders.

International Journal of Eating Disorders, *34*(4), 383–396. https://doi.org/10.1002/eat.10222

Hudson, J., Hirsi, E., Pope Jr, H., & Kessler, R. (2007). The Prevalence and Correlates of Eating Disorders in the National Comorbidity Survey Replication. *Biol Psychiatry*, *61*(3), 348–58.

Kachouie, R., Sedighadeli, S., Khosla, R., & Chu, M.-T. (2014). Socially assistive robots in elderly care: a mixed-method systematic literature review. *International Journal of Human-Computer Interaction*, *30*(5), 369–7318.

Kalisvaart, J. L., & Hergenroeder, A. C. (2007). Hospitalization of patients with eating disorders on adolescent medical units is threatened by current reimbursement systems. *International Journal of Adolescent Medicine and Health*, *19*(2). https://doi.org/10.1515/IJAMH.2007.19.2.155

Kanamori, M., Suzuki, M., Oshiro, H., Tanaka, M., Inoguchi, T., Takasugi, H., … Yokoyama, T. (2003). *Pilot study on improvement of quality of life among elderly using a pet-type robot*. *Computational Intelligence in Robotics and Automation, 2003. Proceedings. 2003 IEEE International Symposium on* (Vol. 1). IEEE.

Katz, L., Autor, D., Ashenfelter, O., & Card, D. (1999). Handbook of labor economics. *Handbook of Labor Economics*, *3*.

Katz, L. F. (1999). Changes in the wage structure and earnings inequality. *Handbook of Labor Economics*, *3*, 1463–4463.

Kaye, W. H., Ebert, M. H., Raleigh, M., & Lake, C. (1984). Abnormalities in cns monoamine metabolism in anorexia nervosa. *Archives of General Psychiatry*, *41*(4), 350–355. Retrieved from http://dx.doi.org/10.1001/archpsyc.1984.01790150040007

Kaye, W. H., Gwirtsman, H. E., George, D. T., & Ebert, M. H. (1991). Altered serotonin activity in anorexia nervosa after long-term weight restoration: Does elevated cerebrospinal fluid 5-hydroxyindoleacetic acid level correlate with rigid and obsessive behavior? *Archives of General Psychiatry*, *48*(6), 556–562. Retrieved from http://dx.doi.org/10.1001/archpsyc.1991.01810300068010

Keski-Rahkonen, A., Hoek, H. W., Susser, E. S., Linna, M. S., Sihvola, E., Raevuori, A., … Rissanen, A. (2007). Epidemiology and course of anorexia nervosa in the community. *American Journal of Psychiatry*, *164*(8), 1259–1265 %@ 0002–953X.

Kidd, C. D., & Breazeal, C. (2008). *Robots at home: Understanding long-term human-robot interaction*. *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*. IEEE.

Kim, E. S., Berkovits, L. D., Bernier, E. P., Leyzberg, D., Shic, F., Paul, R., & Scassellati, B. (2013). Social robots as embedded reinforcers of social behavior in children with autism. *Journal of Autism and Developmental Disorders*, *43*(5), 1038-1049-3257.

Kim, G. H., Jeon, S., Im, K., Kwon, H., Lee, B. H., Kim, G. Y., … Cho, H. (2015). Structural brain changes after traditional and robot-assisted multi-domain cognitive training in community-dwelling healthy elderly. *PloS One*, *10*(4), e0123251 %@ 1932-6203.

Krauth, C., Buser, K., & Vogel, H. (2002). How high are the costs of eating disorders - anorexia nervosa and bulimia nervosa - for German society? *The European Journal of Health Economics*, *3*(4), 244–250. https://doi.org/10.1007/s10198-002-0137-2

la Rie, S., Noordenbos, G., Donker, M., & Van Furth, E. (2008). The quality of treatment of eating disorders: A comparison of the therapists' and the patients' perspective. *International Journal of Eating Disorders*, *41*(4), 307–317 %@ 1098–108X.

Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, *118*(4), 1279–5533.

Libin, A., & Cohen-Mansfield, J. (2004). Therapeutic robocat for nursing home residents with dementia: preliminary inquiry. *American Journal of Alzheimer's Disease & Other Dementias®*, *19*(2), 111-116-3175.

Lock, J., Couturier, J., & Agras, W. S. (2008). Costs of remission and recovery using family therapy for adolescent anorexia nervosa: a descriptive report. *Eating Disorders*, *16*(4), 322-330–266.

Lock, J., & Le Grange, D. %@ 1462523463. (2015). *Treatment manual for anorexia nervosa: A family-based approach*. Guilford Publications.

Malle, B. F. (2016). Integrating robot ethics and machine morality: the study and design of moral competence in robots. *Ethics and Information Technology*. https://doi.org/10.1007/s10676-015-9367-8

Matarić, M. J., Eriksson, J., Feil-Seifer, D. J., & Winstein, C. J. (2007). Socially assistive robotics for post-stroke rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, *4*(1), 5 %@ 1743-0003.

McFarlane, T., Olmsted, M. P., & Trottier, K. (2008). Timing and prediction of relapse in a transdiagnostic eating disorder sample. *International Journal of Eating Disorders*, *41*(7), 587–593 %@ 1098–108X.

Mehler, P. S., & Brown, C. (2015). Anorexia nervosa – medical complications. *Journal of Eating Disorders*, *3*(1), 11. https://doi.org/10.1186/s40337-015-0040-8

Mitchell, J. E., Myers, T., Crosby, R., O'neill, G., Carlisle, J., & Gerlach, S. (2009). Health care utilization in patients with eating disorders. *International Journal of Eating Disorders*, *42*(6), 571–574 %@ 1098–108X.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., … Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4*(1), 1 %@ 2046-4053.

Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. J. V. (2004). Validity of the Eating Disorder Examination Questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behaviour Research and Therapy*, *42*(5), 551-567-7967.

Morgan, J. F., Reid, F., & Lacey, J. H. (1999). The SCOFF questionnaire: assessment of a new screening tool for eating disorders. *Bmj*, *319*(7223), 1467-1468-8138.

Moyle, W., Jones, C., Cooke, M., O'Dwyer, S., Sung, B., & Drummond, S. (2014). Connecting the person with dementia and family: a feasibility study of a telepresence robot. *BMC Geriatrics*, *14*(1), 7 %@ 1471-2318.

Mustelin, L., Raevuori, A., Bulik, C. M., Rissanen, A., Hoek, H. W., Kaprio, J., & Keski-Rahkonen, A. (2015). Long-term outcome in anorexia nervosa in the community. *International Journal of Eating Disorders*, *48*(7), 851–859 %@ 1098–108X.

Nasser, M., Katzman, M., & Gordon, R. (2000). *Cultures in Transition: Eating Disorders as a Global Marker*. Hove: Brunner/Routledge.

NICE. (2003). *National Institute for Health and Care Excellence: Clinical Guidelines*. *National Institute for Health and Care Excellence: Clinical Guidelines [Internet]*. London: National Institute for Health and Care Excellence (UK). Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK11822/

Nimer, J., & Lundahl, B. (2007). Animal-assisted therapy: A meta-analysis. *Anthrozoos*. https://doi.org/10.2752/089279307X224773

O Brien, J. A., & Ward, A. (2003). Cost of providing acute hospitalization for patients with anorexia nervosa and bulimia. *Drug Benefit Trends*, *15*(2), 37-42-5826.

Odetti, L., Anerdi, G., Barbieri, M. P., Mazzei, D., Rizza, E., Dario, P., … Micera, S. (2007). *Preliminary experiments on the acceptability of animaloid companion robots by older people with early dementia*. *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. IEEE.

Packard, M. G., & Knowlton, B. J. (2002). Learning and Memory Functions of the Basal Ganglia. *Annual Review of Neuroscience*, *25*(1), 563–593. https://doi.org/10.1146/annurev.neuro.25.112701.142937

Pajarinen, M., & Rouvinen, P. (2014). Computerization threatens one third of Finnish employment. *ETLA Brief*, *22*(13.1), 2014.

Pederson, K. J., Roerig, J. L., & Mitchell, J. E. (2003). Towards the pharmacotherapy of eating disorders. *Expert Opinion on Pharmacotherapy*, *4*(10), 1659-1678-6566.

Perdereau, F., Faucher, S., Wallier, J., Vibert, S., & Godart, N. (2008). Family history of anxiety and mood disorders in anorexia nervosa: Review of the literature. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*, *13*(1), 1–13. https://doi.org/10.1007/BF03327779

Pereira, T., Lock, J., & Oggins, J. (2006). Role of therapeutic alliance in family therapy for adolescent anorexia nervosa. *International Journal of Eating Disorders*, *39*(8), 677–684 %@ 1098–108X.

Phillipou, A., Rossell, S. L., & Castle, D. J. (2014). The neurobiology of anorexia nervosa: A systematic review. *Australian & New Zealand Journal of Psychiatry*, *48*(2), 128–152. https://doi.org/10.1177/0004867413509693

Powers, T. M. (2006). Prospects for a Kantian machine. *IEEE Intelligent Systems*, *21*(4), 46–1672.

Prange, G. B., Jannink, M. J. a, Groothuis-Oudshoorn, C. G. M., Hermens, H. J., & Ijzerman, M. J. (2006). Systematic review of the effect of robot-aided therapy on recovery of the hemiparetic arm after stroke. *Journal of Rehabilitation Research and Development*. https://doi.org/10.1682/JRRD.2005.04.0076

Preti, A., Rocchi, M., Sisti, D., Camboni, M., & Miotto, P. (2011). A comprehensive meta-analysis of the risk of suicide in eating disorders. *Acta Psychiatr*, *1*, 6–17. https://doi.org/10.1111/j.1600-0447.2010.0164.x

Rabbitt, S. M., Kazdin, A. E., & Scassellati, B. (2015). Integrating socially assistive robotics into mental healthcare interventions: Applications and recommendations for expanded use. *Clinical Psychology Review*, *35*, 35–46. https://doi.org/10.1016/j.cpr.2014.07.001

Riek, L. D. (2012). Wizard of oz studies in hri: a systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, *1*(1 %@ 2163-0364).

Robinson, H., MacDonald, B. A., Kerse, N., & Broadbent, E. (2013). Suitability of healthcare robots for a dementia unit and suggested improvements. *Journal of the American Medical Directors Association*, *14*(1), 34-40-8610.

Rosen, D. S. (2010). Identification and management of eating disorders in children and adolescents. *Pediatrics*, *126*(6), 1240-1253-4005.

Royakkers, L., & van Est, R. (2015). A Literature Review on New Robotics: Automation from Love to War. *International Journal of Social Robotics*. https://doi.org/10.1007/s12369-015-0295-x

Sabanovic, S., Bennett, C. C., Chang, W.-L., & Huber, L. (2013). *PARO robot affects diverse interaction modalities in group sensory therapy for older adults with dementia*. *Rehabilitation Robotics (ICORR), 2013 IEEE International Conference on*. IEEE.

Scassellati, B. (2007). How Social Robots Will Help Us to Diagnose, Treat, and Understand Autism. In *Robotics Research*. https://doi.org/10.1007/978-3-540-48113-3_47

Shibata, T., & Wada, K. (2011). Robot therapy: a new approach for mental healthcare of the elderly - a mini-review. *Gerontologia (Basel)*.

Simut, R. E., Vanderfaeillie, J., Peca, A., de Perre, G., & Vanderborght, B. (2016). Children with Autism Spectrum Disorders make a fruit salad with Probo, the social robot: an interaction study. *Journal of Autism and Developmental Disorders*, *46*(1), 113-126-3257.

Smink, F. R. E., van Hoeken, D., & Hoek, H. W. (2012). Epidemiology of Eating Disorders: Incidence, Prevalence and Mortality Rates. *Current Psychiatry Reports*, *14*(4), 406–414. https://doi.org/10.1007/s11920-012-0282-y

Smith, N. (2017). What's Wrong With Bill Gates' Robot Tax. *BloombergView*. Retrieved from https://www.bloomberg.com/view/articles/2017-02-28/what-s-wrong-with-bill-gates-robot-tax

Soler, M. V., Agüera-Ortiz, L., Rodriguez, J. O., Rebolledo, C. M., Muñoz, A. P., Pérez, I. R., … Chillón, L. C. (2015). Social robots in advanced dementia. *Frontiers in Aging Neuroscience*, *7*.

Srinivasan, S. M., Eigsti, I.-M., Gifford, T., & Bhat, A. N. (2016). The effects of embodied rhythm and robotic interventions on the spontaneous and responsive verbal communication skills of children with Autism Spectrum Disorder (ASD): A further outcome of a pilot randomized controlled trial. *Research in Autism Spectrum Disorders*, *27*, 73-87-9467.

Stahl, B. C., & Coeckelbergh, M. (2016). Ethics of healthcare robotics: Towards responsible research and innovation. *Robotics and Autonomous Systems*. https://doi.org/10.1016/j.robot.2016.08.018

Steinhausen, H.-C. (2002). The outcome of anorexia nervosa in the 20th century. *American Journal of Psychiatry*, *159*(8), 1284–1293 %@ 0002–953X.

Steinhausen, H.-C. (2009). Outcome of eating disorders. *Child and Adolescent Psychiatric Clinics of North America*, *18*(1), 225-242-4993.

Stiglitz, J. E. %@ 0393088693. (2012). *The price of inequality: How today's divided society endangers our future*. WW Norton & Company.

Strober, M., Freeman, R., & Morrell, W. (1997). The long-term course of severe anorexia nervosa in adolescents: Survival analysis of recovery, relapse, and outcome predictors over 10–15 years in a prospective study. *International Journal of Eating Disorders*, *22*(4), 339–360 %@ 1098–108X.

Sullins, J. P. (2006). When is a robot a moral agent. *Machine Ethics*, 151–160.

Sy, R., Ponton, K., De Marco, P., Pi, S., & IsHak, W. W. (2013). Quality of life in anorexia nervosa: a review of the literature. *Eating Disorders*, *21*(3), 206-222-266.

Tapus, A., Ţăpuş, C., & Matarić, M. J. (2009). The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia. In *2009 IEEE International Conference on Rehabilitation Robotics, ICORR 2009*. https://doi.org/10.1109/ICORR.2009.5209501

Toulany, A., Wong, M., Katzman, D. K., Akseer, N., Steinegger, C., Hancock-Howard, R. L., & Coyte, P. C. (2015). Cost analysis of inpatient treatment of anorexia nervosa in adolescents: hospital and caregiver perspectives. *CMAJ Open*, *3*(2), E192–E197. https://doi.org/10.9778/cmajo.20140086

Treasure, J., & Schmidt, U. (2013). The cognitive-interpersonal maintenance model of anorexia nervosa revisited: a summary of the evidence for cognitive, socio-emotional and interpersonal predisposing and perpetuating factors. *Journal of Eating Disorders*, *1*, 13.

https://doi.org/10.1186/2050-2974-1-13

Tronto, J. C. %@ 0415906423. (1993). *Moral boundaries: A political argument for an ethic of care*. Psychology Press.

Ueyama, Y. (2015). A bayesian model of the uncanny valley effect for explaining the effects of therapeutic robots in autism spectrum disorder. *PloS One*, *10*(9), e0138642 %@ 1932-6203.

Van Der Drift, E. J. G., Beun, R.-J., Looije, R., Blanson Henkemans, O. A., & Neerincx, M. A. (2014). *A remote social robot to motivate and support diabetic children in keeping a diary*. *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. ACM.

van Wynsberghe, A. (2013). Designing Robots for Care: Care Centered Value-Sensitive Design. *Science and Engineering Ethics*. https://doi.org/10.1007/s11948-011-9343-6

van Wynsberghe, A. (2016). Service robots, care ethics, and design. *Ethics and Information Technology*. https://doi.org/10.1007/s10676-016-9409-x

van Wynsberghe, A. %@ 1317123166. (2016). *Healthcare robots: Ethics, design and implementation*. Routledge.

Varoufakis, Y. (2017). A Tax on Robots? *Project Syndicate*. Retrieved from https://www.project-syndicate.org/commentary/bill-gates-tax-on-robots-by-yanis-varoufakis-2017-02

Wada, K., & Shibata, T. (2006). *Robot therapy in a care house-its sociopsychological and physiological effects on the residents*. *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*. IEEE.

Wada, K., Shibata, T., & Kawaguchi, Y. (2009). Long-term robot therapy in a health service facility for the aged - A case study for 5 years -. In *2009 IEEE International Conference on Rehabilitation Robotics* (pp. 930–933). IEEE. https://doi.org/10.1109/ICORR.2009.5209495

Wada, K., Shibata, T., Saito, T., Sakamoto, K., & Tanie, K. (2005). *Psychological and social effects of one year robot assisted activity on elderly people at a health service facility for the aged*. *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*. IEEE.

Wada, K., Shibata, T., Saito, T., & Tanie, K. (2004). Effects of robot-assisted activity for elderly people and nurses at a day service center. *Proceedings of the IEEE*, *92*(11), 1780-1788-9219.

Wallach, W., & Allen, C. %@ 0199705968. (2008). *Moral machines: Teaching robots right from wrong*. Oxford University Press.

Wentz, E., Gillberg, I. C., Anckarsäter, H., Gillberg, C., & Råstam, M. (2009). Adolescent-onset anorexia nervosa: 18-year outcome. *The British Journal of Psychiatry*, *194*(2), 168-174-1250.

Wilson, G. T., Grilo, C. M., & Vitousek, K. M. (2007). Psychological treatment of eating disorders. *American Psychologist*, *62*(3), 199 %@ 1433801132.

Wu, Y.-H., Cristancho-Lacroix, V., Fassert, C., Faucounau, V., de Rotrou, J., & Rigaud, A.-S. (2016). The

attitudes and perceptions of older adults with mild cognitive impairment toward an assistive robot. *Journal of Applied Gerontology*, *35*(1), 3-17-4648.

Yager, J., & Andersen, A. E. (2005). Anorexia nervosa. *New England Journal of Medicine*, *353*(14), 1481-1488-4793.

Yu, R., Hui, E., Lee, J., Poon, D., Ng, A., Sit, K., … Woo, J. (2015). Use of a Therapeutic, Socially Assistive Pet Robot (PARO) in Improving Mood and Stimulating Social Interaction and Communication for People With Dementia: Study Protocol for a Randomized Controlled Trial. *JMIR Research Protocols*, *4*(2), e45. https://doi.org/10.2196/resprot.4189

Zeeck, A., & Hartmann, A. (2005). Relating therapeutic process to outcome: are there predictors for the short-term course in anorexic patients? *European Eating Disorders Review*, *13*(4), 245-254-968.

Zhu, A. J., & Walsh, B. T. (2002). Pharmacologic treatment of eating disorders. *The Canadian Journal of Psychiatry*, *47*(3), 227-234-7437.

About the Authors

The interdisciplinary think tank *Thinking Machines, 2016 – 2017* consisted of nine Master's students from Radboud University in Nijmegen, the Netherlands.

Marco Dessì
Philosophy of Mind

Dylan Henssen
Medical Studies

Sophie Horsman
Philosophy & Artificial Intelligence

Rutger Meijers
Medical Studies

Annika Schiefner
Language and Communication

Fenja Schlag
Medical Biology

Luuk Schmitz
Political Science

Vika Shimanskaya
Medical Studies

Ricarda Weiland
Cognitive Neuroscience

The think tank was supervised by:

Noortje ter Berg
Programme Director and Skills Teacher
Radboud Honours Academy

Lucien Engelen
Director REshape Center for Health(care) Innovation
Radboud University Medical Centre