A Virtual Agent for Post-Traumatic Stress Disorder Treatment

Tielman, Myrthe

DOI
10.4233/uuid:21756661-2d92-447d-8ac7-0cd2b1b6dc8b

Publication date
2018

Document Version
Final published version

Citation (APA)
Tielman, M. (2018). *A Virtual Agent for Post-Traumatic Stress Disorder Treatment*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:21756661-2d92-447d-8ac7-0cd2b1b6dc8b

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

A Virtual Agent for Post-Traumatic Stress Disorder Treatment

A Virtual Agent for Post-Traumatic Stress Disorder Treatment

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op donderdag 25 januari 2018 om 15:00 uur

door

Myrthe Lotte TIELMAN

Master of Science in Cognitive Artificial Intelligence, Universiteit Utrecht, Nederland
geboren te Maarssen, Nederland.

Dit proefschrift is goedgekeurd door de

promotor: prof. dr. M.A. Neerincx
copromotor: dr. ir. W.P. Brinkman

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. M.A. Neerincx,	Technische Universiteit Delft
Dr. ir. W.P. Brinkman,	Technische Universiteit Delft
Onafhankelijke leden:	
Prof. dr. C.M. Jonker,	Technische Universiteit Delft
Prof. dr. F.M. Brazier,	Technische Universiteit Delft
Prof. dr. P. Markopoulos,	Technische Universiteit Eindhoven
Prof. dr. A. Rizzo,	Institute for Creative Technologies, USC
Prof. dr. A. Hanjalic,	Technische Universiteit Delft

Printed by: ProefschriftMaken || www.proefschriftmaken.nl

Front & Back: Myrthe Tielman

ISBN 978-94-6295-829-6

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

Contents

Summary

Post-traumatic stress disorder (PTSD) is a mental disorder with a high impact on quality of life, and despite the existence of treatment, barriers still stop many people from receiving the care they need. Autonomous e-mental health (AEMH) systems might take away some of these barriers to care. Such systems can be used at least partly without direct human involvement, moreover, they support the self-management of mental health patients. As such, they provide a cost-effective, accessible and privacy-sensitive solution to many barriers to care. Treatment for PTSD is difficult though, as it requires patients to actively recollect their traumatic memories and it is important that they comply with this task. As PTSD patients generally actively avoid thinking about their memories, personalized assistance to motivate and guide them in the recollection process can therefore be valuable. Virtual agents have been shown to increase compliance and even treatment outcome when incorporated in applications, so they seem very suitable to offer such assistance. Although virtual agents are increasingly being applied to AEMH systems, their application for PTSD therapy is novel. Therefore, this thesis presents a virtual agent for PTSD therapy, and studies the ways in which such a virtual agent can enhance treatment compliance. Given the novelty of a virtual agent for PTSD therapy, work on this thesis started with an exploratory study. Experts from the field of psychotherapy were presented with scenarios and claims about such an agent, giving explicit examples of how such an agent could act. These claims served as starting points for discussions about how an agent should behave and why. The discussions were analyzed based on underlying concepts, which resulted in ten general guidelines for behavior of a virtual agent for PTSD.

Some of these guidelines could be translated into core functions of the agent, namely; it should provide information about the therapy, offer personalized assistance and motivate the user. The analysis also identified some constraints for the design of the agent, namely that the patient should be protected, and the agent should be placed within a full therapy system with additional functions. Therefore, the remainder of the work presented in this thesis can be split into two parts. First, work on the core functions was done by designing these functions and evaluating them in controlled experiments. Second, work on the constraints for the virtual agent was done through developing proof of concepts. The first core function behavior that was considered in this thesis is the task of informing the patient on the goal of therapy by providing psychoeducation. An empirical study was done comparing textual presentation of psychoeducation with verbal presentation by the virtual agent, including the effects on information recollection, alliance to the virtual agent, and adherence. This study showed that after controlling for recollection and alliance, textual presentation resulted in better adherence. The second core function was personalized assistance, which was studied in the context of memory

recollection. This thesis proposes an ontology-based question system which uses a combination of multiple-choice and open questions. By adapting the questions to the experiences of the user, the interaction is personalized. An empirical evaluation of this concept showed that it was capable of eliciting more detailed descriptions of a memory than a non-personalized, list-based system. The third core function studied was the act of motivating. Personalized motivational statements written by experts were categorized and analyzed to reveal what type of motivation is suitable for what situation. With the results from this analysis, a personalized motivational feedback system was developed, capable of generating motivational statements based on a patient's situation. For instance, compliments are more appropriate when a patient feels they are doing well, and empathy when they do not. Empirical tests revealed that motivational feedback was capable of improving motivation, and personalization was particularly useful when therapy was not going well.

Aside from core functions of a virtual agent for PTSD therapy, this thesis also considers some constraints under which such a virtual agent has to operate. Patient safety is an important element for any AEMH intervention, and this thesis therefore presents three theoretical models of detecting and dealing with risk situations in AEMH systems. First, a detection system describes how to detect a situation and how to decide when the risk is severe enough to scale up to human care. Second, an auto-referral system describes how automatic referral can happen. Third, a motivation system describes how users themselves can be motivated to seek contact with a human caregiver. These models can be used by a variety of systems and interventions to guide the implementation of risk management. The second constraint for a virtual agent is that it needs to operate within a full therapy system. Therefore, this thesis presents a therapy system for PTSD incorporating a virtual agent. This system includes general therapy elements such as a session overview, but also exposure environments in the form of a digital diary and 3D world. The virtual agent is included in this system and interacts with the exposure environment by using an ontology-based question system on trauma, of which the questions can be answered in the diary. A study with former PTSD-patients reveals that the trauma-related questions are relevant to their experiences and the system is usable.

The work presented in this thesis provides insight into a virtual agent for PTSD. General design guidelines have been established, and core functions and constraints for the agent have been studied. To enhance treatment compliance in the form of memory recollection, a virtual agent can best present psychoeducation in written text. Furthermore, it can employ an ontology-based question system to elicit more detailed memory recollection, and a motivational feedback system to generate motivation suitable to the patient's situation. Regarding the constraints of a virtual agent, theoretical models have been presented describing how to deal with risk situations, and a full therapy system for PTSD incorporating the agent was found useful and usable by former patients. Taken together, these findings show how a virtual agent should act and how it can be applied. Given the high impact PTSD has on a person's life, such a virtual agent has the potential to make a real difference. And although this thesis has focused on PTSD, all of the findings are potentially relevant for other domains of AEMH systems.

Samenvatting

Post-traumatische stress stoornis (PTSS) is een mentale stoornis met een grote impact op de kwaliteit van leven. Ondanks het bestaan van behandelingen zijn er nog altijd barrières die voorkomen dat mensen benodigde hulp krijgen. Autonome e-health (AEH) systemen voor de geestelijke gezondheidszorg (GGZ) hebben de potentie om enkele barrières weg te nemen. Zulke systemen kunnen, in ieder geval gedeeltelijk, gebruikt worden zonder directe menselijke input. Bovendien ondersteunen ze het zelfmanagement van patiënten. Deze systemen hebben daarom de potentie om een kosteneffectieve, toegankelijke en privacy respecterende oplossing te zijn voor barrières tot zorg. De kern van therapie voor PTSS is het actief ophalen van traumatische herinneringen en het is belangrijk dat patiënten dit zorgvuldig doen. PTSS patiënten vermijden echter over het algemeen juist hun herinneringen. Om deze reden heeft gepersonaliseerde hulp en sturing de potentie om bij het ophalen van deze herinneringen waardevol zijn.

Onderzoek laat zien dat virtuele agents therapietrouw en zelfs therapie uitkomst kunnen verbeteren, ze lijken dus zeer geschikt om zulke hulp te verlenen. Hoewel virtuele agents meer en meer worden toegepast binnen AEH systemen voor de GGZ, is hun toepassing voor PTSS nieuw. Deze dissertatie presenteert daarom een virtuele agent voor PTSS therapie, en onderzoekt op welke manieren deze agent kan handelen om therapietrouw te vergroten.

Gegeven de relatieve onbekendheid over een virtuele agent voor PTSS therapie begint het werk in deze dissertatie exploratief. Scenario's en stellingen over hoe een virtuele agent voor PTSS zou kunnen handelen zijn gepresenteerd aan experts in klinische psychologie. Op basis van deze stellingen is een discussies gevoerd door de experts over hoe een agent zich zou moeten gedragen en waarom. De argumenten en uitspraken in deze discussies zijn geanalyseerd op basis van onderliggende concepten, wat resulteerde in tien algemene richtlijnen voor het gedrag van een virtuele agent voor PTSS

Enkele van deze richtlijnen konden vertaald worden in kernfuncties van de agent, namelijk dat deze informatie over de therapie moet verstrekken, gepersonaliseerde hulp moet bieden en de gebruiker moet motiveren. De analyse identificeerde ook enkele restricties op het ontwerp van de agent, namelijk dat de patiënt beschermd moet worden en dat deze geplaatst moet worden binnen een breder therapiesysteem.

Het werk in deze dissertatie kan daarom opgesplitst worden in twee delen. Om te beginnen is onderzoek gedaan naar de kernfuncties door deze te ontwerpen en in gecontroleerde experimenten te evalueren. Daarnaast is onderzoek gedaan naar de restricties, door conceptbewijzen te ontwikkelen.

Het eerste kerngedrag dat onderzocht is in deze dissertatie, is het informeren van de patiënt over het doel van de therapie bij wijze van psychoeducatie. In een

empirische studie zijn tekstuele presentatie en verbale presentatie door de virtuele agent vergeleken. Hierbij is ook gekeken naar de effecten op hoe goed de informatie onthouden werd, hoe verbonden men zich voelde met de virtuele agent en hoe goed instructies werden nageleefd. Dit onderzoek laat zien dat na het controleren voor deze eerste twee factoren, tekstuele presentatie leidde tot een betere naleving van de instructies.

Het tweede kerngedrag was gepersonaliseerde hulp, wat bestudeerd is in de context van het ophalen van herinneringen. Deze dissertatie presenteert een vraagsysteem gebaseerd op een ontologie, waarin een combinatie van multiple-choice en open vragen gebruikt wordt. Door de vragen aan te passen aan de ervaringen van de gebruiker wordt de interactie gepersonaliseerd. Een empirische evaluatie van dit concept liet zien dat het in staat was tot het oproepen van gedetailleerdere beschrijvingen van herinneringen dan een niet gepersonaliseerde vragenlijst.

De derde kernfunctie die onderzocht is was motiveren. Gepersonaliseerde motiverende berichtjes, geschreven door experts, zijn gecategoriseerd en geanalyseerd om zo te ontdekken welk type motivatie geschikt is voor welke situatie. Met de resultaten van deze analyse is een gepersonaliseerd motivatie systeem ontwikkeld dat motiverende berichten kan genereren op basis van de situatie van een patiënt. Zo zijn complimenten bijvoorbeeld gepaster wanneer de patiënt het idee heeft dat het goed gaat, en is empathie geschikter als ze dit idee niet hebben. Een empirische studie laat zien dat motiverende berichtjes motivatie kunnen verhogen, en dat personaliseren met name nuttig is wanneer de therapie minder goed gaat.

Naast de kernfuncties van een virtuele agent voor PTSS therapie, kijkt deze dissertatie ook naar enkele restricties waaronder een virtuele agent moet opereren. Patientveiligheid is een belangrijk element voor elke AEH interventie voor de GGZ, deze dissertatie presenteert daarom drie theoretische modellen voor het detecteren en afhandelen van risico situaties in zulke interventies. Ten eerste beschrijft een detectie-model hoe risico situaties ontdekt kunnen worden en hoe bepaald kan worden wanneer een situatie ernstig genoeg is om op te schalen naar menselijke hulp. Ten tweede beschrijft een verwijzing-model hoe automatische doorverwijzing zou kunnen gebeuren. Ten derde beschrijft een motivatie systeem hoe gebruikers gemotiveerd kunnen worden om contact te zoeken met een menselijke hulpverlener. Deze modellen kunnen gebruikt worden door verschillende soorten systemen en interventies om de invulling van risicomanagement te sturen.

De tweede restrictie voor een virtuele agent voor PTSS is dat deze handelt binnen een volledig therapie systeem. Deze dissertatie presenteert daarom een therapiesysteem voor PTSS waarin een virtuele agent is opgenomen. Dit systeem bevat enkele algemene onderdelen zoals een sessieoverzicht, maar ook een digitaal dagboek en een 3D wereld waarin herinneringen opgehaald kunnen worden. De virtuele agent heeft een wisselwerking met het dagboek door gebruik te maken van een vragensysteem dat gebaseerd is op een ontologie voor trauma. De vragen gegenereerd door dit systeem worden gesteld door de virtuele agent en kunnen beantwoord worden in het dagboek. Een studie met voormalig PTSS patiënten laat zien dat deze trauma-gerelateerde vragen relevant zijn voor hun ervaringen en dat het systeem bruikbaar is.

Het werk in deze dissertatie geeft inzicht in een virtuele agent voor PTSS. Algemene richtlijnen voor het gedrag van de agent zijn opgesteld, en enkele kernfuncties en restricties zijn onderzocht. Om therapietrouw in de vorm van het ophalen van herinneringen te vergroten kan een virtuele agent psychoeducatie het beste tekstueel presenteren. Een vragensysteem gebaseerd op een ontologie kan gebruikt worden om gedetailleerdere beschrijvingen van herinneringen te ontlokken, en een systeem wat gepersonaliseerde motivatie genereert kan motivatie vergroten. Om de restricties op een virtuele agent in kaart te brengen zijn drie theoretische modellen gepresenteerd die beschrijven hoe met risico situaties om te gaan. Daarnaast is een therapiesysteem voor PTSS therapie inclusief virtuele agent gepresenteerd en getest met voormalig patiënten. Bij elkaar laten deze resultaten zien hoe een virtuele agent zich moet gedragen en hoe het kan worden toegepast. Gegeven de hoge impact van PTSS op een leven heeft een virtuele agent de potentie om een belangrijk verschil te maken. Hoewel deze dissertatie focust op PTSS, hebben alle bevindingen ook de potentie om relevant te zijn voor andere domeinen van AEH systemen voor de GGZ.

1

Introduction

Post-traumatic stress disorder (PTSD) is a mental disorder with a high impact on quality of life, and despite the existence of validated treatment, many patients do not seek help. To reduce the barriers to care, a combination of techniques and methods from artificial intelligence and human-computer interaction are promising. A computerized therapy system, incorporating a personalized virtual agent as a coach, would allow patients to follow therapy autonomously. PTSD therapy is challenging, however, as it requires patients to face their worst memories. One of the main tasks for the virtual agent would, therefore, be to ensure that patients recollect their trauma in detail. Exactly how the virtual agent can act to ensure this is yet unclear, as it would need to adapt its behavior to this distinct user group and the differences in patient's situation and background.

1.1. Background

Post-Traumatic Stress Disorder

Post-traumatic stress disorder is an anxiety disorder following one or more traumatic experiences. Some of the symptoms of PTSD are intrusive memories such as flashbacks or nightmares; avoidance of stimuli related to the trauma; a persistent negative state and high arousal [1]. PTSD patients often suffer from comorbid disorders, such as depression, substance abuse or anxiety disorders. PTSD is a disorder which can greatly influence life, resulting in substantial loss of quality of life [2], and it is associated with increased suicidality rates [3]. Aside from affecting patients themselves, PTSD also affects their family and loved ones [4]. Lifetime prevalence of PTSD is estimated at 7.4% in the Netherlands [5], but can be significantly higher in regions experiencing conflict [6].

Traumatic experiences at the root of PTSD are defined as *actual or threatened death, serious injury or sexual violence* [1]. This covers a broad range of potentially traumatic experiences; however, some experiences are more likely to result in PTSD than others. For instance childhood sexual abuse and physical assault have been

found to predict PTSD and complex PTSD, while for instance death of a close family member or a traffic accident are less likely to predict PTSD [7, 8]. Studies suggest that events that are life-threatening, interpersonal and early-life are more prone to cause PTSD [9]. Another predictor is the number of traumatic experiences, PTSD becoming more likely with more experienced events. The higher exposure to possible traumatic events among those with vocations that increase risk, such as veterans, explains the higher PTSD rates among these groups [1].

The two most common treatment methods for PTSD are trauma-focused cognitive behavioral therapy (CBT) and Eye Movement Desensitization and Reprocessing (EMDR) therapy. CBT focuses on identifying and restructuring unhelpful cognitive patterns. Applied to PTSD therapy, this means actively recollecting those traumatic memories that are normally avoided. If one repeatedly exposes oneself to the traumatic memory and experiences that remembering is possible, the automatic fear response will eventually lessen [10]. EMDR also involves remembering the traumatic memory, but during this recollection the patient is asked to make rapid eye movements. Although the use of the eye movement has been under some debate in the past, it has been shown to contribute to therapy [11]. This is probably because the eye movements occupy the working memory during recollection, suppressing the emotional response [11]. Both trauma-focused CBT and EMDR have been studied extensively and have been proven effective in treating PTSD [12–14]. In both therapies, trauma recollection lies at the heart of the therapy. To explain to patients why and how this unpleasant task is necessary, psychoeducation is often incorporated as well. Psychoeducation refers to informing the patient about their disorder and why and how certain therapeutic tasks will need to be executed. It can occur both in the beginning and during treatment, or both [15].

Despite the existence of therapies for PTSD, there are different barriers to care that prevent people from successfully receiving treatment. Factors such as treatment cost, logistics [16] and availability of therapists [17] can all deter people from seeking help. Another major problem is the stigma that still surrounds mental illness. Despite increasing efforts to combat stigma and discrimination, many people still hesitate to seek help from a mental health professional [18, 19]. People are afraid their social circle will judge them for following mental health treatment, but also worry about adverse consequences for their career [20]. These barriers to care greatly influence the reach of mental health care treatment. In high-income countries, an estimate of between 35% and 50% of people with a severe mental health disorder do not receive help; in low-income countries, these numbers rise to 76% and 85%. At the same time, the global burden of mental health-related diseases is currently 13% and growing [6].

Technology for Mental Health Care

Autonomous e-mental health (AEMH) applications can help solve several of the problems preventing patients from receiving care. Such applications are technological systems aimed at mental health patients which can be used at least partly without a human caregiver on the other side. These applications are often more affordable, can be used anywhere at any time and can alleviate some of the pri-

vacy concerns patients might have [21]. Additionally, they are flexible in terms of personalization and interactivity, and offer opportunities to increase patient engagement [22]. Current systems can be divided into different categories based on their goal, their level of autonomy and their target group [23]. Some different goals are providing information [24], assessing and monitoring people [25, 26], and offering a full therapeutic intervention [27, 28]. Some systems operate fully autonomously [29], but many contain some combination of human and autonomous care, such as weekly phone calls [30], or an app supporting regular care [31]. Although many systems have been developed for common mental disorders such as depression [32, 33] and anxiety [34–37], applications aimed at rarer and more severe disorders are emerging as well [38]. All these different disorders bring their own unique challenges to any AEMH application.

One of the main challenges of AEMH applications for PTSD is that therapy requires patients to recollect their trauma in detail, a very difficult task. Patients avoid thinking about their memories; this is one of the defining factors of PTSD [1]. A commonly used theoretical framework of avoidance defines both automatically and controlled avoidance, and applies these to both information and emotional memory [39]. So one can actively avoid taking the bus if one's memory involves a bus, or actively try to not think of something, these are examples of controlled avoidance of information. On the other hand, patients are also often simply unable to recollect details, as they have automatically repressed their memories, an example of automatically avoidance of information. With regards to emotion, this can apply to actively trying to push away feelings, which is controlled avoidance or the simple incapability of feeling certain emotions, which is automatic avoidance [40]. Given the prevalence of avoidance in PTSD patients, it is therefore not surprising that their memories are often disorganized and fragmented [41]. This makes the process of recollection challenging, despite the strong motivation patients have to reduce their symptoms. Assistance during exposure therapy can, therefore, be very valuable.

Virtual agents provide a way for a computer system to offer a more human-like and intuitive form of interaction to users. These virtual agents commonly have a human-like appearance and are capable of some form of conversation. They are therefore also often called virtual embodied agents, or virtual conversational agents. These virtual agents are often perceived more positively [42], and can help lead attention better [43] than a simple text-based interface. Moreover, virtual agents are often able to enhance compliance [44] and even outcome of the intervention [45, 46]. Some studies into the appearance of virtual agents suggest that a virtual agent which is more like the user is most effective [47], while others suggest that the most attractive agent might be preferable [48], or that the relevant factor is realism [49]. However, the appearance of an agent has less effect on the outcome than its presence [50]. Given their effect on patient compliance and treatment outcome, the addition of virtual agents to AEMHs is very promising.

Virtual agents are increasingly being applied to systems aimed at improving health, including AEMH applications. In some systems the virtual agent is the core of the application, serving as a conversational partner. In other instances, they are part of a broader application in which they fulfill several functions. Two examples

of a virtual agent as a conversational partner are SimCoach and SimSensei, both aimed at informing and screening people who might have PTSD [51, 52]. Other applications for virtual agents as conversational partners are those aimed at battling loneliness among older adults [53, 54]. However, most virtual agents are embedded in a broader application which includes other functionalities. Some systems aim at supporting healthy living, such as an exercise coach [44], or systems aim at supporting general healthy behavior [55–57]. However, systems are also being developed aimed at specific disorders, both for physical [45, 58, 59] and mental health care. Virtual agents have been applied to AEMH applications for substance abuse [60, 61], insomnia [62], dementia [43], social phobia [63] and depression [64, 65].

Although virtual agents have been applied to many different domains in mental health care, to our knowledge a virtual agent that is a part of an AEMH application for PTSD therapy is entirely new. Such a virtual agent faces its very own distinct challenges related to the context it is applied in. PTSD patients are a specific user group, and follow a specific type of therapy unique to their disorder. The heart of this PTSD therapy is trauma recollection, but this is also the most challenging portion as PTSD patients generally avoid their traumatic memories. It is therefore mainly important for a virtual agent to ensure that patients comply with the main goal of therapy, which is memory recollection. A virtual agent might support this memory recollection using a number of different methods and techniques. Virtual agents might also use some of these in other AEMH applications, but given the specific nature of the task, some methods will be unique to the PTSD domain. The focus of this thesis, therefore, lies in developing and testing how a virtual agent can enhance treatment compliance to PTSD therapy.

1.2. Main Research Question & Hypotheses

This thesis, therefore, explores the following main research question:

In what way can a virtual agent that is a part of a computerized PTSD therapy system act to enhance treatment compliance?

To answer this research question, three sub-questions and three hypotheses were developed. Taken together, they provide an answer of how a virtual agent should act, what it should do and in which context it should operate. This section elaborates on the positions taken in this thesis and how they are supported by findings and theories reported in the literature. Additionally, motivations are explored that warrant the sub-questions studied in the thesis. Figure 1.1 provides a conceptual overview of the virtual agent and the context it operates in. This includes both behavioral and functional aspects. The bold titles represent the sub-questions and hypotheses studies in this thesis.

When designing any system for a specific user group, it is important to first establish the specific human-factor requirements The first question this thesis asks is therefore what general guidelines are important for the design of a virtual agent for PTSD. How to say the 'right thing at the right time in the right way' is a common

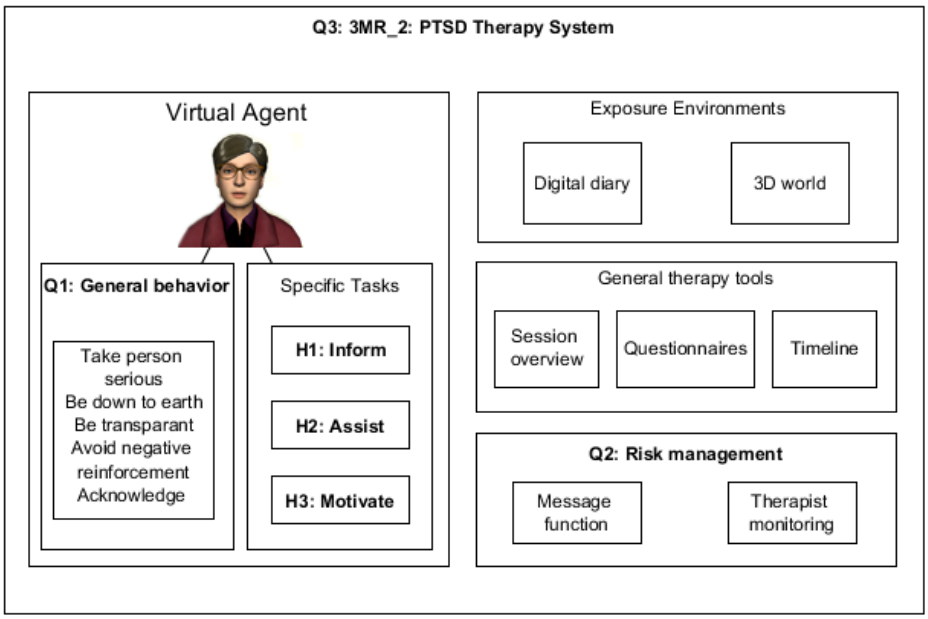

Figure 1.1: Conceptual overview of the virtual agent, it displays specific general behavior and has the tasks of informing, assisting and motivating. It is incorporated in the 3MR_2 PTSD therapy system which has several other components. Such a system should also deal with the risks of at-home treatment of PTSD. The bold titles represent the topics of the research questions and hypotheses.

challenge [66]. However, what exactly the 'right' way is can differ greatly between populations, and PTSD patients have their own specific requirements. Before developing a virtual agent for PTSD, one should, therefore, know the limitations and opportunities this user group offers. Specifically, what guidelines should underlie the general behavior of a virtual agent. Only after these have been established can more specific functions be developed [67]. Therefore, this thesis first studies the general design guidelines for the behavior of a virtual agent for PTSD, following the first sub-question presented below. Only after this first step can decisions be made about how the agent should execute specific tasks.

Question 1: Which design guidelines are important for the design of a virtual agent for PTSD treatment?

One important task for a virtual agent for PTSD is that it should provide psychoeducation. This thesis takes the stance that whether this information is presented orally or by providing written text influences how effective it is in increasing adherence. Psychoeducation is information and context about how and why the therapeutic tasks are necessary [68, 69]. It is a key element of exposure therapy for PTSD [70], and is especially relevant because exposure requires patients to go against their own instincts and habits by recollecting their trauma. Different presentation modes are possible when a virtual agent wishes to provide psychoeducation. It can simply let patients read the text, but it can also present the information aloud by talking. Both options might have their own advantages regarding the eventual goal of enhancing adherence to the virtual agent. Some evidence exists that textual information is recollected better than oral information [71, 72]. However, oral presentation might increase trust in the virtual agent, as it presents itself as an expert [73]. Trust and perceived expertise might, in turn, affect adherence [74–76]. Given these different factors, this thesis hypothesizes that:

Hypothesis 1: Presentation mode of psychoeducation presented by a virtual agent affects adherence.

The core of exposure therapy for PTSD is the detailed and complete recollection of traumatic memories, making assistance in memory recall the second task for a virtual agent for PTSD. This thesis takes the position that a virtual agent can assist with this recollection by asking questions based on a knowledgebase of trauma. Trauma recollection is difficult for patients as memories are often fragmented, and patients tend to avoid thinking about them in detail [40, 41]. The exposure is generally imaginary, meaning that people have to think back and imagine they are back in those moments most disturbing to them [70]. In an AEMH system, this imaginary exposure can be supplemented by digital media such as photos, maps, and sounds [77]. To guide patients through the exposures wherein they need to recollect their memories, personalized and detailed questions can be posed. For a virtual agent to ask these questions, however, it needs some understanding of what questions to ask in what situation. One common method of adding domain knowledge is the use

of ontologies, which define the relationships among important concepts in a domain [78, 79]. This thesis hypothesizes that with an ontology on traumatic memories, a virtual agent can assist in detailed memory recollection, leading to hypothesis 2:

Hypothesis 2: An ontology-based question system enhances memory recollection.

Though shown effective, PTSD treatment is challenging, making motivation during therapy the third task for a virtual agent for PTSD [68]. In this thesis, the stance is taken that a virtual agent enhances motivation by adapting motivational feedback to the current situation. Drop-out rates for PTSD are high even in regular therapy, showing that therapy is difficult to complete [80]. One observation often made by practicing therapists is that symptoms might get worse before they get better, which can be very demotivating [81]. However, symptom progression differs per person, and it is important to take these differences in into account. Work in the area of recommendation systems shows that personalized feedback is important [82]. A motivational system should therefore also reflect on the current situation during therapy. Progress models based on questionnaires are one situational measure a virtual agent can use to provide personalized motivation. This thesis hypothesizes that a virtual agent can enhance trust in the therapy and motivation to continue; it can do this by analyzing the current progress and trust of a patient in a good outcome and adapting motivational feedback to this situation. This leads to hypothesis 3:

Hypothesis 3: By presenting situation-based motivational messages a virtual agent enhances therapy trust and motivation.

Next to general behavior and specific therapeutic functions, it is important to consider that a virtual agent for PTSD interacts with a vulnerable user group. This thesis proposes a set of safety protocols aimed at recognizing when a situation requires support to be scaled up and how to handle this. AEMH systems offer many opportunities due to their autonomy, but because of that they also carry the responsibility of ensuring patient safety [83]. One of the ways to do this is to clearly define the scope of a system, exactly which situations it can and cannot safely deal with. In regular mental healthcare, such practices are already often in place, defining when and how patients should be screened and referred [84]. Suicidality is often a situation warranting referral in regular care [85], but one can imagine that for AEMH many more situations would qualify such as substance abuse, self-harm, and violent behavior. On the other hand, not all users of AEMHs might be willing to be referred to human care, as anonymity is often one of the biggest draws of the technology [86]. It is, therefore, important to carefully consider which situations warrant referral and which do not, and how to ensure patient safety. This question is relevant for all AEMH systems, not just for PTSD. Therefore, this thesis gives general recommendations of how AEMH systems could detect risk situations, and how a virtual agent could successfully refer mental health patients in case their safety

cannot be warranted anymore. These recommendations follow from sub-question 2:

Question 2: How do we support patient safety when they work with an AEMH system?

Aside from considering the specific tasks a virtual agent for PTSD has, it is also important to consider the PTSD therapy system it is a part of. This thesis proposes the blueprint for such a system and identify four key components: 1) a virtual coach, 2) exposure environments, 3) general therapy tools and 4) risk management. Although some virtual agents for mental-health function purely as a stand-alone conversational partner [51, 52], many also function within a broader system with additional functions [63, 64]. For a PTSD therapy system such functions can be especially helpful, as the system needs to facilitate exposure. The most commonly used technology for exposure for PTSD is virtual reality, which has been shown to be effective, but has the problem of being difficult to personalize [87]. For this reason, van der Steen et al. [77] proposed the Multi-Model Memory Restructuring (3MR) system, which incorporates two exposure environments, namely a digital diary, supplemented with a so-called virtual worldbuilder in which patients themselves can recreate their memories in a 3D environment. To make the system easier to use in home therapy, functions such as a memory timeline and a session overview can be added. Finally, by adding general monitoring functions which allow a therapist to monitor progress, patient safety can be improved. In this thesis, we present the 3MR_2 system, which incorporates exposure environments, general therapy tools, risk management and a virtual agent which assists patients during therapy. This system provides the answer to sub-question 3:

Question 3: What is a feasible computerized PTSD therapy system incorporating a virtual agent?

1.3. Research Approach

A distinction in research approach can be made between the hypotheses and research questions presented in the previous section. The research topics formulated by the research questions were mostly explorative, and do not necessarily include an expected outcome. Therefore, these questions were answered using qualitative methods such as expert interviews, focus groups and use-cases, all grounded in existing research. In some cases, the resulting models and systems resulted in new hypotheses, which were then empirically evaluated. The hypotheses do incorporate an expected outcome, which means that these could be empirically tested. The research approach for the hypotheses starts with a base in existing theory and literature, which formed the basis for the formulation of the hypothesis. Following this formulation, systems were developed that would allow testing of the hypothesis, either fully automated or with a human-in-the-loop component. Following the development, the hypotheses were tested in empirical experiments.

The experimental studies described in this thesis focus on evaluating specific elements of a virtual agent for PTSD therapy, such as its motivational messages or personalized questions. Through this component-based evaluation, the effects of specific elements can be studied closely quickly, something that is much more difficult to when testing a whole therapy application [88]. This relates to the technique of abstraction, which is also used in gaming. Through reducing complexity, it becomes easier to identify cause and effect, removes extraneous factors and it reduces time necessary to grasp the concepts involved [89]. Through measuring the underlying concepts of the components such as compliance, motivation and working alliance, a better understanding is reached on exactly how these components will affect behavior [90]. This approach also means that most evaluations were not performed with PTSD patients in therapy, as participants were only presented with elements from a therapeutic intervention. Generalizability of the results was enhanced through taking into account PTSD symptoms as covariate, approaching trauma recollection with the recollection of unpleasant memories, asking participants to imagine they are in certain situations and by recruiting former PTSD patients.

The first research question, exploring what design guidelines are important for the design of a virtual agent that support PTSD treatment, was studied following the scenario based design method with experts [91]. Several scenarios were developed, outlining possible ways for a virtual agent to behave during treatment. Accompanying these scenarios, strong claims about how a virtual agent should behave were established. These scenarios and claims were discussed with experts in the field of PTSD therapy (n=10). All discussions were recorded, and all expert statements were analyzed on underlying principles. From these principles, ten design guidelines were derived, as well as suggestions on how to implement them in a virtual agent for PTSD research. More details can be found in chapter 2.

The first hypothesis states that the presentation mode of psychoeducation presented by a virtual agent affects adherence. This hypothesis was tested in an experimental setting, with a between-subject design. 46 participants were asked by a virtual agent to list and rate their five worst memories. Subsequently, they received psychoeducation, either reading this information on the screen or being told the information by the virtual agent in a text-to-speech form. After receiving the information, participants were asked by the virtual agent to pick one memory to describe in detail, and advised to pick the worst. In this experiment, attitude towards the agent was measured, as well as how well the information was recollected. Three different measures of adherence were used. More information can be found in chapter 3.

The second hypothesis states that an ontology-based question system can enhance memory recollection. This hypothesis was also tested in an experimental setting, but in a within-subject design, every participant (n=24) receiving both experimental conditions. In each condition, participants were asked to describe a holiday memory in a digital diary. In one condition, they were asked personalized questions based on an ontology of holiday memories. In the other conditions, the questions were not personalized and followed a set sequence. Their answers were

analyzed by examining the content and level of detail. More information can be found in chapter 4.

The third hypothesis states that a virtual agent can enhance therapy trust and motivation by motivating based on the situation. To test this hypothesis, a personalized motivational feedback system was developed. In a survey experts (n=13) were asked what they would say to patients in different situations, which were characterized by symptom progression and the patient's trust in a good therapy outcome. These answers were categorized to form a database of statements including the probability of a category occurring in a given situation. An online experiment (n=207) was done to study the effect of the personalized motivational feedback generated by this system. Participants were asked to imagine a situation and presented with personalized feedback, general feedback or no feedback. After the feedback, they were asked to indicate how this changed their motivation to continue and trust in a good therapy outcome. More details can be found in chapter 5.

The second research question is how we can support patient safety while they work with an AEMH system. This question was studied with a combination of qualitative and quantitative methods. The focus in both parts was recognizing when an AEMH system cannot safely deal with the situation anymore, and how to refer patients back to human care. Three models were developed for detection of the problem, automatic referral to human care and motivating patients to seek human care. This third model operates on the principle that it is important to take into account the patients situation and motivation before referring them. This principle was tested in an online experiment, wherein participants (n=160) were asked to envision a certain situation for themselves. They then chatted with an online coach that would take a certain approach in referring them to human care. We measured not only what actions people would take given their condition, but also how satisfied they were with the interaction. More details can be found in chapter 6.

The final research question is what a feasible AEMH system is for PTSD therapy incorporating a virtual agent. To answer this question, we present the Multi-Model Memory Restructuring 2.0 (3MR_2) system. 3MR_2 incorporates a virtual agent, a digital diary and a 3D environment to offer full PTSD therapy wherein a human therapist is only involved in a monitoring capacity. During therapy, patients are guided by the virtual agent in describing their memories in the digital diary, and recreating them in 3D. A small study was done with former PTSD patients (n=4), who tested several parts of the system. The questions posed by the virtual coach were rated on the level of appropriateness, and the whole system was rated on usability. More details can be found in chapter 7.

Chapter 8 discusses the conclusions about how a virtual agent that is a part of PTSD therapy can act to enhance treatment compliance. It reflects on the specific tasks of a virtual agent and the context it operates in. Additionally, it considers the implications of this work for other virtual agent-based systems and AEMH applications. Finally, this chapter provides recommendations for future work in this area based on the results and limitations of this work.

References

[1] A. P. Association, *Diagnostic and Statistical Manual of Nental Disorders, Fifth Edition* (American Psychiatric Association, 2013).

[2] J. Alonso, M. Angermeyer, S. Bernert, R. Bruffaerts, T. Brugha, H. Bryson, G. de Girolamo, R. de Graaf, K. Demyttenaere, I. Gasquet, J. Haro, S. Katz, R. Kessler, V. Kovess, J. Lepine, J. Ormel, G. Polidori, L. Russo, and G. Vilagut, *Disability and quality of life impact of mental disorders in europe: results from the european study of the epidemiology of mental disorders (esemed) project,* Acta Psychiatrica Scandinavica **109**, 38 (2004).

[3] M. Pompili, L. Sher, G. Serafini, A. Forte, M. Innamorati, G. Dominici, D. Lester, M. Amore, and P. Girardi, *Posttraumatic stress disorder and suicide risk among veterans,* The Journal of Nervous and Mental Disease (2013).

[4] T. Galovski and J. A. Lyons, *Psychological sequelae of combat violence: A review of the impact of ptsd on the veteran's family and possible interventions,* Agression and Violent Behavior **9**, 477 (2004).

[5] G.-J. de Vries and M. Olff, *The lifetime prevalence of traumatic events and posttraumatic stress disorder in the netherlands,* Journal of Traumatic Stress **22**, 259 (2009).

[6] WHO, *Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level,* (2012).

[7] P. Hyland, J. Murphy, M. Shevlin, F. Vallières, E. McElroy, A. Elklit, M. Christoffersen, and M. Cloitre, *Variation in post-traumatic response: the role of trauma type in predicting icd-11 ptsd and cptsd symptoms,* Social Psychiatry and Psychiatric Epidemiology **52**, 727 (2017).

[8] O. Frans, P.-A. Rimmo, L. Adberg, and M. Fredrikson, *Trauma exposure and post-traumatic stress disorder in the general population,* Acta Psychiatrica Scandinavica **111**, 291 (2005).

[9] C. M. Ogle, D. C. Rubin, D. Berntsen, and I. C. Siegler, *The frequency and impact of exposure to potentially traumatic events over the life course,* Clin Psychol Sci. **1**, 426 (2013).

[10] D. A. Clark and A. T. Beck, *Cognitive Therapy of Anxiety Disorders* (The Guilford Press, 2010).

[11] I. M. Engelhard, M. A. van den Hout, W. C. Janssen, and J. van der Beek, *Eye movements reduce vividness and emoemotional of "flashforwards",* Behaviour Research and Therapy , 1 (2010).

[12] M. B. Powers, J. M. Halpern, M. P. Ferenschalk, S. J. Gillihan, and E. B. Foa, *A meta-analytic review of prolonged exposure for posttraumatic stress disorder,* Clinical Psychology Review **30**, 635 (2010).

[13] R. Bradley, J. Greene, E. Russ, L. Dutra, and D. Westen, *A multidimensional meta-analysis of psychotherapy for ptsd,* AM J Psychiatry **162**, 214 (2005).

[14] N. Kar, *Cognitive behavioral therapy for the treatment of post-traumatic stress disorder: a review,* Neuropsychiatric Disease and Treatment **7**, 167 (2011).

[15] S. Farooq and F. Naeem, *Tackling nonadherence in psychiatric disorders: current opinion,* Neuropsychiatric Disease and Treatment **10**, 1069 (2014).

[16] J. A. Cartreine, D. K. Ahern, and S. E. Locke, *A roadmap to computer-based psychotherapy in the united states,* Harv Rev Psychiatry **18**, 80 (2011).

[17] R. Kenter, L. Warmerdam, C. Brouwer-Dudokdewit, P. Cuijpers, and A. van Straten, *Gguide online treatment in routine mental health care: an observational study on uptake, drop-outs and effects,* BCM Psychiatry **13**, 13 (2013).

[18] C. Lyons, P. Hopley, and J. Horrocks, *A decade of stigma and discrimination in mental health:plus ça change, plus c'est la même chose (the morethings change, the more they stay the same),* Journal of Psychiatric and Mental Health Nursing **16**, 501 (2009).

[19] J. Jankovic, I. Vidakovic, A. Matanov, M. Schützwohl, D. Ljubotina, D. Lecic-Tosevski, and S. Priebe, *Reasons for not receiving treatment in people with posttraumatic stress disorder following war,* The Journal of Nervous and Mental Disease **199**, 100 (2011).

[20] V. Kantor, M. Knefel, and B. Lueger-Schuster, *Perceived barriers and facilitators of mental health service utilization in adult trauma survivors: A systematic review,* Clinical Psychology Review **52**, 52 (2017).

[21] T. Donker, M. Blankers, E. Hedman, B. Ljótsson, K. Petrie, and H. Christensen, *Economic evaluations of internet interventions for mental health: a systematic review,* Psychological Medicine , 1 (2015).

[22] C. Erbes, R. Stinson, E. Kuhn, M. Polusny, J. Urban, J. Hoffman, J. Ruzek, C. Stepnowsky, and S. Thorp, *Access, utilization, and interest in mhealth applications among veterans receiving outpatient care for ptsd,* Military Medicine **179**, 1218 (2014).

[23] S. Lal and C. E. Adair, *E-mental health: A rapid review of the literature,* Psychiatric Services **65**, 24 (2014).

[24] D. Santor, C. Poulin, J. LeBlanc, and V. Kusumakar, *Online health promotion, early identification of difficulties, and help seeking in young people.* Journal of the American Academy of Child and Adolescent Psychiatry **46**, 50 (2007).

[25] J. Robinson, S. Hetrick, G. Cox, S. Bendall, A. Yung, and J. Pirkis, *The safety and acceptability of delivering an online intervention to secondary students at risk of suicide: Findings from a pilot study,* Early Intervention in Psychiatry **9**, 498 (2015).

[26] J. Becker, H. Fliege, R. Kocalevent, J. Bjorner, M. Rose, O. Walter, and B. Klapp, *Functioning and validity of a computerized adaptive test to measure anxiety (a-cat).* Depression and Anxiety **25**, E182 (2008).

[27] B. Klein, J. Mitchell, J. Abbott, K. Shandley, D. Austin, K. Gilson, L. Kiropoulos, G. Cannard, and T. Redman, *A therapist-assisted cognitive behavior therapy internet intervention for posttraumatic stress disorder: Pre-, post- and 3-month follow-up results from an open trial,* Journal of Anxiety Disorders **24**, 635 (2010).

[28] V. Spek, P. Cuijpers, I. Nyklicek, H. Riper, J. Keyzer, and V. Pop, *Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis,* Psychological Medicine **37**, 319 (2007).

[29] M. Hirai and G. A. Clum, *An internet-based self-change program for traumatic event related fear, distress, and maladaptive coping,* Journal of Traumatic Stress **18**, 631 (2005).

[30] J. Spence, N. Titov, L. Johnston, M. Jones, and B. F. Dear, *Internet-based trauma-focused cognitive behavioural therapy for ptsd with and witwith exposure components: A randomised controlled trial,* Journal of Affective Disorders **162**, 73 (2014).

[31] K. Possemato, E. Kuhn, E. Johnson, J. Hoffman, and E. Brooks, *Development and refinement of a clinician intervention to facilitate primary care patient use of the ptsd coach app,* Translational Behavioral Medicine **7**, 116 (2017).

[32] H. M. Christensen, P. J. Batterham, and B. O'Dea, *E-health interventions for suicide prevention,* International journal of environmental research and public health **11**, 8193 (2014).

[33] E. Karyotaki, H. Riper, J. Twisk, n. A. Hoogendoor, A. Kleiboer, A. Mira, A. Mackinnon, B. Meyer, C. Botella, E. Littlewood, G. Andersson, H. Christensen, J. Klein, J. Schröder, J. Bretón-López, J. Scheider, K. Griffiths, L. Farrer, M. Huibers, R. Phillips, S. Gilbody, S. Moritz, T. Berger, V. Pop, V. Spek, and P. Cuijpers, *Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: A meta-analysis of individual participant data.* JAMA Psychiatry **1;74**, 351 (2017).

[34] C. Pasarelu, G. Andersson, L. Bergman Nordgren, and A. Dobrean, *Internet-delivered transdiagnostic and tailored cognitive behavioral therapy for anxiety and depression: a systematic review and meta-analysis of randomized controlled trials,* Cognitive Behaviour Therapy **46**, 1 (2017).

[35] I. Kampmann, P. Emmelkamp, and N. Morina, *Meta-analysis of technology-assisted interventions for social anxiety disorder,* Journal of Anxiety Disorders **42**, 71 (2016).

1

[36] G. Andrews, J. Newby, and A. Williams, *Internet-delivered cognitive behavior therapy for anxiety disorders is here to stay,* Current psychiatry reports **17**, 533 (2015).

[37] M. Sijbrandij, I. Kunovski, and P. Cuijpers, *Effectiveness of internet-delivered cognitive behavioral therapy for post-traumatic stress disorder: a systematic review and meta-analysis,* Depression and Anxiety **33**, 783 (2016).

[38] U. Schmidt and T. Wykes, *E-mental health - a land of unlimited possibilities,* Journal of Mental Health **21**, 327 (2012).

[39] T. Dalgleish, A. Mathews, and J. Wood, *Inhibition processes in cognition and emotion: A special case?* (John Wiley & Sons, 1999) Chap. 13.

[40] L. Andrews, S. Joseph, N. Troop, T. V. Rooyen, B. D. Dunn, and T. Dalgleish, *The structure of avoidance following trauma: Development and validation of the posttraumatic avoidance scale (pas),* Traumatology **19**, 126 (2013).

[41] M. Schauer, F. Neuner, and T. Elbert, *Narrative Exposure Therapy* (Hogrefe, 2011).

[42] A. Ortiz, M. del Puy Carretero, D. Oyarzun, J. J. Yanguas, C. Buiza, M. F. Gonzalez, and I. Etxeberria, *Elderly users in ambient intelligence: Does an avatar improve the interaction?* in *Lecture Notes in Computer Science*, Vol. 4397 (2007) pp. 99–114.

[43] M. M. Morandell, A. Hochgatterer, S. Fagel, and S. Wassertheurer, *Avatars in assistive homes for the elderly a user-friendly way of interaction?* in *4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society* (2008).

[44] T. W. Bickmore, R. A. S. K. Nelson, D. M. Cheng, M. Winter, L. Henault, and M. K. Paasche-Orlow, *A randomized controlled trial of an automated exercise coach for older adults,* Journal of the American Geriatrics Society **61**, 1676 (2013).

[45] A. D. Andrade, R. Anam, C. Karanam, P. Downey, and J. G. Ruiz, *An overactive bladder online self-management program with embedded avatars: A randomized controlled trial of efficacy,* Urology **85** (2015).

[46] A. L. Baylor and S. Kim, *Designing nonverbal communication for pedagogical agents: When less is more,* Computers in Human Behavior **25**, 450 (2009).

[47] H. van Vugt, J. Bailenson, J. Hoorn, and E. Konijn, *Effects of facial similarity on user responses to embodied agents,* ACM Transactions on Computer-Human Interaction **17** (2010).

[48] Y. Shiban, I. Schelhorn, V. Jobst, A. Hörnlein, F. Puppe, P. Pauli, and A. Múhlberger, *The appearance effect: Influences of virtual agent features on performance and motivation,* Computers in Human Behavior **49**, 5 (2015).

[49] A. van Wissen, C. Vinkers, and A. van Halteren, *Developing a virtual coach for chronic patients: A user study on the impact of similarity, familiarity and realism,* in *Int. Conf. on Persuasive Technology* (2016).

[50] N. Yee, J. N. Bailenson, and K. Rickertsen, *A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces,* in *CHI 2007* (2007).

[51] A. A. Rizzo, B. Lange, J. G. Buckwalter, E. Forbell, J. Kim, K. Sagae, J. Williams, B. O. Rothbaum, J. Difede, G. Reger, T. Parsons, and P. Kenny, *An intelligent virtual human system for providing healthcare information and support,* Medicine meets Virtual Reality (2011).

[52] D. DeVault, R. Artstein, G. Benn, T. Dey, E. Fast, A. Gainer, K. Georgila, J. Gratch, A. Hartholt, M. Lhommet, G. Lucas, S. Marsella, F. Morbini, A. Nazarian, S. Scherer, G. Stratou, A. Suri, D. Traum, R. Wood, Y. Xu, A. Rizzo, and L.-P. Morency, *Simsensei kiosk: A virtual human interviewer for healhealth decision support,* in *AAMAS* (2014).

[53] L. P. Vardoulakis, L. Ring, B. Barry, C. Sidner, and T. Bickmore, *Designing relational agents as long term social companions for older adults,* in *Proceedings of the International Conference on Intelligent Virtual Agents (IVA) 2012* (2012).

[54] H. Huang, Y. Takeda, K. Kiyoshi, and K. Kawagoe, *A framework of an always-with companion agent for the isolated elderly,* in *Gerontechnology* (2014).

[55] O. A. Blanson-Henkemans, C. A. van der Mast, P. J. van der Boog, M. A. Neerincx, J. Lindenberg, and B. J. Zwetsloot-Schonk, *An online lifestyle diary with a persuasive computer assistant providing feedback on self-management,* Technology and Health Care **17**, 253 (2009).

[56] D. Schulman, T. Bickmore, and C. Sidner, *An intelligent conversational agent for promoting long-term health behavior change using motivational interviewing,* To appear.

[57] J. Ren, D. Schulman, J. Brian, and T. Bickmore, *Supporting longitudinal change in many health behaviors,* in *CHI* (2014).

[58] C. Ritchie, J. Richman, H. Sobko, E. Bodner, B. Phillips, and T. Houston, *The e-coach transition support computer telephony implementation study: Protocol of a randomized trial,* Contemporary Clinical Trials **33**, 1172 (2012).

[59] K. Kimani, T. Bickmore, H. Trinh, L. Ring, M. Paasche-Orlow, and J. Magnani, *A smartphone-based virtual agent for atrial fibrillation education and counseling,* in *Intelligent Virtual Agents (IVA)* (2016).

[60] U. Yasavur, C. Lisetti, and N. Rishe, *Let's talk! speaking virtual counselor offers you a brief intervention,* Journal on Multimodal User Interfaces **8**, 381 (2014).

[61] S. Zhou, B. T., A. Rubin, K. Yeksigian, R. Lippin-Foster, M. Heilman, and S. Simon, *A relational agent for alcohol misuse screening and intervention in primary care,* in *A Relational Agent for Alcohol Misuse Screening and Intervention in Primary Care CHI'17 Workshop on Interactive Systems in Healthcare (WISH)* (2017).

[62] C. Horsch, *A virtual sleepcoach for people suffering from insomnia*, Ph.D. thesis, Delft University of Technology (2016).

[63] D. Hartanto, W. P. Brinkman, I. L. Kampmann, N. Morina, P. G. Emmelkamp, and M. A. Neerincx, *Home-based virtual reality exposure therapy with virtual health agent support.* Pervasive Computing Paradigms for Mental Health **204**, 85 (2015).

[64] A. Bresó, J. Martínez-Miranda, and J. M. García-Gómez, *Leveraging adaptive sessions based on therapeutic empathy through a virtual agent,* in *ICAART - Doctoral Consortium* (2014).

[65] L. Ring, T. Bickmore, and P. Pedrelli, *An affectively aware virtual therapist for depression counseling,* in *ACM SIGCHI Conference on Human Factors in Computing Systems (CHI) workshop on Computing and Mental Health* (2016).

[66] G. Fischer, *User modeling in human-computer interaction,* User Modeling and User-Adapted Interaction **11**, 1 (2000).

[67] M. Neerincx and J. Lindenberg, *Situated cognitive engineering for complex task environments,* (Schraagen, 2008) Chap. 18, pp. 371–387.

[68] M. L. Tielman, W.-P. Brinkman, and M. Neerincx, *Design guidelines for a virtual coach for post-traumatic stress disorder patients,* in *Intelligent virtual agents* (2014).

[69] U. D. of Health, H. Services, N. I. of Health, and N. C. Institute, *Theory at a glance - a guide for health promotion practice,* (2005).

[70] S. P. Cahill, E. B. Foa, E. A. Hembree, R. D. Marshall, and N. Nacash, *Dissemination of exposure therapy in the treatment of posttraumatic stress disorder,* Journal of Traumatic Stress **19**, 597 (2006).

[71] Webb and Wallon, *Comprehension by reading versus hearing,* The Journal of Applied Psychology **40**, 237 (1956).

[72] D. L. Rubin, T. Hafer, and K. Arata, *Reading and listening to oral-based versus literate-based discourse,* Communication Education **49**, 121 (2000).

[73] B. Fogg, J. Marshall, O. Laraki, A. Osipovich, C. Varma, N. Fang, J. Paul, A. Rangnekar, J. Shon, P. Swani, and M. Treinen, *What makes web sites credible? a report on a large quantitative study,* in *CHI 2001* (2001).

[74] K. Fiscella, S. Meldrum, P. Franks, C. Shields, P. Duberstein, S. McDaniel, and R. Epstein, *Patient trust: is it related to patient-centered behavior of primary care physicians?* Medical Care (2004).

[75] A. S. B. Bohnert, K. Zivin, D. E. Welsh, and A. M. Kilbourne, *Ratings of patient-provider communication among veterans:serious mental illnesses, substance use disorders, and themoderating role of trust,* Health Communication **26**, 267 (2011).

[76] D. Safran, D. Taira, W. Rogers, M. Kosinski, J. Ware, and A. Tarlov, *Linking primary care performance to outcomes of care.* Journal of Family Practice (1998).

[77] M. van den Steen, W.-P. Brinkman, E. Vermetten, and M. Neerincx, *Design and usability evaluation of a multi-modal memory restructuring system for the treatment of combat-related ptsd,* in *ECCE 2010 Workshop on Cognitive Engineering for Technology in Mental Health Care and Rehabilitation* (2010).

[78] N. F. Noy and D. L. McGuinnes, *Ontology Development 101: A Guide to Creating Your First Ontology*, Tech. Rep. (Stanford Knowledge Systems Laboratory, 2001).

[79] T. Bickmore, D. Schulman, and C. Sidner, *A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology,* Journal of Biomedical Informatics **44**, 183 (2011).

[80] M. A. Schottenbauer, C. R. Glass, D. B. Arnkoff, V. Tendick, and S. H. Gray, *Nonresponse and dropout rates in outcome studies on ptsd: Review and methodological considerations,* Psychiatry: Interpersonal and Biological Processes **71**, 134 (2008).

[81] M. Tielman, M. Neerincx, and W. Brinkman, *Generating situation-based motivational feedback in a post-traumatic stress disorder e-health system,* in *IVA* (2017).

[82] M. Kunaver and T. Požrl, *Diversity in recommender systems - a survey,* Knowledge-based systems **123**, 154 (2017).

[83] D. D. Luxton, *Recommendations for the ethical use and design of artificial intelligent care providers,* Artificial Intelligence in Medicine **62**, 1 (2014).

[84] A. L. Siu and the US Preventive Services Task Force, *Screening for depression in adults us preventive services task force recommendation statement,* JAMA **315**, 380 (2016).

[85] B. H. Belnap, H. C. Schulberg, F. He, S. Mazumdar, C. F. R. III, and B. L. Rollman, *Electronic protocol for suicide risk management in research participants,* Psychosomatic Research **78**, 340 (2015).

[86] G. Lucas, J. Gratch, A. King, and L.-P. Morency, *It's only a computer: Virtual humans increase willingness to disclose,* Computers in Human Behavior **37**, 4 (2014).

[87] K. Meyerbröker and P. Emmelkamp, *Virtual reality exposure therapy in anxiety disorders: A systematic review of process-and-outcome studies,* Depression and Anxiety **27**, 933 (2010).

[88] T. M. Baker, D. H. Gustafson, and D. Shah, *How can research keep up with ehealth? ten strategies for increasing the timeliness and usefulness of ehealth research,* Journal of Medical Internet Research **16** (2014).

[89] K. M. Kapp, *The Gamification of Learning and Instruction*, edited by R. Taff (ASTD and Pfeiffer, 2012).

[90] P. Klasnja, S. Consolvo, and W. Pratt, *How to evaluate technologies for health behavior change in hci research,* in *CHI* (2011).

[91] J. M. Carroll, *Making Use: Scenario-Based Design of Human-Computer Interactions* (MIT Press, 2000).

2

Design guidelines for a virtual coach for Post-Traumatic Stress Disorder patients

Patients with Post Traumatic Stress Disorder (PTSD) often need to specify and relive their traumatic memories in therapy to relieve their disorder, which can be a very painful process. One new development is an internet-based guided self-therapy system (IBGST), where people work at home and a therapist is remotely involved. We propose to enrich an IBGST with a virtual coach to motivate and assist the patient during the therapy. We have created scenarios and requirements for an IBGST coach and discussed these with 10 experts in structured interviews. From these interviews, we have identified 10 important guidelines to assist with the design of a virtual coach assisting with PTSD treatment.

This chapter has been published in international conference on Intelligent Virtual Agents. (2014) [1].

2.1. Introduction

An important upcoming role for virtual agents is coaching, the task of motivating and assisting people to achieve their goals. Blanson-Henkemans et al. [2] show that a virtual coach can motivate and support people successfully. Rizzo et al. [3] present a virtual coach which can be used to assist Post-Traumatic Stress Disorder (PTSD) patients by guiding them towards information.

PTSD is a mental disorder following one or more traumatic experiences with symptoms such as intrusive memories of the traumatic event, dissociative reactions and irritable behavior [4]. One of the most practiced treatments for PTSD is Cognitive Behavioral Therapy (CBT) with exposure, which is the process of exposing the patient to stimuli which are related to the traumatic memory and will elicit a fear response.

A new development within the treatment of PTSD is an internet-based guided self-therapy (IBGST) system, where people work at home and a therapist is only remotely involved. In this paper we will focus on the Multi-Model Memory Restructuring (3MR) system [5], which allows patients to structure memories and follow exposure-therapy on their own PC through creating memories on a visual timeline and adding media such as photos, music and text. An additional functionality is the possibility to recreate personal memories in a 3D environment. One difficulty is that exposure treatments can be very painful and possibly demotivating at times. For this reason, we believe that a virtual coach would be a very useful addition to the system. Such a coach could be capable of offering personalized and motivational assistance during the therapy process, increasing trust in the therapy and hope in a positive outcome.

In this paper, we present the first steps towards developing such a virtual coach for patients with PTSD working with the IBGST 3MR system. Because of differences between PTSD patients we have chosen to focus on two specific patient groups, namely victims of childhood sexual abuse (CSA) and military veterans.

2.2. Structured Interviews

To determine the specific user requirements for PTSD patients, we have designed scenarios with discussed these in structured interviews with 10 experts.

We adopted a scenario-based approach to inform the experts on the context in which our coach would be operating. The scenarios represented the types of sessions a patient would follow during treatment including a possible way in which the virtual coach could assist and also described aspects of behavior of the virtual coach which were based on literature on motivation [6] and medical communication guidelines [7]. From the scenarios we identified requirements for the coach, dealing with topics such as facial expressions, giving explanations on the therapy and motivation. From these requirements we formulated strong claims such as 'It would pose a problem if the virtual coach had the same characteristics with each patient' to stimulate a discussion. In structured interviews, we presented the scenarios and claims to 10 experts specialized in trauma treatment. The interviews were conducted in 6 sessions, each with 1 to 3 experts present. Each session discussed

Table 2.1: Guidelines for a virtual coach for PTSD

Guideline & Suggestions for use by virtual coach
Motivate: Compliment patients, when they have a hard time, explain that this is normal, remind them of personal goals.
Take patient seriously: Let the patients decide if they are ready to end therapy, do not explain what they already know.
Protect patient: Build in relaxation exercises for when exposure becomes too much, give relapse-prevention.
Be down to earth: Be very factual, never act shocked, do not show many emotions, use simple, unaffected language.
Personalize: When motivating, mention specific things about this person and progress, adapt gender and age to fit patient.
Transparency: Be clear on what is going to happen and the coach's capabilities, explain why the patient needs to do things.
Avoid negative reinforcement: Do not reflect too much on negative things, never give negative feedback, never act disappointed.
Protect patients from themselves: Too much choice can lead to dissatisfaction, discourage avoidance behavior.
Psychoeducation: Put the progress and therapy into a theoretical frame, explain why things need to be done and why they work.
Acknowledge: Acknowledge the tension and emotional state of the patient, let them know the coach has heard them.

between 6 and 24 claims.

All interviews were recorded and the statements of the experts were written down. After this, we determined the underlying guideline and grouped the statements. From this analysis, we derived the guidelines as shown in Table 2.1. From the contexts, we can also present some suggestions for the use of each guideline. Motivating a patient was mentioned often, so we also considered what the experts said on motivation in more detail and identified six ways of motivating, as shown in Table 2.2.

2.3. Discussion & Conclusion

A first observation we can make of our guidelines is that they agree with the principles identified by Miller and Rollnick and Wouda at al. [6, 7]. Another thing to consider is the role of our virtual coach. Based on the guidelines identified we can say that the coach can be a safe-guard, an educator and a motivator. These three roles correspond to roles a human therapist would also have, but a virtual coach does have its own limitations and possibilities which set it apart. Opportunities for a virtual coach may lie with being down to earth and personalization. A virtual coach will never have strong emotional reactions to traumatic memories which would need to be kept in check and a virtual coach has the unique quality that appearance, age and gender can be adapted specifically for a patient. Another difference which is

Table 2.2: A decomposition of the guideline of Motivation for a virtual coach for PTSD

Motivational theme	**Explanation for use by virtual coach**
Compliment	Give concrete, but many compliments
Empathy	Show empathy, but never include strong feelings
Evaluate	Set concrete goals to evaluate progress and outcome
Doing badly is not possible	There is no 'wrong' way to do things
Personalize	Motivate in a personalized way, be as concrete as possible
Be positive	Give positive feedback and reflect most on positive things

Table 2.3: Examples of behavior for a virtual coach for PTSD, based on the guidelines.

Potential behavior of a virtual coach coach	**Guideline**
When patients fill in how they feel, mention that you have heard them, i.e. 'I notice you are feeling distressed now.'	Acknowledge
In the beginning of a session, explain what is going to happen.	Transparency
If the patient indicates wishing to quit, remind them of their goals.	Motivate
Let the patient choose the gender of their coach.	Personalization
Be factual in complementing, so do not say 'Great job, amazing, you are wonderful', but 'You completed this task, well done'.	Be down to earth
Never express a negative emotion. Whenever the patient fails to do something, do not punish but ask them why they did not do it.	Do not reinforce the negative
If distress is very high, explain that this is normal and will decrease.	Psychoeducation
Monitor symptoms, if these keep increasing, notify a therapist.	Protect patient
Whenever explaining something for the second time, state this explicitly. 'I know you have heard this before, but...'	Take patient seriously
Do not let patients decide when to start with exposure by themselves, assist them in setting a date.	Protect patient from themselves

connected to personalization is the support of memory retrieval. Therapists typically ask questions to assist patients with remembering, and it is important that these questions are well tailored to the patient. For humans natural language is the key ingredient, which is challenging for a virtual coach. However, because the coach is embedded into a system it has more modalities available. A coach could react to items which are placed in the application, for instance ask who is present in a photograph or a video. Another example is that a coach could notice how many items are added when specific questions are asked to keep track of which questions work well for which patient.

To conclude this paper, we would like to give examples of some of the ways in which our guidelines could be applied to specific functions and behavior of a virtual coach. These are presented in Table 2.3.

References

[1] M. L. Tielman, W.-P. Brinkman, and M. Neerincx, *Design guidelines for a virtual coach for post-traumatic stress disorder patients,* in *Intelligent virtual agents* (2014).

[2] O. A. Blanson-Henkemans, C. A. van der Mast, P. J. van der Boog, M. A. Neerincx, J. Lindenberg, and B. J. Zwetsloot-Schonk, *An online lifestyle diary with a persuasive computer assistant providing feedback on self-management,* Technology and Health Care **17**, 253 (2009).

[3] A. Rizzo, B. Lange, J. G. Buckwalter, E. Forbell, J. Kim, K. Sagae, J. Williams, J. Difede, B. O. Rothbaum, G. Reger, T. Parsons, and P. Kenny, *Simcoach: an intelligent virtual human system for providing healthcare information and support,* in *Proc. 8th Intl Conf. Disability, Virtual Reality & Associated Technologies* (2010).

[4] A. P. Association, *Diagnostic and Statistical Manual of Nental Disorders, Fifth Edition* (American Psychiatric Association, 2013).

[5] W.-P. Brinkman, E. Vermetten, M. van den Steen, and M. Neerincx, *Cognitive engineering of a military multi-modal memory restructuring system,* Journal of CyberTherapy and Rehabilitation **4**, 83 (2011).

[6] W. R. Miller and S. Rollnick, *Motivational interviewing: Preparing people to change addictive behavior* (New York: Guilford Press, 1991).

[7] J. Wouda, H. van de Wiel, and K. van Vliet, *Medische communicatie, Gespreksvaardigheden voor de arts* (de Tijdstroom, 2000).

3

How should a virtual agent present psychoeducation? Influence of verbal and textual presentation on adherence.

With the rise of autonomous e-mental health applications, virtual agents can play a major role in improving trustworthiness, therapy outcome and adherence. In these applications, it is important that patients adhere in the sense that they perform the tasks, but also that they adhere to the specific recommendations on how to do them well. One important construct in improving adherence is psychoeducation, information on the why and how of therapeutic interventions. In an e-mental health context, this can be delivered in two different ways: verbally by a (virtual) embodied conversational agent or just via text on the screen. The aim of this research is to study which presentation mode is preferable for improving adherence. This study takes the approach of evaluating a specific part of a therapy, namely psychoeducation. This was done in a non-clinical sample, to first test the general constructs of the human-computer interaction. We performed an experimental study on the effect of presentation mode of psychoeducation on adherence. In this study, we took into account the moderating effects of attitude towards the virtual agent and recollection of the information. Within the paradigm of expressive writing, we asked participants (n = 46) to pick one of their worst memories to describe in a digital diary after receiving verbal or textual psychoeducation.

This chapter was published in Technology & Healthcare (2017) [1].

We found that both the attitude towards the virtual agent and how well the psychoeducation was recollected were positively related to adherence in the form of task execution. Moreover, after controlling for the attitude to the agent and recollection, presentation of psychoeducation via text resulted in higher adherence than verbal presentation by the virtual agent did.

3.1. Introduction

Imagine a virtual embodied conversational agent asking you to confront your worst memories. How would the agent convince you to do this? This is the question we wish to address in this paper. In our research we studied a talking virtual embodied agent, a human-like graphical interface, controlled by a computer, that can help people with recollecting their worst memories. This virtual agent was a part of an autonomous e-mental-health (AEMH) application, which can be used to give expressive writing or Post-Traumatic Stress Disorder (PTSD) therapy. AEMH applications are designed to monitor, assist and treat mental health problems without any direct human involvement. With rising costs in health-care [2], long travel distances to clinics and access to care in rural areas [3], and stigma on seeking help from a therapist [4] all preventing people from seeking care, these applications can fill an important gap. Aside from being available remotely and being privacy sensitive [5], they are also cost-effective in use [6], and efforts are also being made to reduce cost in the development of these interventions [7]. Virtual agents can play a central role in AEMH applications, contributing to the effectiveness and acceptability of the system and intervention [8–11]. One of the primary functions of our virtual agent is to maximise adherence, i.e. ensure that people recollect their most negative memories in as much detail as possible.

Adherence is a major factor in health care. The World Health Organization identifies it as a primary determinant of the effectiveness of treatment and defines it as *"the extent to which a person's behaviour - taking medication, following a diet, and/or executing lifestyle changes, corresponds with agreed recommendations from a health care provider."* [12]. This importance of adherence is widely accepted. However, measuring to which extent behaviour follows recommendations is challenging and the exact way to do so it differs greatly. For this reason, many regular mental health interventions only consider whether a session was done at all, or rely heavily on self-report measures [13]. Applied to e-health, adherence refers to the extent to which the system's recommendations are followed, because a health-care provider is not directly involved. This can be performing a session, showing a certain behaviour or even taking medication if the system advises to do so. In a review of the impact of adherence on e-therapies, Donkin et al. (2011) found many different measures of adherence, such as completed modules, the number of visits to a page, the number of log-ins and self-report measures [14]. The question arises though, if any of these measures fully describe adherence, as there is a difference in passively viewing information and actively applying its contents [15]. For this reason, we adopt a two-fold definition for adherence to a virtual agent. The first aspect of adherence is if recommendations of the agent are followed, i.e., whether tasks are executed (this could also mean the suppression of something). The second aspect is how well these tasks are performed, i.e., if they are not glossed over, and details are left out.

Psychoeducation is an important method to increase adherence in therapy for mental disorders. Psychoeducation can be an intervention of its own, informing patients on their disorder and how to deal with it [16–18]. It is, however, also used as part of more comprehensive interventions. In this case, it often entails informa-

tion on how the therapy works and why patients will have to perform certain tasks [19, 20]. In the remainder of the article, we will use the term *psychoeducation* to refer to this latter form of information included in an intervention. Psychoeducation is especially relevant for therapies where patients need to perform difficult tasks such as changing habits or facing fears, as is the case in expressive writing or PTSD treatment. It aims to increase adherence by firstly explaining how to perform the tasks correctly. Additionally, it explains why performing these tasks will be beneficial [21]. Virtual agents can present psychoeducation in different ways. One possibility is that the virtual agent asks patients to read the psychoeducation text on the screen. Another option would be for the virtual agent to give the information verbally. Of course, human therapists generally present the information verbally, accompanied by a written brochure. Likewise, it is possible to include the option to read back the information on the screen in an AEMH application as well. Still, the choice for the initial presentation mode is an important one. Both options might have their advantages on maximising adherence to the virtual agent, which leads to our first hypothesis: *presentation mode of psychoeducation influences adherence.*

Several factors exist through which the mode of presentation might affect adherence. Firstly, we predict that textual presentation of psychoeducation will maximise recollection of the information. By better remembering the psychoeducation, specifically how and why actions need to be taken, adherence improves. Secondly, we predict that verbal presentation of psychoeducation will maximise a positive attitude towards the virtual agent, leading to improved adherence. These two factors can be considered mediators of the relationship between the presentation mode and adherence. Because attitude to the agent is predicted to be higher for verbal psychoeducation and recollection for textual psychoeducation, their combined result on adherence is unclear. This leads to our second hypothesis: *the effect of presentation mode of psychoeducation on adherence is mediated by how well the information is recollected and the attitude towards the virtual agent.* Together with hypothesis 1, we can visualise the two hypotheses as shown in 3.1.

The first mediating factor predicted by our hypothesis is recollection, how well the given information is remembered. Webb & Wallon work already showed a difference between listening and reading, indicating that text which is read is remembered better [22]. Later work by Rubin et al. showed similar results, also measuring mental effort, which was higher for reading [23]. In the domain of human-computer interaction, verbal-visual learning aids were found to be more beneficial for retention than verbal aids [24]. However, work with virtual agents shows that people often prefer verbal delivery above text [25, 26]. Important to note though is that most work with virtual agents has focused on short assignments and messages while psychoeducation usually consists of much longer texts. It is also important to consider that the speech of a virtual agent might not be as easy to understand as that of a human because many agents use a text-to-speech system [27]. This study aims to provide insight into the effect of a text-to-speech voice on the recollection of longer texts.

The second mediating factor we predict is the attitude towards the virtual agent. Presenting psychoeducation verbally can increase credibility and trust in a virtual

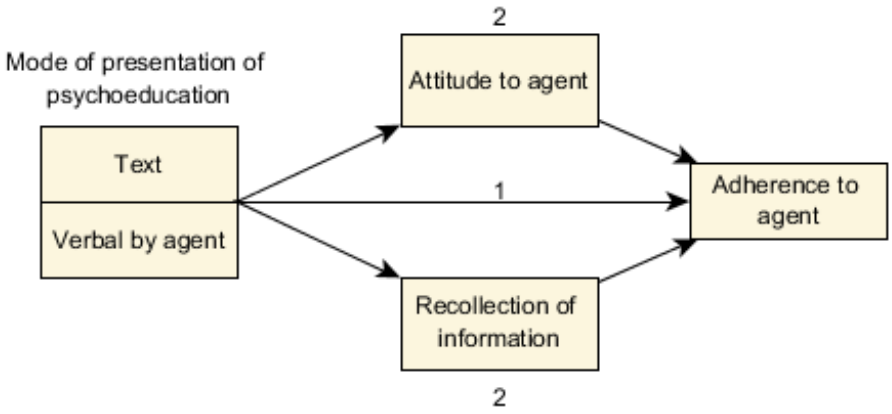

Figure 3.1: Hypotheses on how mode of presentation will influence adherence via the mediators of attitude towards the virtual agent and recollection of the information.

agent, as it presents itself as an expert [28]. Trust in a virtual agent is an important aspect of patient-provider health communications and patient adherence [29–31]. Moreover, both trust and perceived expertise are important aspects of therapeutic alliance [32]. This refers to the bond between a therapist and patient, which is important for both therapy outcome [32] and adherence [33, 34]. Therapeutic alliance has been shown to be relevant for virtual agents [35, 36] as well as human therapists. However, it is not exactly clear what effect mode of presentation of psychoeducation will have on trust and alliance, because of the many aspects involved. Studies show that factors such as the type of language used [37], interpersonal encounter skills [38], and introduction style [39] all influence alliance. However, these studies have been done with human therapists and might not translate to a virtual agent. Agent competence has been shown to affect trust in the agent, but this effect needs to grow over time [40]. Another factor is that virtual agents often use digital voices and do not necessarily bring the same social presence as a human would. Although some indications have been found that even the addition of a text-to-speech voice can increase trust, this effect has been found in the context of consumer trust, not health-care [27].

In this study, we aim to investigate how presentation mode of psychoeducation affects adherence to a virtual agent. Specifically, we compare verbal presentation via a digital voice spoken by a virtual embodied agent with presentation through text. We also consider the effect of the mode of presentation on how well the information is recollected, and the attitude towards the virtual agent. In the remainder of this paper, we present the methods of the experiment we conducted and conclude with the results of our experiment and discussion.

3.2. Method

The gold standard in evaluating e-health interventions is a coherent set of dedicated evaluations in the development stage to test program components, followed by a Randomized Controlled Trial (RCT) when the intervention is stable [41, 41–43]. One recommendation for these evaluations in the development stage is the use of proximal outcome measures such as self-efficacy and adherence [44]. These measures serve as a quicker and efficient way to test if system components have the proposed effect. These studies are usually performed with a non-clinical sample because clinical samples cannot be presented with only a small part of an intervention for ethical reasons. The validity of this choice is supported by the continuum hypothesis [45], which defines mental states on a continuous scale and has been supported for disorders such as psychosis [46], depression [47] and PTSD [48]. This means that many symptoms of clinical populations may also be present in the healthy population, be it to a lesser degree. In this paper, we wish to evaluate how a virtual agent can best present psychoeducation to maximise adherence. Doing this, we took the approach of evaluating a small part of a therapy (psychoeducation), using a proximal outcome measure (adherence) in a non-clinical sample. The psychoeducation was part of a larger intervention, included in the beginning and meant to increase adherence in the latter part.

A between-subjects experiment with two conditions was conducted. In one condition the psychoeducation was delivered verbally by the virtual agent via text-to-speech. In the other, the psychoeducation was presented as text on the screen. The virtual agent was present for the rest of the experiment in both conditions, and was the one prompting the participant to pick and describe a stressful memory. The design of this experiment was approved by the human research ethics committee of Delft University of Technology.

3.2.1. Participants

46 participants (38 male, age M 22.9, SD 3.8) were recruited from the university staff and student population and completed the experiment. All participants had Dutch as their first language. All participants were given experiment information and consent form prior to the experiment. One participant in the text condition voluntarily dropped out due to personal reasons related to the experimental task (not included in the 46).

3.2.2. Therapeutic Task

The subject of psychoeducation differs per therapy and application, as the goal is to inform patients on what that specific therapy is about. To test the effect of presentation mode on adherence, a therapeutic task was chosen which could be performed by a non-clinical person for whom it was yet unpleasant to perform so that a ceiling effect in adherence was avoided. This task was *expressive writing*. Expressive writing is a therapeutic tool [49] where healthy participants have to write about the most negative events in their life, which evoked the most stress. Expressive writing can have several beneficial effects such as reduced stress, a

decrease in healthcare use, and an increase in study performances [50–53]. In the traditional format, individuals write about their memories for 15 minutes on three consecutive days, although studies have also shown it to work for shorter time spans and with less time between sessions [54, 55].

3.2.3. Therapy system, Virtual Agent & Psychoeducation

The therapy application participants used was an adapted version of the Multi-Modal Memory Restructuring (3MR) system [56]. This application was originally developed for PTSD therapy, where the therapy is mainly focused on the recall and description of traumatic memories. The 3MR application is therefore very suitable for an expressive writing task, as it also revolves around describing negative memories.

The 3MR system has a diary function where people can describe their memories. At the top of the screen, a timeline of the user's life is displayed. Memories can be added to this timeline and described in a digital diary. The user of the diary can, besides written text, also include several types of media such as pictures, music or videos. The images, video and music can be added from the computer or via internet sites such as Google image search and YouTube. Locations can be further described using Google maps and emotions can be added in the form of an emotion word. All media are added to the diary in the shape of a small thumbnail, which can be moved around on the screen.

An important aspect of the 3MR system is a virtual agent. Two versions of the virtual agent were used, one female and one male. Gender was matched with the participant to avoid possible effects of interacting with someone with a mismatched gender [57–60]. The agents were rendered by Unity and displayed general idle behaviour and mouth movement when speaking. The voices of the virtual agent were generated with the text-to-speech engine Fluency, which specialises in Dutch voices . Figure 3.2 shows a screenshot of a diary with the timeline at the top and the virtual agent on the right.

The psychoeducation included in this experiment was based on existing research on expressive writing and can be found in [61] and Appendix I in this thesis. The psychoeducation started with a small introduction, followed by possible benefits, how to write to attain the best effect and finished with a short conclusion. The brief introduction stated that extensive research was done on expressive writing and its possible benefits. After this, these benefits were listed: e.g., personal growth, less anxiety, a decrease in health-care utilisation and improved academic performances. After this, instructions were provided on how to maximise the benefits of expressive writing. For example, it was mentioned that it only works for negative memories which evoke stress, works best when writing about your deepest emotions and that the person had to link the memory to current and past events. All examples for both positive effects and instructions were chosen after reviewing previous studies on expressive writing [50–52, 54, 62]. The psychoeducation concluded with linking expressive writing to the digital diary, and people were asked to keep the information in mind during the remainder of the experiment.

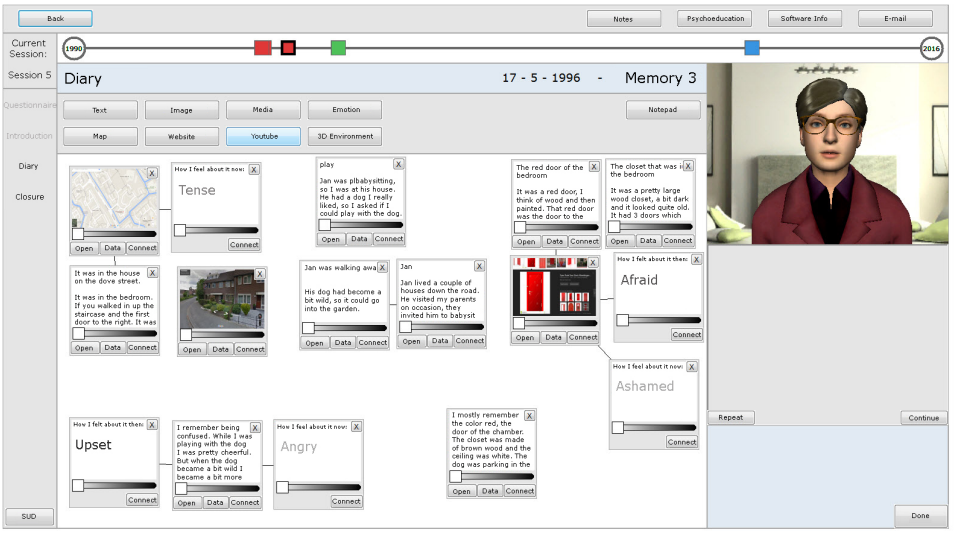

Figure 3.2: The digital diary filled in with an example memory. The items show text, maps, emotion words and images. The bar at the top shows the memory on a timeline. The virtual agent is shown on the right. This screenshot is translated into English; all participants worked with the original Dutch version.

3.2.4. Procedure

At the start of the experiment, all participants received information on the privacy of their data. This included the information that their memory descriptions would be stored to gather statistics such as word count, but that nothing would be read by other people. This was done because we did not want participants to be influenced by privacy concerns during the task. Participants were also informed that they would be able to see exactly what statistics would be stored at the end of the experiment and that they could still withdraw consent at that time. After the information procedure, there was a 3.5-minute video on how the 3MR system worked. In this video, the functions of the diary were explained. Afterwards, there was a short practice session with the diary led by the experimenter. The virtual agent was not yet present during this exercise.

After the practice session, the experimenter left the room, and the virtual agent took over. The virtual agent stayed present during the rest of the experiment, guiding the participant through the assignments with the system that formed the intervention.

The first assignment for the participants was to list their five most negative memories and add a one-sentence description. They were asked to number the memories from most to least stressful, and additionally, assign a number between 0 and 100 to every memory to indicate how stressful they were. In the remainder of the experiment, participants chose one of these five memories to describe in detail. This first task was inserted to get a measure of how stressful the eventually selected memory was. After five memories were listed, the participants received

the psychoeducation. In one condition the virtual agent presented the information verbally via text-to-speech, in the other condition the information was present in the form of text on the screen and the virtual agent simply asked participants to read the text.

Next, participants had to pick one of their five memories to write about in detail. The virtual agent specifically asked the participants to pick one memory, and advised them to pick the most negative. It reminded them that expressive writing works best for the most negative memories, those that evoke stress. This is the recommendation for which we wish to maximize adherence. As a final step, the participants created a diary page for the chosen memory and were asked by the virtual agent to write about the memory in detail in the diary. This was the actual expressive writing task, slightly adapted from the original [49] as our participants only followed one session of writing for 15 minutes instead of three.

At the end of the 15 minutes, the virtual agent asked the participants to fill out several questionnaires on paper and afterwards call in the experimenter. The experimenter closed the experiment by allowing the participants to see which data would be stored, namely the diary descriptions and several statistics such as the word count and the number of emotion items. The data from the diary was shown to the participants in such a way that the experimenter could not see the content. No participants objected to their data being stored for analysis at this point.

3.2.5. Measures

Several measures were collected using both behavioural data and questionnaires. The original Dutch version and an English translation of all questionnaires can be found in [63] and Appendix II in this thesis.

PRIMARY MEASURES

Suggestion adherence was the measure of how stressful the memory chosen for the detailed description was. This measure describes how well participants complied with the virtual agent's suggestion to describe a very negative memory that evoked stress. The initial assignment asked to list the five most stressful memories experienced, of which only one was described in detail in the diary. To judge how negative the chosen memory was, the participants were asked to order the memories from least stressful to most stressful. This gave a ranking score to each memory. Additionally, the stressfulness of each memory was rated from 0 (not stressfull at all) to 100 (maximum stresfulness level) by the participant. This stresfulness score was added to control for interpersonal differences in the stressfulness of past experiences.

Task adherence refers to how well participants performed the task they were given by the virtual agent, namely describing their memory in detail. All participants spent 15 minutes describing their chosen memory in detail in the virtual diary. From this description, the total number of words typed, which emotion items were added to the diary and the total number of items added were recorded. Taken together these measures formed an index for task adherence.

Subjective adherence was measured with a questionnaire asking for the partic-

ipants' experiences in recollecting the memory. The questionnaire asked if people had truly described their deepest feelings, how confronting the memory was, how much they linked it to other memories and if it affected how they saw the memory now. These were all recommendations given in the psychoeducation, which the virtual agent asked the participant to follow. All four items were posed in the form of statements such as *I truly described my deepest emotions and feelings*. All questions were answered on a continuous scale from *don't agree* to *fully agree* in the form of a cross mark on a line. The line had a distinct middle point and incremental marks.

MEDIATING MEASURES

Recollection of the psychoeducation was measured with three open questions. The first question asked for the term used for writing about negative experiences, the second in which areas it could have positive effects and the third what was important during the writing. All of these asked for information given as psychoeducation at the beginning of the experiment. For the first question, a correct answer yielded 2 points, a wrong answer 0. For the second and third question, all correct answers yielded 1 point, and for all concepts written down which were not correct, 1 point was deducted. This resulted in a score between 0 and 5 for questions two and three. The sum of these scores was used as a measure for how well the information was recollected.

Agent attitude was measured with a questionnaire asking for the participant's attitude towards the virtual agent. The questions were on the topics of trust, expertise, realism, perceived amiableness and the adherence of the participant to the agent. All 22 questions were answered on a 7-point Likert scale.

EXPLORATORY MEASURES

Psychoeducation attitude was measured to check for influences of how the information was received. A questionnaire was posed with the topics of trust in the information, how interesting it was, how it was presented and the degree of clarity. All 17 questions were answered on a 7-point Likert scale.

Function usefulness measures were taken for the psychoeducation, the virtual agent, the diary functions and the agent speaking. These four questions asked if these functions added something to the task or not. All questions were posed in the form of statements such as *The presence of the virtual agent added something to this exercise*. All questions were answered on a continuous scale from *don't agree* to *fully agree* in the form of a cross mark on a line. The line had a distinct middle point and incremental marks. All lines were measured to determine where the cross stood and transformed to a value between -1 and 1.

3.2.6. Data Preparation & Analysis

Suggestion adherence was calculated from both the stressfulness and ranking score of the chosen memory. These two scores were found to be positively correlated $r_s = 0.62$, $n = 46$, $p<0.001$. To get a single score for suggestion adherence z scores were calculated. The stressfulness scores were transformed into z-scores

per participant to diminish the inter-person variability. The ranking of the chosen memory was also converted into a z-score, but taken over all participants. The mean of these z-scores formed the *suggestion adherence* measure.

To create a single *task adherence* index, z scores were calculated for the number of diary items added, the number of specific emotion-word items and the total number of words written in the diary over all participants. These three scores were averaged to form a single index score.

Subjective adherence was measured on a continuous scale with four questions. These four questions measured the different aspects of adherence to the virtual agent. All lines were measured to determine where the cross stood and transformed to a value between -1 and 1. The *subjective adherence* was the mean of these scores.

For the questionnaire on attitude towards the virtual agent and attitude towards the psychoeducation, the scores for the negatively phrased questions were inverted. Both questionnaires were tested with Cronbach α, resulting in good internal validity scores of α = 0.89 for the agent attitude questionnaire and α = 0.83 for the psychoeducation attitude questionnaire. The average scores for both questionnaires were used.

SPSS version 22 was used for data management and analysis. A Kolmogorov-Smirnov test for normality showed that suggestion adherence and the function usefulness measures deviated significantly (p>.05) from a normal distribution. For this reason bootstrapping (1000 samples) was applied to all tests including these measures. For the attitude towards the agent and attitude to the psychoeducation, a one-sample t-test was performed to test for differences from the middle value of 4. For the usefulness measures and subjective adherence, a one-sample t-test was performed to test for differences from the middle value of 0. To test for differences between conditions, independent-sample t-tests were performed for all measures. A correlation was done for the outcome measures, secondary measures and psychoeducation attitude to examine related measures. An additional ANCOVA test was performed to study the effect of mode of presentation on adherence with the mediators as covariates, treating them as extraneous variables. The original dataset can be found in [64].

3.3. Results

Table 1 shows the descriptive statistics for all measures collected. One-sample t-tests indicate that agent attitude, psychoeducation attitude, usefulness of the diary, and usefulness of the psychoeducation were all significantly higher than the neutral point on the test-scale. For subjective adherence, usefulness of the virtual agent, and usefulness of the virtual agent speaking, no significant difference from the neutral position was found. The values of usefulness of the agent and usefulness of the agent speaking were found to be highly correlated r_s = 0.89, n = 46, p<0.001. A closer inspection of the data showed a u-shaped histogram for both questions, indicating that two subgroups exist (Figure 3.3). One group that considered the agent and voice a useful addition, another group that did not. For both suggestion adherence and task adherence, the original measures are also included in Table

Table 3.1: Descriptive statistics & one-sample t-test with bootstrapping results

					One-sample t-test		
						95% CI	
Measure	***Type***	***Mean***	***SD***	***Neutral point***	***sig***	***Lower***	***Higher***
Agent attitude	*7 pt. Likert*	4.38	0.75	4	0.002	0.17	0.61
Psychoeducation attitude	*7 pt. Likert*	4.96	0.63	4	0.001	0.79	1.14
Function use Diary	*Analog*	0.42	0.37	0	<0.001	0.32	0.53
Function use Psychoeducation	*Analog*	0.33	0.47	0	<0.001	0.20	0.45
Function use agent	*Analog*	0.05	0.60	0	0.586	-0.14	0.21
Function use speak.	*Analog*	0.06	0.62	0	0.513	-0.13	0.22
Subjective adherence	*Analog*	0.02	0.36	0	0.756	-0.08	0.12
Suggestion adherence	*Behavior*	0.50	0.74				
Chosen score	*Behavior*	81.17	16.60				
Chosen rank	*Behavior*	4.54	0.81				
Task adherence	*Behavior*	0.00	0.65				
Nr. items made	*Behavior*	13.91	5.50				
Nr. emo-items made	*Behavior*	4.61	3.24				
Wordcount	*Behavior*	419.74	340.03				
Recollection	*Behavior*	3.65	2.14				

3.1.

Independent t-tests were done to check for differences between conditions for the primary, secondary and exploratory measures. The full results are shown in Table 3.2. A trend was found for agent attitude, which was more positive in the verbal condition. A trend was also found for perceived use of the diary, which was more positive in the text condition. However, no significant differences were found. This means that no effect of mode of presentation was found on agent attitude, recollection or adherence. As these are all premises for hypothesis 2, a mediation analysis was unnecessary to conclude that our data did not support hypothesis 2.

Although no effect of mode of presentation on adherence was found, other aspects might still have an effect. To explore the relationships between the different measures, a correlation was performed on all primary and mediating measures, as well as on the attitude towards the psychoeducation. Results show that the suggestion adherence was significantly positively correlated with subjective adherence. Subjective and suggestion adherence were not, however, correlated with task adherence. This means that the negativity of the chosen memory was not related to how well that chosen memory was described. Agent attitude and psychoeducation attitude were also positively correlated. A correlation between recollection and adherence was found for task adherence. Similarly, agent attitude was significantly correlated with task adherence. Full statistics for the correlation analysis can be found in Table 3.3.

Given the correlations between recollection and task adherence, and between

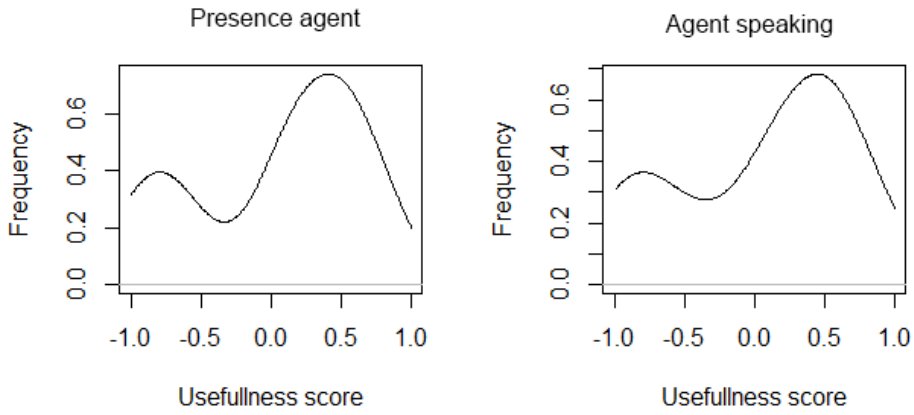

Figure 3.3: Frequency Histograms of the perceived usefulness of the virtual agent and the fact that it spoke, as quizzed with the statement *The virtual agent is a useful addition to this exercise* and *The fact that the virtual agent spoke was a useful addition to this exercise*. -1 corresponds to *don't agree at all* while 1 corresponds to *fully agree*.

Table 3.2: Descriptive statistics & independent-sample t-test with bootstrapping checking for differences between conditions.

	Mean			***95% CI***	
Measure	***Verbal***	***Textual***	***sig***	***Lower***	***Upper***
Subjective adherence	0.01	0.03	0.86	-0.23	0.20
Suggestion adherence	0.48	0.51	0.90	-0.46	0.40
Task adherence	-0.11	0.11	0.24	-0.59	0.13
Agent attitude	4.58	4.18	0.08	-0.02	0.81
Recollection	3.60	3.70	0.66	-1.46	1.11
Psychoed. attitude	4.98	4.93	0.19	-0.39	0.39
Function use diary	0.52	0.32	0.06	-0.003	0.39
Function use psychoeducation	0.34	0.32	0.90	-0.26	0.29
Function use agent	0.06	0.04	0.90	-0.31	0.39
Function use speaking agent	0.15	-0.03	0.36	-0.21	0.52

Table 3.3: Correlations between measures.

Measure	*Subjective adherence*	*Suggestion adherence*	*Task adherence*	*Agent attitude*	*Recollection*	*Psychoed. attitude*
Subjective adherence	1	.46**	.26	.11	.05	.25
Suggestion adherence	.46**	1	.18	.17	.07	.14
Task adherence	.26	.18	1	.31*	.39**	.22
Agent attitude	.11	.17	.31*	1	-.04	.64**
Recollection	.05	.07	.39**	-.04	1	.13
Psychoed. attitude	.25	.14	.22	.64**	.13	1

* the 95% Confidence Interval does not include 0

** the 99% Confidence Interval does not include 0

agent attitude and task adherence, these were considered extraneous variables affecting task adherence. Because presentation mode did not influence either recollection nor agent attitude, an ANCOVA with covariates recollection and attitude to the agent was performed on task adherence. The analysis found a significant effect for presentation mode on task adherence ($F(1.42)=4.2$, $p=0.047$, $\eta^2 = 0.91$), the adherence being better when the information was presented as text (corrected M= 0.17) than when offered verbally by the virtual agent (corrected M=-0.17).

3.4. Discussion & Conclusion

Overall participants held a positive attitude towards the psychoeducation and the virtual agent, as shown by the questions about their usefulness. However, room for improvement exists for both aspects, with average scores being closer to the neutral point than to the maximum. Another finding was the subjective usefulness of several functions during the given task. Firstly, both the digital diary and psychoeducation were considered useful additions to the task. More interestingly, there was no significant consensus on whether the virtual agent and the fact that it spoke were useful. The results indicated a divide between people finding the agent either useful or not useful, with few people in the middle. This indicates that a virtual agent might not be a suitable addition for everybody. Therefore inter-personal differences might have to be considered in the decision to add a virtual agent or not. However, as preference did not correlate with any of the adherence measures or recollection of the psychoeducation, the behavioural consequences of a preference for the agent are less clear. One option would be to let people themselves decide if they wish to have a virtual agent or plain text when working with an autonomous e-mental health application.

Considering the two hypotheses formulated, a relationship was found between agent attitude and task adherence, as well as between recollection and task adherence. After controlling for these variables, presentation mode had an effect on task adherence. Textual presentation resulted in better adherence than verbal

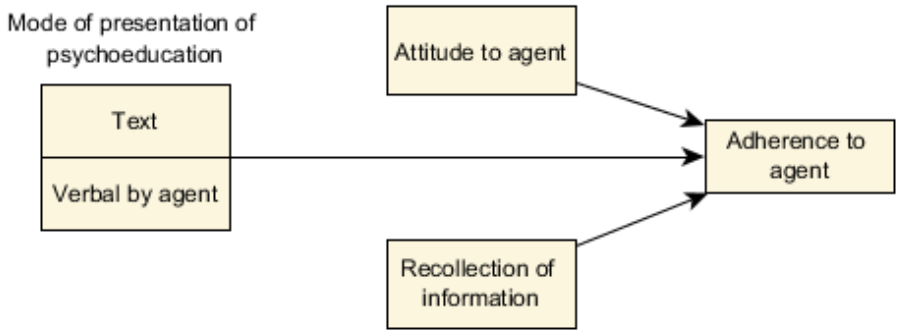

Figure 3.4: Model showing the result that presentation mode effects adherence to the agent, and attitude to the agent and recollection are extraneous variables.

presentation, confirming the first hypothesis. However, no such effect was found for the other two outcome measures: suggestion adherence and subjective adherence. This suggests that the condition had no effect on how difficult the chosen task was, but only on how well the eventual task was executed. More work is necessary to understand the relationship between different types of adherence fully. No correlation was found between suggestion adherence and task adherence. This indicates that these might have to be viewed as formative indicators for general adherence instead of reflexive indicators, e.g. that they are factors which together explain general adherence, even though they are not correlated [65]. The relationship between different types of adherence and treatment outcome also needs to be studied in more detail to know exactly what type of adherence has what effect.

No effect was found for presentation mode on either agent attitude or recollection. In other words, the difference found for presentation mode in adherence to the agent could not be explained by changes the mode had caused in people attitude towards the agent or in their ability better recollect the psychoeducation. This means that hypothesis 2 cannot be confirmed. Given the results, we can modify our initial model as shown in Figure 3.4. Both attitudes towards the virtual agent and recollection of the psychoeducation effect task adherence and are extraneous variables and no longer included as explaining or mediating factors for presentation mode effect on people's adherence to the task set by the agent.

The relationship between agent attitude and adherence corresponds with the prevailing notion that therapist alliance is related to adherence [33, 34]. However, despite the relationship between displaying expertise and being credible [28, 40], this study found no support for the theory that agent attitude can be influenced by presentation mode of psychoeducation. One reason for this is that expertise needs to be displayed continuously over time to have an effect on trust. Further work is necessary to study exactly how agent attitude could be maximised. First steps in this process might be to study which aspects of therapeutic alliance in human therapists could be translated to a virtual agent.

Our findings also show that recollection of the psychoeducation is positively related to task adherence. This matches with the main goal of psychoeducation, namely increasing adherence through explanation and information [21]. No relationship was found between presentation mode of psychoeducation and recollection. This might be explained by different processes being in place. While traditionally, text is remembered better than audio [22], for agents a preference of voice over text has been reported [26]. More research is necessary to understand the relationship between presentation mode of information by a virtual agent and recollection. The presentation of longer texts by a virtual agent, in particular, has not received much attention.

The finding that textual presentation of psychoeducation is preferable above verbal presentation cannot be explained by any of the measures taken in this experiment. None of the measures collected correlated with both presentation mode and task adherence. This means that the mechanism that explains why presentation mode changes people adherence to agents should be sought outside the factors studied. One possible explanation could be that written psychoeducation is more persuasive in getting people to execute the task thoroughly. Although many factors play a role, the written modality has been found to have a greater persuasive effect than audio or video for complicated persuasive messages [66]. This would also fit with the finding that presentation mode only had an effect after controlling for attitude towards the agent, which has been documented to influence the persuasiveness of messages [67]. Further research is necessary to test this hypothesis and to study the impact of other factors.

Although virtual agents in AEMH systems provide an opportunity to deliver care to patients which are not always reached by regular care, they also come with their own risks and responsibilities. Firstly, an AEMH system should be able to identify situations in which a patient is at risk, and ensure patient safety. If it is not possible to put this in place fully, patients should be screened by human caregivers before use with an AEMH, or monitored during use. Another concern unique to virtual agents is that people might develop an emotional bond with the agent and develop certain expectations. It is always important to be very clear about the abilities and limitations of a virtual agent and to manage the expectations of users. Finally, an AEMH system should be conscious of patient privacy, especially in systems collecting sensitive information [68].

Aside from these general concerns, we can identify some limitations to this particular study design which are necessary to appreciate the findings. Firstly, it only considers a computerised voice for the virtual agent. Another option is a recorded human voice, which could have a different effect. Moreover, as text-to-speech systems improve, results might change. The exact effect of a more realistic voice is yet unknown. Similarly, we did not compare the virtual agent to a video of a human presenting the psychoeducation. Research has shown these mediums to be different [69], but the effect of this difference on adherence to psychoeducation is still unknown to the best of our knowledge. Another limitation is the short duration of this experiment. Although it gives a good idea of a short exercise, many e-mental health applications are for more long-term use. This gives

a different dynamic in the attitude towards the virtual agent and recollection of information given in the beginning [70]. This study was also not conducted in a double-blinded setting, although the experimenter was only present shortly, she was aware of the experimental condition during the initial program instructions. Finally, we found that the virtual agent was not necessarily considered a useful addition to the exercise, despite the general positive agent attitude. For this finding, it should be noted that the virtual agent had a very basic role in this experiment. It guided the participants through the experiment by telling them where to click and what to do, and in half of the cases told the psychoeducation. It had no interactive or personalising features such as in Gilani et al. [71], and did not assist in the actual recollection of the memory such as in Tielman et al. [72]. A more interactive and personalised virtual agent might also be perceived as more useful.

Psychoeducation is a tool used in a wide range of interventions for mental health problems to increase efficacy. Given the societal burden of mental health problems that lead to both direct and indirect cost [73], as well as loss of quality of life [12], good and effective mental health applications are very valuable. As more virtual agents are deployed in the field of mental-health, the need for good psychoeducation procedures therefore grows. This is relevant particularly for those virtual agents that have an active coaching, guiding or conversational role, taking on some of the tasks of a regular therapist, such as psychoeducation. Interventions with such virtual agents have been developed for a range of mental-health problems [74] such as insomnia [75], depression [76], social phobia [77] and PTSD [78]. For all of these, psychoeducation that maximises adherence is an interesting component. However, the question of how to present information is relevant even for applications outside of the scope of mental health. General health-care information has similarities to psychoeducation. Patients often need to know what medication to take, what exercises to do, or even need to change attitudes. Virtual agents are being developed for these uses [79, 80] and as with mental health care adherence to the agent is an important aspect [35]. Moreover, the results of this study can be relevant even for fields outside of health-care such as education or marketing. With agents teaching, explaining or convincing users, adherence to information is essential.

The main contribution of this study lies in identifying several aspects of psychoeducation that influence adherence. Firstly, how well the information is recollected has an effect. Secondly, the attitude towards the virtual agent, which asks people to perform the task, affects adherence. Thirdly, if the psychoeducation is presented in text it results in a better adherence than if a virtual agent, with a text-to-speech voice, offers it verbally. Additionally, from this study, we can infer that the perceived usefulness of a virtual agent differs considerably between people dividing them into a pro and a con camp, but no indications of any effect of preference on adherence is found. These findings are a first step in determining how exactly psychoeducation in mental health applications with virtual agents influences adherence.

References

[1] M. Tielman, M. A. Neerincx, M. van Meggelen, I. Franken, and W. Brinkman, *How should a virtual agent present psychoeducation? influence of verbal and textual presentation on adherence,* Technology & Healthcare , 1 (2017).

[2] M. Olfson, R. Mojtabai, N. Sampson, I. Hwang, B. Druss, P. Wang, K. Wells, H. Pincus, and R. Kessler, *Dropout from outpatient mental health care in the united states.* Psychiatric Services **60**, 898 (2009).

[3] P. S. Wang, M. Lane, M. Olfson, H. A. Pincus, K. B. Wells, and R. C. Kessler, *Twelve-month use of mental health services in the united states,* Twelve-Month Use of Mental Health Services in the United States **62**, 629 (2005).

[4] C. Lyons, P. Hopley, and J. Horrocks, *A decade of stigma and discrimination in mental health:plus ça change, plus c'est la même chose (the morethings change, the more they stay the same),* Journal of Psychiatric and Mental Health Nursing **16**, 501 (2009).

[5] J. A. Cartreine, D. K. Ahern, and S. E. Locke, *A roadmap to computer-based psychotherapy in the united states,* Harv Rev Psychiatry **18**, 80 (2011).

[6] T. Donker, M. Blankers, E. Hedman, B. Ljótsson, K. Petrie, and H. Christensen, *Economic evaluations of internet interventions for mental health: a systematic review,* Psychological Medicine , 1 (2015).

[7] M. Zhang, E. Cheow, C. S. Ho, B. Y. Ng, R. Ho, and C. C. S. Cheok, *Application of low-cost methodologies for mobile phone app development,* JMIR MHealth and UHealth **2**, e55 (2014).

[8] N. Yee, J. N. Bailenson, and K. Rickertsen, *A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces,* in *CHI 2007* (2007).

[9] O. A. Blanson-Henkemans, C. A. van der Mast, P. J. van der Boog, M. A. Neerincx, J. Lindenberg, and B. J. Zwetsloot-Schonk, *An online lifestyle diary with a persuasive computer assistant providing feedback on self-management,* Technology and Health Care **17**, 253 (2009).

[10] R. Looije, M. A. Neerincx, and F. Cnossen, *Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors,* International Journal of Human-Computer Studies **68**, 386 (2010).

[11] A. D. Andrade, R. Anam, C. Karanam, P. Downey, and J. G. Ruiz, *An overactive bladder online self-management program with embedded avatars: A randomized controlled trial of efficacy,* Urology **85** (2015).

[12] WHO, *Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level,* (2012).

[13] S. A. Shumaker, J. K. Ockene, and K. A. Riekert, eds., *The handbook of Health Behavior Change* (Springer, 2009).

[14] L. Donkin, H. Christensen, S. Naismith, B. Neal, I. Hickie, and N. Glozier, *A systematic review of the impact of adherence on the effectiveness of e-therapies,* Journal of Medical Internet Research **13** (2011).

[15] R. Gould and G. Clum, *A meta-analysis of self-help treatment approaches,* Clinical Psychology Review **13**, 169 (1993).

[16] K. Coulthard, D. Patel, C. Brizzolara, R. Morriss, and S. Watson, *A feasibility study of expert patient and community mental health team led bipolar psychoeducation groups: implementing an evidence based practice,* BMC Psychiatry **13** (2013).

[17] M. W. B. Zhang and R. C. M. Ho, *Tapping onto the potential of smartphone applications for psycho-education and early intervention in addictions,* Frontiers in Psychiatry **7** (2016).

[18] E. Taylor-Rodgers and P. J. Batterham, *Evaluation of an online psychoeducation intervention to promote mental health help seeking attitudes and intentions among young adults: Randomised controlled trial,* Journal of Affective Disorders **168**, 65 (2014).

[19] S. Farooq and F. Naeem, *Tackling nonadherence in psychiatric disorders: current opinion,* Neuropsychiatric Disease and Treatment **10**, 1069 (2014).

[20] F. Eker and S. Harkin, *Effectiveness of six-week psychoeducation program on adherence of patients with bipolar affective disorder,* Journal of Affective Disorders **138**, 409 (2012).

[21] U. D. of Health, H. Services, N. I. of Health, and N. C. Institute, *Theory at a glance - a guide for health promotion practice,* (2005).

[22] Webb and Wallon, *Comprehension by reading versus hearing,* The Journal of Applied Psychology **40**, 237 (1956).

[23] D. L. Rubin, T. Hafer, and K. Arata, *Reading and listening to oral-based versus literate-based discourse,* Communication Education **49**, 121 (2000).

[24] E. Sanchez and H. Garcia-Rodicio, *The use of modality in the design of verbal aids in computer-based learning environments,* Interacting with Computers **20**, 545 (2008).

[25] J. Fox and J. G. Ahn, *Recommendations for designing maximally effective and persuasive health agents,* in *Int. Conf. on Intelligent Virtual Agent* (2014).

[26] Walter, Ortback, and Niehaves, *Designing electronic feedback - analyzing the effects of social presence on perceived feedback usefulness,* (2015).

[27] Qui and Benbasat, *Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems,* (2009).

[28] B. Fogg, J. Marshall, O. Laraki, A. Osipovich, C. Varma, N. Fang, J. Paul, A. Rangnekar, J. Shon, P. Swani, and M. Treinen, *What makes web sites credible? a report on a large quantitative study,* in *CHI 2001* (2001).

[29] K. Fiscella, S. Meldrum, P. Franks, C. Shields, P. Duberstein, S. McDaniel, and R. Epstein, *Patient trust: is it related to patient-centered behavior of primary care physicians?* Medical Care (2004).

[30] A. S. B. Bohnert, K. Zivin, D. E. Welsh, and A. M. Kilbourne, *Ratings of patient-provider communication among veterans:serious mental illnesses, substance use disorders, and themoderating role of trust,* Health Communication **26**, 267 (2011).

[31] D. Safran, D. Taira, W. Rogers, M. Kosinski, J. Ware, and A. Tarlov, *Linking primary care performance to outcomes of care.* Journal of Family Practice (1998).

[32] J. C. Muran and J. P. Barber, eds., *The Therapeutic Alliance: An Evidence-Based Guide to Practice* (Guilford, 2010).

[33] P. Czobor, R. A. V. Dorn, L. Citrome, R. S. Kahn, W. W. Fleischhacker, and J. Volavka, *Treatment adherence in schizophrenia: A patient-level meta-analysis of combined catie and eufest studies,* European Neuropsychopharmacology **25**, 1158 (2015).

[34] E. Leclerc, C. Noto, R. A. Bressan, and E. Brietzke, *Determinants of adherence to treatment in first-episode psychosis: a comprehensive review,* Revista Brasileira de Psiquiatria **37**, 168 (2015).

[35] S. Baker, D. Richards, and P. Caldwell, *Relational agents to promote ehealth advice adherence,* in *PRICAI* (2014).

[36] T. Bickmore and A. Gruber, *Relational agents in clinical psychiatry,* Harv Rev Psychiatry **18**, 119 (2010).

[37] K. Laska, T. Smith, A. Wislocki, T. Minami, and B. Wampold, *Uniformity of evidence-based treatments in practice? therapist effects in the delivery of cognitive processing therapy for ptsd,* Journal of Counseling Psychology **60**, 31 (2013).

[38] T. Anderson, B. Ogles, C. Patterson, M. Lambert, and D. Vermeersch, *Therapist effects: Facilitative interpersonal skills as a predictor of therapist success,* Journal of Clinical Psychology (2009).

[39] S. Priebe, C. Palumbo, S. Ahmed, N. Strappelli, J. J. Gavrilovic, and S. Brenner, *How psychiatrists should introduce themselves in the first consultation: an experimental study,* The British Journal of Psychiatry **202**, 459 (2013).

[40] P. Kulms and S. Kopp, *The effect of embodiment and competence on trust and cooperation in human-agent interaction,* in *Int. Conf. on Intelligent Virtual Agents* (2016).

[41] R. Looije, M. A. Neerincx, and K. V. Hindriks, *Specifying and testing the design rationale of social robots for behavior change in children,* Cognitive Systems Research (2016).

[42] M. A. Jacobs and A. L. Graham, *Iterative development and evaluation methods of mhealth behavior change interventions,* Current Opinion in Psychology **9**, 33 (2016).

[43] W.-P. Brinkman, *Current and future research directions in mental health computing,* in *Int. conf. on Computation and Communication advancement* (2013).

[44] T. M. Baker, D. H. Gustafson, and D. Shah, *How can research keep up with ehealth? ten strategies for increasing the timeliness and usefulness of ehealth research,* Journal of Medical Internet Research **16** (2014).

[45] H.-J. Lee, S.-H. Lee, H.-S. Kim, S.-M. Kwon, and M. J. Telch, *A comparison of autogenous/reactive obsessions and worry in a nonclinical population: a test of the continuum hypothesis,* Behaviour Research and Therapy **43**, 999 (2005).

[46] J. Murphy, M. Shevlin, J. Houston, and G. Adamson, *A population based analysis of subclinical psychosis and help-seeking behavior,* Schizophrenia Bulletin **38**, 360 (2012).

[47] P. Watson, S. M. Sawrie, R. L. Greene, and R. Arredondo, *Narcissism and depression: Mmpi-2 evidence for the continuum hypothesis in clinical samples,* Journal of Personality Assessment **79**, 85 (2002).

[48] J. F. McDermut, D. A. Haaga, and L. Kirk, *An evaluation of stress symptoms associated with academic sexual harassment,* Journal of Traumatic Stress **13**, 397 (2000).

[49] J. W. Pennebaker and S. K. Beall, *Confronting a traumatic event: Toward an understanding of inhibition and disease,* Journal of Abnormal Psychology **95**, 274 (1986).

[50] K. A. Baikie and K. Wilhelm, *Emotional and physical health benefits of expressive writing,* Advances in Psychiatric Treatment **11**, 338 (2005).

[51] J. Frattaroli, M. Thomas, and S. Lyubomirsky, *Opening up in the classroom: Effects of expressive writing on graduate school entrance exam performance,* Emotion **11**, 691 (2011).

[52] A. H. Harris, *Does expressive writing reduce health care utilization? a meta-analysis of randomized trials,* Journal of Consulting and Cli **74**, 243 (2006).

[53] D. Morisano, J. B. Hirsch, J. B. P. R. O. Pihl, and B. M. Shore, *Setting, elaborating and reflecting on personal goals improves academic performance,* Journal of Applied Psy **95**, 255 (2010).

[54] C. M. Burton and L. A. King, *Eeffect of (very) brief writing on health: The two-minute miracle,* British Journal of Health Psychology **00**, 1 (2007).

[55] C. K. Chung and J. W. Pennebaker, *Variations in the spacing of expressive writing,* British Journal of Health Psychology **13**, 15 (2008).

[56] M. van den Steen, W.-P. Brinkman, E. Vermetten, and M. Neerincx, *Design and usability evaluation of a multi-modal memory restructuring system for the treatment of combat-related ptsd,* in *ECCE 2010 Workshop on Cognitive Engineering for Technology in Mental Health Care and Rehabilitation* (2010).

[57] E. Lee, C. Nass, and S. Brave, *Can computer-generated speech have gender? an experimental test of gender stereotype,* in *CHI* (2000).

[58] A. Baylor and E. Plant, *Pedagogical agents as social models for engineering: The influence of agent appearance on female choice,* in *Conference on Artificial Intelligence in Education* (2005).

[59] A. Baylor, *Promoting motivation with virtual agents and avatars: role of visual presence and appearance,* Philosophical Transactions of the Royal Society B **364**, 3559 (2009).

[60] R. Guadagno, J. Blascovich, J. Balienson, and C. McCall, *Virtual humans and persuasion: The effects of agency and behavioral realism,* Media Psychology **10**, 1 (2007).

[61] M. Tielman, M. Neerincx, M. van Meggelen, I. Franken, and W. Brinkman, *How should a virtual agent present psychoeducation? - dataset part i - psychoeducation,* (2017), doi: 10.4121/uuid:8231a591-6a51-40f5-bd3c-54435a4fc61c.

[62] S. H. Hemenover, *The good, the bad, and the healthy: Impacts of emotional disclosure of trauma on resilient self-concept and psychological distress,* Personality and Social Psychology Bulletin **2003**, 1236 (2003).

[63] M. Tielman, M. Neerincx, M. van Meggelen, I. Franken, and W. Brinkman, *How should a virtual agent present psychoeducation? - dataset part ii - questionnaires,* (2017), doi: 10.4121/uuid:30db1be5-5070-4c7f-87af-210e09bda1c4.

[64] M. Tielman, M. Neerincx, M. van Meggelen, I. Franken, and W. Brinkman, *How should a virtual agent present psychoeducation? - dataset part iii - dataset,* (2017), doi: 10.4121/uuid:828177fa-3a6d-4ecf-83d0-90d84414f7d7.

[65] N. J. Blunch, *Introduction to Structural Equation Modeling* (Sage, 2013).

[66] S. Chaiken, *Communication modality as a determinant of message persuasiveness and message comprehensibility,* Journal of Personality and Social Psychology **34**, 605 (1976).

[67] S. Chaiken, *Social influence,* in *The Ontario Symposium* (1987).

[68] D. D. Luxton, *Recommendations for the ethical use and design of artificial intelligent care providers,* Artificial Intelligence in Medicine **62**, 1 (2014).

[69] S.-H. Kang, A. W. Feng, M. Seymour, and A. Shapiro, *Study comparing video-based characters and 3d-based characters on mobile devices for chat,* in *MIG '16 Proceedings of the 9th International Conference on Motion in Games* (2016).

[70] T. Bickmore and R. Picard, *Establishing and maintaining long-term human-computer relationships,* in *ACM Transactions on Computer Human Interaction (ToCHI)* (2005).

[71] S. Gilani, K. Sheetz, G. Lucas, and D. Traum, *What kind of stories should a virtual human swap?* in *Int. Conf. Intelligent Virtual Agents* (2016).

[72] M. L. Tielman, M. van Meggelen, M. A. Neerincx, and W.-P. Brinkman, *An ontology-based question system for a virtual coach assisting in trauma recollection,* in *Int. Conf. on Intelligent Virtual Agents* (2015).

[73] R. Ho, K. Mak, A. Chua, C. Ho, and A. Mak, *The effect of severity of depressive disorder on economic burden in a university hospital in singapore,* Expert Review of Pharmacoeconomics & Outcomes Research **13**, 549 (2013).

[74] S. Provoost, H. M. Lau, J. Ruwaard, and H. Riper, *Embodied conversational agents in clinical psychology: A scoping review,* JOURNAL OF MEDICAL INTERNET RESEARCH **19**, e151 (2017).

[75] R. J. Beun, W. P. Brinkman, S. Fitrianie, F. Griffioen-Both, C. Horsch, J. Lancee, and S. Spruit, *Improving adherence in automated e-coaching,* Persuasive Technology **9638**, 276 (2016).

[76] A. Bresó, J. Martínez-Miranda, and J. M. García-Gómez, *Leveraging adaptive sessions based on therapeutic empathy through a virtual agent,* in *ICAART - Doctoral Consortium* (2014).

[77] D. Hartanto, W. P. Brinkman, I. L. Kampmann, N. Morina, P. G. Emmelkamp, and M. A. Neerincx, *Home-based virtual reality exposure therapy with virtual health agent support.* Pervasive Computing Paradigms for Mental Health **204**, 85 (2015).

[78] A. A. Rizzo, B. Lange, J. G. Buckwalter, E. Forbell, J. Kim, K. Sagae, J. Williams, B. O. Rothbaum, J. Difede, G. Reger, T. Parsons, and P. Kenny, *An intelligent virtual human system for providing healthcare information and support,* Medicine meets Virtual Reality (2011).

[79] R. Edwards, T. Bickmore, L. Jenkins, M. Foley, and J. Manjourides, *Use of an interactive computer agent to support breastfeeding,* Maternal and Child Health Journal **17**, 1961 (2013).

[80] J. Ren, D. Schulman, J. Brian, and T. Bickmore, *Supporting longitudinal change in many health behaviors,* in *CHI* (2014).

4

An Ontology-based Question System for a Virtual Coach Assisting in Trauma Recollection

*Internet-based guided self-therapy systems provide a novel method for Post-Traumatic Stress Disorder patients to follow therapy at home with the assistance of a virtual coach. One of the main challenges for such a coach is assisting patients with recollecting their traumatic memories, a vital part of therapy. In this paper, an ontology-based question system capable of posing appropriate and personalized questions is presented. This method was tested in an experiment with non-patients (*n *= 24), where it was compared with a non-ontology-based system. Results show that people take more time answering questions with the ontology-based system and use more words describing properties, such as adjectives. This indicates that the ontology-based system facilitates more thoughtful and detailed memory-recollection.*

This chapter has been published in International conference on Intelligent Virtual Agents. (2015) [1].

4.1. Introduction

Post-Traumatic Stress Disorder (PTSD) is a mental disorder caused by one or more traumatic experiences [2]. Several treatments for PTSD are available, with the most common element being exposure, which is the process of exposing patients to their traumatic memories [3]. One problem for PTSD treatment is that there is often a barrier to talk about problems and a stigma on seeking help from the mental health-care system. One new method for exposure treatment for PTSD which addresses this issue is a self-therapy system with a virtual coach [4, 5]. With such a system, patients follow their therapy at home behind their computer with the assistance of a virtual coach and a human therapist is only remotely involved. One of the main challenges for a virtual coach in such a self-therapy system is providing the assistance PTSD patients need for exposure sessions. PTSD patients often have fragmented memories of their trauma and are very reluctant to recall them, requiring detailed questions to stimulate memory retrieval. To provide this assistance, a virtual coach should therefore be able to communicate with the patient and have an understanding of the topic, namely the traumatic experience. Further, because every trauma is unique, it needs some understanding of the patient to personalize the communication. It is also important to get it right, because the topic is highly sensitive and difficult for patients to re-experience. To solve this problem, we propose an ontology-based conversational system with minimal natural-language processing with which a virtual coach can pose relevant and personalized questions to assist individuals with memory recollection.

For our system to be effective, we envision that a virtual coach is able to assist the human patient in a similar manner as a human therapist. Virtual avatars have been around for some time, and are being used more and more in health-care applications. These applications rely on the anthropomorphism of these characters. Even though people know they're interacting with a digital agent, they will still behave as if the avatar were human [6]. This means virtual characters can follow the same methods for motivation and behavioral change as human therapists and coaches would. Blanson-Henkemans et al. [7] showed that a virtual coach with emotional facial expressions can motivate and support people to live a healthier life, and Bickmore et al. [8] showed the effectiveness of an application with virtual character to elicit healthier behavior in older adults. For mental health-care, virtual characters have been employed for complex user groups such as people suffering from depression [9]. For PTSD, Rizzo et al. [10] developed the SimCoach, a virtual coach guiding veterans who potentially have PTSD towards treatment. Even though virtual avatars have some limitations compared to human coaches, such as the lack of full language abilities, they also have their own advantages, like full-time availability. Moreover, the anonymous nature of a virtual character can increase self-disclosure by patients, which is crucial in many health-care applications [11].

Despite a lack of full natural language capabilities, virtual agents which can communicate with humans in a meaningful way have been developed. One example is by Qu et al. [12], who showed that the answers of the human conversational partner can be steered through priming with images or video. Similarly, Ter Heijden and Brinkman [13] proposed a system in which the agent steered the conversation

and used limited speech recognition, providing a low-level method to understanding responses. The draw-backs of these systems is that they do not interpret input from users or store patient information in a meaningful way, while both are required for a more meaningful conversation [14]. One study which addresses this was done by Schulman et al. [15], who developed a conversational agent using Motivational Interviewing (MI, [16]) for health-behavior change. Their method relied on multiple-choice and free text input based on which specific dialogue acts for MI were selected. Also considering MI, Friedrichs et al. [17] developed a system which repeats back utterances of the user and employs multiple-choice input to personalize the content. Both these systems have been evaluated with users, showing that even without natural language understanding a system can hold a personalized and meaningful dialogue with a user and elicit behavior change.

Ontologies are often used in dialogue systems to add additional meaning and world knowledge. For example, Bickmore et al. [18] developed an ontology-based counseling framework which described a patient's mental states and therapist's actions affecting these states. The use of ontologies was also shown by Beveridge and Fox [19], who employed an ontology to formulate the high-level representations in their dialogue system. These systems show that an ontology is a helpful tool which can be used to go from user input to a response which serves a certain goal. Another possible use of ontologies is to add meaning to the speech of the agent itself, something often used for chatbots. For example, Al-Zubaide et al. [20] used ontologies to form a relational database which could be used to drive chat interactions. The advantage of this system was that the ontology behind it could be changed for different domains without any other adaptations being necessary. Similarly, Augello et al. [21] employed an ontology to add world knowledge and flexibility to a chatbot. Both these studies show that ontologies can be used in dialogue systems to add knowledge and meaning. In this paper, we propose to use ontologies in such a way that they assist in interpreting the user input, to steer the conversation towards its goal, and giving meaning to the dialogue of the virtual agent.

4.2. Ontology

Several definitions of the term *ontology* have been proposed. In this paper we use the working definition as formulated by Noy & McGuiness [22], where an ontology is *"a formal explicit description of concepts in a domain of discourse (classes) properties of each concept describing various features and attributes of the concepts (slots), and restrictions on properties"*. When these classes and properties are instantiated, a knowledge base is formed. Ontologies give meaning to words, making them to some extent understandable to machines. They also enable a structured knowledge base of a domain. For trauma, this means that the ontology allows a system to have an understanding of traumatic events. Through this understanding, which information is still unknown and should therefore be asked for can be determined. For instance, whenever an abuse victim never mentions a perpetrator, something is missing from the discourse and the system should ask an appropriate question.

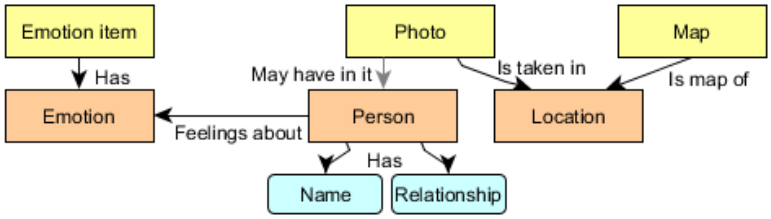

Figure 4.1: Section of an ontology based on diary items in the 3MR system.

The first thing to consider when designing an ontology for a question system is the modularity of the expected answers. For this paper, we consider the possibilities of the self-therapy Multi-Modal Memory Restructuring system (3MR) [4]. This system allows users to employ different types of media, namely text, images, music, video, google maps, websites, and emotion labels. For a user, all these modularities are available to expose them to their memory and a question system can employ these as possible answers. An example would be the question *Where did this happen?*, which might be answered in the form of text, but also through adding a map. Because this is one of the unique characteristics of the 3MR system, the ontology is based around these modularities.

Figure 4.1 shows a section of an ontology, based on the modularities Emotion, Photo and Map. A photo has two links in this example, namely the people on the photo and where it is taken. People have three properties, namely their name, their relationship to the person filling in the diary, and how this person feels about them. This last property is an instance of the Emotion class, and can therefore be represented by an emotion item. The location of a photo is an instance of the Location class, which can be represented by a map item. In this way, every diary item has certain properties, which can again be instances of other classes again having their own properties. Which classes and properties these are is determined by the type of memory one wants to retrieve. For a war veteran, the location of the trauma in a foreign country is very relevant, while for abuse victims this might be the type of room they were in.

4.3. Question System

Together with the ontology, a question dialogue based on this ontology needs to be in place. For this study, the ontology led the design of the natural language questions. Whenever a specific item was entered, such as a photo, the resulting questions could be derived from the ontology. Whenever the ontology was not specific enough to decide on the order of the questions (for instance which type of item to start with), a basic *when, where, who, what*, paradigm was employed. We furthermore decided to ask for two types of memories, those relating to a general time period and those relating to a specific moment in time. For PTSD patients, both types are an important aspect of exposure therapy and with the 3MR system, the therapy follows a gradual exposure paradigm in that people first

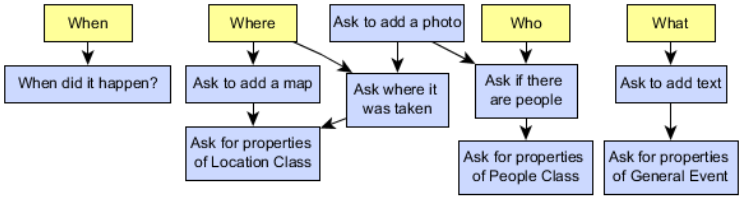

Figure 4.2: Outline conversation general period

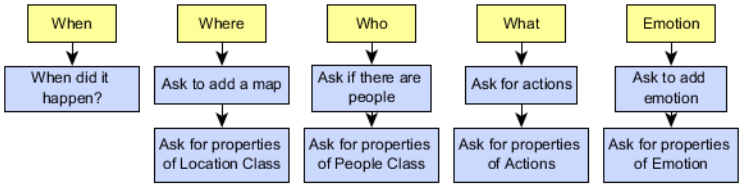

Figure 4.3: Outline conversation specific moment

confront their general memories before working on the trauma itself. These types of memories call for slightly different types of questions and therefore ontologies. In Figure 4.2 and 4.3, an outline is given of both conversations. An example of the difference between those conversations for war veterans would be that in the general conversation the *where* question would be *Where was your mission?*, while in the specific conversation this would be *At which exact spot were you in that moment?*

Although these outlines are based on an ontology, a dialogue without ontology could follow the exact same pattern. Personalization begins when the ontology is filled in and specific questions are asked. This could happen for instance, by knowing if a photo was taken in an inside or outside location. These two location types have different properties and only those properties for the correct type of location would be asked. While a system not based on an ontology could still ask for properties of the location, it could only ask for those applying to all locations and give examples. To illustrate this concept, Figure 4.4 shows a section of an ontology of a location of a holiday memory. Figures 4.5 shows the dialogue based on the ontology in Figure 4.4. The corresponding dialogue not based on the ontology would be the question *Can you describe your location? If you were outside, think of what your surroundings looked like and what the weather was like, and if you were inside, what type of place you were in and what it looked like?* followed by *If google streetview is available for this spot, could you find it and add it to the diary?* In this example it is clear that the non-ontology based conversation consists of fewer, more general and longer questions.

Figure 4.5 shows that the addition of an ontology allows for asking questions which are much more in-depth and specific. Another advantage is that with an

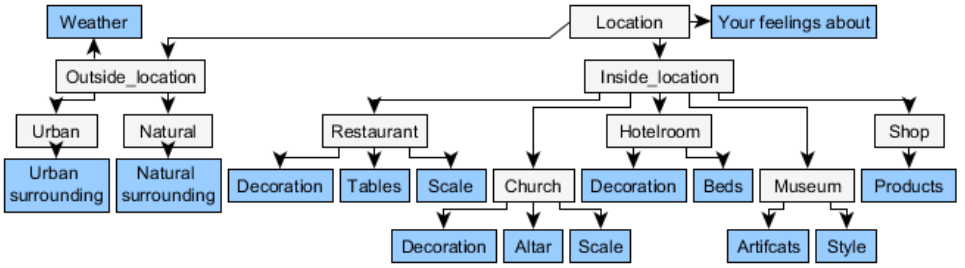

Figure 4.4: Ontology of holiday moment locations. The light boxes are the classes, the darker ones their properties.

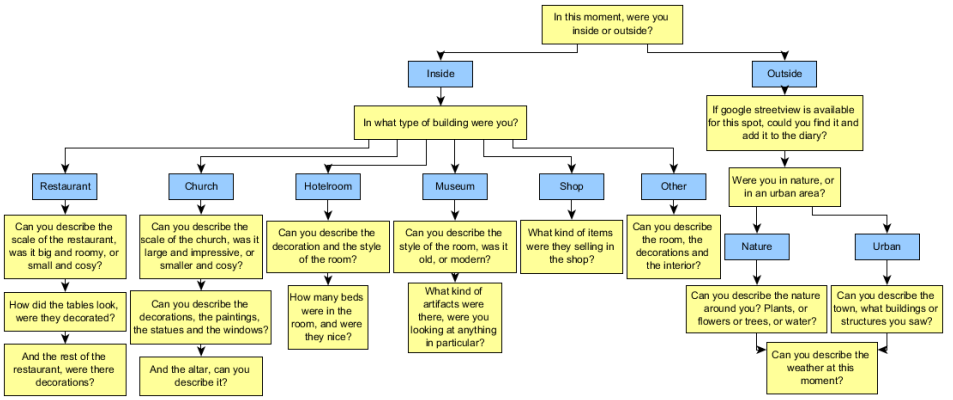

Figure 4.5: Dialogue tree on location of an experience using an ontology. The large boxes are the questions, the small ones are multiple-choice answer options.

ontology, one has a clear overview of all the topics of the questions. If the ontology contains classes and properties for every concept in a certain memory, so if it is complete, we can state that the resulting question dialogue is also complete. Another expected advantage of such an ontology-based question dialogue is that it is more personal. It can be expected, therefore, that such a system also assists people better when recollecting memories than a non-ontology based system. To test this hypothesis, an experiment was set up comparing an ontology-based to a non-ontology based question dialogue.

4.4. Experiment

A within-subject experiment with two conditions was conducted. We wished to know if an ontology-based system would allow people to recollect their memories better. For this purpose, our ontology-based system was compared with a non-ontology-based question system with the same topics and order of questions, but without any of the personalization. The effect this difference had on people's opinions and experiences with the system was tested, as well as the effect it had on the level of detail with which the questions were answered.

4.4.1. Participants

Giving exposure sessions to PTSD patients without providing a full treatment was not considered ethically appropriate. For this reason 24 healthy participants (10 female, mean age 28.4, SD 3.1) were recruited from the University staff and student population. Eight Participants performed the experiment in their native language (English or Dutch), all others in their second language (English). Because the participants did not have traumatic experiences, the memories they had to recollect were holiday memories. This topic was chosen because it was universal for this sample and could be modelled with an ontology quite well. Furthermore, in exposure treatment PTSD patients would also start with a positive memory to get familiar with the system.

4.4.2. Question System

The question system was based on two ontologies, one for a general holiday memory and one focusing on a specific moment within that same holiday. For the non-ontology based system, the questions followed the exact same order and topics as the ontology-based system, such as destination and travelling companions. An example of the ontology and the difference between the systems can be seen in Section 3. Figure 4.5 also shows a number of multiple-choice options the participants could choose from. This use of multiple-choice answers ensures that the ontology-based system could react to answers appropriately. The non-ontology based system did not include any multiple-choice options. Finally, in the ontology-based system it was possible to pose a constraint on the type of answers possible, in this case the length of the answers. For the questions asking for descriptions, the answer needed to be at least six words long. Whenever this was not the case, a follow-up question would ask the participant to tell more.

4.4.3. Wizard of Oz & Procedure

A Wizard of Oz procedure was followed. The full dialogue of questions was written in advance and the procedure was fully specified to avoid any influence from the human wizard on the dialogue. The order of the questions was set and participants could signal they were finished answering through a button. The wizard was in the same room as the participants, but they could not see the wizard controlling the system. All questions appeared on the screen of the participant as typed text. The whole question would appear at once, along with multiple-choice options if applicable.

Prior to the experiment, all participants were asked to bring media (photos, video and music) from four holidays. The two holiday memories which were used in the experiment were randomly chosen from this set. All participants started with a small introductory exercise to get familiar with the system. After this, the agent posing the questions gave a short introduction, in which it explained its function and that it could communicate through text only. After this, two dialogue sessions followed in which participants were asked to describe two holiday memories, once with the ontology-based and once with the non-ontology based question system. The order of the two dialogue sessions was counter-balanced. Each dialogue session consisted of 10 minutes of questions on the general experience and 10 minutes of questions on a specific moment. Participants had the option to take a short break between the two dialogue sessions. The experiment was approved by the University ethics committee.

4.4.4. Measures

Both subjective experience and the amount of detail in participant's answers were studied. The subjective experience was examined in three ways. The first was the emotion experienced when thinking back to the memory. In particular, if the subjective feeling changed when the memory was recollected. Pre and post-measures were taken with the 5-point Self-Assesment Manikin scale (SAM) scale [23] of Arousal and Valence for both memories recollected. The second subjective measure was how well people felt the system helped them in recollecting their memory. The third way was how people experienced the conversation with the system. These two measures were examined with six questions each in a questionnaire answered on a 7-point Likert scale. Examples of these questions are *The questions assisted me well in recollecting my memory* for the memory recollection and *The conversation did not run smoothly* for the conversational experience. This questionnaire was presented directly after each dialogue session. Finally, each participant answered four questions on their overall preference of one system over the other. The first was on which system helped recollect the memory best, the second which system was most pleasant to work with, the third on which asked the best questions to trigger the memory and the fourth on which system they would use again. The objective measure considered was the amount of detail in the answers of the patients. On a general level, the number of words typed and the number of question topics posed were checked. The number of question topics could differ per participant because the dialogue sessions had a fixed time, i.e. some participants

Table 4.1: Annotations and definition

Annotation	Consists of
Objects	All nouns, except those referring to a person (or multiple, such as *people*), and those referring to a period of time (e.g. *day* or *moment*).
People	All nouns referring to people (e.g. *girl, tourists*) and names of people.
Descriptives	Adjectives (including terms as *very* and *three*), as well as words describing aspects of something (e.g. *cold*, in *the room is cold*). Excluding adjectives of feelings. Double adjectives were counted apart (*a very cold room*, and *a large cold room* both counting 2 descriptives)
Feelings	All words referring to feelings (e.g. *excited*), including *looking forward to* and *tired*, as well as all adjectives of feelings (*very excited* counting both words).
Time	All nouns referring to time, such as *month* or *period*.

would answer only questions about location and travel, while others were quicker and would also answer questions on travel companions. To consider the amount of detail present in the texts, all answers were annotated and the number of objects, people, descriptives, feelings and time references were counted. The description of the categories in this annotation can be found in Table 4.1. A second annotator annotated 1235 words to ascertain reliability of the rating. Interrater reliability was assessed with Cohen's Kappa and showed a good agreement between annotators $\kappa = 0.86$, $p<.001$.

4.4.5. Data preparation & Analysis

Two questionnaires were designed specifically for this experiment, namely the one measuring how well people felt the system helped them in memory recollection, and the one measuring how people experienced the conversation with the system. The validity of these questionnaires was tested with Cronbach's alpha, after which one question was removed from the recollection questionnaire and two from the conversation questionnaire to improve internal validity. Internal validity after this was acceptable to good (α 0.72 to 0.81 for the recollection and α 0.54 to 0.63 for the conversation questionnaire). For objective measures we considered the answers given to the questions. For one of the participants, the answers to the questions were lost due to a technical error. Because of this, the answers of only 23 participants were taken into account. When considering the amount of detail in the answers, we only considered the comparable answers. Here *comparable*, means the answers to question topics which were actually posed and answered in both conditions. This gives a measure where we can compare, as if per question, how detailed the actual answers were. As it is possible that more questions were posed in one condition than the other, comparing all texts could result in comparing answers of, for instance, answers to five questions to answers to three questions. This would give a distorted image of how detailed the response to each question actually was.

Table 4.2: Comparison between the number of words and number of topics for the ontology-based and non-ontology-based system.

	Mean(SD)					
Measurement	***Ontology***	***Non-Ontology***	***t***	***df***	***p***	***Cohen's d***
Nr. of Words	237 (114)	285 (11)	-2.37	22	0.027	0.33
Nr. of Topics	8 (3)	10 (4)	-3.98	22	0.001	-0.54

4.5. Results

4.5.1. Questionnaires

The first questionnaire measured if recollecting a memory changed peoples arousal and/or valence regarding the memory. A doubly multivariate repeated measures was done for both arousal and valence, with moment of measurement (pre/post) and system (ontology/non-ontology based) as within-subject factors. No significant results were found ($p > .05$). For both the recollection and the conversation questionnaire, a paired samples t-test was done to compare scores between conditions. Neither of these two questionnaires yielded any significant results between conditions (Recollection: $t(23) = -.38$, $p = .71$, Conversation: $t(23) = -.27$, $p = .79$). On the overall preference, a single-sample t-test showed no significant difference between the result of any of the 4 questions and the middle position on the scale (50), signifying that there was no significant preference for one system over the other (Recollection: $t(23) = 1.26$, $p = .22$, Pleasant: $t(23) = 1.82$, $p = .81$ Questions: $t(23) = 1.17$, $p = .25$ Use again: $t(23) = .1.57$, $p = .13$).

4.5.2. Answers

A paired t-test was performed on the amount of words typed in answers and the number of question topics answered in both conditions, the results of which are presented in Table 4.2. The table shows that there is a significantly higher number of total words in the answers with non-ontology based system compared to the answers with the ontology-based system. The result for the number of topics is similar, a significantly higher number of topics was covered with the non-ontology-based system compared to the ontology-based system.

Finally, the amount of detail in comparable texts in the two conditions was considered. An omnibus test was done on the annotations of the total number of comparable texts, showing a trend of a higher number of words in the ontology-based system, but no significant result $F(5, 18) = 23.63$, $p = .084$, $\eta^2 = .87$. Table 4.3. shows the univariate analysis for the individual annotation categories. Here we see that there was a significantly higher number of descriptive words used in the ontology-based system even after a Bonferroni correction which sets the α level at 0.01. None of the other categories showed significant results.

Table 4.3: Comparison between the ontology-based and non-ontology-based system based on the total number of objects, people, descriptives, feelings and time references in the participants comparable texts.

	Mean(SD)				
Category	***Ontology***	***Non-Ontology***	***F(1,22)***	***p***	η^2
Objects	21 (14)	20 (13)	1.00	0.328	0.044
People	7 (7)	6 (4)	0.66	0.425	0.029
Descriptive	22 (13)	16 (9)	8.91	**0.007**	0.288
Feeling	5 (3)	6 (4)	0.34	0.567	0.015
Time	4 (3)	3 (3)	0.32	0.263	0.057

4.6. Discussion & Conclusion

The first conclusion we can draw based on the results is that no subjective difference between an ontology-based and a non-ontology-based system was found. This indicates that people have no preference for one type of system over the other. The second conclusion is that people answered the questions more quickly with the non-ontology based system. This is shown by a higher number of topics answered with the non-ontology-based system, while both conditions lasted equally long. This is promising, as it suggests that people put more effort into answering the questions from the ontology-based system. When making statements about effort in memory recollection it is, however, also important to consider the amount of detail in the answers and not just the time taken. Concerning this detail, we see that that there is a significantly higher number of descriptive terms for the ontology-based system. From this we can conclude that people describe memories in more detail with this system. Taken together with the result that people take more time, we can conclude that people recollect their memories in more detail with an ontology-based system. This study also has some limitations, the main drawback being that the participants tested were healthy individuals, and not PTSD patients. We believe, however, that our results do provide a valid insight in memory recollection with an ontology-based system as it shows that such a system can assist in detailed memory recollection. Future work will have to study the effect of an ontology-based system on the recollection of memories which people would rather forget. One contribution of this study has been to show that aside from high-level planning [18], and adding domain knowledge [20], ontologies can also be used to store specific knowledge of the user and steer the conversation based on this. It has also shown that the use of multiple-choice options to personalize the conversation [15, 17] can be combined with such an ontology. Finally, we have shown that an ontology-based question system is effective in assisting users with detailed memory recollection, as necessary in PTSD exposure therapy [3].

References

[1] M. L. Tielman, M. van Meggelen, M. A. Neerincx, and W.-P. Brinkman, *An ontology-based question system for a virtual coach assisting in trauma recollection,* in *Int. Conf. on Intelligent Virtual Agents* (2015).

[2] *Diagnostic and statistical manual of mental disorders*, 5th ed. (American Psychiatric Association, Washington, DC, 2013).

[3] E. Foa, E. Hembree, and B. Olaslov Rothbaum, *Prolonged Exposure Therapy for PTSD: Emotional Processing of Traumatic Experiences Therapist Guide (Treatments That Work)* (Oxford University Press, 2007).

[4] W.-P. Brinkman, E. Vermetten, M. van der Steen, and M. A. Neerincx, *Cognitive engineering of a military multi-modal memory restructuring system,* Journal of CyberTherapy and Rehabilitation **4**, 83 (2011).

[5] M. Tielman, W.-P. Brinkman, and M. A. Neerincx, *Design guidelines for a virtual coach for post-traumatic stress disorder patients,* in *Int. Conf. IVA* (2014).

[6] B. Reeves and C. Nass, *The Media Equation* (CSLI Publications, 1996).

[7] O. Blanson-Henkemans, C. van der Mast, P. van der Boog, M. Neerincx, J. Lindenberg, and B. Zwetsloot-Schonk, *An online lifestyle diary with a persuasiv computer assistant providing feedback on self-management,* Tech. & Health Care **17**, 253 (2009).

[8] T. W. Bickmore, R. A. S. K. Nelson, D. M. Cheng, M. Winter, L. Henault, and M. K. Paasche-Orlow, *A randomized controlled trial of an automated exercise coach for older adults,* Journal of the American Geriatrics Society **61**, 1676 (2013).

[9] J. Martínez-Miranda, A. Bresó, and J. M. García-Gómez, *Look on the bright side: A model of cognitive change in virtual agents,* in *Int. Conf. IVA* (2014).

[10] A. Rizzo, B. Lange, J. G. Buckwalter, E. Forbell, J. Kim, K. Sagae, J. Williams, B. O. Rothbaum, J. Difede, G. Reger, T. Parsons, and P. Kenny, *An intelligent virtual human system for providing healthcare information and support,* Medicine Meets Virtual Reality **18**, 503 (2011).

[11] G. Lucas, J. Gratch, A. King, and L.-P. Morency, *It's only a computer: Virtual humans increase willingness to disclose,* Computers in Human Behavior **37**, 4 (2014).

[12] C. Qu, W.-P. Brinkman, P. Wiggers, and I. Heynderickx, *The effect of priming pictures and videos on a question-answer dialog scenario in a virtual environment.* Presence: Teleoperators and Virtual Environments. **22**, 91 (2013).

[13] N. ter Heijden and W.-P. Brinkman, *Design and evaluation of a virual reality exposure therapy system withh automatic free speech interaction,* Journal of CyberTherapy and Rehabilitation **4**, 41 (2011).

[14] T. Bickmore and T. Giorgino, *Health dialog systems for patients and consumers,* Journal of Biomedical Informatics **39**, 556 (2006).

[15] D. Schulman, T. Bickmore, and C. Sidner, *An intelligent conversational agent for promoting long-term health behavior change using motivational interviewing,* To appear.

[16] W. R. Miller and S. Rollnick, *Motivational interviewing: Preparing people to change addictive behavior* (Guilford Press, New York, 1991).

[17] S. Friederichs, C. Bolman, A. Oenema, J. Guyaux, and L. Lechner, *Motivational interviewing in a web-based physical activity intervention with an avatar: Randomized controlled trial,* Journal of Medical Internet Research **16**, e48 (2014).

[18] T. Bickmore, D. Schulman, and C. Sidner, *A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology,* Journal of Biomedical Informatics **44**, 183 (2011).

[19] M. Beveridge and J. Fox, *Automatic generation of spoken dialogue from medical plans and ontologies.* Journal of Biomedical Informatics **39**, 482 (2006).

[20] H. Al-Zubaide and A. A. Issa, *Ontbot : Ontology based chatbot,* in *Innovation in Information & Communication Technology* (2011).

[21] A. Augello, G. Pilato, G. Vassallo, and S. Gaglio, *Chatbots as interface to ontologies,* Advances in Intelligent Systems and Computing **260**, 285 (2014).

[22] N. F. Noy and D. L. McGuinness, *Ontology Development 101: A Guide to Creating Your First Ontology*, Report (Stanford Knowledge Sys. Lab. Tech. Report, 2001).

[23] P. Lang, *Technology in mental health care delivery systems,* (Ablex, 1980) Chap. Behavioral treatment and bio-behavioral assessment: Computer applications.

5

Design and Evaluation of Personalized Motivational Messages by a Virtual Agent that assists in Post-Traumatic Stress Disorder Therapy

Background *E-mental health care systems incorporating virtual agents can play a major role, as barriers to care still prevent some patients from receiving the help they need. To properly assist the users of these systems, the virtual agent needs to promote motivation, something which can be done by offering motivational messages.*

Objective *This paper presents the design and evaluation of a motivational message system for a virtual agent assisting in post-traumatic stress disorder (PTSD) therapy. A database of motivational statements was built with expert (*n*=13) input on what types of statements to use in what user situation. Using this database, the system generates personalized motivational messages.*

Methods *To investigate if the motivational message system improves motivation to continue and trust in a good therapy outcome, an online study was performed (n=207). Participants were asked to imagine they were in a certain*

A part of the work presented in this chapter has been published in the International Conference on Intelligent Virtual Agents (2017) [1].

situation, and received either a personalized motivational message as generated by the system, a general motivational message, or a message without motivation. They were asked how this message changed their motivation and trust, as well as how much they felt being heard by the agent.

Results *Overall, receiving a motivational message improved motivation to continue, trust in a good therapy outcome and the feeling of being heard by the agent. Moreover, this feeling of being heard was further improved if the motivational message was personalized to the user's situation. This personalization was also shown to be important in those situations where the symptoms were getting worse. In these situations, personalized messages outperformed general messages both in terms of motivation to continue and trust in a good therapy outcome.*

Conclusions *Based on expert's input, a personalized motivational message system was developed, which can improve motivation and trust in PTSD therapy. Given the confrontational nature of PTSD therapy, this system has the potential to make an important difference. As both motivation and trust play an important role in many other e-mental health systems as well, personalized motivational message systems such as these can be very valuable.*

5.1. Introduction

Virtual agents are increasingly applied in e-mental health care, as they can improve the efficacy of interventions [2]. An important objective of these virtual agents is to improve the patient's motivation to keep going by offering short motivational messages. But due to the large number of domains virtual agents are applied to, the exact content of these messages differs a lot and it is difficult to re-use content for new applications. Human therapists, however, harbor a wealth of domain knowledge. Moreover, they have experience with offering motivation and personalizing it to a patient's situation. This paper presents a motivational message system for a virtual agent that assists in post-traumatic stress disorder (PTSD) therapy. This system uses a database of motivational statements for different situations, developed using knowledge from human experts [1]. The goal of this message system is to increase the patient's motivation to continue and trust in a good therapy outcome.

5.1.1. Background

Untreated mental disorders account for 13% of the total global burden of disease, and the gap between treatment needed and treatment received is wide [3]. E-health can help remove some of the barriers to care for mental health, offering a cost-effective, accessible and privacy-sensitive solution to mental health problems [4, 5]. A broad range of e-mental-health systems exists, differing in purpose and domain. Some systems are meant for monitoring only [6], while others offer full therapeutic interventions [7, 8]. Especially the systems offering full interventions might be classified as behavior change support systems (BCSS), aiming to change unhealthy behavior patterns. Such systems can be aimed at a range of different disorders, such as substance abuse [9], anxiety [10], depression [11] or PTSD [12].

PPTSD is a disorder where motivation plays an especially important role. Patients have experienced one or more traumatic experiences, and have symptoms such as intrusive memories, a persistent negative state and active avoidance of anything to do with their memories [13]. One of the most common types of therapy for PTSD is cognitive behavioral therapy (CBT) including an exposure component. During this exposure, patients need to break the pattern of avoidance by actively recollecting their traumatic experiences [14]. The idea behind exposure is that although tension will initially rise, the automatic fear reaction will eventually fade away [15]. This initial tension means, however, that the situation might temporarily get worse, something that is reflected in treatment manuals [16] and patient reports [17]. This potential worsening also highlights the importance of trust; it is important that patients still believe they will get better if they persevere. Although shown to be effective [18], exposure therapy is therefore also challenging. This means that in an e-health therapy system for PTSD such as in [12], motivational messages can be very welcome.

A therapy system for PTSD could present motivational messages in several ways. Increasingly, BCSS's for mental health care include virtual agents to present their content to patients [19]. These agents are virtual characters that communicate with the users of a system in some way, and they have been applied for instance for dementia [20], substance abuse [21] and PTSD [22]. Virtual agents are per-

ceived more positively by users than a text-based interface [23]. Moreover, they have the potential to positively affect treatment compliance and outcome [24]. This implies that virtual agents can successfully motivate users simply by being present. Their impact might partially be explained by self-determination theory (SDT), which describes the concept of motivation [25]. It distinguishes between extrinsic motivation, coming from external sources, and intrinsic motivation, which stems from an internal drive, and is the more powerful of the two. Motivation from a BCSS system is always extrinsic, but extrinsic motivation might be internalized, a process that is supported by relatedness[26]. When a person feels more related to the source of the extrinsic motivation, the motivation is more likely to be internalized. Virtual agents have the potential to increase relatedness above simple text-messages [27], and are therefore very suitable to present motivational messages.

Aside from the presentation mode of motivational messages, their content is also an important factor. The first question when designing this content is how to tailor it to the user so it is sufficiently effective. This tailoring can be done based on several different factors. One possible factor is readiness to change, which is the underlying concept of the trans-theoretical model (TTM) [28]. This is a motivational model that identifies six stages of change for health behavior, from where people are not ready yet to make a change, to where a change has occurred but needs to be maintained. To change behavior, TTM states a person should move from one stage to the next. Motivational interviewing (MI) is a motivational tactic also based on the readiness to change concept, focusing on increasing motivation by highlighting the discrepancies between the current and the ideal situation [29]. In this way readiness to change is slowly increased. Both MI and TTM have been applied to virtual agents in various applications [30–33]. However, these models are best considered as full motivational strategies, as they focus less on the content of individual motivational messages. Moreover, research showed that people in different stages of change do not seem to also prefer different motivational messages [34], making this a less suitable factor for tailoring.

Another possible way to tailor motivation to the user is to consider personality traits. Studies have shown that this influences what type of motivational messages is found preferable [35]. Similar results were found for the effect of masculinity and femininity [36]. Inter-personal differences seem to influence what type of motivation is preferred, however, effect sizes in these studies are small. Other factors might therefore also be useful to tailor messages in an e-mental health system for PTSD. Moreover, studies with personality traits are commonly done on the general population, and are not focused on very specific situations. In PTSD therapy, however, motivation is necessary for a very specific task. It seems more useful, therefore, to tailor the messages to the specific health situation of the user instead.

To tailor motivational messages for PTSD therapy to the user's situation, specific parameters for a situation need to be established. In the remainder of this paper, situation will, therefore, be defined in terms of the progression of PTSD symptoms and the user's trust in a good therapy outcome. These two factors can be related to several theories on health behavior and motivation. Protection motivation theory

(PMT) proposes threat appraisal and coping appraisal as main predictors of health behavior [37]. Progression of PTSD symptoms relates to threat appraisal, as progression allows one predict how successful therapy will be. Trust in a good therapy outcome, on the other hand, relates to coping appraisal, how much a person feels they will be able to handle what is coming. Similar concepts are found in other theories, such as the health belief model [38], which describes factors such as perceived threat, perceived barriers, and self-efficacy.

To actually develop situation-based motivational messages for a virtual agent assisting in PTSD therapy, their specific content needs to be established. Some systems base the content of their motivational messages on concepts from the literature [39]. However, literature appropriate to the domain is not always available. If larger amounts of data are available on the motivational messages used by therapists, machine-learning techniques could in principle also be used to classify what messages are used when. However, such data is rare given the private nature of conversations between therapist and patient. Another way to build a motivational system is to code human-human interactions for the use of a virtual agent [40]. This assumes, however, that the virtual agent has the same input from the user as a human would, which is usually not the case as it includes visual cues such as facial expression and posture. A final option to gather content for motivational messages is to rely on experts to write the statements [10]. This results in an appropriate list, but experts need to write many full statements to allow the system to vary its messages. A way to bypass this final problem would be to categorize the messages written by the experts, breaking them up into smaller sentences that can be re-combined. Moreover, by also asking experts to write messages for specific situations, the categorized messages could be matched to those situations in which they are most suitable.

This paper proposes a flexible, situation-based motivational system in the domain of PTSD therapy. Experts in this domain were asked to write motivational messages and these statements were used to build a system that can adapt its motivational messages to the patient's situation. This paper first presents the methods for creating the situation-based motivational message system based on expert input. Furthermore, an experiment was conducted to study the effect of personalized motivational messages on a user's motivation to continue and trust in a good therapy outcome.

5.1.2. Motivational feedback

Vision

This paper envisions a situation-based motivational message system for a virtual agent for PTSD therapy, as shown in Figure 5.1. This system generates motivational messages based on a patient's situation, which in this context is defined by the PTSD symptom trend and the patient's trust in a good therapy outcome. To enable the system to process PTSD symptom trend data, it is expressed in terms of a score on the brief PTSD-Check-List (PCL) [41]. Both PCL trend and trust level are defined in three levels. Trust in a good therapy outcome can be either low, medium or high, and PCL scores can be dropping, stable, or rising. This results in a total of nine

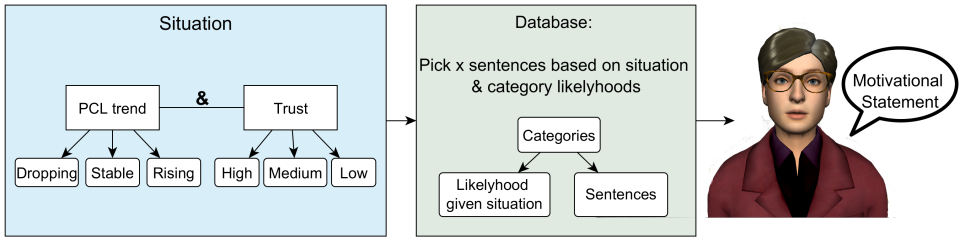

Figure 5.1: Vision the situation-based motivational message system.

possible situations.

Given these situations, a motivational message can be composed using a database with motivational statements. Each motivational statement has a certain type, and the database includes likelihoods that express how likely a type of statement is to occur given the situation. Both the specific statements and the likelihood ratios are based on input from experts familiar with PTSD therapy. By combining specific sentences from each type, the system can generate motivational messages suitable for a patient's situation, usable by the virtual agent.

The motivational message system is meant to increase a patients' motivation to continue with therapy, their trust in a good therapy outcome and their feeling of being heard by the virtual agent. This paper firstly hypothesizes that motivational messages can increase these factors. Moreover, it hypothesizes that by personalizing the motivational messages to a user's situation, the motivation, trust and feeling of being heard will improve even further. This leads to the following two hypotheses, to be tested in this paper:

- H1: Motivational messages improve trust in a good outcome, motivation to continue and the feeling of being heard more than messages without motivational content.
- H2: Personalized motivational messages as generated by the motivational system improve trust in a good outcome, motivation to continue and the feeling of being heard more than general motivational messages.

Building the database

To test the hypotheses, the envisioned motivational message system was created. A database was built including motivational statements, their types and the likelihood they would be used by experts in a given situation. 13 therapists (5 male, 8 female) were recruited from six different mental health clinics. All experts had experience in treating patients with PTSD on a professional basis. They were presented with nine different situations expressed by patient trust in therapy outcome and a graph representing PCL trend. Every situation was shown twice, using two different graphs for each PCL trend, as shown in Figure 5.2. For every situation, the expert was asked to write what they would say to the patient in that situation to increase their motivation to continue and their trust in a good therapy outcome. As context, the

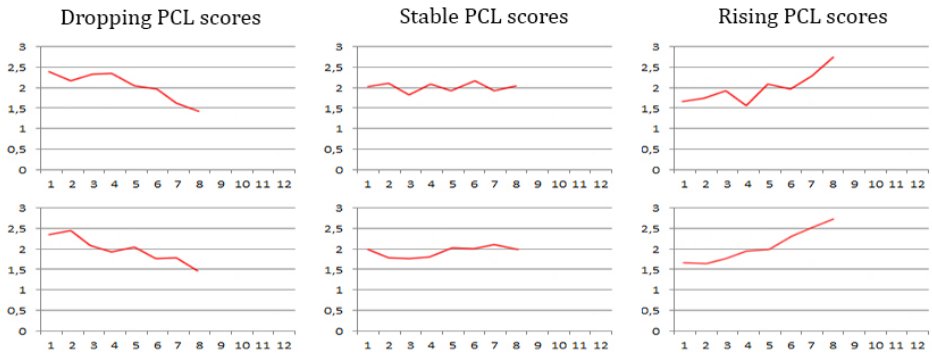

Figure 5.2: Graphs describing PTSD symptoms that were presented to the experts

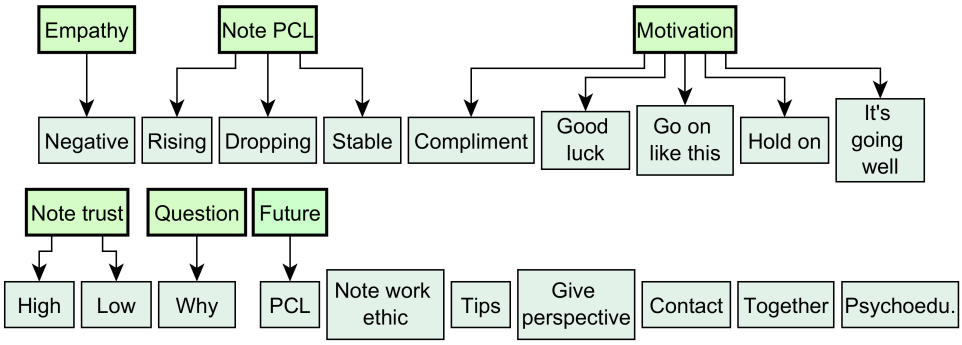

Figure 5.3: Types of motivational statements found in the expert survey. Thicker bordered green types are super types.

experts were given the example of the PTSD therapy system presented in [12]. All situations described a patient in session 8 out of 12, which is right in the middle of the most intense phase of treatment. The design of this study was approved by the university ethics committee, ID number 134.

The answers given by the experts were between 1 and 143 words long (Mean=40.40, SD=24.68). These answers were split into statements based on subject, resulting in a total of 844 statements. The mean number of statements was three for the *dropping PCL-high trust, dropping PCL -medium trust* and *stable PCL-high trust* situations, and four for the others. All statements were categorized using statement types which arose during the analysis; the resulting types are shown in Figure 5.3. A subset of 32 statements was rated by a second coder, interrater reliability was substantial (κ =0.73, P=0). All statement types with less than ten statements were removed. After this selection, 97% of the statements were included in the analysis.

The data were analyzed with R 3.3. First, it was noted whether a type of statement was present for every answer given by the experts. A multi-level analysis with expert as random intercept was done, showing that adding statement type as fixed factor significantly improved the model predicting if a statement type oc-

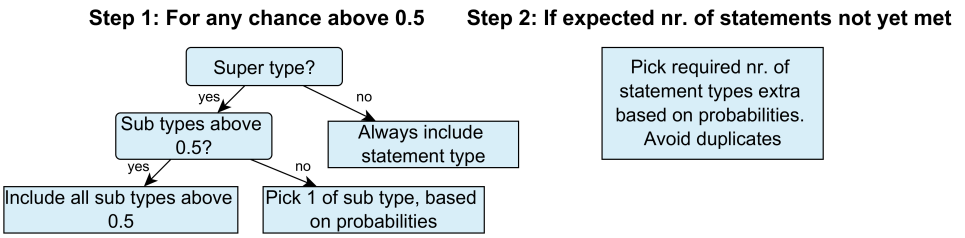

Figure 5.4: Algorithm for choosing the statement types for the motivational message system.

curred in the expert answers (F(49,11638)= 88.86, *P*<.001). This indicates that there are differences in how much different statement types were used. Second, an additional analysis showed that the model containing only statement type was significantly improved by adding the three-way interaction effects for PCL trend, trust and statement type as fixed factor (χ^2_3= 1954.80, *P*<.001). As this shows that whether a statement type is used is influenced by the situation, a second multilevel logistic regression was done for each statement type separately, predicting if a statement type occurred in the expert's answers for that situation. This analysis used expert as random intercept and the two-way interaction between PCL graph trend and trust as fixed effect. This resulted in the probability of each statement type occurring in a given situation. Together with the motivational statements given per type, these probabilities form the database that lies at the base of the motivational message system. The full probability table can be found in [42] and Appendix III of this thesis.

Implementation message system

To generate the full motivational messages from the database, either three or four statements were combined, depending on the mean number of statements given by the experts for that situation. When generating a motivational message for a specific situation, the statement types are chosen first. Any statement type with a probability above .5 is always included. If this is a super-type, all sub-types with a probability above .5 are included, if none met this criterion one is chosen based on probability. This probability-based selection could, for instance, be done by adding all probabilities and assigning every statement type an interval equal to their probability. By generating a random number between zero and the sum of probabilities, the statement type belonging to the interval the random number falls in is chosen. If after this selection the expected number of statements is not met, statements are added one by one based on probability, while duplicate types are avoided. The full algorithm is described in Figure 5.4. Finally, for every statement type included a specific sentence is selected at random from the database. So even if the statement types selected for a certain situation are always the same, the specific sentences making up the motivational message can differ.

An example of the selection and result of a personalized motivational message for the situation with *rising PCL, low trust* would be as follows. First, the types *mo-*

tivation, give perspective and *note PCL* will be included as their probability is higher than 0.5. For *motivation* one of the sub-types will be chosen based on probability as none are higher than 0.5, in this example, we will choose the sub-type *hold on*, as it is the highest for this situation. For the *note PCL* type, this will always be *dropping* because it is over 0.5. Because we now only have three statements while the mean number for this situation was 4, one extra type will be chosen based on probability, *empathy* is selected for this example, as it has the highest probability of any statement type not yet included. Taken together, this could result in the following motivational message (translated from the original Dutch).

I see you indicate that your complaints have gotten substantially worse. (note rising PCL) *I'm sorry to hear that* (Empathy). *However, it's always hard work before we see any results* (Give perspective). *Hold on!* (Motivation hold on).

5.2. Methods

To test the two hypotheses presented in section 1, an empirical study was done. The personalized motivational messages were compared to general motivational messages and messages without motivational content in a mixed 3x3x3 design.

The motivational message system is designed for PTSD patients following therapy assisted by a virtual agent. However, to evaluate the message system with patients in therapy would mean a full clinical trial, which traditionally mainly focuses on clinical outcome. Instead, e-health systems are often evaluated in phases [43], and separate components are first tested using proximal outcome measures such as motivation [44]. Testing components separately also allows for a clearer picture of exactly what effect a certain manipulation, such as motivational message type, has on specific outcome variables, such as motivation to continue and trust in a good therapy outcome. This abstraction strategy is for instance also used in game design to reduce surrounding and complicating factors and speed up the process [45]. The evaluation of the motivational message system takes this component-focused approach. Because it was tested in isolation instead of incorporated a full therapy, participants were asked to imagine they were in a certain situation during PTSD therapy, as defined by symptom progression and trust in therapy outcome. One can imagine, however, that people who have experience with PTSD exposure therapy will interpret the situations differently from people who have not. Participants were therefore also asked for their experience with PTSD symptoms and exposure therapy so that these factors could be taken into account as potential covariates.

The message system was evaluated in a 3x3x3 mixed design. Message type was measured between-subject, while the three graphs describing PTSD symptom trend and the three levels of initial trust in therapy outcome were presented to all participants within-subject. Afterwards, participants saw a virtual agent presenting a personalized motivational message, a general motivational message or a message without motivational content. Because the situations were imaginary, participants were asked to indicate changes in their motivation to continue and trust in a good therapy outcome. The design of this study was approved by the University ethics

committee, ID number 184.

5.2.1. Participants

To study the effect of the situation-based motivational message system participants were recruited via Amazon Mechanical Turk. An a priori power analysis was performed to determine the necessary number of participants[1]. Given a one way ANOVA with three groups, a medium effect size of 0.25 and power of 0.9, a preferred sample size of 207 was calculated. A total of 529 participants started the experiment. Participants were excluded if they could not confirm reading all situations and questions (n=189), if they did not complete the survey (n=128), or in case of administrative errors (n=5). Participants were excluded directly after completing the study, which kept running until the required number of 207 included participants was met. Of the excluded participants, 25.5% were in the personalized motivation condition, 27.6% in the general motivation condition, 29.5% in the no motivation condition and 30.1% stopped before being assigned a condition. Of the 207 included participants, 34.3% were in the personalized motivation condition, 34.3% in the general motivation condition and 31.4% in the no motivation condition. All participants were over 18 and native English speakers[2].

5.2.2. Measures

Five different measures were collected via questionnaires, two of them repeatedly for every situation that was presented. A list of all questions used can be found in [42] and in Appendix III of this thesis.

Primary measures

Trust in a good therapy outcome was measured in terms of change. After each situation, the question would be posed how much the comments by the virtual agent changed trust in a good therapy outcome. Answers ranged on an analog scale from -10 (*it decreased a lot*) via the neutral point of 0 (*nothing changed*) to 10 (*it increased a lot*). This question was repeated at the end of the study, asking for the overall change in trust.

Motivation to continue was measured similarly to trust, in terms of change caused by the comments of the virtual agent, asked for in every situation. The analog scale also ranged from -10 (*it decreased a lot*) via the neutral point of 0 (*nothing changed*) to 10 (*it increased a lot*), and this question too was repeated at the very end asking for the overall change in motivation.

Feeling of being heard (FBH) was measured only at the very end, and included to get a measure of how much the participants felt the agent took them seriously and really heard them. This questionnaire consisted of item 1 and 4 of the short patient satisfaction questionnaire [46], 5 and 8 from the trust in physician scale [47], questions were slightly adapted to be positively phrased statements. Three additional questions were added to complete the questionnaire, stating that *the virtual agent took you seriously, the virtual agent replied appropriately to you,* and

[1] using G*Power 3.1

[2] Due to an administrative error, specific age and gender data of participants were not collected.

Figure 5.5: Virtual Agent

the virtual agent listened to your preferences.

Descriptive measures
To get an indication of whether participants had experience with PTSD symptoms, a brief version of the PCL was presented at the beginning of the study, including questions 2, 3, 6, 7, 16 and 17. People were asked to keep in mind their most stressful or traumatic experience when answering these questions. Participants were also asked if they had experience in following exposure therapy, and if so for what disorder.

5.2.3. Procedure

After accepting the assignment on Amazon Mechanical Turk, participants were redirected to Qualtrics for the main experiment. They were first presented with an information form outlining the procedure of the experiment and the consent form. Afterwards, they were given the PCL and asked if they had ever followed exposure therapy. Before the main part of the experiment, all participants saw an example situation with a PCL graph that was not re-used in the remainder of the experiment, a clip from the virtual agent introducing itself and the motivation and trust questions. This example included the explicit instruction that a rising PCL graph meant symptoms were getting worse and vice versa. After the example, participants were randomly divided into the three conditions. The virtual agent always appeared as a video clip and displayed idle and mouth movement, speaking with a computer-generated text-to-speech voice. The virtual agent is shown in Figure 5.5.

In all conditions, participants were asked to repeatedly imagine they were in therapy and had just answered the PCL questionnaire resulting in the situation shown. All nine situations representing the combinations of PCL trend (rising, stable, dropping), and trust level (no trust, doubt, high trust) were shown in randomized order. The PCL graphs used in this experiment are shown in Figure 5.6. After

Table 5.1: table:examples

Type	Message
Personalized to: *Stable PCL & Doubts*	Your complaints show little consistency. Sometimes they get less, and then they get worse again. That makes sense, you're working on your past and that's tough. Stick to it now! Let's see together what you need to continue.
General	Your complaints show little consistency. Sometimes they get less, and then they get worse again. That makes sense, you're working on your past and that's tough. Stick to it now! Let's see together what you need to continue.
No motivation	Thank you for filling in the questionnaires. We'll now continue with the session.

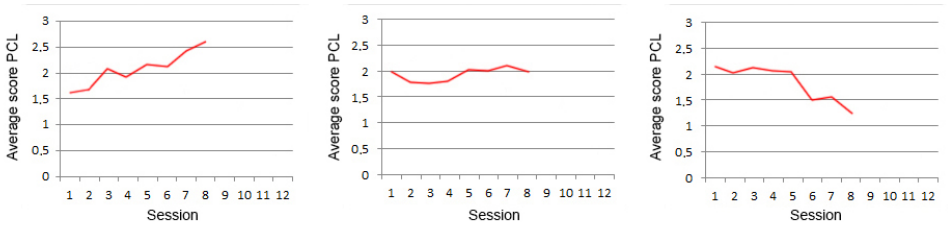

Figure 5.6: Situations used in the experiment. From right to left: Rising PCL, Stable PCL, Dropping PCL.

reading the situation, participants could view the reaction from the virtual agent. In the condition with the personalized messages, one of three comments generated by the system for that particular situation was used, all translated from the original Dutch to English. For the condition with general messages, nine different statements were written by three different experts and translated into English. Comments were matched with the situations in three different variations, each comment occurring once for each PCL trend and each trust situation. What combination participants were presented with was randomized. The same type of randomization was performed for the messages without motivational content. In this condition, the virtual agent merely thanked the participant for filling in the PCL and announced the continuation of the session, using nine different formulations. Three examples of messages are shown in Table 5.1, all messages for the different conditions can be found in [42].

After the message, participants were presented with the question how much their trust in a good outcome and motivation to continue had changed. Both the situation and the virtual agent were still visible at this point, so participants had the option to review either or both. After all situations, these two questions were repeated for the overall experience, and the FBH questionnaire was presented.

5.2.4. Data preparation & Analysis

Data was analyzed with R. version 3.3; the full data table can be found in [42]. Following the guidelines for the PCL [41], a participant was classified as potentially having PTSD if question 2 or 3 was 3 or higher, and question 6 or 7 was 3 or higher, and 16 or 17 was 3 or higher. Reliability for the FBH scale was high with a Cronbach's alpha of α=0.96, so answers to the items of the FBH scale were averaged to form one score.

Covariates

To examine if showing an *indication of PTSD* and having *experience with exposure* as potential covariate factors in the analysis, ANOVAs were done with these variables as predictor of overall change in motivation, trust and FBH score. As these analyses did not find any effects, neither PTSD indication or exposure experience were further taken into account as covariates.

Overall Measures

For all overall measures, a linear regression model with only an intercept was run to establish deviation from zero. Type of message was added to the model to establish if it affected the overall outcome measures. When this analysis showed a significant effect of type of message, individual post-hoc analyses were run on data containing only two message types to test the differences between them.

Per situation

To analyze the data per situation, every situation was treated as a separate datapoint. A multi-level regression was run with participant as random intercept. For both change in motivation to continue and trust in a good outcome, the null model with only a fixed intercept was compared to model 1, including message type as fixed effect. Model 2 built on model 1 and added PCL trend as fixed effect; model 3 added initial trust level. Further models were built adding first the two-way interaction effects and finally the three-way interaction effect. Post-hoc analyses were run on subsets of the data to find the differences between each of the three message types. This was done for the three PCL trend types, the three levels of initial trust, and for the 3x3 combination of these measures resulting in all nine situations.

5.3. Results

5.3.1. Descriptive measures

Table 5.2 shows the descriptive measures including their influence on the overall outcome measures. Scores on the PCL show that almost a third of participants showed an indication of having PTSD. 14 participants had experience with exposure therapy. Out of the seven with experience with exposure for PTSD, five also showed an indication of PTSD. The other specific disorders mentioned were obsessive compulsive disorder (OCD) and both specific and general anxiety disorders. Table 5.2 also shows that an indication of PTSD or exposure experience were not

Table 5.2: Descriptive characteristics of participants.

Characteristics	n(%)	Motivation		Trust		FBH	
		F(1,205)	*P*	F(1,205)	*P*	F(1,205)	*P*
PTSD Potential	66(31.88)	0.99	.32	0.13	.72	0.04	.83
Exposure experience	14(6.76)	1.48	.22	1.18	.28	0.34	.56
PTSD	7 (3.38)						
Anxiety	6 (2.90)						
OCD	1 (0.48)						

Table 5.3: Description overall measures & effects of condition

Variable	Mean	SD	Deviance from 0		Effect of feedback type	
			t	*P*	F(2,204)	*P*
Motivation	1.06	4.05	3.76	<.001	20.58	<.001
Trust	0.74	4.16	2.56	.011	16.18	<.001
FBH	-0.66	4.83	-1.96	.051	50.97	<.001

found to affect overall change in motivation, trust or FBH score. There was, therefore, no indication found that experience with PTSD or exposure therapy influenced the interpretation of the scenarios and effect of the virtual agent.

5.3.2. Overall experiences

Table 5.3 shows the mean values of all overall outcome variables measured at the end of the experiment, as well as their deviance from the neutral point of zero. Both change in motivation and trust were rated significantly above neutral, indicating that overall motivation and trust improved. The feeling of being heard did not deviate significantly from zero. For all three outcome measures, linear regression analyses revealed a significant effect for message type.

Post-hoc analyses revealed the differences between each of the three message types for the overall variables. Figure 5.7 visualizes the results from this analysis, showing that motivational messages significantly improves overall increase in motivation and trust. Moreover, the feeling of being heard increases significantly not only if a motivational message is present, but even more if this message is personalized to the participant's situation. The confidence intervals show that if a personalized message was presented, motivation, trust and the feeling of being heard improved, while no motivation always resulted in a decrease.

5.3.3. Per situation

Table 5.4 shows the effects of message type, PCL trend and initial trust on the outcome measures of motivation to continue and trust in a good outcome, as measured per situation. This table also includes the effects of the two and three-way interactions.

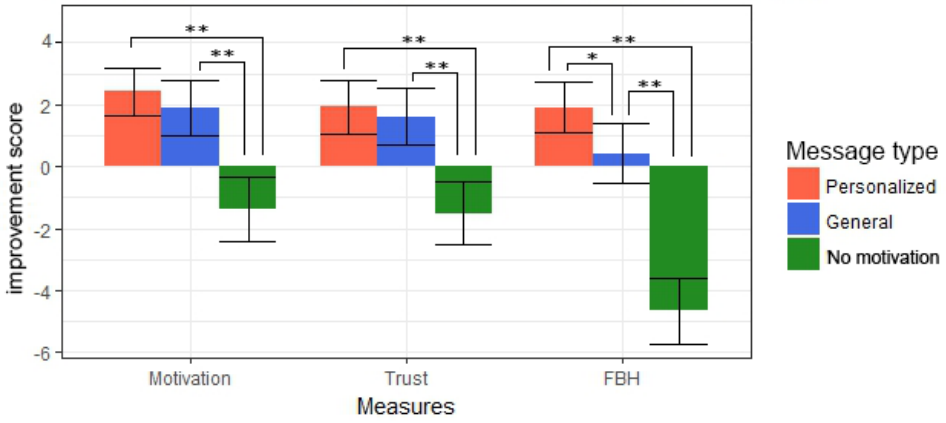

Figure 5.7: Effect of message type on the overall outcome measures. * =P<.05, ** = P<.01. Error bars indicate 95% Confidence Intervals.

Table 5.4: Influence of situation and feedback type on motivation to continue and trust in therapy outcome.

Model comparison	χ^2	P
Motivation to continue		
add Message type (M_0 vs M_1)	$\chi^2_5 = 29.07$	<.001
add PCL Trend (M_1 vs M_2)	$\chi^2_7 = 27.54$	<.001
add Initial Trust (M_2 vs M_3)	$\chi^2_9 = 40.46$	<.001
add Initial Trust x PCL Trend(M_3 vs M_4)	$\chi^2_{13} = 9.35$	.053
add Message type x PCL Trend (M_4 vs M_5)	$\chi^2_{17} = 25.17$	<.001
add Message type x Initial Trust (M_5 vs M_6)	$\chi^2_{21} = 11.16$	.025
add Message type x Initial Trust x PCL Trend (M_6 vs M_7)	$\chi^2_{29} = 17.11$	.029
Trust in Outcome		
add Message type (M_0 vs M_1)	$\chi^2_5 = 22.92$	<.001
add PCL Trend (M_1 vs M_2)	$\chi^2_7 = 55.70$	<.001
add Initial Trust (M_2 vs M_3)	$\chi^2_9 = 59.92$	<.001
add Initial Trust x PCL Trend(M_3 vs M_4)	$\chi^2_{13} = 5.30$	.258
add Message type x PCL Trend (M_4 vs M_5)	$\chi^2_{17} = 20.46$	<.001
add Message type x Initial Trust (M_5 vs M_6)	$\chi^2_{21} = 7.41$	.116
add Message type x Initial Trust x PCL Trend (M_6 vs M_7)	$\chi^2_{29} = 12.73$	.121

n=1863

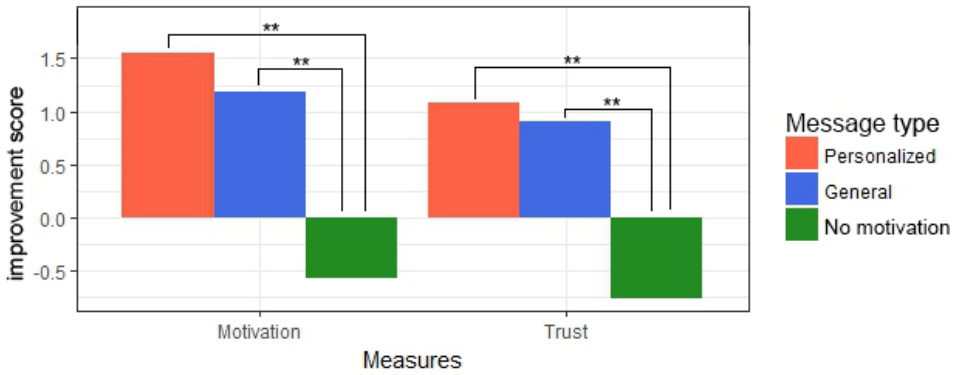

Figure 5.8: Influence of message type on motivation & trust per situation. * =*P*<.05, ** = *P*<.01.

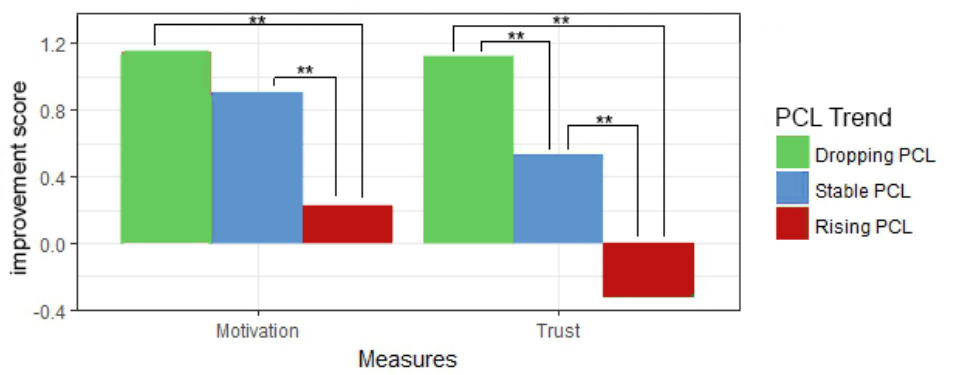

Figure 5.9: Influence of PCL trend on motivation & trust per situation. * =*P*<.05, ** = *P*<.01.

This table shows that both message type and the two situational factors influence motivation and trust as measured per situation. Figure 5.8 visualizes the pair-wise comparison of the effect of message type, showing that motivational messages outperform messages with no motivational content. The differences between the three levels of PCL trend and initial trust are displayed in Figures 5.9 and 5.10 respectively, showing how these situational factors affected the outcome measures.

Aside from individual effects, Table 5.4 also shows the existence of an interaction effect between message type and PCL trend for both motivation and trust, and a further effect between message type and initial trust for motivation. Given these results, a pair-wise analysis was done comparing each of the message types given the three possible values of initial trust and PCL trend, shown in Figures 5.11 and 5.12 respectively. Figure 5.11 shows that initial trust level effects how much motivation and trust improve. Moreover, if the initial trust level is doubtful of low, not receiving motivation seems to even reduce motivation to continue and trust in a good outcome. Figure 5.12 shows that when PCL scores are rising and the situation

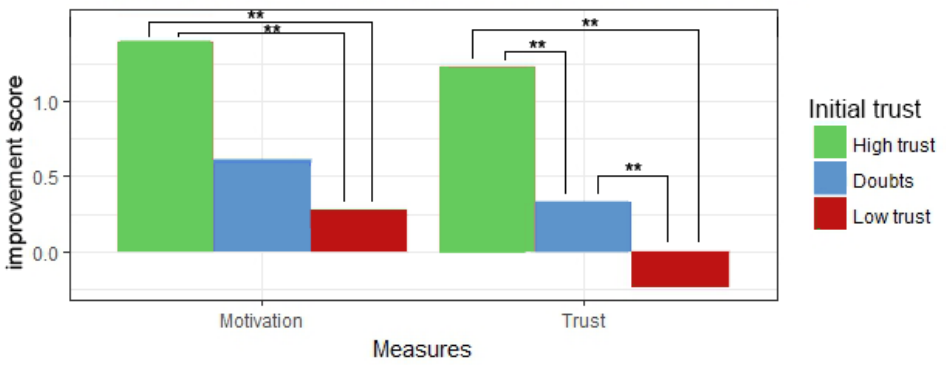

Figure 5.10: Influence of initial trust on motivation & trust per situation. * =P<.05, ** = P<.01.

Figure 5.11: Effect of message type on motivation and trust, per initial trust level. * =P<.05, ** = P<.01.

therefore is getting worse, personalized motivational messages significantly outperformed general motivational messages in terms of both motivation to continue and trust in therapy outcome.

Finally, Table 5.4 shows the existence of a three-way interaction effect of message type, initial trust and PCL trend on motivation to continue. Figure 5.13 shows the effect of message type on motivation to continue for all nine situations, combining PCL trend and initial trust level. These graphs show that the exact effect of message type differs per situation. In the situations where symptoms are getting worse but initial trust is doubtful or high, personalized motivational messages show an advantage over general motivational messages. The same is true for the situation with dropping PCL scores and high trust, which shows not only that personalized motivation is best, but also that general motivational messages do not even outperform messages without motivational content. In most other situations, motivational messages of some type outperform messages without motivational content. The only exception is the situation with dropping PCL scores and low

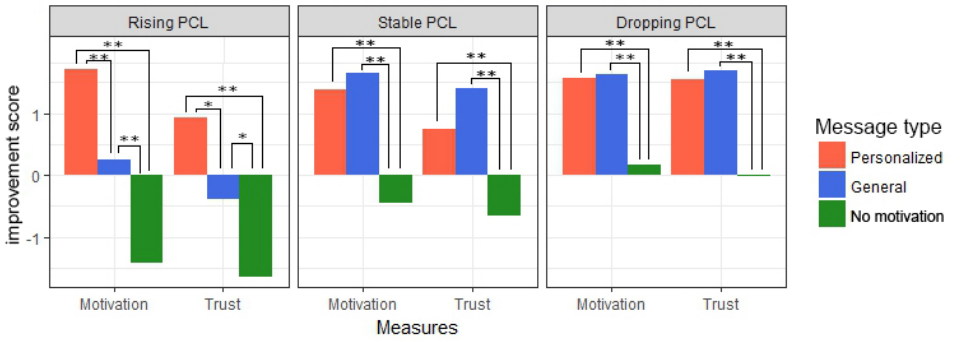

Figure 5.12: Effect of message type on motivation and trust, per PCL trend. * =P<.05, ** = P<.01.

trust, where motivation is improved most by general motivational messages.

5

5.4. Discussion & Conclusion

5.4.1. Primary results

This paper presents a personalized motivational message system for PTSD therapy. Based on input from experts, a database of motivational statements was built, including probabilities of certain categories of statements occurring in certain situations. By considering these situations, which are defined in terms of symptom progress and trust in a good therapy outcome, the system is capable of generating personalized motivational messages.

The effect of these motivational messages on motivation to continue and trust in a good therapy outcome was studied in an experiment. As participants in the experiment were asked to imagine situations in which they were following exposure therapy for PTSD, it is relevant to note whether experience with exposure therapy or PTSD influences the results. Although participants were recruited from the regular population, a third showed an indication of PTSD. Moreover, a subset of the participants also had experience with the process of exposure therapy and facing one's fears. Neither of these factors was found to influence the overall outcome measures, however. This supports the generalizability of the results to populations that do have experience with these factors, such as PTSD patients following exposure therapy.

Results from the experiment show effects of motivational message type on all outcome measures, indicating the relevance of the content of these messages. More specifically, both the overall experiences of participants, as the evaluations in specific situations show that motivational messages outperform messages without motivational content in terms of motivation to continue, trust in a good therapy outcome and the feeling of being heard. This confirms hypothesis 1, and is in line with recommendations by experts that motivational messages are important [48].

The motivational message system presented in this paper aims to further improve motivational messages through personalization to the user's situation. For

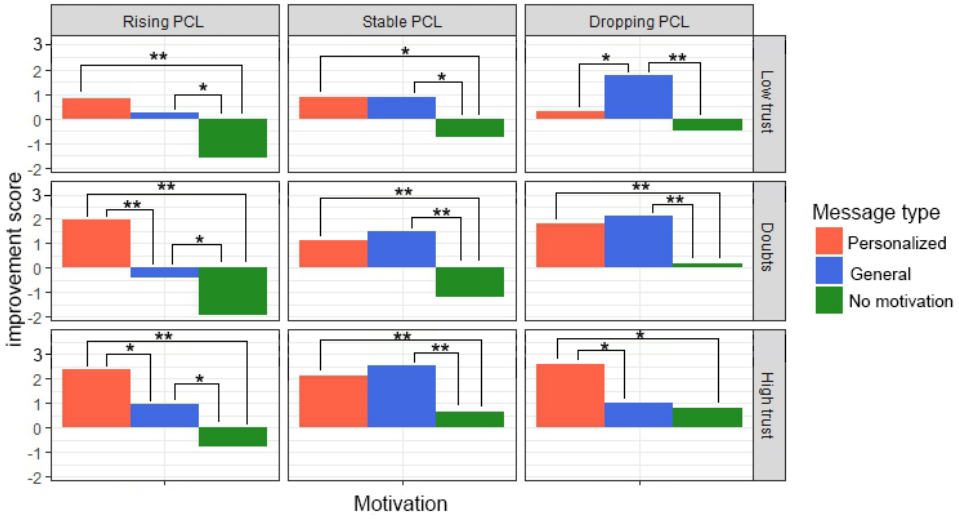

Figure 5.13: Effect of message type on motivation in all separate situations. * =P<.05, ** = P<.01.

the overall feeling of being heard by the virtual agent, the personalization achieves this aim. Given the importance of relatedness on the internalization of extrinsic motivation [26], this feeling of being heard might further increase the effect of motivational statements. An examination of the different situation types sheds further light on the effect of personalization on motivation to continue and trust in a good outcome. Especially when PTSD symptoms are worsening, personalization leads to improvements in both motivation and trust. However, an indication was found that to fully understand the impact of personalization on motivation to continue, the combined effect of PTSD symptom trend and initial trust needs to be considered. Results show that for situations with worsening symptoms, personalization improved motivation when the initial trust level was doubtful or high. The situations where PTSD symptoms are reducing also show a clear impact of initial trust level on the effect of personalization. While personalization improves motivation in situations with high trust, situations with low trust show an opposite effect, with general motivational messages outperforming personalization. Although the overall results, as well as the results from other situations, confirm hypothesis 2, this final result suggests a rejection. This conclusion might, however, be too strong. Instead, it might be more fitting to argue that for this specific situation of dropping PTSD scores and low trust, general motivation is more suitable. One possible reason for this result is that the personalized messages will point out that scores are dropping, which is already a good sign. If trust level is low, however, people might instead prefer to get reassurance that it is normal to find therapy difficult. Nearly all general motivational messages included this reassurance, while the personalized messages did not, as things were actually going well.

5.4.2. Limitation

To fully appreciate the system and results presented in this paper, it is important also to consider the limitations. Although the motivational message system is based on patient situation, it only considers two parameters, namely symptom progression and trust. It is, however, imaginable that other factors also influence how to motivate a person. Interpersonal differences such as personality and gender might, for instance, play a role [35]. Moreover, the system currently only considers a patient in one stage of therapy, but a person's stage in treatment might also affect what motivational message is most suitable [49]. Another limitation is that the experiment was not performed with PTSD patients currently in therapy. However, by taking into account the participant's level of experience with similar situations, results should still be generalizable to clinical practice. It is more difficult to say what the effect of the message system would be for a full therapy, as participants were only presented with situations in a single session. Further research would have to be done to study what results repeated motivation during the course of therapy would have.

5.4.3. Contribution

Although the motivational message system presented in this paper is catered towards PTSD therapy, the results are applicable to other domains as well. The method used to create the message system has the advantage that no data on patient-therapist interactions is necessary. Because this data is often not available, many current systems are based exclusively in theory [39]. The system described in this paper gives an example of how to effectively employ expert knowledge to create a flexible and effective motivational message system. Moreover, motivating statements are used for a broad range of mental health applications [10, 24, 50], so similar systems could be applied to many different interventions.

5.4.4. Conclusion

This paper presents a motivational message system. Applied to situations in PTSD therapy, this system can generate motivational messages personalized on the users' situation. An empirical evaluation revealed that the messages from this system can successfully increase a user's trust in a good outcome and motivation to continue. Moreover, it outperforms general motivational messages in several specific situations, as well as making the user feel more heard by the system. Given the importance of motivation in mental health applications, a motivational message system has the potential to make an important difference.

References

[1] M. Tielman, M. Neerincx, and W. Brinkman, *Generating situation-based motivational feedback in a post-traumatic stress disorder e-health system,* in *IVA* (2017).

[2] A. D. Andrade, R. Anam, C. Karanam, P. Downey, and J. G. Ruiz, *An over-*

active bladder online self-management program with embedded avatars: A randomized controlled trial of efficacy, Urology **85** (2015).

[3] WHO, *Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level,* (2012).

[4] J. A. Cartreine, D. K. Ahern, and S. E. Locke, *A roadmap to computer-based psychotherapy in the united states,* Harv Rev Psychiatry **18**, 80 (2011).

[5] T. Donker, M. Blankers, E. Hedman, B. Ljótsson, K. Petrie, and H. Christensen, *Economic evaluations of internet interventions for mental health: a systematic review,* Psychological Medicine , 1 (2015).

[6] J. Kasckow, S. Zickmund, A. Rotondi, A. Mrkva, J. Gurklis, M. Chinman, L. Fox, M. Loganathan, B. Hanusa, and G. Haas, *Development of telehealth dialogues for monitoring suicidal patients with schizophrenia: Consumer feedback,* Community Mental Health Journal **Brief Report**, 1 (2013).

[7] H. Kordy, M. Backenstrass, J. Hüsing, M. Wolf, K. Aulich, M. Bürg, B. Puschner, C. Rummel-Kluge, and H. Vedder, *Supportive monitoring and disease management through the internet: An internet-delivered intervention strategy for recurrent depression,* Contemporary Clinical Trials **36**, 327 (2013).

[8] I. L. Kampmann, P. M. Emmelkamp, D. Hartanto, W.-P. Brinkman, B. J. Zijlstra, and N. Morina, *Exposure to virtual social interactions in the treatment of social anxiety disorder: A randomized controlled trial,* Behaviour Research and Therapy **77**, 147 (2016).

[9] U. Yasavur, C. Lisetti, and N. Rishe, *Let's talk! speaking virtual counselor offers you a brief intervention,* Journal on Multimodal User Interfaces **8**, 381 (2014).

[10] D. Hartanto, W. P. Brinkman, I. L. Kampmann, N. Morina, P. G. Emmelkamp, and M. A. Neerincx, *Home-based virtual reality exposure therapy with virtual health agent support.* Pervasive Computing Paradigms for Mental Health **204**, 85 (2015).

[11] A. Bresó, J. Martínez-Miranda, and J. M. García-Gómez, *Leveraging adaptive sessions based on therapeutic empathy through a virtual agent,* in *ICAART - Doctoral Consortium* (2014).

[12] M. L. Tielman, M. A. Neerincx, R. Bidarra, B. Kybartas, and W.-P. Brinkman, *A therapy system for post-traumatic stress disorder using a virtual agent and virtual storytelling to reconstruct traumatic memories,* Journal of Medical Systems **41** (2017).

[13] A. P. Association, *Diagnostic and Statistical Manual of Nental Disorders, Fifth Edition* (American Psychiatric Association, 2013).

[14] E. Foa, E. Hembree, and B. Olaslov Rothbaum, *Prolonged Exposure Therapy for PTSD: Emotional Processing of Traumatic Experiences Therapist Guide (Treatments That Work)* (Oxford University Press, 2007).

[15] S. G. Hofmann and M. W. Otto, eds., *Cognitive Behavioral Therapy for Social Anxiety Disorder* (Routledge, 2008).

[16] E. B. Foa, E. A. Hembree, and B. O. Ro, *Prolonged Exposure therapy for PTSD* (Oxford University Press, 2007).

[17] N. Hundt, T. Barrera, A. J., and S. M.A., *"it's worth it in the end": Veterans' experiences in prolonged exposure and cognitive processing therapy,* Cognitive and Behavioral Practice **24**, 50 (2017).

[18] M. B. Powers, J. M. Halpern, M. P. Ferenschalk, S. J. Gillihan, and E. B. Foa, *A meta-analytic review of prolonged exposure for posttraumatic stress disorder,* Clinical Psychology Review **30**, 635 (2010).

[19] S. Provoost, H. M. Lau, J. Ruwaard, and H. Riper, *Embodied conversational agents in clinical psychology: A scoping review,* JOURNAL OF MEDICAL INTERNET RESEARCH **19**, e151 (2017).

[20] M. M. Morandell, A. Hochgatterer, S. Fagel, and S. Wassertheurer, *Avatars in assistive homes for the elderly a user-friendly way of interaction?* in *4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society* (2008).

[21] S. Zhou, B. T., A. Rubin, K. Yeksigian, R. Lippin-Foster, M. Heilman, and S. Simon, *A relational agent for alcohol misuse screening and intervention in primary care,* in *A Relational Agent for Alcohol Misuse Screening and Intervention in Primary Care CHI'17 Workshop on Interactive Systems in Healthcare (WISH)* (2017).

[22] A. A. Rizzo, B. Lange, J. G. Buckwalter, E. Forbell, J. Kim, K. Sagae, J. Williams, B. O. Rothbaum, J. Difede, G. Reger, T. Parsons, and P. Kenny, *An intelligent virtual human system for providing healthcare information and support,* Medicine meets Virtual Reality (2011).

[23] N. Yee, J. N. Bailenson, and K. Rickertsen, *A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces,* in *CHI 2007* (2007).

[24] T. W. Bickmore, R. A. S. K. Nelson, D. M. Cheng, M. Winter, L. Henault, and M. K. Paasche-Orlow, *A randomized controlled trial of an automated exercise coach for older adults,* Journal of the American Geriatrics Society **61**, 1676 (2013).

[25] R. M. Ryan and E. L. Deci, *Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being.* American Psychologist **55**, 68 (2000).

[26] R. M. Ryan and E. L. Deci, *A self-determination theory approach to psychotherapy: The motivational basis for effective change,* Canadian Psychology **49**, 186 (2008).

[27] Walter, Ortback, and Niehaves, *Designing electronic feedback - analyzing the effects of social presence on perceived feedback usefulness,* (2015).

[28] J. Prochaska and C. DiClemente, *The transtheoretical approach,* (Oxford University Press, 2005) pp. 147–171.

[29] W. R. Miller and S. Rollnick, *Motivational interviewing: Preparing people to change addictive behavior* (New York: Guilford Press, 1991).

[30] S. A. Friederichs, A. Oenema, C. Bolman, J. Guyaux, H. M. van Keulen, and L. Lechner, *Motivational interviewing in a web-based physicalactivity intervention: questions and reflections,* Health Promotion International (2013).

[31] O. A. Blanson-Henkemans, C. A. van der Mast, P. J. van der Boog, M. A. Neerincx, J. Lindenberg, and B. J. Zwetsloot-Schonk, *An online lifestyle diary with a persuasive computer assistant providing feedback on self-management,* Technology and Health Care **17**, 253 (2009).

[32] D. Schulman, T. Bickmore, and C. Sidner, *An intelligent conversational agent for promoting long-term health behavior change using motivational interviewing,* To appear.

[33] C. Horsch, *A virtual sleepcoach for people suffering from insomnia*, Ph.D. thesis, Delft University of Technology (2016).

[34] R. de Vries, K. P. Truong, and V. Evers, *Crowd-designed motivation: combining personality and the transtheoretical model,* in *International Conference on Persuasive Technology* (2016).

[35] R. de Vries, K. Truong, C. Zaga, J. Li, and V. Evers, *A word of advice: how to tailor motivational text messages based on behavior change theory to personality and gender,* Personal and Ubiquitous Computing **21**, 675 (2017).

[36] M. Busch, E. Mattheiss, M. R. Reisinger, and M. Tscheligi, *More than sex: The role of femininity and masculinity in the design of personalized persuasive games,* in *Persuasive Technology* (2006).

[37] J. E. Maddux and R. W. Rogers, *Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change,* Journal of Experimental Social Psychology **19**, 469 (1983).

[38] K. Glanz, B. K. Rimer, and K. Viswanath, eds., *Health Behavior and Health Education* (Jossey-Bass, 2008).

[39] E. Grigore, A. Pereira, and B. Scassellati, *Modeling motivational states in adaptive robot companions,* in *AAAI Fall Symposium Series* (2015).

[40] L. Süssenbach, N. Riether, S. Schneider, I. Berger, F. Kummert, I. Lütkebohle, and K. Pitsch, *A robot as fitness companion: towards an interactive action-based motivation model,* in *The 23rd IEEE International Symposium on Robot and Human Interactive Communication* (2014).

[41] F. Weathers, B. Litz, T. Keane, P. Palmieri, B. Marx, and P. Schnurr, *The PTSD Checklist for DSM-5 (PCL-5)* (2013), scale available from the National Center for PTSD at www.ptsd.va.gov.

[42] M. Tielman, M. A. Neerincx, and W.-P. Brinkman, *Dataset belonging to the paper: Generating personalized motivational messages for a virtual agent that assists in post-traumatic stress disorder therapy,* (2017), doi: 10.4121/uuid:3396c380-6102-4043-9693-847a3e482e03.

[43] W.-P. Brinkman, *Current and future research directions in mental health computing,* in *Int. conf. on Computation and Communication advancement* (2013).

[44] T. M. Baker, D. H. Gustafson, and D. Shah, *How can research keep up with ehealth? ten strategies for increasing the timeliness and usefulness of ehealth research,* Journal of Medical Internet Research **16** (2014).

[45] K. M. Kapp, *The Gamification of Learning and Instruction*, edited by R. Taff (ASTD and Pfeiffer, 2012).

[46] L. C. Zandbelt, E. M. Smets, F. J. Oort, M. H. Godfried, and H. C. de Haes, *Satisfaction with the outpatient encounter,* J GEN INTERN MED **19**, 1088 (2004).

[47] M. A. Hall, B. Zheng, E. Dugan, F. Camacho, and K. E. Kidd, *Measuring patients' trust in their primary care providers,* Medical Care Research and Review **59**, 293 (2002).

[48] M. L. Tielman, W.-P. Brinkman, and M. Neerincx, *Design guidelines for a virtual coach for post-traumatic stress disorder patients,* in *Intelligent virtual agents* (2014).

[49] W. Wang, *Self-management Support System for Renal Transplant Patients*, Ph.D. thesis, TU Delft (2017).

[50] R. J. Beun, W. P. Brinkman, S. Fitrianie, F. Griffioen-Both, C. Horsch, J. Lancee, and S. Spruit, *Improving adherence in automated e-coaching,* Persuasive Technology **9638**, 276 (2016).

6

Considering patient safety in autonomous e-mental health systems – Detecting risk situations and referring patients back to human care

Background
Autonomous e-mental health (AEMH) have the opportunity to tackle some of the barriers to care in mental health today, such as cost, availability and stigma. However, because they lack the involvement of human care-givers, they also carry the responsibility of ensuring patient safety. Many current AEMH systems still employ a human to detect and deal with risk situations. Before AEMH systems can take over these tasks, it is important they have clear safety protocols in place.
Methods
Given the many different AEMH systems, it is important that any generic protocol for detecting risk can be adapted to both the modality of communication within the AEMH, and the type of risks involved. We therefore present a detection model that takes these two adaptation requirements into account and gives recommendations on how to detect risks accordingly . After a risk is detected, patients might need to be referred back to human care. This can happen automatically, the conditions and procedure are described in an auto-referral model. However, auto-referral is not always preferable or even

possible. We therefore present a third model, which is aimed at convincing the patient to contact a human caregiver. This motivation model assumes that both system referral tactic and patient situation are important when trying to convince them to seek help. This model was tested in an online study (n = 160). After reading a scenario of the situation, participants were presented with a virtual agent that facilitated referral, tried to persuade the user to seek care or accepted that they did not wish human care. After seeing the virtual agent, participants were asked for their intentions.

Results

Three models concerning risk in AEMH systems were developed. Results from an experimental study show that the tactic a virtual agent follows in convincing users to seek care, influences their intention to seek care, intention to interact with the agent again and the feeling of being heard by the agent.

Condlusion

In this paper we present safety protocols for detecting risk situations in AEMH systems, and how to refer users back to human care when necessary. Given the great rise in AEMH systems and the vulnerability of their user group, it is important that such safety protocols are in place. We envision that future AEMH systems benefit from the protocols as starting point for their own safety procedures.

6.1. Background

Advances in technology are creating opportunities for computer systems to provide health care. These systems are increasingly used in mental health care as well as regular health care. Human involvement in the care processes will change and might decrease, due to the rise of autonomous e-mental health (AEMH) systems. These systems provide unique opportunities, but also come with their own practical and ethical challenges.

The WHO states that mental disorders account for 13% of the global burden of disease and hinder economic development on a global level. On a personal level, mental disorders can cause severe problems such as unemployment (with rates up to 90%) and high mortality risks [1]. Yet between 35 and 50% of people in high-income countries with a severe mental disorder receive no treatment. In low-income countries, these numbers rise to 76 and 85%. Several factors contribute to these high numbers, such as the cost of health care, the availability of therapy and the accessibility [2]. Another important issue is the stigma on mental-health care, which stops people from seeking help [3]. AEMH's are interesting, as they provide a way of breaking down these barriers to care, helping people at home or on their mobile phone, for comparatively low costs [4].

AEMH's also provide a challenge, as human caregivers are minimally involved. This means that the AEMH themselves need to be able to monitor and assess risks. Moreover, if a situation is detected that the system is not equipped to deal with, it should take appropriate action [5]. This fits in with the concept of stepped care. The AEMH's are a low-level entry to health care, but might scale-up to full human-based interventions only if necessary, thus saving cost and saving human-resources for the more serious cases [6]. Current interventions incorporating AEMH systems have a range of ways to deal with situations beyond the scope of the system, often depending on the target audience, the technological possibilities and the main purpose of the system.

Safety in AEMH - current procedures

Some AEMH systems are specifically aimed at getting people to seek help from health care professionals. These systems generally employ a combination of questionnaires and free text interaction aimed at assessment. They are designed to provide the user with information and a low-level entry point with the ultimate goal to refer them to a human care-giver. They reply to either text or physiological input from the user and provide them information on their state [7, 8], as well as informing them of the possibilities of professional human care [9]. As the main goal of these systems is to refer people, their will try to convince people to contact a professional in most situations.

After assessment, the next step is usually therapeutic treatment or medication. AEMH's offering therapies are being developed for several disorders. Safety procedures are mostly seen in therapies for disorders with increased suicide risk. AEMH systems generally focus on automatization of the intervention, but the safety procedures often still include a human component. An example of a system that does fully automate the safety procedure is the Help4Mood project, which focuses on

developing a home-therapy system for patients with depression [10]. This system is equipped to screen for suicide risk. If there is a small chance of risk, the system will provide the user with a list of options to deal with the situation. These options range from low-level options such as taking a bath, to a phone number of the 24/7 suicide help line with human caregivers on the other side. Similar work was done with SUMMIT, which focused on reducing depressive episodes [11]. However, in this case a human therapist is involved, screening the online forum that is a part of the system. 24/7 help options are made available to users and contact is monitored and screened. In serious cases, a caregiver is automatically alerted. This screening by human caregivers is also applied in an application to reduce suicide risk [12]. In a therapy system for PTSD [13], patients are given the option to always contact a human caregiver, and this same caregiver is able to monitor the questionnaire scores of all patients. Another example of a system for PTSD incorporates weekly phone calls to a therapist as check-up [14]. Even when patients will mainly take medication and not follow therapy, AEMH's might come into play for monitoring reasons. The Health Buddy, a monitoring tool for schizophrenic suicidal patients, is a mobile device in which patients answer questions. Answers are monitored by human clinicians who decide if interference is necessary [15, 16].

Patient safety is important for mental health care interventions, and in many AEMH systems there is still a human in the loop who does the risk assessment. These blended-care solutions are very promising, as they already reduce therapist resources while keeping some of the advantages of human care. However, these blended-care solutions might not always be ideal, or even possible. One of the goals of AEMH systems is to reduce the cost and accessibility of therapy. Scalability is crucial to achieve both these reductions fully, but having a human in the loop severely limits the scalability of an intervention. Another important advantage of AEMH systems is that they might be used anonymously, lowering the threshold for patients afraid of stigmatization. Web-based applications are promising because they combine this anonymity with being widely accessible [17, 18]. However, a web service that can be used anonymously cannot incorporate a human caregiver to check safety. It is therefore important to also have a better understanding on how these AEMH systems themselves should deal with situations beyond their scope. How to determine if a situation requires scaling up to human care, and how to ensure that this scaling up actually occurs.

Safety protocols

Risk situations occur in all types of mental-health care, and to understand how AEMH's should deal with them regular care can be a good starting point. The main trouble in translating regulations that exist in regular care to an AEMH, however, is that these regulations are often not specific enough. They are formulated as guidelines that still rely heavily on the human ability to interpret the content in relation to the context . An example is the guideline that a caregiver should explicitly ask for suicidality when a person expresses despair [19]. An AEMH, however, would then first need an exact formalization of the expression of despair. Another possible starting point are study protocols aimed at risk prevention, as these need to be

very explicit and unambiguous for scientific purposes. Also useful are systematic analyses of human risk screening, as they attempt to formalize how care givers screen for dangerous situations.

Study protocols generally identify several different steps in the risk detection process. A first step is information gathering, as identified by Knight et al. [20] and Sands et al. [21] in studies considering practices in risk assessment by phone. Belnap et al. [22] describe a study protocol for suicide screening and incorporate triggers as a first step. In all cases, this first step is meant to identify whether a problem might exist. In phone interviews information gathering is generally described as conversation and open-ended questions, which are less suitable for translation into formal protocols. However, study protocols as described by both Belnap et al. and Rollman et al. [23] identify two possible manners of initial information gathering for suicide detection; namely spontaneous mention of the risk or a flag on routine screening with scales. After this first step, Belnap et al., Sands at al. and Knight et al. all identify the decision making stage, or the actual risk assessment. In this stage the professional makes the decision whether a risk actually exists. Sands et al. determine several factors important in making this decision, including self-report measures of behaviour, mood and thought, as well as the duration of the complaint and a person's history. In the case no professional is included to ask open questions, as in Belnap et al., a specific scale is used to determine risk. If the decision is made that action needs to be taken, the final step is to meet the patient's needs. This can be to automatically contact a clinician, as advised by Belnap et al. in case of suicide risk, but also to simply refer a patient, as mentioned by Sands et al. It is important to note that when referring a patient, the patient does need to be informed of this plan first, and needs to consent. Knight et al. describe recommendations instead of referral, asking the patient to contact a human caregiver. They also note that the caregiver might offer advice for symptom management. As these studies all consider the formalisation of risk detection for mental health, they give valuable insights into what formal protocols for risk assessment might look like for AEMH systems. Dividing risk detection into an information gathering and a decision making stage, using scales to screen for disorders, and informing patients before automatically referring them, are all aspects that are incorporated into the protocols presented in this paper.

The following section of this paper proposes three theoretical models describing how an AEMH can detect risk and subsequently refer patients to human care. This referral can be done automatically by the system, or the system can ask the patient themselves to take action. This last framework takes the standpoint that it is important to take into consideration a person's stance towards seeing a human caregiver and the severity of their situation. This situation should be taken into account by the system when choosing a tactic. In some situations persuasion is important, in some this is not necessary and human care can directly be facilitated, and in some situations it might be best to accept that the user really does not wish for human care. An experiment was conducted to study the hypotheses behind this model. Section 3 presents the results of this study, followed by a general discussion and conclusion.

6.2. Theoretical models

Structured interviews were held as starting point to developing theoretical models for safety procedures for AEMH systems. Eight experts took part, with a scientific background in clinical psychology or human-computer interaction, or with experience as therapist using technology for treatment. The conversations took suicide risk in PTSD therapy as starting point, but also allowed for discussions about other types of risks and systems. The main discussion points that arose from these interviews were presented in a workshop on e-health for virtual agents [24], and were further discussed with the audience. The results of these discussions were combined with previous research on safety protocols for mental health care, as discussed in the previous section. This resulted in three theoretical models that describe risk detection and the following actions by AEMH systems. These models were specifically designed to be applicable to a broad range of mental health problems, risks and systems. As such, the models are general and do not, for instance, cover instruments to measure specific risks. This means that they should be taken as a starting point and need to be operationalized for every specific AEMH they are applied to. This might mean filling in the blanks, but in some cases also adapting the protocol to better fit the solution. The theoretical models serve as a starting point to implement safety procedures.

The following sections describe three theoretical models of how to deal with risk situations in AEMH systems. The first model describes how to detect that a situation might warrant the attention of a human caregiver. After such a situation is detected there are two options. One would be to automatically contact a caregiver. The second model describes the procedure this would entail. The other option is described in the third model, wherein the patient themselves are asked to seek human care. This option is especially relevant in systems where not all patient data are known, such as anonymous web services, as auto referral is not an option in these systems. However, it is important that the system motivates the patient to seek help as effectively as possible. The motivation model describes this process.

6.2.1. Detecting Risk

The *detection model* is presented in Figure 6.1. It is important to note that this model does not identify exactly what risk situation is to be detected. Before the detection model can be applied, it is important to establish exactly what risks exist in the user group. The most common example is suicide, which for instance is a risk for PTSD patients [25], but some systems might also wish to detect other risks, such as substance abuse [26]. After this has been established, the detection model identifies two steps. The first is detecting that a possible risk exists, depicted in the top half of the model. The second step is to determine if that risk situation is severe enough to warrant referral, depicted in the bottom half of the model.

Exactly how to detect a risk situation depends on several factors that might differ per AEMH system. The first factor is what type of input the AEMH system gets from the user. Some systems allow free text input from the user [8] and work as a type of chatbot, while others work with closed questions and questionnaires [13]. A system with free text input is capable of screening the input on keywords and

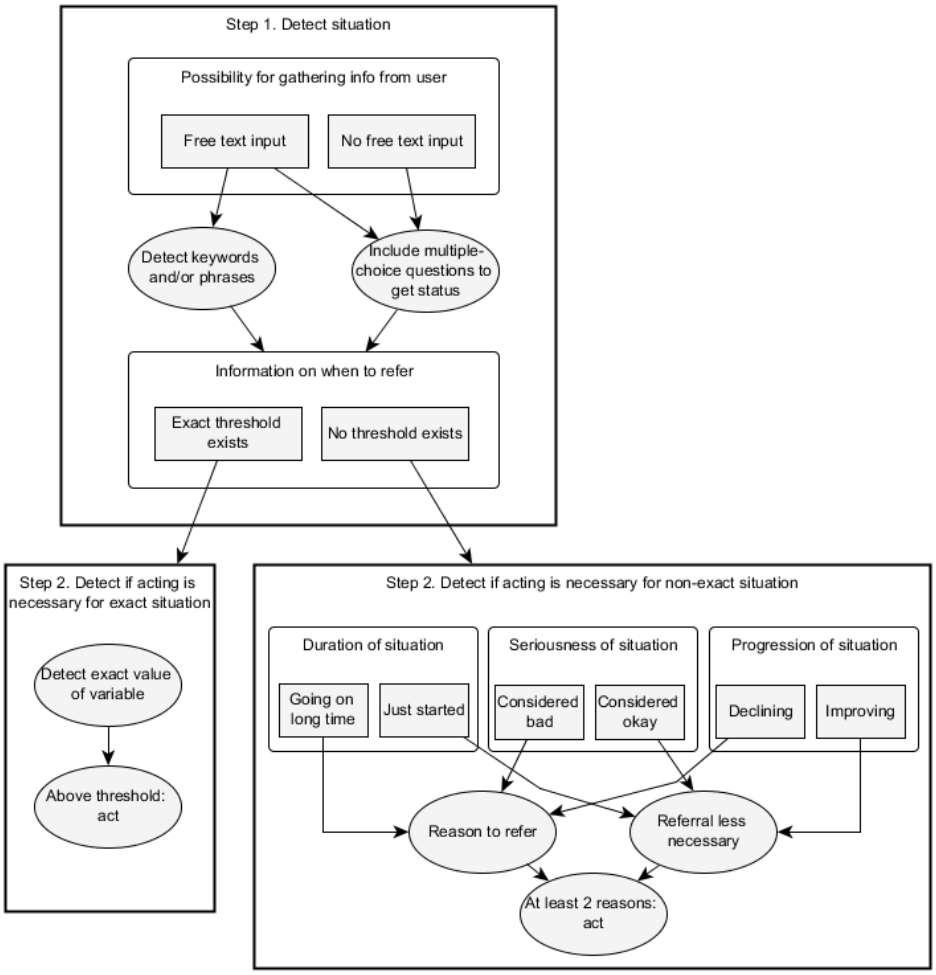

Figure 6.1: Detection model. Step 1 is to detect if a risk might exist, which can be done by text detection or by including specific questions, depending on the possibilities within the system. Step 2 is to detect if the situation is severe enough to refer to human care. If a threshold value on a questionnaire exists, this can be used for the decision. If not, the duration, severity and progression of the situation could be taken into account to make this decision. Although dependent on exact implementation and situation type, the model gives the guideline of referring if at least two of these values are negative.

phrases related to risk situations. Additionally, it might include specific questions for risk screening. A system without free-text input would need to rely on the second screening method.

The second factor determining how to detect a risk is if there is a measure for the risk situation that can be used to establish a threshold for referral. Questionnaires exist for many situations and their scores may be used to determine when a situation warrants human intervention. For instance, question 9 of the patient-health questionnaire for depression has been shown to be correlated with suicide [27]. The exact threshold score needs to be established in concord with clinical specialists. It might also be necessary to adapt the threshold during use if it becomes clear that too many or too few situations are detected. It is important to note that posing questionnaires to establish risk thresholds might only be appropriate after a possible risk situation has already been detected, which is why this action is not located in the detection part of this model. Especially for suicide, posing unwarranted questions about suicidal ideation might not be the preferred approach. Suicide contagion is a well-studied phenomenon where hearing about suicide prompts suicide people to commit suicide [28]. This means that care should be taken to only pose suicide questionnaires when the situation already indicates this might be a problem.

Because questionnaires are not always available, this model also describes a second approach to determining if a risk situation is severe enough for referral. This approach is meant as a guideline of what to take into account while making this decision. However, the exact procedure might have to be changed to reflect specific risk factors and situations. For instance, suicide will be a cause for referral quicker than sleeping problems, and when implementing this model care should be taken to reflect such differences. This model does not claim to be directly applicable to all types of system s and risks, but does offer a starting point in situations where exact scales and questionnaires are not applicable. The model considers three factors in determining if a situation warrants referral to human care. The first is duration of the situation, how long has this situation been going on. The second factor is the severity of the situation. The third is the progression, is the situation getting worse or getting better. As a default, the model recommends that if at least two of these three factors are negative, the user should be referred to human care. Negative in this case is if the situation has been going on for a long time, it is severe and it is getting worse. Obviously, exactly what warrants as a severe situation, or what is a long duration, should be determined per system and situation. This also holds for situations that are in the middle, such as a situation that neither improves nor gets worse. Depending on the system and risks involved, the number of negative factors necessary for referral can also be expanded or reduced to better reflect the situations that might occur.

6.2.2. Auto referral

In some AEMH systems it is possible to automatically refer patients back to a human caregiver. In any case where automatic referral can take place, it is important that the patient is informed beforehand and has approved this procedure [5]. This ties

in with the ethical guidelines to protect the patient's confidentiality and privacy, and to respect their autonomy [29]. Figure 6.2. outlines the procedure for an automatic referral.

The recommended actions for automatic referral firstly depend on whether a human caregiver is available to directly step in, for instance by taking over from a chatbot [30]. If the system is able to directly escalate the interaction this way, the AEMH system merely needs to inform the patient that they will be interacting with a human, before letting the caregiver take over. Such escalation is not always possible, however, for instance because a caregiver is not always available or the system does not allow direct intervention. In this case it is important that the users of the AEMH system are known to caregivers before they start [11, 31], as contact details of the user need to be available to the caregiver. In these cases, the model distinguishes between direct crisis situations and situations without direct crisis. In the first case, the system directly contacts the human caregiver involved, informs the patient of that fact, and also states when the patient can expect contact. It is important to manage this expectation as caregivers might not always be able to react quickly. In this case care should be taken that the patient does not believe they are forgotten by their caregiver. Additionally, the system should provide options for the patient to contact the caregiver themselves, as they might be able to respond quicker themselves. This is also relevant in cases of system or network failures. If the patient's network fails it might be impossible for the system to send out a message and the patient should know who to contact and how. Finally, the system should provide some short-term options for the patient to reduce the risk, in case it takes a little while for the caregiver to respond. In case no direct crisis is detected, the model recommends that the patients themselves are asked to contact a human caregiver. Additionally, a message is still sent to the caregiver in case the patient does not comply with the request. Only if the patient does not contact the caregiver within a certain time frame will the care giver seek contact.

6.2.3. Motivate to seek care

In some cases, an automatic referral to human care is not preferable, or even possible. Patient anonymity is often an advantage to AEMH systems because of the stigma that still surrounds mental health care [3]. In such anonymous situations the AEMH system does not know anything about the patients, except what they entered into the system themselves. If such a system detects a risk situation, it should generally try to persuade the patients to undertake action themselves and contact a human care giver. Figure 6.3 presents recommendations on how the system should adapt its strategy for motivation to the users situation.

The model presented in Figure 6.3. has three overall goals, the first is to get as many users as possible to seek human care. The second goal is to give the user the feeling that they are heard and taken seriously by the system. This ties in with the third goal, which is to get users to return to the system if they experience problems again in the future, or if their problems worsen. The model proposes that to achieve these goals, the system can adapt its tactic to the situation of the user. Three different situation types are distinguished by how likely users are to

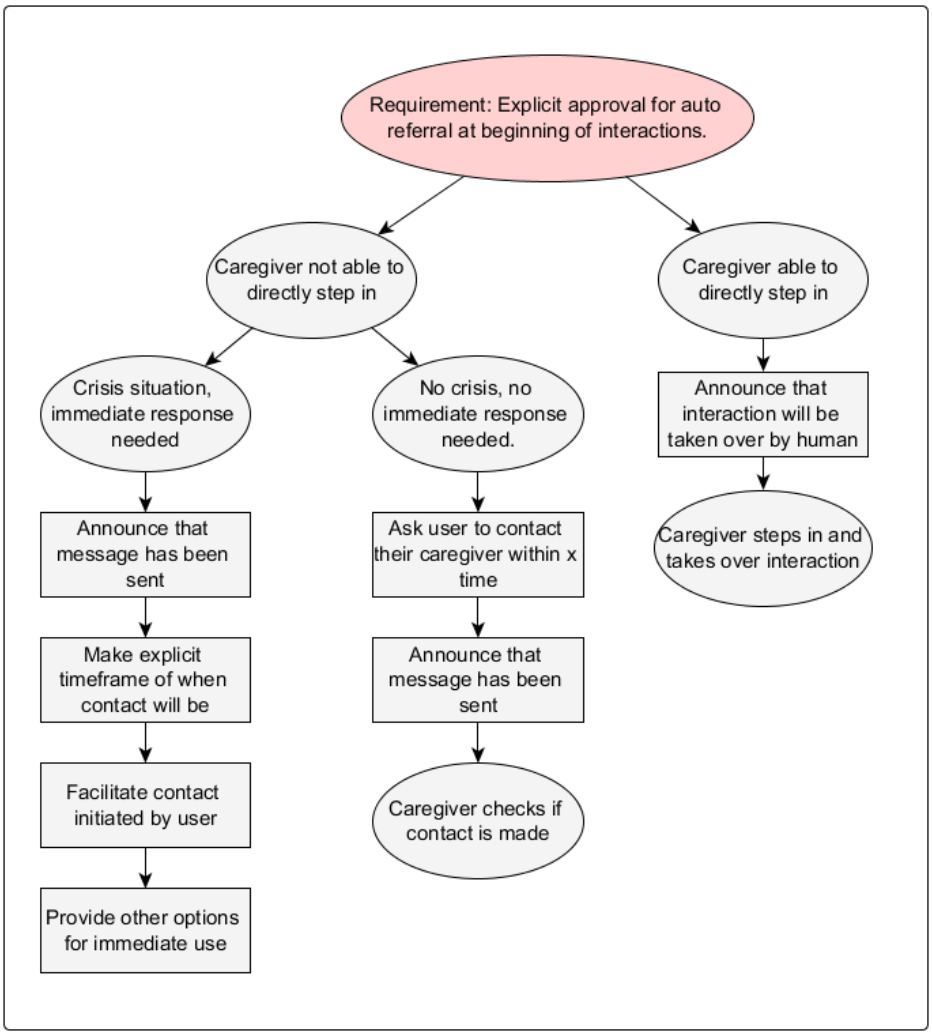

Figure 6.2: Auto referral model. To automatically refer patients it is important that they have given consent to do so beforehand. If a caregiver is able to directly step in, this will only have to be announced before they take over the interaction. If this is not possible, the model distinguishes between crisis and no crisis situations. In a direct crisis situation, the caregiver is alerted and prompted to act, while the user is also given contact information. If no direct crisis is present, the incentive for seeking contact lies with the user, while the caregiver is still informed as safety net.

seek human care and how important it is that they do so. In other words: how their stance is towards seeing a human and how severe their situation is. In every situation, the model proposes a different tactic. Both this division into situation-based tactics and the specific actions within each tactic are based on several theories of behaviour change. Social judgement theory first identified the concept of latitude of acceptance, which defines the space of options a person finds acceptable [32]. If a suggestion is made in a person's latitude of acceptance, the user will accept this suggestion. Aside from this latitude of acceptance a user also has a latitude of noncommitment and latitude of rejection. Any suggestion made in the latitude of rejection will fail, and additionally make the user even more opposed to the idea. The latitude of noncommitment lies in-between. The three situation types identified by the model each place the user in one of these three latitudes. The first situation type is accept care, which is defined by a very positive stance towards seeking human care, placing the user in the latitude of acceptance. The second situation type is care potential, which places the user in the latitude of noncommitment, where the suggestion to seek care still has potential. These situations are defined by doubt about seeking human care and either a negative or severe situation, or a negative stance and a severe situation. The final situation type is care rejected, which is defined by a not very severe situation and doubt about seeking care, or a negative stance towards seeking care and a medium or low severity situation. These place the user in the latitude of rejection, assuming that the suggestion to seek human care will be not be successful. The model assumes a different tactic for each of the situation types.

When users are in the accept care situation, the model follows the facilitate tactic. This tactic is based on two different theories of behaviour change. Fogg (2009) describes a theory where the combination of motivation and ability need to be high enough for a trigger to succeed [33]. Similarly Michie (2011) describe the behaviour change wheel that states that motivation, capability and opportunity all play a role [34]. To change behaviour, a tactic needs to change one of those three factors to be successful. In this accept care situation the motivation to seek care is high. The model will therefore increase the capability of the user to seek care by facilitating contact to a human caregiver, for instance by providing contact details. This facilitation also serves as trigger to provide the user with the opportunity to seek care.

In the care potential situation a trigger is less likely to directly succeed, as motivation is lower. The model will therefore try to persuade the user in an attempt to increase the users motivation. After this, it will announce the facilitation of contact. If the user does not actively object, the model will automatically continue with triggering the user by providing opportunity to seek care. This means that the default option is to facilitate, although the user might still stop the system if they were not convinced by the motivation. This implementation of a default tactic is done because people have a tendency to go with a default option when given a choice [35].

For users in the *reject care* situation a different approach is taken. Because users are in the latitude of rejection, the suggestion of seeking human care might

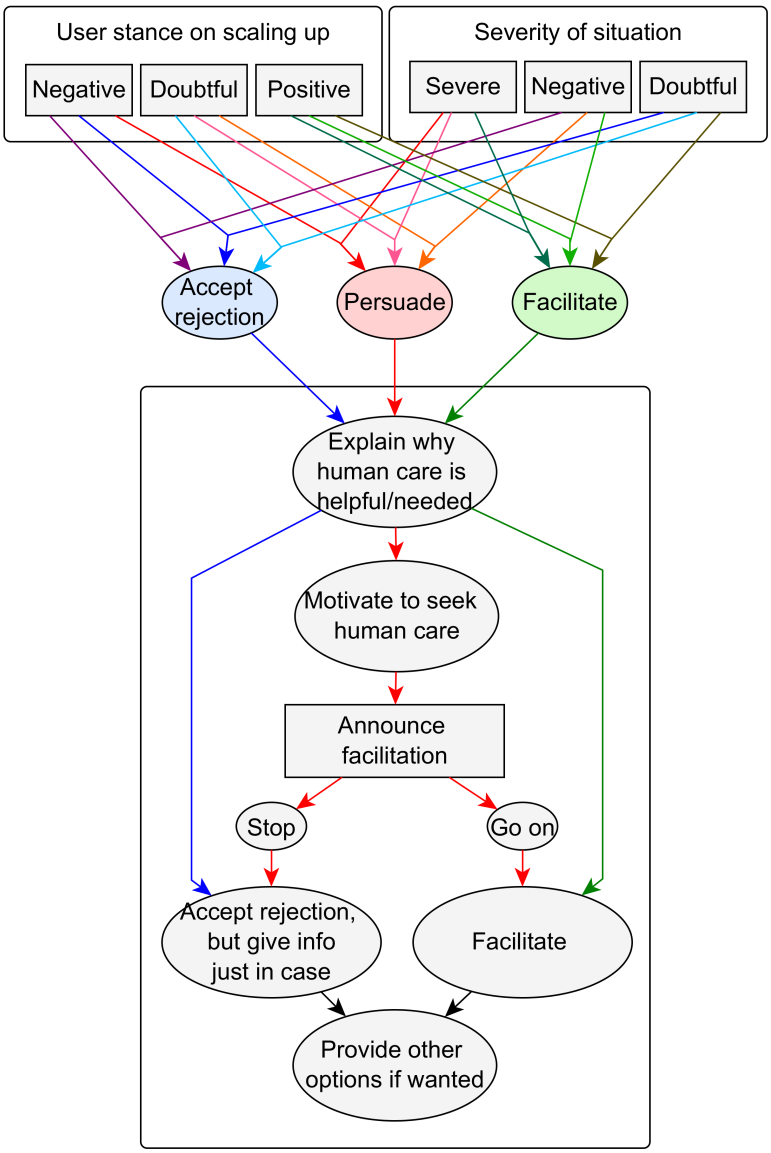

Figure 6.3: Motivate patient to seek human care.

Table 6.1: Situations set by health severity and user stance on scaling up.

Severity of the situation		*User stance on scaling up*		
		Negative	Doubtful	Positive
	Severe	Situation 1: care potential System tactic: persuade H2: increase seeking care	Situation 2: care potential System tactic: persuade H2: increase seeking care	Situation 3: accept care System tactic: facilitate H3: seeking health care likely
	Negative	Situation 4: care rejected System tactic: accept rejection H4: increase return and feeling being heard	Situation 5: care potential System tactic: persuade H2: increase seeking care	Situation 3: accept care System tactic: facilitate H3: seeking health care likely
	Doubtful	Situation 7: care rejected System tactic: accept rejection H4: increase return and feeling being heard	Situation 8: care rejected System tactic: accept rejection H4: increase return and feeling being heard	Situation 9: accept care System tactic: facilitate H3: seeking health care likely

have an adverse effect. Therefore the tactic is to *explicitly accept* that the user rejects human care, but to still give users the opportunity to contact a human in case they change their minds. In this way, the model aims at increasing the users intention to return to the system in the future, and how much the user feels the system really heard them. It should be noted that situation 1 as defined by Table 6.1 is not a *reject care* situation despite the low stance on scaling up. The reason for this is that in the situation is severe in this case, meaning that there is little to lose by trying to persuade the user despite the lower chance of that succeeding.

Table 6.1 shows the different situations as defined by stance on scaling up and situation severity, along with the situation type, the matching tactics and underlying hypotheses. These four hypotheses of the motivation model are as follows:

- H1: Agent personalisation in terms of its communication tactic, based on the situations characterised by user stance on seeking care and the severity of health risk situation, has an effect on the user's intention to seek health care, their likeliness of contacting the agent again, and their feeling of being heard by the agent.
- H2: In situations with care potential, users are more inclined to seek health care if the agent provides persuasive messages instead of given no messages, or providing only referral information.
- H3: In situations where users are likely to accept health care, they are inclined to seek health care if the agent facilitates referral.

- H4: In situations where users are likely to reject advise to seek (additional) health care, users are more likely to contact the agent again, and feel that the agent has heard better them if the agent accepts this rejection instead of persuading to seek care or only facilitating referral.

6.3. Method

The motivation model is particularly suited for empirical evaluation, given the four hypotheses that underlie it. Both the detection and the auto-referral model are less suitable for empirical evaluation, however, as they focus on user response. These models provide a framework for the implementation of procedures and technology. Whether these procedures are suitable for a specific AEMH system would have to be determined per intervention. Therefore this section describes an experiment performed to study the hypotheses presented in the previous section.

An experiment with a 3×3×3 design was performed, studying the effect of situation severity, stance on seeking human care and agent tactic on intention to seek care, intention to return to the agent and the feeling of being heard. Hypothesis 1 was tested over all data, while hypotheses 2, 3 and 4 were tested on subsets of the data representing the relevant situations. Participants were recruited online from the general population and were asked to imagine scenarios representing the situation types. The domain chosen for this experiment was that of sleeping problems as many people have had at least some measure of experience with sleeping problems, and about a third meet at least one criteria for clinical insomnia in their lifetime [36]. The AEMH system used for this experiment was a virtual agent with chat function.

6.3.1. Participants

229 participants were recruited via Amazon's Mechanical Turk and paid 1$ for their time. Participants were excluded if they had not read the consent form properly, if they could not confirm having seen the virtual agent, if they did not complete the survey, or if they did not complete the chats with the virtual agent. 160 participants were eventually included in the study, all were native English speakers. Age, gender and insomnia scores for all participants can be found in Table 6.2.

6.3.2. Measures

Primary outcomes

Intention to seek care (ISC): Participants were asked to indicate how likely they would be to seek human care after the conversation with the agent, given their imagined situation. This question was answered on a 7-point scale ranging from *Extremely unlikely* via *Neutral* to *Extremely likely*.

Intention to contacting agent again (ICAA): The second primary outcome measure was studied with the question how likely the participant would be to contact the virtual agent again if they would have sleeping problems in the future. This question was answered on a 7-point scale ranging from *Extremely unlikely* via *Neu-*

tral to *Extremely likely*.

Feeling being heard (FBH): Finally, a measure was in place to study how much participants felt the virtual agent heard them and took them seriously. This was measured with a seven item questionnaire, all questions answered on a 7-point scale ranging from not at all via neutral to very much. This questionnaire contained two questions from the patient satisfaction questionnaire (PSQ) [37] (number 1 and 4), two from the trust in physician scale [38](number 5 and 8) and three additional questions. Questions from the PSQ and trust in physician scale were slightly adapted to apply to virtual agents instead of doctors, were phrased as statements and negative statements were phrased positively to avoid double negatives in the scale answers. Reliability for this scale was measured for all 9 situations as participants filled in the scale three times for different situations. Cronbach's alpha was between α = 0.93 and α = 0.97, showing high reliability.

Descriptive measures
Two descriptive measures were in place. Firstly, the insomnia severity index (ISI) was presented to all participants at the start of the experiment to check if experience with insomnia made a difference in the behaviour of the participants. Secondly, all participants were asked what their main reason would be to not seek human care. The options were money, travel time and social stigma. These options were later used to manipulate the situations, so if the answer was money the situation would refer to the potential cost of therapy.

Manipulation check
During the interaction, the virtual agent asked the participant how severe they would rate their situation, and how they felt about seeking human care during their interaction. These questions were in place to check if the participant's interpretation of their situation was as intended in the scenario they were presented with. They were implemented as multiple-choice options with three options so the answers could be compared to the three-tiered design of the scenarios.

6.3.3. Virtual agent chat
The virtual agent chat was realized with the Roundtable Authoring tool developed by the USC Institute for Creative Technologies [1] .The chats included questions asking after the severity of the situation and motivation of the participant to seek human care. A female virtual agent was displayed on the left of the screen while the chat was displayed on the right. Figure 6.4 shows the virtual agent and examples of the text in all three tactics.

6.3.4. Procedure
Because it was not known beforehand if participants suffered from sleeping problems, let alone how severe these were or how people would feel about seeing a

[1] https://authoring.simcoach.org

General format:

To help me understand you even better, could you tell me how serious you would rate your problem?

A severe problem, very serious.

Thank you. Given what you've told me about your problem, I think it would be best if you could talk to a human therapist abou this. I'm only a virtual coach, so I can't always give the optimal help. How would you feel about seeing a human therapist?

I'd like that very much

For now, would you like me to give some tips to help you through?

Yes

Some things that might help you sleep a bit better are the folloowing: Try not to use a computer or phone in the hour before you go to bed, avoid caffeine in the evening and if thoughts keep running through your head write them down.

Facilitate

Thank you. The therapist I would reccomend for youis Dr. Smith. You can immediately contact them with this email: ecoach_refer@smiththerapy.com. They will know I referred you and immediately schedule your appointment. It's best if you contact them directly after our chat.

Persuade

Almost everyone has some reservations about seeing a therapist. What would your biggest reason be for not seeing a therapist, money, travel time or what the people around you would say?

Money

Thanks for telling me. That's a valid concern.
If a therapist is too expensive for you, you can often apply for financial aid. Depending on your financial situation it's possible to get a part of the therapy refunded. This can be arranged for you by the clinic you visit.
Does this ease your concerns a bit?

No

The clinic also lets you try one session first before you decide for the rest. If it's not worth the money, that's also an option.

I'll give you the contact information of the therapist you can see about your problem. The therapist I would reccomend for youis Dr. Smith. You can immediately contact them with this email: ecoach_refer@smiththerapy.com. They will know I referred you and immediately schedule your appointment. It's best if you contact them directly after our chat.

Accept rejection

Thank you. I understand seeing a therapist is not for everyone, that's okay. You can also come back to chat with me.
If you ever change your mind you can email ecoach_refer@smittherapy.com to schedule an appointment with a human therapist.

Figure 6.4: Virtual agent and the different tactics. Please note that small variations are possible. E.g. if people answer 'no' to the tips, these would not be given.

human therapist about it, participants were asked to imagine scenarios. These scenarios described a situation on the topics of severity of the sleeping problem and how likely the person was to seek human care. This second factor was personalized. At the start of the experiment, participants were asked if time, money or stigma would be the most likely factor to stop them from seeking help from a therapist. For those answering time or money, the next question would be how much time/money would be no problem, perhaps a problem or definitely a problem. In the case of stigma, participants were asked if telling their friends, family or boss would be the biggest problem. In the following scenarios, these variables were used to describe a situation in which seeing a human therapist would be no problem, perhaps a problem or definitely a problem. For example; if the participant's answers were money and *50 dollar would definitely be a problem*, the scenario describing a negative stance towards seeing a therapist would state that the only possible clinic would cost at least 50 dollar. Each participant was randomly presented with three different scenarios describing a level of severity and stance towards scaling up (See Appendix I for example of the scenario). For each scenario, participants were first asked to read the text and imagine the situation. Afterwards, they were redirected to a link where they could chat with a virtual agent. The virtual agent would either accept that they rejected care (even if they did not), try to persuade them to seek help, or immediately facilitate referral to human care. In this way, all possible 9 (3 motivation × 3 severity) situations were mixed with all three tactics. Which combinations a participant saw and in what order were randomized. After the chat, participants were asked a number of questions about how likely they would be to seek human care, how likely they would be to return to the agent and if they felt the agent really heard them.

6.3.5. Data preparation & Analysis

Data was analysed with R version 3.3. For both the feeling of being heard scale and the ISI scale scores to the questions were averaged to create a single score. For the FBH scale, the ICAA and ISC the answers were transformed into a score between -3.5 and 3.5 so deviation from the neutral 0 could be calculated. The full dataset can be found in: [39].

Descriptive outcomes

The primary outcome measures were studied to study whether users generally intended to seek care, to return to the agent and if they felt being heard. Multilevel analyses were done taking participant as random intercept to control for the multiple measures per participant. Null models were run to reveal the deviance from the neutral point of zero for all dependent variables, while situation severity and motivation to seek help were each added as fixed effect to reveal if they further affected the dependent variables.

Manipulation check

Every participant imagined three out of nine possible situations, varying in the severity of their problem and their motivation to seek human care. To ensure if these

scenarios were interpreted correctly, participants were also asked by the virtual agent how they would rate these factors. Cumulative linked mixed models with participant as random intercept were run to study if the situations in the scenarios correctly predicted how the situation was interpreted.

Covariate check

Covariate checks were performed for participant age, gender, ISI scores and the reason why they would not seek help from a human therapist. These reveal that only gender had an effect, namely on FBH score ($F(2, 156) = 4.01$ p=.02), male participants giving higher scores (Male M=1.15 SD=1.53, Female M=0.54 SD=1.70, Other M=0.29 SD=1.38).

Hypotheses

Multilevel analyses were conducted for each of the nine situations, including only a subset of two agent tactics (e.g. persuasion and reference only) on the three outcome measures: ISC, ICAA, and FBH. Null models included participant as random intercept and for FBH gender was included as fixed factor. The extended model also included agent tactic as fixed factor. The null models and extended models were compared to test the effect of agent tactic.

The first hypothesis was tested using a multilevel analysis on ISC, ICAA and FBH scores. This null model only included a fixed intercept and participant as random intercept. Model 1 built on the null model and added situation as fixed effect. Model 2 also built on the null model but added agent tactic instead. Model 3 built on model 2 and added situation as fixed effect as well. Finally model 4 was the full model that built on model 3 and included two-way interaction between situation and agent tactics as fixed effect. Comparing these models tested the added contribution of fixed effects.

The second hypothesis was tested with a multilevel analysis on intention to seek care data by only including the subset of situations with care potential, i.e. situations 1, 2 and 5. The null model included participant as random intercept and fixed intercept. Model 1 built on model 0 and added agent tactic. After comparing model 0 and 1, the analysis was repeated on two subsets: one that include only persuasion and no action as agent tactic, and the other subset that included only persuasion and referral only as agent tactic.

The third hypothesis was tested with a single multilevel analysis on data only including situations 3, 6 and 9, in which the agent followed the facilitate tactic. This analysis only included a null model on intention to seek care that included fixed intercept effect and participants as random intercept effects. The hypothesis was examined by considering whether fixed intercept significant deviated from zero.

The fourth and final hypotheses were tested with two multilevel analyses, for both how likely participants are to contact the agent again, and how much they felt being heard by the agent. A null model was run including fixed intercept effect and participants as random intercept. The hypothesis was examined by comparing the null models to a model also including agent tactic. Furthermore, null models and models including agent tactic were also compared in two subsets: one that include

Table 6.2: Descriptive statistics & one-sample t-test with bootstrapping results

Characteristics		F(1.318)	*p* value
Age in years, mean (SD)	36(10.41)		
Gender, n(%)			
Female	91 (56.88)		
Male	67 (41.88)		
Other	2 (1.25)		
Insomnia severity index score, mean (SD)	17 (6.53)		
No insomnia, n (%)	10 (6.31)		
Sub-threshold, n (%)	54 (33.46)		
Moderate severity insomnia, n (%)	55 (34,57)		
Severe insomnia, n (%)	41 (25,65)		
Main concern not seeking care			
Money, n (%)	128 (79.92)		
No problem, Mean in $ (SD)	24.76 (20.66)		
Maybe a problem, Mean in $ (SD)	39.68 (22.27)		
Problematic, Mean in $ (SD)	57.01 (27.59)		
Travel time, n (%)	15 (9.29)		
No problem, Mean in hours (SD)	0.67 (0.47)		
Maybe a problem, Mean in hours (SD)	1 (0.58)		
Problematic, Mean in hours (SD)	1.85 (1.07)		
Social stigma, n (%)	17 (10.78)		
Family, n (%)	5 (29.41)		
Friends, n (%)	3 (17.65)		
Boss, n (%)	12 (70.59)		
Intention to seek care, mean (SD)	0.76 (1.85)	52.28	<.001
Intention to contact agent again, mean (SD)	-0.03 (1.98)	0.05	.824
Feeling of being heard by agent, mean (SD)	0.79 (1.66)	52.09	<.001

only persuasion and no action as agent tactic, and the other subset that included only no action and referral only as agent tactic.

6.4. Result

Descriptive outcomes

Table 6.2 shows the demographic characteristics of participants. This table shows that the main concern for not seeking care was money. Average insomnia scores were high, with one-fourth of participants meeting the criteria for severe insomnia and one-thirds for moderate severity insomnia. For the dependent variables, ISC and FBH were both significantly above neutral, participants indicating that they would seek care in this situation, and that they felt being heard by the agent. ICAA scores do not show this pattern.

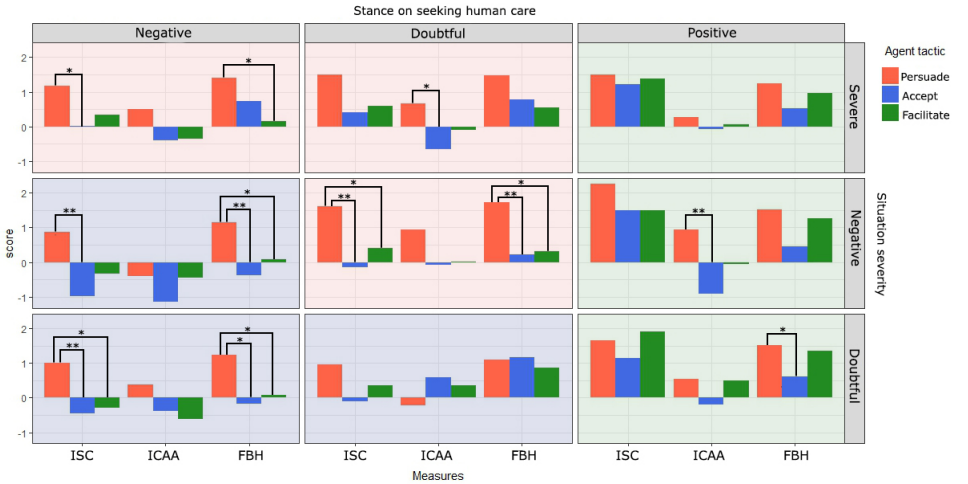

Figure 6.5: Comparison of agent tactics over the 9 different situations. * Represents p<.05, ** p<.01.

Manipulation check

Analysis shows that scenario value of the situation significantly predicted the subjective interpretation for both situation severity (χ^2_1 =142.98 p<.0001) and motivation to seek human care (χ^2_1=184.49 p<.0001). Pseudo R^2 was .14 for severity and .18 for motivation as calculated following [40], effect size was calculated using the χ^2test for goodness of fit was large for both severity at w=.55 and motivation at w= .62 [41]. This indicates that largely, situations as presented in the scenarios were interpreted correctly by the participants. For this reason, only the values as presented in the scenarios were considered in the remaining analyses.

Hypothesis 1: Personalisation agent communication tactics

Table 6.3 shows the results from the analysis for hypothesis 1. This table shows that both situation and agent tactic significantly influenced intention to seek care, intention to contact the agent again, and the feeling of being heard. The interaction between situation and tactic was not found to affect any of these measures. To further study interaction, this test was repeated on a datasets excluding one of the tactics, revealing that for *accept rejection* and *facilitate only*, an interaction was present for feeling of being heard (χ^2_8=18.46 p<.018), as for persuade and facilitate only (χ^2_8=18.91 p<.015). Examining Figure 6.5 it seems that especially for the *facilitate only* tactic, the interaction between tactic and situation influences how much participants feel being heard by the system, this tactic working better when stance on scaling up is higher.

Hypothesis 2: Persuasion in situations with care potential

Table 6.4 shows that in situations with care potential, the agent tactic significantly effects intention to seek care. Further analysis shows that the persuasion tactic significantly differs from the other two tactics, Figure 6.5 showing that persuasion

Table 6.3: Descriptive statistics & one-sample t-test with bootstrapping results

Model comparison	*n*	χ^2	*p*
Intent to seek care			
Situation (M_0 vs M_1)	477	$\chi^2_8 = 69.84$	<.001
Agent tactics (M_0 vs M_2)	477	$\chi^2_2 = 33.80$	<.001
Situation × Agent tactics (M_3 vs M_4)	477	$\chi^2_{16} = 13.67$	.623
Intent to contact the agent again			
Situation (M_0 vs M_1)	477	$\chi^2_8 = 18.30$	.019
Agent tactics (M_0 vs M_2)	477	$\chi^2_2 = 28.26$	<.001
Situation × Agent tactics (M_3 vs M_4)	477	$\chi^2_{16} = 11.80$	.758
Feeling of behing heard			
Situation (M_0 vs M_1)	477	$\chi^2_8 = 17.51$	.025
Agent tactics (M_0 vs M_2)	477	$\chi^2_2 = 47.02$	<.001
Situation × Agent tactics (M_3 vs M_4)	477	$\chi^2_{16} = 19.18$	.260

Table 6.4: Descriptive statistics & one-sample t-test with bootstrapping results

Model comparison	*n*	χ^2	*p*
Agent tactics (M_0 vs M_1)	173	$\chi^2_2 = 69.84$	<.001
Sub set agent tactics			
Persuasion vs Accept rejection (M_0 vs M_1)	111	$\chi^2_1 = 33.80$	<.001
Persuasion vs Facilitate s (M_0 vs M_1)	116	$\chi^2_1 = 13.67$	<.001

leads to higher scores.

Hypothesis 3: Providing referral information in situation where user likely to accept health care.
In situations 3, 6 and 9 as described in Table 6.1 , intention to seek care after interaction with the virtual agent that facilitated contact significantly deviated from the neutral value of zero ($F(1,147) = 59.04$, $p<.0001$). Figure 6.5. shows that participants would seek care in the *accept care* situation after the agent used the *facilitate* tactic

Hypothesis 4: No messages in situation where users reject advise to seek health care.
Table 6.5 shows the results from the analysis for hypothesis 4. In the *reject care* situation, agent tactic did not affect the participant's intention to contact the agent again. Tactic did have an effect on FBH score, Figure 6.5. shows that the persuasion tactic outperformed the accept rejection tactic.

Discussion results
The demographic results show that more than half of the participants met the criteria for clinical insomnia. These high numbers could be partly caused by a potential

Table 6.5: Descriptive statistics & one-sample t-test with bootstrapping results

Model comparison	*n*	χ^2	*p*
Intent to contact the agent again			
Agent tactics (M_0 vs M_1)	152	$\chi^2_2 = 1.00$	.605
Sub set agent tactics			
Accept rejection vs Persuasion (M_0 vs M_1)	99	$\chi^2_1 = 0.59$	.441
Accept rejection vs Facilitate (M_0 vs M_1)	102	$\chi^2_1 = 0.03$	.874
Feeling agent has heard them			
Agent tactics (M_0 vs M_1)	152	$\chi^2_2 = 16.60$	<.001
Sub set agent tactics			
Accept rejection vs Persuasion (M_0 vs M_1)	99	$\chi^2_1 = 12.03$	<.001
Accept rejection vs Facilitate (M_0 vs M_1)	102	$\chi^2_1 = 0.003$	.953

response bias, specifically the Hawthorne effect, where participants behave differently because they are in an experimental setting. The ISI is generally only applied after sleeping problems have already been reported, and people with minor sleeping problems might rate these as more severe simply because they anticipated this as more desirable for the experimenter. Even taking a conservative stance and expecting that severity of insomnia problems to be lower, it seems likely that some level of insomnia was existing in this sample indicating that the sleep domain was fitting for this participant group. Score on the ISI did not affect the outcome measures in the experiment, which indicates that having sleeping problems does not affect the interpretation of a scenario about sleeping problems.

Overall, both situation and agent tactic affected intention to seek care, intention to return to the agent and the feeling of being taken seriously. However, results show no interaction effect between situation and three tactics. Still, when comparing two tactics only, an interaction effect was found on FBH as long as the *facilitate tactic* is included. It showed that especially for the *facilitate only* tactic specifically, the feeling of being heard was influenced by the interaction between situation and tactic. This might indicate that situation does matter for the effect of the *facilitate* tactic, but less for the other two. These results therefore only partly support hypothesis 1. When considering the individual situation types, results show that users in the *care potential* situation are more likely to seek care when the agent uses the *persuade* tactic than if it uses *accept rejection* or *facilitate*. So when participants are in doubt whether they wish human care, persuading them to seek care is an effective tactic, confirming hypothesis 2. In the *accept care* situation, results show that participants are inclined to seek care after the *facilitate tactic*. This supports hypothesis 3, and indicates that if a user is motivated to seek care, simply providing information is enough to get them to seek contact.

The final situation is the reject care situation, in which motivation to seek care is low. The focus of the model lies on increasing the intent to contact the agent again and the feeling of being heard through the *accept rejection* tactic. Results show, however, that the *persuade* tactic has a better effect on the feeling of being heard,

while no difference between tactics is found for the intention of contacting the agent again, thus not confirming, or even giving grounds for rejecting, hypothesis 4. This rejection might, however, be too strong a conclusion. Hypothesis 4. was based on the assumption that if motivation is low the user is in the latitude of rejection for the suggestion to seek care. That the results do not reflect this assumption might have two causes. The first would be that motivation was indeed low, but for the suggestion to seek care, the latitude of rejection simply does not exist. In this case the model would need to be adapted to reflect this. Another option is that the manner in which motivation was manipulated in this experiment was not actually low enough to achieve the latitude of rejection. Motivation was framed in terms of money, time and how many people would be told of therapy. All of these concerns could be partially taken away by the agent in the *persuasion* tactic, however. In real-live situations, low motivation might be more difficult to dispel. The trans-theoretical model of behaviour change [42] identifies a stage where people do not agree their problem warrants action, in which case it would be more difficult to persuade people to seek care. Also, not all reasons to not seek care could be taken away as easily as done in the scenarios in this experiment. This could mean that the simulated low motivation situation in the experiment did not accurately reflect a real-life low motivation.

6.5. Discussion

In this paper three models are presented, which together cover how to detect risk situations in AEMH applications and how to act once those risks are detected. Based on existing models of risk detection, theoretical models from behaviour change and expert interviews, these models provide a base for designing risk detection systems. The final model, on how to motivate users to seek human care, was also empirically evaluated. This evaluation revealed agent tactic is relevant for intention to seek care, intention to visit the agent again and the feeling of being heard.

In the design of the models, care was taken to ensure they would be generalizable to a broad range of different AEMH systems. The models can be used by systems that differ in target group, goal, platform and technical capabilities. However, the models do not fill in all the details, as those would need to be determined for every system separately. Given the broad range of AEMH systems in use [43, 44] this generalizability is important for any risk framework.

To fully appreciate the models presented in this paper some limitations should be considered. The detection model and auto-referral model were not evaluated in this paper. These need to be applied for a specific system and intervention, and experts would need to evaluate their fit for different purposes. However, by basing these models on literature, existing models and specific input from experts the models form a solid base to start from. In the evaluation of the motivation model participants were asked to imagine a situation, and indicate on a subjective scale what their actions would be. Although a more feasible methodology than recruiting participants with actual sleeping problems and studying their behaviour, this should be taken into account. The results from this study confirmed some of the hypotheses, but not all. Especially the reject care scenarios did not seem

to recreate situations where the suggestion to seek care falls in the latitude of acceptance. Further research is necessary to study if these situations exist, and what agent tactic is most appropriate.

Several other directions for future research can be established. Application of the detection model in different types of AEMH systems could reveal if the model fully covers the important decisions in detecting risk in AEMH. This model could also be extended with more specific instructions for specific risk situations. The first situation to cover would be suicide detection, as this is still a large problem within mental health care [45, 46]. These suggestions could take the form of what keywords to detect, what questionnaires to use, what threshold values to use, etc. Similarly, the auto-referral model could be evaluated by application into systems. This model currently operates on the premise that the user is aware they can be automatically referred. However, some risks might be so severe that safety takes priority over consent. Exactly how this would change the model is another avenue for further study. Finally, as technology advances it would also be important to re-evaluate the models to ensure that they stay applicable for new systems.

6.6. Conclusion

This paper shows the possibility of formal and generalizable models for risk detection in AEMH systems. These models represent a first step towards risk detection and management in e-mental health care. The detection model shows two important steps; namely detecting a possible risk situation, and deciding if this situation is severe enough to seek human care. The second model considers auto-referral, and stresses the importance of expectation management and the role the user can still play in seeking contact. The third model considers motivating users to seek care themselves. An empirical evaluation shows that agent tactic is important for intention to seek care, intention to return to the agent and the feeling of being heard. Given the importance of patient safety and the growing research into AEMH systems, these models provide a valuable foundation on which to build risk detection and management systems.

References

[1] WHO, *Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level,* (2012).

[2] J. A. Cartreine, D. K. Ahern, and S. E. Locke, *A roadmap to computer-based psychotherapy in the united states,* Harv Rev Psychiatry **18**, 80 (2011).

[3] C. Lyons, P. Hopley, and J. Horrocks, *A decade of stigma and discrimination in mental health:plus ça change, plus c'est la même chose (the morethings change, the more they stay the same),* Journal of Psychiatric and Mental Health Nursing **16**, 501 (2009).

[4] T. Donker, M. Blankers, E. Hedman, B. Ljótsson, K. Petrie, and H. Christensen,

Economic evaluations of internet interventions for mental health: a systematic review, Psychological Medicine , 1 (2015).

[5] D. D. Luxton, *Recommendations for the ethical use and design of artificial intelligent care providers,* Artificial Intelligence in Medicine **62**, 1 (2014).

[6] P. Bower and S. Gilbody, *Stepped care in psychological therapies: access, effectiveness and efficiency,* The British Journal of Psychiatry **186**, 11 (2005).

[7] J. Gratch, L.-P. Morency, S. Scherer, G. Stratou, J. Boberg, S. Koenig, T. Adamson, and A. Rizzo, *User-state sensing for virtual health agents and telehealth applications,* Studies in Health Technology Informaticts **184**, 151 (2013).

[8] D. DeVault, R. Artstein, G. Benn, T. Dey, E. Fast, A. Gainer, K. Georgila, J. Gratch, A. Hartholt, M. Lhommet, G. Lucas, S. Marsella, F. Morbini, A. Nazarian, S. Scherer, G. Stratou, A. Suri, D. Traum, R. Wood, Y. Xu, A. Rizzo, and L.-P. Morency, *Simsensei kiosk: A virtual human interviewer for healhealth decision support,* in *AAMAS* (2014).

[9] A. A. Rizzo, B. Lange, J. G. Buckwalter, E. Forbell, J. Kim, K. Sagae, J. Williams, B. O. Rothbaum, J. Difede, G. Reger, T. Parsons, and P. Kenny, *An intelligent virtual human system for providing healthcare information and support,* Medicine meets Virtual Reality (2011).

[10] C. Pagliari, C. Burton, B. McKinstry, A. Szentatotai, D. David, A. S. blanco, L. Ferrini, S. Albertini, J. C. Castro, S. Estevez, and M. Wolters, *Psychosocial implications of avatar use in supporting therapy for depression,* Stud Health Technol Inform. **181**, 329 (2012).

[11] H. Kordy, M. Backenstrass, J. Hüsing, M. Wolf, K. Aulich, M. Bürg, B. Puschner, C. Rummel-Kluge, and H. Vedder, *Supportive monitoring and disease management through the internet: An internet-delivered intervention strategy for recurrent depression,* Contemporary Clinical Trials **36**, 327 (2013).

[12] J. Robinson, S. Hetrick, G. Cox, S. Bendall, A. Yung, and J. Pirkis, *The safety and acceptability of delivering an online intervention to secondary students at risk of suicide: Findings from a pilot study,* Early Intervention in Psychiatry **9**, 498 (2015).

[13] M. L. Tielman, M. A. Neerincx, R. Bidarra, B. Kybartas, and W.-P. Brinkman, *A therapy system for post-traumatic stress disorder using a virtual agent and virtual storytelling to reconstruct traumatic memories,* Journal of Medical Systems **41** (2017).

[14] J. Spence, N. Titov, L. Johnston, M. Jones, and B. F. Dear, *Internet-based trauma-focused cognitive behavioural therapy for ptsd with and witwith exposure components: A randomised controlled trial,* Journal of Affective Disorders **162**, 73 (2014).

[15] J. Kasckow, S. Zickmund, A. Rotondi, A. Mrkva, J. Gurklis, M. Chinman, L. Fox, M. Loganathan, B. Hanusa, and G. Haas, *Development of telehealth dialogues for monitoring suicidal patients with schizophrenia: Consumer feedback,* Community Mental Health Journal **Brief Report**, 1 (2013).

[16] J. Kasckow, S. Zickmund, A. Rotondi, A. Welch, J. Gurklis, M. Chinman, L. Fox, and G. Haas, *Optimizing scripted dialogues for an e-health intervention for suicidal veterans with major depression,* (2014), in Press.

[17] G. Andrews, J. Newby, and A. Williams, *Internet-delivered cognitive behavior therapy for anxiety disorders is here to stay,* Current psychiatry reports **17**, 533 (2015).

[18] H. M. Christensen, P. J. Batterham, and B. O'Dea, *E-health interventions for suicide prevention,* International journal of environmental research and public health **11**, 8193 (2014).

[19] A. van Hemert, A. Kerkhof, J. de Keijser, B. Verwey, C. van Boven, J. Hummelen, M. de Groot, P. Lucassen, J. Meerdinkveldboom, M. Steendam, B. Stringer, A. Verlinde, and G. van de Glind, *Multidisciplinaire richtlijn diagnostiek en behandeling van suicidaal gedrag,* (2012).

[20] K. Knight, BAppSci, A. Kenny, and R. Endacott, *Assessing clinical urgency via telephone in rural australia,* Nursing and Health Sciences **17**, 201 (2015).

[21] N. Sands, S. Elsom, M. Gerdtz, K. Henderson, S. Keppich-Arnold, N. Droste, R. K. Prematunga, and Z. W. Wereta, *Identifying the core competencies of mental health telephone triage,* Journal of Clinical Nursing **22**, 3203 (2013).

[22] B. H. Belnap, H. C. Schulberg, F. He, S. Mazumdar, C. F. Reynolds, and B. L. Rollman, *Electronic protocol for suicide risk management in research participants,* Journal of Psychosomatic Research **78**, 340 (2015).

[23] B. L. Rollman, B. H. Belnap, S. Mazumdar, K. Z. Abebe, J. F. Karp, E. J. Lenze, and H. C. Schulberg, *Telephone-delivered stepped collaborative care for treating anxiety in primary care: A randomized controlled trial,* Journal of General Internal Medicine **32**, 245 (2017).

[24] M. Tielman, C. Pagliari, and W. Brinkman, *Risk detection in smart mental-health applications,* IVA Workshop on Virtual Health Agents (2015).

[25] M. Pompili, L. Sher, G. Serafini, A. Forte, M. Innamorati, G. Dominici, D. Lester, M. Amore, and P. Girardi, *Posttraumatic stress disorder and suicide risk among veterans,* The Journal of Nervous and Mental Disease (2013).

[26] B. F. Grant, F. S. Stinson, D. A. Dawson, S. P. Chou, M. C. Dufour, W. Compton, R. P. Pickering, and K. Kaplan, *Prevalence and co-occurrence of substance use disorders and independentmood and anxiety disorders results from the national epidemiologic survey on alcohol and relatedconditions,* Arch Gen Psychiatry **61**, 807 (2004).

[27] G. E. Simon, C. M. Rutter, D. Peterson, M. Oliver, U. Whiteside, B. Operskalski, and E. J. Ludman, *Do phq depression questionnaires completed during outpatient visits predict subsequent suicide attempt or suicide death?* Psychiatr Serv. **64**, 1195 (2013).

[28] M. Gould, *Suicide and the media,* in *Annals of the New York Academy of Sciences*, Vol. 932 (2001).

[29] G. Koocher and P. Keith-Speigel, *Ethics in psychology and the mental health professions: standards and cases (oxford textbooks in clinical psychology)* (Oxford University Press, 2008).

[30] A. Shevat, *Designing Bots*, edited by A. Rufino (O'Reilly, 2015).

[31] I. L. Kampmann, P. M. Emmelkamp, D. Hartanto, W.-P. Brinkman, B. J. Zijlstra, and N. Morina, *Exposure to virtual social interactions in the treatment of social anxiety disorder: A randomized controlled trial,* Behaviour Research and Therapy **77**, 147 (2016).

[32] H. W. Simons and J. G. Jones, *Persuasion in Society* (Routledge, 2011).

[33] B. Fogg, *A behavior model for persuasive design,* in *Proceedings of the 4th International Conference on Persuasive Technology* (2009).

[34] S. Michie, M. van Stralen, and R. West, *The behaviour change wheel: A new method for characterising and designing behaviour change intervention,* Implementation Science **6**, n.42 (2011).

[35] R. H. Thaler and C. R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* (Yale University Press, 2008).

[36] M. M. Ohayon, *Epidemiology of insomnia: what we know and what we still need to learn,* Sleep Medicine Reviews **6**, 97 (2002).

[37] A. J. Thayaparan and E. Mahdi, *The patient satisfaction questionnaire short form (psq-18) as an adaptable, reliable, and validated tool for use in various settings,* Med Educ Online **18** (2013).

[38] M. A. Hall, B. Zheng, E. Dugan, F. Camacho, and K. E. Kidd, *Measuring patients' trust in their primary care providers,* Medical Care Research and Review **59**, 293 (2002).

[39] M. Tielman, M. A. Neerincx, C. Pagliari, A. Rizzo, and W.-P. Brinkman, *Dataset belonging to the paper: Considering patient safety in autonomous e-mental health systems - detecting risk situations and referring patients back to human care,* (2017), doi: 10.4121/uuid:d158b291-e41d-430a-82d0-e91da98f6b0f.

[40] D. McFadden, *Frontiers in econometrics,* (New York: Academic Press, 1974) Chap. Conditional logit analysis of qualitative choice behavior, pp. 105–142.

[41] J. Cohen, *A power primer,* Quantative Methods in Psychology **112**, 155 (1992).

[42] J. Prochaska and C. DiClemente, *The transtheoretical approach,* (Oxford University Press, 2005) pp. 147–171.

[43] S. Provoost, H. M. Lau, J. Ruwaard, and H. Riper, *Embodied conversational agents in clinical psychology: A scoping review,* JOURNAL OF MEDICAL INTERNET RESEARCH **19**, e151 (2017).

[44] U. Schmidt and T. Wykes, *E-mental health - a land of unlimited possibilities,* Journal of Mental Health **21**, 327 (2012).

[45] K. Zivin, H. M. Kim, J. F. McCarthy, K. L. Austin, K. J. Hoggatt, H. Walters, and M. Valenstein, *Suicide mortality among individuals receiving treatment for depression in the veterans affairs health system: Associations with patient and treatment setting characteristics,* Am J Public Health **97**, 2193 (2007).

[46] G. Turecki and D. A. Brent, *Suicide and suicidal behaviour,* Lancet **387**, 1227 (2016).

7

A Therapy System for Post-Traumatic Stress Disorder

Although post-traumatic stress disorder (PTSD) is well treatable, many people do not get the desired treatment due to barriers to care (such as stigma and cost). This paper presents a system that bridges this gap by enabling patients to follow therapy at home. A therapist is only involved remotely, to monitor progress and serve as a safety net. With this system, patients can recollect their memories in a digital diary and recreate them in a 3D World-Builder. Throughout the therapy, a virtual agent is present to inform and guide patients through the sessions, employing an ontology-based question module for recollecting traumatic memories to further elicit a detailed memory recollection. In a usability study with former PTSD patients (n=4), these questions were found useful for memory recollection. Moreover, the usability of the whole system was rated positively. This system has the potential to be a valuable addition to the spectrum of PTSD treatments, offering a novel type of home therapy assisted by a virtual agent.

This chapter has been published in the Journal of Medical Systems **41**, (2017) [1].

7.1. Introduction

Post-Traumatic Stress Disorder (PTSD) is a mental disorder following one or more traumatic experiences. It is characterized by recurring intrusive memories, avoidance of reminders of the trauma and a persistent negative mood [2]. Cognitive Behavioral Therapy (CBT) with exposure is one of the most widely used treatments for PTSD. It relies on active recollection of the memory of the trauma to reduce the automatic fear response and facilitate cognitive restructuring [3–5]. Despite the existence of well-documented treatment, many PTSD patients do not seek help. Stigma on mental health-care is high, especially amongst veterans [6], and issues such as travel times and cost can form further barriers to care. A stand-alone home-therapy system can therefore fill an important gap by providing a treatment which is easily accessible, privacy sensitive and cost-effective. Although many e-solutions for mental health are being developed, fully autonomous systems offering exposure therapy for PTSD are rare. The multi-model memory restructuring (3MR) system for home therapy for PTSD is one such system [7, 8]. In this paper we present 3MR version 2 (3MR_2), which incorporates several new elements such as a virtual agent and a question system for trauma recollection.

The original 3MR system was designed for in-clinic use, where the patient works together with a therapist in a face-to-face setting. Its goal was to support this therapy by facilitating trauma recollection and storytelling. The system was developed using expert input and included three main functionalities, namely a timeline, diary and 3D world editor. Memories could be added to the timeline to form an overview, while pictures, text and maps could be added to a diary, and a basic 3D version of the memory could be created in the 3D editor [7, 8]. The patient could work on these environments with the therapist, but also at home as part of homework assignments. By creating a visual representation the therapist would also get a better understanding of the patient's experiences. This combination of exposure platforms was novel to 3MR, but several other technology-driven methods exist for exposure therapy. The most common is virtual reality (VR), which has had some very promising results [9, 10]. The main drawback of traditional VR therapy for PTSD is that the virtual environments are pre-created, and therefore difficult to match with personal memories. Although feasible for a specific group of veterans, most other patient groups generally do not share similar memories. In order to better personalize VR therapy, several platforms have been developed wherein therapists can build a virtual environment for their patients [11–13]. However, the original 3MR system shifts this task from therapist to patient, changing the main therapeutic component from the experience to the creation of a virtual world, which requires triggering the memory in an active way. As patients create their own 3D world, it facilitates the creation of a very personal autobiographic virtual environment. This shift also makes these environments suitable for home-use, as well as use in a stand-alone therapy where no therapist is present. The new 3MR_2 system expands on the concept of personal memory recreation by further developing both the diary and the 3D editor.

As 3MR_2 is designed for home-use instead of the original in-clinic use with a therapist, two additional main functionalities are necessary. Firstly, because the

system is fully autonomous, it requires some form of procedure to ensure patient safety [14]. With the rise of technology in mental-health care, new ethics guidelines also need to be in place. Patient safety is a particular concern for systems that display some level of automation. Many systems have safety checks not in the system itself, but in the procedure surrounding its use. These checks can take the form of exclusion criteria, but also regular email or phone contact with a clinician [15–18]. In this way, it is still a human who provides the safety support. An alternative is to include all safety checks in the system itself without any human in-the-loop. These checks can take the form of questionnaires and crisis management options within the system [17]. A combination of these two solutions is to facilitate monitoring by a clinician through the system. In this situation, the system itself monitors the patient but a therapist uses the information gathered to ensure patient safety. An example of this situation is given by Robinson et al., where a clinician monitors distress scores in an application for students at risk for suicide [19]. A similar approach is taken by the 3MR_2 system; questionnaire scores can be monitored by clinicians, who make the call to interfere if patient safety is in question.

Secondly, a home-therapy system requires some form of interaction and guidance. The first goal of this interaction is to inform the patient and provide the rationale behind the assignments. The second goal is to assist the patient with memory recollection in a personalized way. Virtual agents have been gaining popularity in health applications as a way to add a social and personal aspect to systems. The addition of virtual agents has been shown to have a positive effect on attention [19], adherence [20, 21] and likability [22, 23] of applications. To assist with memory recollection during exposure, some knowledge of traumatic events is necessary. Ontologies provide a good way to add domain knowledge to computer systems [24, 25]. Through combining multiple-choice and open questions, knowledge can be gathered from the patient while still retaining a natural interaction between patient and system [26]. A combination of these paradigms has been shown to result in a question system which can elicit greater detail in the responses [27]. In the 3MR_2 system, a virtual agent is added to provide assistance and personalization, employing an ontology and a structured dialogue to pose the correct questions while assisting in memory recollection.

In the rest of this paper, we will present the 3MR_2 system in more detail. We will first expand on the monitoring function and the additions to the diary and virtual environment already present in the original 3MR. Secondly, we will describe the virtual agent and specifically the question system it employs to assist the patient during exposure. This paper will conclude with the results of a usability study of the system with former PTSD patients.

7.2. 3MR_2 System

The 3MR_2 system offers therapy for PTSD. It includes the exposure environments introduced in the original 3MR, namely a timeline with memories which can be both described in a digital diary and re-created in a 3D WorldBuilder. The 3MR_2 system introduces monitoring within the system via questionnaires, improvements to the virtual environment design and a virtual agent to guide users through therapy and

assist in memory recollection. It is specifically aimed at either victims of childhood sexual abuse (CSA) or war veterans. Differences consist of wording in some texts (e.g. 'When you think back on your deployment' vs. 'When you think back on the period of the abuse'), the possible content of the virtual environment (e.g. models of tanks in the war version, children's beds in the CSA version) and the concepts in the ontology on which the question system is based.

Figure 7.1 shows an outline of the different components within 3MR_2. During a session, the patient is guided through the session components to perform the different therapy tasks. The general components can be accessed any time if the patient so wishes, but are not included in the therapy flow. These include general information on the system, a possibility to read back the psychoeducation and an e-mail function to contact the helpdesk or therapist with questions. The patient is not required or advised to use this e-mail function, but it is included to give the patient the option to reach out. The system always opens on the session overview screen, where patients can see a list of their sessions and their planned dates. Via this screen, they can start their next session. The patient always starts with the questionnaires. These are the post-traumatic stress disorder checklist (PCL) [28] and the patient-health questionnaire (PHQ) for depression [29]. These questionnaires are taken every session and visible for the therapist who monitors the patient. After the questionnaire, the patient is asked to read a brief introduction to the session explaining what the content is, and to select which feeling is predominant at the moment. Next, the patient is presented with the period overview, which lists the periods they have worked on in the diary. At the top of the screen these periods are also represented on a timeline. During the sessions where patients work on their memories, they can use this list to go to the appropriate diary page. In this diary page, the virtual agent asks questions to guide the patient through filling in the diary. After completing the diary, patients can use the WorldBuilder to recreate or review the 3D version of the memory they just described. This is the final exercise of the session. Patients then close the session by again selecting which feeling is predominant now, to illustrate changes between the beginning and end of the session. After that, they can return to the start page to close the program. Slight variations in this procedure exist, for instance the last session focuses on relapse prevention and does not actively include working on the memory. The full therapy consists of 12 sessions, starting with two sessions to get familiar with the system and memory recollection. The following eight sessions gradually introduce the exposure elements, working on three traumatic memories, which are increasing in impact. The final two sessions are aimed at review, reflection and a brief relapse-prevention. Exactly when and how the sessions are scheduled depends on the clinical setting that the system is applied in.

7.2.1. Monitoring

Although the 3MR_2 system is aimed at stand-alone home therapy, a clinician still has a monitoring role. Patients fill in the PHQ and PCL at the start of every session. Additionally, the system asks the patient to enter their subjective unit of discomfort (SUD) score [30] before, during and after each exposure session. Finally, the system

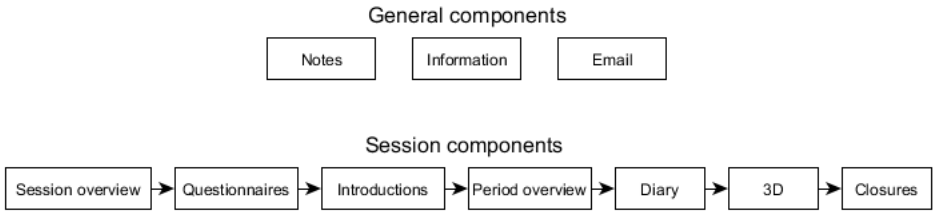

Figure 7.1: Outline of the components in the 3MR_2 system. In a typical session, the patient starts with the session overview page before answering several questionnaires. Afterwards, they receive an introduction to the session and goes to the period overview, showing a list of the added memories. In the diary, one memory can be described in detail, after which it is also recreated in the 3D WorldBuilder. After the memory is described and recreated, the session is closed. The notes, general information and e-mail pages can be accessed by the patient at any time during a session.

collects activity data from the diary and 3D environment, including the number of items added or viewed during a session. A monitoring tool shows these scores in graphs, as shown in Figure 7.2. If the therapist deems the scores serious enough to intervene they can do so. This can, for instance, be when the depression scores rise very high, or if the patient does not show activity in the diary when they should have performed a session. Additionally, the patient has the option to send a secure message to the therapist through the same server, if they would wish. The therapist views these messages with the same monitoring tool.

7.2.2. Exposure Environments

The 3MR_2 system has two main exposure environments, a digital diary and a 3D tool, the WorldBuilder. Both are based on the environments in the original 3MR, but have been completely re-designed to improve usability, and add additional functionality to increase the possibilities within the environment. The digital diary can be filled with text, images, media, emotions and web-items such as maps and YouTube clips. The emotion items are a new addition compared to the original 3MR system and appear as words in the diary. This function was added because describing emotions and feelings is an important part of writing about memories [31]. Included in the emotion function is an option to relate the emotion to feelings during the memory or feelings in the present. As in the original 3MR system, all diary items can be added via a menu at the top and appear as a movable thumbnail on the screen, which can be enlarged to focus solely on that item. If items are too confrontational for permanent display they can be darkened. A new function has been added to allow connections to be made between items, indicating a relationship. The diary is used to describe the memory in detail. When working on the traumatic memories, the virtual agent guides patients through filling the diary with the ontology-based question system. If patients wish to add items after the questions they are always free to do so though. Figure 7.3 shows a diary filled with items.

The WorldBuilder is 3D tool in which patients can build a virtual environment, recreating their memories. Where the original 3MR only used a birds-eye view, two

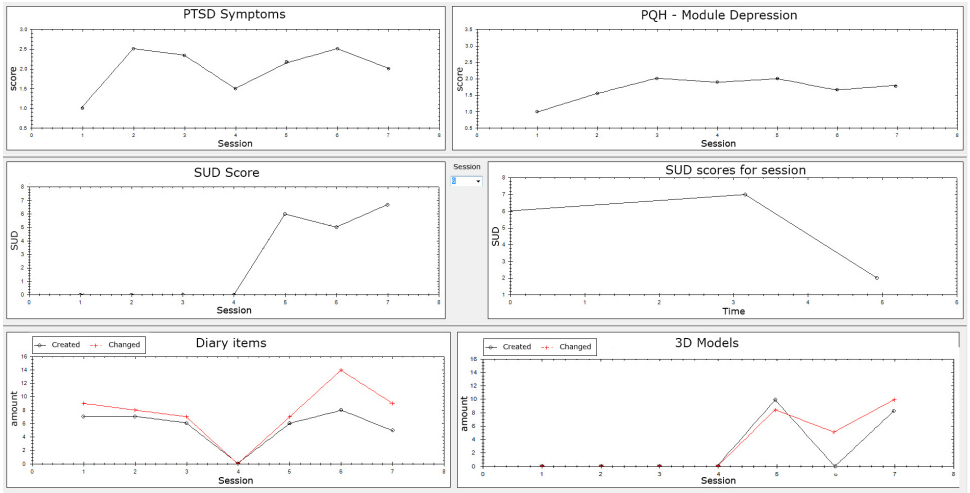

Figure 7.2: Screenshot therapist system showing example data for PTSD symptom scores, Depression symptom scores, SUD scores and activity data of diary.

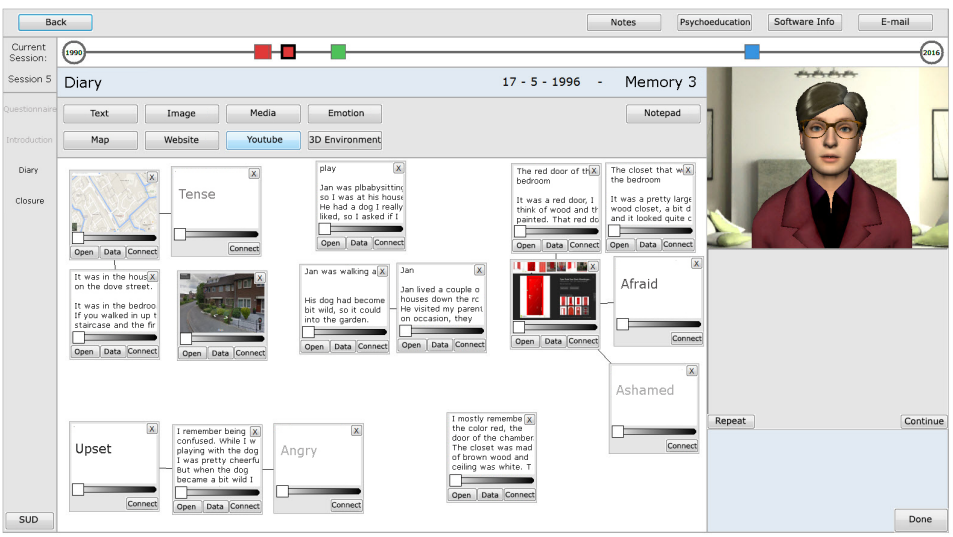

Figure 7.3: Diary, what it could look like when filled. Translated from the original Dutch

perspectives have been added to the new WorldBuilder. Using a collection of 3D-models, patients can make a top-down sketch of their recollections. In addition to this top-down edit view, patients can view the scene from a birds-eye perspective, which allows for zooming and viewing the scene from different angles. When building the scene, the patients can select 3D models from a menu on the right and drop them in the scene, rotating and moving them to put them in place. The interface is designed such that people with basic computer skills can use the tool. A brief instruction video is included as well, to further familiarize users with the interface. The CSA version is developed for indoor scenes, the scale of items reflecting the size of a large to smaller room. The version for veterans on the other hand is designed for outdoor scenes. Both versions have a selection of textures for the walls and ground, including both indoor textures such as wooden or carpet floors, and outside textures such as dirt and grass. Additionally, buildings are offered. The CSA version uses European-style buildings specifically, while the veteran version also includes a broad range of eastern-styled buildings. Similarly, the CSA version includes regular vehicles such as a car and bike, while the veteran version includes a range of army vehicles such as trucks and tanks. Regarding furniture, the CSA version is equipped with a range of objects to build bedrooms, studies and bathrooms, while the veteran version only includes basic furniture. The human models in both versions come in different ethnicities, both including a range of male, female and child models. The veteran version also includes a selection of Dutch army models. Finally, both versions are equipped with a range of general objects such as toys and books for the CSA version, and crates and roadblocks for the veteran version. Aside from some additions in the model library, the new 3MR_2 system has several other novel functions. These are the function to watch the environment through the eyes of the different people in-scene, and the option to create different scenes of the same memory. These scenes can be used to represent a timeline of events.

Through changing the environment over scenes patients can tell their story. By clicking through the scenes, patients can watch the events play out in a very controlled manner. Patients work on one memory in two different sessions. In the first they are asked to create a static scene, placing floors and big objects first, small objects and people last. In the second session they are asked to review the memory and create the different scenes. Figure 7.4 shows various worlds created with the two WorldBuilder versions and the different perspectives. Figure 7.5 shows an example of a story told through three scenes.

7.2.3. Virtual Agent

At the beginning of the first session, patients pick their virtual agent. Four different agents exist, but all patients get a choice between two, depending on gender and patient group (Figure 7.6). This pre-selection is made for several reasons, firstly so that the female CSA victims will not be confronted with a male agent, but also to enhance later satisfaction with the agent, as more choice may lead to lower satisfaction [32]. Aside from picking the gender and appearance, the patient also chooses a voice for the agent. The voices are generated by the Dutch text-to-speech system Fluency. All virtual agents display general idle behaviour and mouth

Figure 7.4: WorldBuilder. Top images are from the CSA version, in-scene person perspective (left) and edit (right). Bottom images are the war version. Left the in-scene Person perspective, right birds-eye perspective. Translated from the original Dutch version.

Figure 7.5: An example of three scenes telling a story of a jeep exploding in the WorldBuilder.

Figure 7.6: Possible virtual agents. Female abuse victims will be able to choose between the two female agents, male abuse victims between the two agents on the right. Male veterans are able to choose between the two male agents, female veterans between the agent on the top left and bottom right.

movement while talking. The agent talks to the patient, but a repeat option exists which also displays the text. The user only responds with actions, or in the case of a multiple-choice question, with selecting a pre-set answer. The virtual agent has several key functions. First, it acts as a guide through the system and the therapy, telling the patient what to do and where to click. For instance, the virtual agent will ask the patient to fill in the questionnaires. Second, it provides background information on the therapy-concept, explaining why certain tasks need to be performed. For instance, when asking the patient some reflective questions after building the 3D environment, the agent also reminds people that they should not avoid thinking deeply about their memories. Third, it assists the patients during the exposure by asking personalized questions within the diary environment.

The virtual agent is not present while patients work on the 3D WorldBuilder, but does have a couple of functions related to this environment. First, if patients wish to take a break during the exposure they can select the option to have a relaxation exercise that is led by the virtual agent. Additionally, the virtual agent asks a number of questions after the 3D world has been finished. Several questions are general and focused on the experience, such as what the patient hears and feels when looking at their world. However, the virtual agent also notices some significant events in the 3D environment. It notices when objects move between scenes, when people

models are laying down instead of standing, and when explosion models are added. On these occasions, it specifically asks the patient what this event meant for them.

The goal of exposure within PTSD therapy is that patients confront their memories and experience that thinking about these moments is possible. It is important that people think back in detail, so personalized and detailed questions are very helpful. To do this, the virtual agent employs an ontology-based question system. Past research has shown that such an ontology-based question system is able to elicit more detailed descriptions in memory recollection [27]. The ontology in the 3MR system represents knowledge about traumatic memories, differing slightly for either the veterans or CSA victims. Four different versions exist of the military ontology, one for Afghanistan, one for Bosnia, one for Libya and one not related to a specific deployment. These locations are chosen to best represent the Dutch missions.

The ontology is based around the topics of location, objects, people, actions, senses and emotions. Each of these concepts is represented by a hierarchy of classes in the ontology. For example, the top-class location has subclasses for inside and outside locations. For the CSA version, the inside location has classes such as house, school and church, while the war version focuses more on outdoor locations and includes a road and base camp. The virtual agent introduces each topic with a multiple-choice question to determine which class is relevant. Using that answer, it asks open questions corresponding with each of the properties of this class, which the patient can answer in the diary. For instance, if the location is a school the agent will ask if the person also went to this school, and if they can remember places in the school where they often went. If it is a road, the question will ask where the road goes and if they were here often, etc. Figure 7.7 shows a schematic example of this process. This ontology-based question system is used to fill the diary for the trauma recollection, and guides the patient through this process in two sessions. In the first session, the virtual agent asks about location, objects and people. Although answers to the open questions are entered in the form of text, the agent will also ask to add maps or photos if these are available. In the second session, the agent will ask about what happened in the scene, what the patient smelled and heard, and finally what emotions they felt. Regarding what happened, the agent asks the patient to select verbs about what they themselves were doing, and what the others were doing and follow up from there. For instance, if the person would select 'shooting', the agent will ask what that person was shooting at, but also what happened right before and after this action. For emotions, the agent will ask how the person felt during the memory, but also to describe how they feel about it now and what has changed and why.

7.3. Evaluation

Usability was evaluated in two stages. An initial usability test with healthy participants was performed for the virtual agent and diary environment. Based on this test, improvements were made to the system. An additional study was done with former PTSD patients studying both the usability of the system and its usefulness for recollecting traumatic memories.

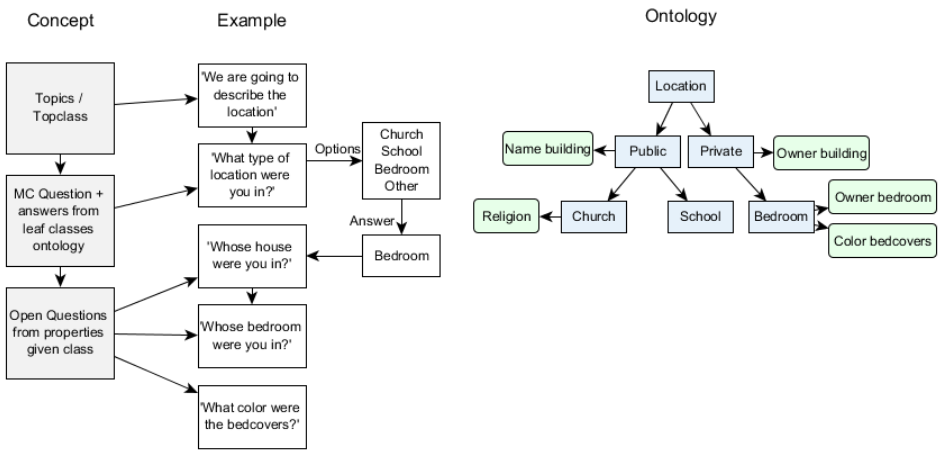

Figure 7.7: The process used within the question system. Based on the simplified ontology on the right, blue represent the classes (concepts) and green the properties of those classes. Given the topic of location, the first question will be multiple-choice, with as possible answers the leaf classes in the ontology. If the given answer would be 'other', the question is posed again with the options one level higher in the ontology (public/private in this case). Given the answer 'Bedroom', the following questions are open and correspond to the three properties of the Bedroom class.

7.3.1. First usability test

Three healthy participants were recruited for an initial usability test, all were researchers or students at the Computer Science department, two had a background in psychology, one in computer science. They performed the first therapy session, in which a positive memory is described. This session did not yet include questions from the question module or the 3D WorldBuilder. No explicit instructions were given. Participants were asked to 'think aloud', and an experimenter was present to note all comments and what went wrong. Based on these usability tests, small improvements such as button placement were made to the system. Two instructional videos were made using the usability input, one describing the general system, and one specifically for the 3D WorldBuilder.

7.3.2. Second usability test

A second usability test was performed with former PTSD patients. Its goal was threefold, firstly to study the general usability of the system and its components. The second goal was to study how much the system elements contributed to therapy according to users. The final goal was to study how useful and appropriate the questions generated by the question module were in recollecting traumatic memories. The design of this study was approved by the ethics committee of Delft University of Technology.

Participants
4 participants were recruited via practicing therapists. All participants had in the past followed therapy for PTSD. Participants 1 and 2 were war-veterans (both male), participants 3 and 4 had experienced childhood sexual abuse (both female).

Procedure
All participants first received general information on the 3MR_2 system, what a full therapy would look like and what was expected of them during the experiment. After this, participants followed the entire first therapy session, which is aimed at familiarizing oneself with the system by describing a positive memory. Following the first session, participants skipped ahead and followed parts of two sessions in which one traumatic memory is described. During these sessions, participants were asked to keep one personal memory in mind. They were requested to answer all multiple-choice questions, but none of the open questions. These were only rated in terms of usefulness. The experiment ended with a general questionnaire.

Measures
The first measure was how useful the questions generated by the system were for remembering the trauma. All questions were rated on an analog scale ranging from *works against recollection* to *helps a lot recollecting*. The center point was marked as question has no effect. All questions were rated immediately after asking. The second measure was how useful and understandable the functions in the program were. Two questions were posed on an analog scale, firstly asking how useful the function was from *detrimental* to *helps a lot*. The second asked how understandable the function was, ranging from *confused me* to *very understandable*. The final questionnaire was the Dutch version System Usability Scale [33], applied to the whole 3MR_2 program, answered on a 5pt Likert scale.

Analysis
All statistical analyses were carried out with R version 3.3. Before the analysis, the usefulness scores for the questions were transformed to range from -50 to 50, with 0 as the neutral point so deviation from 0 could be tested. Multilevel analyses taking participant as a random intercept were conducted on the usefulness scores for the questions, the usefulness scores of the system components and the usability scores of the system components. The analysis of the usefulness of the questions only included a fixed intercept. In the analysis on the usefulness and usability of the system, system component was included as a fixed effect to study if this factor influenced the result.

Results & Discussion
The analysis of the usefulness ratings of the questions revealed that on average, participants found the questions helped them to recall their memory (Mean = 15.11, SD = 22.93, $F(1,140) = 10.03$, $p = 0.002$). A significant variation between participants was found, however (SDrandom intercept = 8.37, 95% CI [3.72, 18.83]). Figure 7.8 shows a density plot of the given scores for each participant, showing that participant 1,3 and 4 have a relatively similar pattern. Participant 2, however,

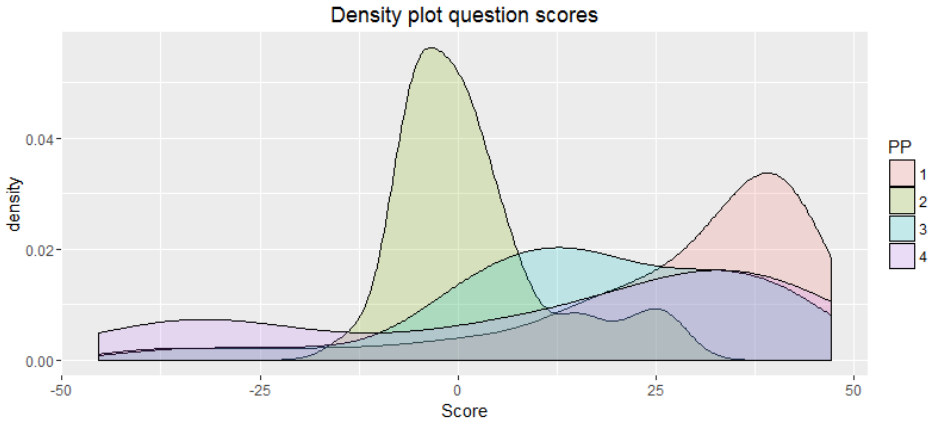

Figure 7.8: Density plot of the scores given to the questions posed by the question system. Participants 1, 3 and 4 scored the majority of questions above 0, but a couple got quite low scores. Participant 2 scored nearly all questions around 0, a couple slightly higher but none very low.

gave nearly every question a score surrounding 0, with none below -25 and a few above 25. The questions receiving low scores were individually reconsidered and revised.

The analysis of the usefulness scores revealed that on average, participants found the system useful (Mean = 16.32, SD = 16.17, $F(1,20) = 13.98$, $p = 0.001$). No difference was found between the different system components ($F(5,15) = 1.34$, $p = 0.30$). However, ratings varied significantly between participants ($SD_{\text{random intercept}} = 6.12$, 95% CI [1.54, 24.36]). A closer inspection showed that only participant 2 found functions to be detrimental to the therapy, namely the virtual agent and the instruction video. This might partly be explained by the fact that he had worked with a system without either virtual agent or question module in the past, and was therefore used to working with the diary alone. This might have resulted in impatience with the virtual agent and questions. Nevertheless, this preference for a system with less guidance might also exist in people without prior experience with 3MR.

The usability scores reveal a similar pattern to the usefulness scores. On average, the system was rated as well usable (Mean = 21.98, SD = 20.71, $F(1,20) = 7.26$, $p = 0.014$). No significant difference was found between system components ($F(5,15) = 0.61$, $p = 0.69$), while participants did significantly vary in their ratings ($SD_{\text{random intercept}} = 14.96$, 95% CI [6.78, 32.99]). The rating on the System Usability Scale ranged between 73 and 75 for participants 1, 2 and 3, which can be labeled as above average. The exception was participant 3 who gave a rating of 55. This is probably caused by a bug only this participant experienced in the 3D WorldBuilder (the scale of the 3D models was wrong).

7.4. Conclusion & Discussion

In this paper we described the 3MR_2 system, a therapy system for PTSD patients. The system contains two exposure environments, a digital diary and a 3D World-Builder in which memories can be recreated. During a 12-session therapy, a virtual agent guides and assists patients with their therapy tasks, employing an ontology-based question module. A human therapist is involved only to monitor progress.

Initial evaluations revealed that the system was usable by both non-patients and former PTSD patients. These evaluations did reveal small usability concerns, which were then resolved. They also exposed some differences in personal preferences, one participant strongly preferred working at his own pace without much guidance. The current system is less appropriate for patients with such a working style, which should be kept in mind for future use. The questions generated by the question module were deemed useful. The evaluations presented in this paper were only concerned with usability and the appropriateness of the generated questions, but did not look into therapeutic effectiveness. A benchmark study is currently being set-up to test if the 3MR_2 system is successful in significantly reducing PTSD symptoms.

Although the 3MR_2 system is specifically designed to treat PTSD, components are also relevant for other domains. Firstly, providing safety and reducing human resources are very important in e-mental-health [34–36]. The 3MR_2 gives an example of limited human monitoring, which achieves both goals. Secondly, the exposure environments present in 3MR_2 provide a novel view on technology-assisted exposure. Many studies have been done with VR environments for exposure therapy [10], but the concept of patients themselves recreating these worlds is novel, and might solve the difficulty of building worlds relevant to different patients [9]. Although embedded in a home-therapy in 3MR_2, both the diary and the 3D World-Builder could also be used as tools in regular therapy. Thirdly, the 3MR_2 system incorporates a virtual agent that can pose questions aimed at memory recollection. Although exposure to memories is specific to PTSD therapy, other health applications for this technology do exist. One example is expressive writing, a therapeutic tool aimed at writing about negative memories [31, 37]. Another possible application might be found in the field of life review therapy, wherein reminiscence is used to alleviate mental health symptoms, for instance in older adults with depression [38].

The 3MR_2 system provides opportunities to reduce the barriers to care. However, it is not suitable for all PTSD patients. Depression is a common comorbid disorder to PTSD, and suicide rates among PTSD patients are high [39]. Because patients work alone with the 3MR_2 system, it might be less suitable for patients with a history of suicidality. Similar concerns arise for patients which have substance problems [2], or which severely dissociate [40]. Another limitation to the 3MR_2 system is that patients cannot follow therapy while ignoring the virtual agent. Some people prefer more guidance than others, and some might not respond well to following a set course. Where a therapist might be able to steer the patient back on track, this is more challenging for a virtual agent.

Despite these challenges, we believe that the 3MR_2 system is a valuable addi-

tion to the spectrum of PTSD treatments. Through the use of new technologies, it offers a novel type of therapy which is convenient to patients and costs very little in therapist resources. Given the societal impact of PTSD, it may have a great positive effect on society.

References

[1] M. L. Tielman, M. A. Neerincx, R. Bidarra, B. Kybartas, and W.-P. Brinkman, *A therapy system for post-traumatic stress disorder using a virtual agent and virtual storytelling to reconstruct traumatic memories,* Journal of Medical Systems **41** (2017).

[2] A. P. Association, *Diagnostic and Statistical Manual of Nental Disorders, Fifth Edition* (American Psychiatric Association, 2013).

[3] R. Bradley, J. Greene, E. Russ, L. Dutra, and D. Westen, *A multidimensional meta-analysis of psychotherapy for ptsd,* AM J Psychiatry **162**, 214 (2005).

[4] T. Ehring, R. Welboren, N. Morina, J. M. Wicherts, J. Freitag, and P. M. Emmelkamp, *Meta-analysis of psychological treatments for posttraumatic stress disorder in adult survivors of childhood abuse,* Clinical Psychology Review **34**, 645 (2014).

[5] M. B. Powers, J. M. Halpern, M. P. Ferenschalk, S. J. Gillihan, and E. B. Foa, *A meta-analytic review of prolonged exposure for posttraumatic stress disorder,* Clinical Psychology Review **30**, 635 (2010).

[6] C. Lyons, P. Hopley, and J. Horrocks, *A decade of stigma and discrimination in mental health:plus ça change, plus c'est la même chose (the morethings change, the more they stay the same),* Journal of Psychiatric and Mental Health Nursing **16**, 501 (2009).

[7] M. van den Steen, W.-P. Brinkman, E. Vermetten, and M. Neerincx, *Design and usability evaluation of a multi-modal memory restructuring system for the treatment of combat-related ptsd,* in *ECCE 2010 Workshop on Cognitive Engineering for Technology in Mental Health Care and Rehabilitation* (2010).

[8] W.-P. Brinkman, E. Vermetten, M. van den Steen, and M. Neerincx, *Cognitive engineering of a military multi-modal memory restructuring system,* Journal of CyberTherapy and Rehabilitation **4**, 83 (2011).

[9] K. Meyerbröker and P. Emmelkamp, *Virtual reality exposure therapy in anxiety disorders: A systematic review of process-and-outcome studies,* Depression and Anxiety **27**, 933 (2010).

[10] T. E. Motraghi, R. W. Seim, E. C. Meyer, and S. B. Morissette, *Virtual reality exposure therapy for the treatment of posttraumatic stress disorder: A methodological review using consort guidelines,* Journal of Clinical Psychology **0**, 1 (2013).

[11] R. M. Banos, C. Botella, , V. Guillen, A. García-Palacios, S. Quero, J. Bretón-López, and M. Alcaniz, *An adaptive display to treat stress-related disorders: Emma's world,* British Journal of Guidance and Counselling (2009).

[12] C. Botella, R. M. Banos, A. García-Palacios, S. Quero, V. Guillen, and M. Alcaniz, *An adaptive display for the treatment of diverse trauma ptsd victims,* Cyberpsychology, Behaviour, and Social Networking **13**, 67 (2010).

[13] G. Riva, L. Carelli, A. Gaggioli, A. Gorini, C. Vigna, D. Algeri, C. Repetto, S. Raspelli, R. Corsi, G. Faletti, and L. Vezzadini, *Neuro vr 1.5 - a free virtual reality platform for the assessment and treatment in clinical psychology and neuroscience,* Studies in health technology and informatics **144**, 57 (2009).

[14] D. D. Luxton, *Artificial intelligence in psychological practice: Current and future applications and implications.* Professional Psychology: Research and Practice **45**, 332 (2014).

[15] B. Klein, J. Mitchell, K. Gilson, K. Shandley, D. Austin, L. Kiropoulos, J. Abbott, and G. Cannard, *A therapist-assisted internet-based cbt intervention for post-traumatic stress disorder: Preliminary results,* Cognitive Behaviour Therapy **38**, 121 (2009).

[16] B. Klein, J. Mitchell, J. Abbott, K. Shandley, D. Austin, K. Gilson, L. Kiropoulos, G. Cannard, and T. Redman, *A therapist-assisted cognitive behavior therapy internet intervention for posttraumatic stress disorder: Pre-, post- and 3-month follow-up results from an open trial,* Journal of Anxiety Disorders **24**, 635 (2010).

[17] H. Kordy, M. Backenstrass, J. Hüsing, M. Wolf, K. Aulich, M. Bürg, B. Puschner, C. Rummel-Kluge, and H. Vedder, *Supportive monitoring and disease management through the internet: An internet-delivered intervention strategy for recurrent depression,* Contemporary Clinical Trials **36**, 327 (2013).

[18] J. Spence, N. Titov, L. Johnston, M. Jones, and B. F. Dear, *Internet-based trauma-focused cognitive behavioural therapy for ptsd with and witwith exposure components: A randomised controlled trial,* Journal of Affective Disorders **162**, 73 (2014).

[19] J. Robinson, S. Hetrick, G. Cox, S. Bendall, A. Yung, and J. Pirkis, *The safety and acceptability of delivering an online intervention to secondary students at risk of suicide: Findings from a pilot study,* Early Intervention in Psychiatry **9**, 498 (2015).

[20] A. D. Andrade, R. Anam, C. Karanam, P. Downey, and J. G. Ruiz, *An overactive bladder online self-management program with embedded avatars: A randomized controlled trial of efficacy,* Urology **85** (2015).

[21] T. W. Bickmore, R. A. S. K. Nelson, D. M. Cheng, M. Winter, L. Henault, and M. K. Paasche-Orlow, *A randomized controlled trial of an automated exercise*

coach for older adults, Journal of the American Geriatrics Society **61**, 1676 (2013).

[22] M. M. Morandell, A. Hochgatterer, S. Fagel, and S. Wassertheurer, *Avatars in assistive homes for the elderly a user-friendly way of interaction?* in *4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society* (2008).

[23] N. Yee, J. N. Bailenson, and K. Rickertsen, *A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces,* in *CHI 2007* (2007).

[24] T. Bickmore, D. Schulman, and C. Sidner, *A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology,* Journal of Biomedical Informatics **44**, 183 (2011).

[25] N. F. Noy and D. L. McGuinnes, *Ontology Development 101: A Guide to Creating Your First Ontology*, Tech. Rep. (Stanford Knowledge Systems Laboratory, 2001).

[26] S. A. Friederichs, A. Oenema, C. Bolman, J. Guyaux, H. M. van Keulen, and L. Lechner, *Motivational interviewing in a web-based physicalactivity intervention: questions and reflections,* Health Promotion International (2013).

[27] M. L. Tielman, M. van Meggelen, M. A. Neerincx, and W.-P. Brinkman, *An ontology-based question system for a virtual coach assisting in trauma recollection,* in *Int. Conf. on Intelligent Virtual Agents* (2015).

[28] F. Weathers, B. Litz, T. Keane, P. Palmieri, B. Marx, and P. Schnurr, *The PTSD Checklist for DSM-5 (PCL-5)* (2013), scale available from the National Center for PTSD at www.ptsd.va.gov.

[29] K. Kroenke, R. L. Spitzer, and J. B. Williams, *The phq-9 validity of a brief depression severity measure,* Journal of General Internal Medicine **16**, 606 (2001).

[30] D. M. Kaplan, T. Smith, and J. Coons, *A validity study of the subjective unit of discomfort (sud) score.* Measurement and Evaluation in Counseling and Development **27**, 195 (1995).

[31] K. A. Baikie and K. Wilhelm, *Emotional and physical health benefits of expressive writing,* Advances in Psychiatric Treatment **11**, 338 (2005).

[32] K. Diehl and C. Poynor, *Great expectations?! assortment size, expectations, and satisfaction,* Journal of Marketing Research **47**, 312 (2010).

[33] J. Brooke, *Sus: A "quick and dirty" usability scale,* (Taylor & Francis, 1996) pp. 189–194.

[34] D. D. Luxton, *Recommendations for the ethical use and design of artificial intelligent care providers,* Artificial Intelligence in Medicine **62**, 1 (2014).

[35] J. A. Cartreine, D. K. Ahern, and S. E. Locke, *A roadmap to computer-based psychotherapy in the united states,* Harv Rev Psychiatry **18**, 80 (2011).

[36] T. Donker, M. Blankers, E. Hedman, B. Ljótsson, K. Petrie, and H. Christensen, *Economic evaluations of internet interventions for mental health: a systematic review,* Psychological Medicine , 1 (2015).

[37] J. W. Pennebaker and S. K. Beall, *Confronting a traumatic event: Toward an understanding of inhibition and disease,* Journal of Abnormal Psychology **95**, 274 (1986).

[38] G. J. Westerhof and E. T. Bohlmeijer, *Celebrating fifty years of research and applications in reminiscence and life review: State of the art and new directions,* Journal of Aging Studies **29**, 107 (2014).

[39] M. Pompili, L. Sher, G. Serafini, A. Forte, M. Innamorati, G. Dominici, D. Lester, M. Amore, and P. Girardi, *Posttraumatic stress disorder and suicide risk among veterans,* The Journal of Nervous and Mental Disease (2013).

[40] M. A. Hagenaars, A. van Minnen, and K. A. Hoogduin, *The impact of dissociation and depression on the efficacy of prolonged exposure treatment for ptsd,* Behaviour Research and Therapy **48**, 19 (2010).

8

Conclusion

8.1. Conclusions

The research presented in this thesis focuses on the development of a virtual agent for post-traumatic stress disorder (PTSD) treatment, and studies how the agent can act to enhance treatment compliance. It provides general guidelines (Ch. 2) for a virtual agent and how it should act in terms of informing (Ch. 3), assisting (Ch. 4) and motivating (Ch. 5) patients. Moreover, it formalizes how autonomous e-mental health (AEMH) systems might identify possible risk situations in (Ch. 6), and finally presents a computerized PTSD therapy system including virtual agent (Ch. 7). The main research question of this thesis is:

In what way can a virtual agent that is a part of a computerized PTSD therapy system act to enhance treatment compliance?

From this main research question, three sub questions and three hypotheses were derived. These questions and hypotheses focus on the different aspects of the behavior of a virtual agent for PTSD and the context in which this agent operates.

Q1: Which design guidelines are important for the design of a virtual agent for PTSD treatment?

Q2: How do we support patient safety when they work with an autonomous e-mental health (AEMH) system?

Q3: What is a feasible computerized PTSD therapy system incorporating a virtual agent?

H1: Presentation mode of psychoeducation presented by a virtual agent affects

adherence.

H2: An ontology-based question system enhances memory recollection.

H3: By presenting situation-based motivational messages a virtual agent enhances therapy trust and motivation.

The studies presented in this thesis aim to answer these research questions and investigate the hypotheses. The conclusions from these studies are as follows.

Q1: Which design guidelines are important for the design of a virtual agent for PTSD treatment?
To find address the first question, scenarios were created that described possible roles and behaviors for a virtual agent for PTSD. In these scenarios, interactions between agent and patient were described. Ten experts, all experienced with treating PTSD patients, were asked to read and discuss the scenarios. From these discussions, key statements were categorized depending on underlying concepts. This analysis resulted in a list of ten guidelines for the behavior of a virtual agent for PTSD patients, including examples of how such guidelines could be implemented. These guidelines are: motivate, take patients seriously, protect the patient, be down to earth, personalize, be transparent, avoid negative reinforcement, protect patients from themselves, psychoeducation, and acknowledge.

Q2: How do we support patient safety when they work with an autonomous e-mental health (AEMH) system ?
The answer to the second question was presented in the form of three models describing how an AEMH system should deal with risk situations. These models were based on expert interviews, existing safety procedures in regular health care, and principles from behavior change research. The first model describes how risk can be detected, via keyword detection or questionnaires. After a risk is detected, a system can either use a threshold on a scale to decide if the risk warrants intervention from a human care-giver, or consider the duration, severity and progression of a situation. The second model describes automatic referral to human care, taking into account the severity of a situation. If a crisis is present, the model recommends that both the user and care-giver are asked to seek contact with each other. If no direct crisis exists, the model recommends to ask the user to seek contact, while still informing the care-giver of the situation. The final model describes how an AEMH system might motivate the user to seek contact themselves. This model proposes to adapt the motivational tactic to the situation of the user to promote both help-seeking behavior and the users opinion of the system, either directly referring a user, persuading them to seek care or accepting that they do not wish to seek help. This final model was empirically evaluated in an online study where participants (n=160) were asked to imagine they had a sleeping problem and afterwards chat with an agent following the motivation model. Results show that the motivational tactic of the agent influences a users' intention to seek care, intention

to seek the agent again and their feeling of being heard by the agent.

Q3: What is a feasible computerized PTSD therapy system incorporating a virtual agent?
To answer the third question this thesis presents a computerized PTSD therapy system, the multi-model memory restructuring system version 2 (3MR_2). This system combines components previously used in PTSD therapy, such as virtual environments [1] and a virtual agent [2], but is novel in bringing them all together to form a system with which patients can autonomously follow therapy. The original 3MR system (CognitiveEngineeringMilitary2011Brinkman) was developed for in-clinic use and introduced a digital diary and 3D world editor for patients to recreate their memories. The new 3MR_2 system is intended instead for home use, and introduces both a patient monitoring function and a virtual agent, as well as a redesign and update of the existing components. This new 3MR_2 system also includes some general therapy components, such as a session overview and questionnaire. The patient monitoring system is used by therapists and shows the scores of patients so they can be monitored for risk situations. The main therapeutic components of 3MR_2 are the exposure environments and a virtual agent. The two exposure environments are a digital diary and a 3D worldbuilder in which memories can be recreated. The virtual agent acts as a coach and guides patients through the therapy. Furthermore, the agent employs an ontology-based question system to ask personalized and detailed questions about the traumatic experiences. The whole 3MR_2 system was evaluated on usability by both non-patients and former PTSD patients, showing that the system was well usable. Additionally, the questions generated by the ontology-based system were evaluated on usefulness by former patients. This evaluation showed that the questions were found useful in recollecting traumatic experiences.

H1: Presentation mode of psychoeducation presented by a virtual agent affects adherence.
This hypothesis was formulated on the premise that psychoeducation could affect adherence in two ways. Firstly, that verbal presentation by a virtual agent improves alliance to that agent, and this in turn improves adherence. Secondly, that textual presentation improves how well the information is recollected, and this in turn improves adherence. Because these premises both give advantage to another presentation mode, the combined effect was unclear, resulted in the unidirectional hypothesis. To study this hypothesis, an experiment was conducted on the effect of presentation mode of psychoeducation on adherence. Participants (n = 46) were asked to briefly list five of their worst memories, and both rate and rank these on level of how distressful they were. Following this exercise, a virtual agent would either ask the participant to read the psychoeducation from text on the screen, or tell the participant the information with a text-to-speech voice. Finally, the virtual agent asked the participant to describe one of their memories in detail in a digital diary for 15 minutes, advising them to pick the worst memory. The analysis did not find that presentation mode affected the participants' attitude towards the agent or

how well they recollected the information. However, both attitude and recollection did influence adherence in the form of task execution: how much was written in the diary and how well emotions were described. Moreover, after controlling for these two factors, textual presentation of psychoeducation showed a greater task adherence than verbal presentation, confirming the hypothesis.

H2: An ontology-based question system enhances memory recollection.
The second hypothesis was confirmed in an experiment comparing an ontology-based question system with a list of broad questions. Participants (n = 24) were asked to bring pictures of four holiday memories. During the experiment, two of these memories were chosen at random to be described in a digital diary. For one of the memories, an ontology-based question system posed questions, for the other a simple list of questions with the same topics was used. The ontology-based system contains a knowledgebase of concepts relevant for a holiday memory. For instance it had concepts of different types of locations. Linked to these concepts were properties, which translated to open questions to the participants. For instance, if the option a shop were selected as location, a follow-up question would ask what the shop was selling. An analysis of the diary input of the participants showed that the answers to the questions contained more detail for the ontology-based system than for the question list, confirming the hypothesis.

H3: By presenting situation-based motivational messages a virtual agent enhances therapy trust and motivation.
To study the third hypothesis, a personalized motivational system was developed. By using input from experts (n = 13), a system was developed that is capable of generating motivational messages based on a patient's situation, as defined by PTSD symptom progression and trust in a good therapy outcome. This system was evaluated in an online study (n = 207) and compared with general motivational messages and a message without motivational content. The results show that motivational messages increase motivation to continue, trust in a good therapy outcome, and how much participants feel being heard by the system in general. Moreover, this feeling is improved further when the motivational message is personalized based on the progression of PTSD symptoms and people's trust in the success of the treatment. For example, in the situation in which PTSD symptoms were getting worse, the personalized motivational messages increased both motivation and trust more than general messages, indicating that adapting to the users' situation is especially relevant.

8.2. Limitations

To fully appreciate the findings presented in this thesis, it is important to consider the limitations of the studies described. One of the main limitations of this work is the lack of experimental data gathered from PTSD patients. Due to ethical reasons, the experiments did not specifically solicit participation of people currently suffering from PTSD. Instead, participation was sought from the general public, therapists,

and from former patients. To address the ability to generalize the findings to a clinical population, several strategies were used. First of all, as the treatment includes exposure to traumatic memories which PTSD patients often avoid doing, healthy participants were asked to recollect negative memories. In this way, an attempt was made to study adherence to a task people normally would feel reluctant doing. Next, in the online studies neither indication of having insomnia (60% of participants) nor indication of having of PTSD (32% of participants) were found to affect participant's answers to questionnaires. As these studies involved imagining situations, this indicates that the situations were interpreted similarly by people who had personal experience and those who did not. Third, participation of former PTSD patients was sought as they are in position to imagine how actual patients might behave and feel towards the PTSD therapy system. Although these strategies aimed at establishing findings that approximate results obtain from actual PTSD patients, caution should be taken when generalizing the presented findings to a clinical population.

A second limitation of the work is the lack of insight it provides into the virtual agents' ability to actually improve clinical outcome. However, to study this question a usable and safe therapy system is required. This did not exist when this research was started. Therefore, the work instead focused on studying proxy measures which have been identified to relate to clinical outcome. One example of such a measure is adherence, one of the key predictors for outcome in various e-health applications [3, 4]. Another important factor is working alliance, which has been shown to influence outcome for PTSD in an online setting [5]. Motivation was also studied, as factors such as trust in outcome and intrinsic motivation are related to therapy results for various disorders including PTSD [6–8]. Finally, usability was studied as it has been shown to influence trust [9, 10], which in turn is an important aspect of the aforementioned working alliance (Developmentandvalidation1989Horvath).

The third limitation to consider is that the different behaviors of the virtual agent have only been studied in isolation. Although the work gives some idea of their singular effect, they might also influence each other. For instance, when psychoeducation is presented effectively, patients might also answer the questions from the virtual agent in even more detail. However, these exact relations between the effect of behaviors are very difficult to predict, and would require a more thorough investigation to be properly studied. As a single study into combined behaviors would not show which behavior had which effect, this thesis choose to study the components separately first. In this way, these different behaviors might be used more reliably to achieve a specific outcome when individually applied to different applications and domains.

The fourth limitation is that this thesis did not consider the nonverbal behavior of the virtual agent. Other work has shown that nonverbal behavior can be useful for virtual agents. Gestures have been shown to improve communication [11], and gaze direction has the potential to positively affect disclosure [12] and guide attention [13]. Showing emotion also has the potential to increase disclosure [14] and alliance [15]. However, emotions can also be perceived negatively [16], so care

should be taken in designing these behaviors. Early on in discussion with therapists it was established that the agent would need to have a positively neutral expression most of the time. It was therefore deemed more important to focus on the verbal behavior in this thesis. However, a larger range of expressions, as well as more gestures and eye gaze behavior might further improve the agent.

Finally, the work reveals that a virtual agent might not be a suitable medium for all patients. Some people simply did not like the presence of an agent, or listening to its voice. Some also did not like having to follow step-by-step instructions with the virtual agent. Both of these misgivings might be lessened by further developments to the agent, such as improvement in text-to-speech systems and more dynamic dialogue systems. Nevertheless, it is important to consider the distinction between people who like a virtual agent, and people who do not. More research is needed to study how this distinction can be made, but its presence might explain the relatively low effect sizes found in studies comparing agent interfaces to those without [17].The work on this thesis started on the premise that a virtual agent is present, studying only what its behavior should be. It did not look into when a virtual agent should be present or not, and how the level of presence of an agent could be personalized to the user.

8.3. Contributions

8.3.1. Scientific

The main scientific contributions of this thesis lie in the insights gained into the behavior of a virtual agent for PTSD therapy, as well as about safety in AEMH systems and a system for PTSD therapy. Although virtual agents for PTSD patients existed [18], the application of an agent in a therapy system is a novel one. Given the specific nature of this context for a virtual agent, it was important to first establish the general user requirements. In this thesis, several design guidelines have been established specifically for a virtual agent for PTSD therapy. These guidelines can be a useful starting point for future work in agents for PTSD, as this specific user group also requires specific behaviors for a virtual agent.

Additionally, three specific behaviors for a virtual agent for PTSD have been studied, namely informing, assisting and motivating. Regarding information, this thesis compared oral and textual presentation of psychoeducation. Because there is both evidence that reading is preferable for retention [19], and that oral presentation increases alliance to a virtual agent [20], the exact effect of presentation mode by a virtual agent was unclear. The experiment presented in this thesis found that textual presentation might be preferable over oral presentation in increasing adherence to the agent. Given that psychoeducation is relevant for many different disorders, this finding could be applied to virtual agents for other mental health applications. This thesis did not, however, consider the combination of textual and oral presentation, nor the effect of further modalities such as gestures, and facial expressions. Further work would be necessary to study the effect of these combinations on adherence.

To provide assistance, it was necessary for the virtual agent to ask personalized

questions suitable to the experience of the user to enhance detailed memory recollection. This thesis presents an ontology-based question system, which successfully elicits a more detailed response to autobiographic memory related questions. Although both the combination of open and multiple choice questions [21] and the use of ontologies to formalize domain knowledge [22] have been applied to e-health before, their combination to elicit memory recollection is novel. This system, therefore, presents a new method to ask users questions without a full text-to-speech system being necessary. Aside from trauma recollection, similar systems could be useful for therapies such as expressive writing [23] or life review therapy [24].

Finally, a situation-based motivational message system was developed and tested. This system demonstrates a method for developing a system that generates personalized motivational statements based on input that does not require a large database for training. Moreover, this thesis argues that this particular system can be used successfully to motivate people, and is more effective than general motivational messages especially in situations that are more severe. Given the importance of motivation in behavior change [25, 26], these findings are relevant for many researchers in the field of behavior change support systems.

Aside from the general and specific behaviors of a virtual agent for PTSD, this thesis also considers both patient safety and the context of a virtual agent for PTSD in the form of a therapy system. Given the rise in AEMH applications, it is important to carefully consider the course of action if the patient's situation grows beyond the systems capabilities. Although general ethical guidelines for AEMH applications exist [27], specific procedural recommendations were lacking. This thesis therefore presents models on how to detect risks, and how to either automatically refer a patient or motivate them to seek help themselves. Given the high prevalence of serious risks such as suicide for some mental disorders [28, 29], such recommendation models seem very welcome. Because of the broad scope of these recommendations, they seem applicable to many different systems in the field of mental health care.

Finally, this thesis presents a full computerized therapy system for PTSD incorporating a virtual agent. This system is novel in combining exposure environments, therapy components and a virtual agent using an ontology-based question system into one therapy application. This makes it the first system of its kind, and by embedding the virtual agent in a complete therapy, it is now ready to be used and tested for clinical effectiveness.

8.3.2. Societal

Aside from the scientific contributions, this thesis is also relevant for a broader societal audience. As it describes a therapy system for PTSD, the main societal contributions are for PTSD patients, but also for therapists working in this field, as well as for developers and designers of AEMH systems.

Patients

This thesis describes the first steps of a home therapy system for PTSD including a virtual agent. PTSD has a large impact on quality of life [30], and following therapy can therefore make a great difference. Barriers to care such as travel times, cost,

and stigma currently still prevent people from seeking care [31]. Being able to follow a therapy at home can therefore be a large advantage to patients in itself, as it saves on travel cost and time, as well as being a very privacy-sensitive form of therapy. The addition of a virtual agent which increases compliance seems very beneficial, as this compliance might in turn lead to a greater therapeutic effect. In this way, the behaviors for a virtual agent described in this thesis have the potential to eventually lead to a reduction in PTSD symptoms and improvement to patient's quality of life.

Healthcare professionals

Aside from patients, the work presented in this thesis might also be beneficial to professionals providing mental-health care. Firstly, the development of a virtual agent and home therapy system for PTSD means that a certain group of patients might not need face-to-face treatment anymore. Given the waiting lists that currently exist within the Dutch mental health-care system of two up to three months [32], this can be a positive thing for the health-care system. It means that therapist efforts can be fully directed to those patients who do require a face-to-face treatment. A second contribution of this work is that it takes a closer look at exactly what behaviors have what effect on people. Many therapists learn their trade through practice and experience, and although this works well for human learners, it does mean that they cannot always pinpoint exactly what actions should be used in what situation beforehand. Because a virtual agent does require this precise and exact knowledge, its development leads to further insights into the therapeutic process. Finally, the work presented in this thesis gives an insight into the possible future roles of therapists. With e-health on the rise and the number of remote and online system growing, it seems reasonable to expect that the work of therapists will start to involve more remote monitoring tasks. This thesis presents one example of what this role could look like.

This thesis introduces a virtual agent and therapy system for PTSD therapy. Although this system is currently still in the clinical evaluation stage, the eventual goal would be to apply such systems in clinical practice. With the application of the 3MR_2 and other AEMH systems, new options for research also open up. The data collected through such systems could be used to further improve theories and models that underlie mental health-care interventions. These improvements could benefit future AEMH systems, but also research into regular clinical interventions.

Developers & Designers

The work presented in this thesis contributes to the knowledge of how a virtual agent should behave during PTSD therapy. It gives several general guidelines for behavior, and specific examples on providing information, assistance and motivation. These results can be useful for designers of virtual agents as part of PTSD treatment, but might also be applicable to other domains. For instance, psychoeducation is used for many mental health disorders such as bipolar disorder [33], addiction [34], and depression [35]. Moreover, motivation during therapy is not restricted to PTSD either. Aside from these knowledge-based contributions, this thesis also presents an ontology-based question system and motivation system. Both the ontology of traumatic memories [36] and the database of motivational statements

(see [37] and Appendix III) can be beneficial for future e-health projects on PTSD. Developers and designers could therefore benefit from the findings presented in this thesis for different projects in the field of virtual agents for health care.

8.4. Future Work

The first obvious step of future work for the virtual agent and PTSD therapy system described in this thesis would be on evaluating clinical effectiveness (CI1-12-S028-1). Although this work focused on compliance, the eventual goal of both agent and system would be to reduce PTSD symptoms. Aside from these clinical evaluations, work could be done on how to adapt the system to other modalities of use. The virtual agent described in this thesis is meant for patients who work alone, at home. However, the therapy application might also be used in settings with a therapist, or intermediate forms where a patient performs only parts of the therapy alone. The exact role of a virtual agent would change in this case, but exactly in what way would require more research. Another different way of applying the virtual agent and system would be to consider different groups of PTSD patients. At the moment, the agent is focused specifically on either childhood sexual abuse victims or war veterans. Although the only difference in the agent behavior currently lies in the content of their questions, more distinctions could be made. The military has a very specific culture, which could be catered to more. This acknowledgement of cultural background would also become important if the system were applied to people from different countries, for example asylum seekers with war related psychological trauma [38]. These people could potentially greatly benefit from a system if it provides them with support in their own language and be adapted to their trauma related context. Furthermore, because the system operates remotely it could offer support to people in remote refugee camps with their relative high concentration of people with PTSD symptoms [39, 40] and only limited access to mental health care. Modification to system might, however, be necessary to adapt to culture differences [41]. More work would also need to be done to study if the human monitoring currently in the system could be applied to such a remote context.

Although the therapy system is strongly based in cognitive behavioral therapy (CBT), it currently focuses most of its efforts on the behavioral aspects, letting patients experience that they can think about their trauma. The virtual agent would be a particularly useful interface to incorporate the cognitive restructuring common in CBT. It could, for instance, reflect more on how patients feel about recreating their memories, and reflect on unhelpful cognitions. A very common example for PTSD patient is the feeling of guilt. By using CBT, the virtual agent could play a role in explicitly asking people to reconsider this guilt question in an appropriate manner.

Another aspect of the virtual agent which was not studied relates to memory. Although the agent fills the ontology with concepts, it does not yet use these answers in the rest of the therapy much. However, because of the organized structure of ontologies, these have the potential to be used later on. One clear way to apply this knowledge would be in the 3D worldbuilder. By linking the concepts in the on-

tology to 3D models, this world might even be partly automatically built. Moreover, the agent might notice when things are missing in the 3D world that were discussed before.

8.5. Take-home message

Post-traumatic stress disorder can lead to a significant reduction in quality of life, and although treatments for PTSD exist, barriers to care still prevent many people from getting the help they need. Autonomous e-mental health applications have the potential to take away some of these barriers as they are easily accessible, private and cost-effective. However, as PTSD therapy requires patients to face the very traumatic memories they wish to avoid, assistance in the form of a virtual agent has the potential to make a difference. Therefore, the aim of this thesis was to study in what ways a virtual agent for PTSD therapy can enhance treatment compliance. This resulted in a set of guidelines for behavior of a virtual agent, as well as recommendations on how to present psychoeducation, an ontology-based question system that elicits more detail in answers, and a motivational system that motivates patients based on their situation. Furthermore, a set of recommendation models has been developed on how AEMH systems should deal with situations beyond their scope. Finally, this thesis presents a full therapy system for PTSD incorporating a virtual agent. Taken together, these contributions paint the picture of how a virtual agent for PTSD should act, and in what context it should operate. They show how a virtual character can shape people's behavior through adapting its own, and how a computer can support mental health in a human-like manner.

References

[1] C. Botella, B. Serrano, R. Banos, J. M. García-Gómez, and A. Palacios, *Virtual reality exposure-based therapy for the treatment of post-traumatic stress disorder: a review of its efficacy, the adequacy of the treatment protocol, and its acceptability,* Neuropsychiatric Disease and Treatment **11**, 2533 (2015).

[2] A. A. Rizzo, B. Lange, J. G. Buckwalter, E. Forbell, J. Kim, K. Sagae, J. Williams, B. O. Rothbaum, J. Difede, G. Reger, T. Parsons, and P. Kenny, *An intelligent virtual human system for providing healthcare information and support,* Medicine meets Virtual Reality (2011).

[3] C. Horsch, J. Lancee, R. Beun, M. Neerincx, and W. Brinkman, *Adherence to technology-mediated insomnia treatment: A meta-analysis, interviews, and focus groups,* Journal of medical internet research **17**, e214 (2015).

[4] L. Donkin, H. Christensen, S. Naismith, B. Neal, I. Hickie, and N. Glozier, *A systematic review of the impact of adherence on the effectiveness of e-therapies,* Journal of Medical Internet Research **13** (2011).

[5] B. Wagner, J. Brand, W. Schulz, and C. Knaevelsrud, *Online working alliance predicts treatment outcome for posttraumatic stress symptoms in arab war-*

traumatized patients, Depression and Anxiety **29**, 646 (2012), cited By (since 1996)0.

[6] S. Klag, P. Creed, and F. O'Callaghan, *Early motivation, well-being, and treatment engagement of chronic substance users undergoing treatment in a therapeutic community setting.* Substance Use and Misuse **45**, 1112 (2010).

[7] M. Price, J. Maples, T. Jovanovic, S. Norrholm, M. Heekin, and B. Rothbaum, *An investigation of outcome expectancies as a predictor of treatment response for combat veterans with ptsd: Comparison of clinician, self-report, and biological measures,* Depression and Anxiety **32**, 392 (2015).

[8] C. Hunt, K. Rooney, L. Humphreys, D. Harding, M. Mullen, and J. Kearney, *Prediction of outcome for veterans with post-traumatic stress disorder using constructs from the transtheoritical model of bahaviour change,* Australian and New Zealand Journal of Psychiatry **41**, 590 (2007).

[9] D. Salanitri, C. Hare, S. Borsci, G. Lawson, S. Sharples, and B. Waterfield, *Relationship between trust and usability in virtual environments: An ongoing study,* in *International Conference on Human-Computer Interaction* (2015).

[10] A. Fruhling and S. Lee, *The influence of user interface usability on rural consumers' trust of e-health services,* International Journal of Electronic Healthcare **2**, 305 (2006).

[11] H. Li, *Patterns of backchannel responses in canadian-chinese conversations,* in *Paper presented at the annual meeting of the International Communication Association, TBA, San Francisco, CA Online* (2007).

[12] S. Andrist, B. Mutlu, and M. Gleicher, *Conversational gaze aversion for virtual agents,* in *Intelligent Virtual Agents* (2013).

[13] S. D'Mello, A. Olney, C. Williams, and P. Hays, *Gaze tutor: A gaze-reactive intelligent tutoring system,* International Journal of Human-Computer Studies **70**, 377 (2012).

[14] M. Skowron, M. Theunis, S. Rank, and A. Kappas, *Affect and social processes in online communication-experiments with an affective dialog system,* in *IEEE Transactions on Affective Computing* (2013).

[15] C. Creed, R. Beale, and B. Cowan, *The impact of an embodied agent's emotional expressions over multiple interactions,* Interacting with Computers **27**, 172 (2015).

[16] T. Liew, N. Mat Zin, N. Sahari, and S.-M. Tan, *The effects of a pedagogical agent's smiling expression on the learner's emotions and motivation in a virtual learning environment,* International Review of Research in Open and Distance Learning **17**, 248 (2016).

[17] N. Yee, J. N. Bailenson, and K. Rickertsen, *A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces,* in *CHI 2007* (2007).

[18] A. Rizzo, B. Lange, J. G. Buckwalter, E. Forbell, J. Kim, K. Sagae, J. Williams, B. O. Rothbaum, J. Difede, G. Reger, T. Parsons, and P. Kenny, *An intelligent virtual human system for providing healthcare information and support,* Medicine Meets Virtual Reality **18**, 503 (2011).

[19] Webb and Wallon, *Comprehension by reading versus hearing,* The Journal of Applied Psychology **40**, 237 (1956).

[20] Qui and Benbasat, *Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems,* (2009).

[21] S. Friederichs, C. Bolman, A. Oenema, J. Guyaux, and L. Lechner, *Motivational interviewing in a web-based physical activity intervention with an avatar: Randomized controlled trial,* Journal of Medical Internet Research **16**, e48 (2014).

[22] T. Bickmore, D. Schulman, and C. Sidner, *A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology,* Journal of Biomedical Informatics **44**, 183 (2011).

[23] K. A. Baikie and K. Wilhelm, *Emotional and physical health benefits of expressive writing,* Advances in Psychiatric Treatment **11**, 338 (2005).

[24] G. J. Westerhof and E. T. Bohlmeijer, *Celebrating fifty years of research and applications in reminiscence and life review: State of the art and new directions,* Journal of Aging Studies **29**, 107 (2014).

[25] R. M. Ryan and E. L. Deci, *A self-determination theory approach to psychotherapy: The motivational basis for effective change,* Canadian Psychology **49**, 186 (2008).

[26] B. Fogg, *A behavior model for persuasive design,* in *Proceedings of the 4th International Conference on Persuasive Technology* (2009).

[27] D. D. Luxton, *Recommendations for the ethical use and design of artificial intelligent care providers,* Artificial Intelligence in Medicine **62**, 1 (2014).

[28] M. Pompili, L. Sher, G. Serafini, A. Forte, M. Innamorati, G. Dominici, D. Lester, M. Amore, and P. Girardi, *Posttraumatic stress disorder and suicide risk among veterans,* The Journal of Nervous and Mental Disease (2013).

[29] G. Turecki and D. A. Brent, *Suicide and suicidal behaviour,* Lancet **387**, 1227 (2016).

[30] J. Alonso, M. Angermeyer, S. Bernert, R. Bruffaerts, T. Brugha, H. Bryson, G. de Girolamo, R. de Graaf, K. Demyttenaere, I. Gasquet, J. Haro, S. Katz, R. Kessler, V. Kovess, J. Lepine, J. Ormel, G. Polidori, L. Russo, and G. Vilagut, *Disability and quality of life impact of mental disorders in europe: results from the european study of the epidemiology of mental disorders (esemed) project,* Acta Psychiatrica Scandinavica **109**, 38 (2004).

[31] V. Kantor, M. Knefel, and B. Lueger-Schuster, *Perceived barriers and facilitators of mental health service utilization in adult trauma survivors: A systematic review,* Clinical Psychology Review **52**, 52 (2017).

[32] *Wachttijden in de geestelijke gezondheidszorg (ggz),* Online report (2016).

[33] K. Coulthard, D. Patel, C. Brizzolara, R. Morriss, and S. Watson, *A feasibility study of expert patient and community mental health team led bipolar psychoeducation groups: implementing an evidence based practice,* BMC Psychiatry **13** (2013).

[34] M. W. B. Zhang and R. C. M. Ho, *Tapping onto the potential of smartphone applications for psycho-education and early intervention in addictions,* Frontiers in Psychiatry **7** (2016).

[35] H. Conradi, E. Bos, J. Kamphuis, and P. de Jonge, *The ten-year course of depression in primary care and long-term effects of psychoeducation, psychiatric consultation and cognitive behavioral therapy,* Journal of Affective Disorders **217**, 60 (2017).

[36] M. Tielman, *Ontologies traumatic memories,* (2017), doi: 10.4121/uuid:7ba977ef-891f-419b-983a-3aa8126e3a2e.

[37] M. Tielman, M. A. Neerincx, and W.-P. Brinkman, *Dataset belonging to the paper: Generating personalized motivational messages for a virtual agent that assists in post-traumatic stress disorder therapy,* (2017), doi: 10.4121/uuid:3396c380-6102-4043-9693-847a3e482e03.

[38] S. Sandhu, N. V. Bjerre, M. Dauvrin, S. Dias, A. Gaddini, T. Greacen, E. Ioannidis, U. Kluge, N. K. Jensen, M. Lamkaddem, R. P. i Riera v Zsigmond Kósa, U. Wihlman, M. Stankunas, C. Straŝmayr, K. Wahlbeck, M. Welbel, and S. Priebe, *Experiences with treating immigrants: a qualitative study in mental health services across 16 european countries,* Social Psychiatry and Psychiatric Epidemiology **28**, 105 (2013).

[39] F. Kazour, N. R. Zahreddine, M. G. Maragel, M. A. Almustafa, M. Soufia, R. Haddad, and S. Richa, *Post-traumatic stress disorder in a sample of syrian refugees in lebanon,* Comprehensive Psychiatry **72**, 41 (2017).

[40] K. I. Marwa, *Psychosocial sequels of syrian conflict,* Journal of Psychiatry **19** (2016).

[41] U. Schnyder, R. Bryant, A. Ehlers, E. Foa, A. Hasan, G. Mwiti, C. Kristensen, F. Neuner, M. Oe, and W. Yule, *Culture-sensitive psychotraumatology,* European Journal of Psychotraumatology **7**, 31179 (2016).

Appendix I: Psychoeducation content

Original Dutch version

In dit experiment ga je een hele negatieve herinnering uitwerken. Hier heb je wellicht niet heel veel zin in, het is immers niet leuk om terug te denken aan negatieve ervaringen. Het is begrijpelijk dat het in eerste instantie beter lijkt om vooral niet aan beladen momenten te denken. Waarschijnlijk heb je echter ook wel eens gemerkt dat het niet lukt om negatieve herinneringen te vergeten, en zelfs als deze op de achtergrond spelen kunnen ze toch grote effecten hebben. Het is dus vaak beter om toch stil te staan bij zulke ervaringen. Een goede manier om dit te doen is door deze herinneringen van je af te schrijven. Dit kan hele positieve effecten hebben. Er is in de afgelopen jaren veel onderzoek gedaan naar het van je af schrijven van negatieve ervaringen. Dit wordt ook wel 'Expressief schrijven' genoemd. Expressief schrijven is een methode waarbij mensen schrijven over stressvolle en aangrijpende situaties. Hierbij wordt gevraagd of iemand de meest stressvolle en negatieve ervaring die ze hebben willen beschrijven. Uit verschillende onderzoeken blijkt dat dit expressief schrijven een hele positieve uitwerking kan hebben. Hoewel het onderzoek naar Expressief schrijven nog zeker niet klaar is, is er gebleken dat het op vele vlakken positief kan werken. Het kan je bijvoorbeeld nieuwe perspectieven geven, en je anders naar jezelf laten kijken. Zo is bijvoorbeeld gebleken dat het opschrijven van negatieve herinneringen kan leiden tot persoonlijke groei, een positiever zelfbeeld, minder angst en minder negatieve gevoelens. Het is belangrijk hierbij op te merken dat het schrijven over neutrale gebeurtenissen deze effecten niet heeft. Het is bij het schrijven belangrijk om echt te focussen op de hele negatieve momenten uit je leven. Behalve de psychologische uitwerkingen kan Expressief schrijven nog meer voordelen hebben. Zo is bijvoorbeeld gebleken dat mensen die gewerkt hebben aan hun negatieve herinneringen in de periode hierna minder gezondheidszorg nodig hebben. Expressief schrijven kan dus ook helpen of fysieke klachten te verminderen. Verder is aangetoond dat het zelfs kan helpen om academische prestaties te verbeteren. Belangrijk bij Expressief schrijven is dat je echt je diepste emoties en gedachten onderzoekt. Terwijl je bezig bent met je herinnering, kan je deze ook linken aan andere levensgebeurtenissen. Bijvoorbeeld hoe het relateert tot je jeugd, je ouders, je relatie met anderen en je school of werk. Je kan het linken met wie je in de toekomst wil zijn, wie je vroeger was of wie je nu bent. Je gaat je herinnering uitwerken met het dagboek wat je eerder gezien hebt. Je kan beginnen met tekst toevoegen, en schrijven over je herinnering. Probeer ook andere mogelijkheden erbij te betrekken, voeg bijvoorbeeld een kaart toe, of emoties. Je kan ook bijvoorbeeld via internet foto's

opzoeken die relatie hebben met hetgeen je beschrijft. Probeer tijdens de rest van het experiment goed in gedachten houden wat je hier hebt gehoord. Hoe het van je af schrijven van gebeurtenissen een heel positief effect kan hebben, juist als het gaat over hele negatieve herinneringen. Probeer tijdens het beschrijven echt je diepste emoties en gedachten te onderzoeken.

English version, translated from the original Dutch

In this experiment you'll work on a very negative memory. You're probably not looking forward to this. After all, it's not fun to remember very negative experiences. It's understandable that it initially seems better to not recall these loaded moments. However, you've probably also noticed that trying to forget negative memories doesn't always work and even in the background they can have great influence. It's therefore often better to not forget these memories, but give them some attention. A good way to do this is to write about memories. This can have very positive effects. In past years, a lot of research has been done into writing about your negative memories. This process is also called 'Expressive writing'. Expressive writing is a method where people write about stressful and impactful situations. A person is asked to write about the most stressful and negative experience they have. Several studies show that this expressive writing can have very positive effects. Although the research into expressive writing is certainly not done, it has shown to have positive effects in many areas. It can, for instance, provide new perspective, and make it look at yourself in new ways. Writing about negative memories can lead to personal growth, a more positive self-image, less anxiety and less negative feelings. It is important to note that writing about neutral experiences doesn't have these effects. In writing, it is important to really focus on the very negative moments in your life. Except for the psychological effects, expressive writing can have more advantages. It has been shown, for instance, that people who've worked on their negative memories require less health care in the period afterwards. So expressive writing can also help to reduce physical complaints. Additionally it has even been shown to help improve academic results. Important in expressive writing is that you really investigate your deepest emotions and thoughts. While you're working on the memory, you can also link it to other life events. For example how it relates to your youth, your parents, your relationship with others and your school or work. You can link it to who you want to be in the future, who you were in the past or who you are now. You're going to work on your memory with the diary you've seen before. You can start with adding text, and write about your memory. Also try to involve the other possibilities, for instance add a map, or emotions. You can also use the internet to look up pictures that have a relationship with what you're describing. During the rest of the experiment, try to keep in mind what you've heard here. How writing about your memories can have a positive effect, especially when it's about very negative memories. Try to really investigate your deepest emotions and feelings while writing.

Appendix II: Questionnaires Chapter 3

Original Dutch version

Vragenlijst 1 Answered on 7 pt. Likert scale from 'Do not agree at all' to 'Completely agree'.

De volgende vragen gaan over de virtuele coach, en jullie samenwerking tijdens de sessie die je in het experiment gedaan hebt. Kies voor elke vraag welke stelling hier het beste bij past. Het eerste rondje correspondeert met 'Helemaal niet mee eens', het 7e rondje met 'Helemaal mee eens'. Als je mening hier ergens tussen ligt, kruis dan het rondje aan wat het meest overeenkomt met jou mening.

- Ik had het gevoel slechts een plaatje van een coach te aanschouwen
- De virtuele coach kwam echt op mij over
- Ik had het gevoel dat er een intelligent iemand aanwezig was als coach
- Ik geloof wat de virtuele coach mij vertelt
- Ik vertrouw de virtuele coach
- Ik voelde dat de virtuele coach niet helemaal eerlijk was betreffende zijn/haar gevoelens ten opzichte van mij
- Ik kan me voorstellen dat de interactie met virtuele coach zo ook in het echt kan plaatsvinden
- Ik denk dat de virtuele coach kennis heeft van zaken
- Ik denk dat de virtuele coach weet waar hij/zij het over heeft
- Ik geloof dat de virtuele coach mij aardig vond
- Ik denk dat de virtuele coach deskundig is
- Ik had het gevoel dat ik interactie had met een echte coach
- Ik geloof dat de virtuele coach echt bekommerd was om mijn welzijn
- De virtuele coach kwam betrouwbaar over

- Ik voelde me ongemakkelijk bij de virtuele coach
- Mijn motivatie om de opdrachten van de virtuele coach uit te voeren was erg laag
- Ik voelde dat de virtuele coach mij apprecieerde
- Ik wilde graag alle opdrachten van de virtuele coach goed uitvoeren
- Ik vertrouw dat de virtuele coach het beste met me voor heeft
- Ik denk dat de virtuele coach een expert is op het gebied van je herinneringen beschrijven
- Ik vond het niet belangrijk om de opdrachten van de virtuele coach goed uit te voeren
- Ik heb erg mijn best gedaan om te doen wat de virtuele coach van mij vroeg

Vragenlijst 2 Answered on 7 pt. Likert scale from 'Do not agree at all' to 'Completely agree'.

De volgende vragen gaan over de informatie die je gekregen hebt van de virtuele coach over het schrijven over negatieve gebeurtenissen. Selecteer bij elke vraag wel antwoord het beste past bij de stelling. Het eerste rondje correspondeert met 'Helemaal niet mee eens', het 7e rondje met 'Helemaal mee eens'. Als je mening hier ergens tussen ligt, kruis dan het rondje aan wat het meest overeenkomt met jou mening.

- De informatie kwam van een expert
- Het was moeilijk om de aandacht bij de informatie te houden
- De informatie was te lang
- Ik vond de informatie onpersoonlijk gepresenteerd
- De informatie was onduidelijk
- Ik vond de informatie prettig gepresenteerd
- De informatie was saai
- De informatie werd afstandelijk gepresenteerd
- De informatie hield de aandacht goed vast
- Ik was overtuigd door de informatie
- Ik vond de informatie betrouwbaar

- De informatie was te simpel
- De informatie werd fijn gepresenteerd
- De informatie was ingewikkeld
- De informatie werd vriendelijk gepresenteerd
- De informatie was goed te begrijpen
- Ik vertrouw de informatie

Vragenlijst 3 Answered with free writing in text box

De volgende vragen zijn bedoeld om te testen hoe goed je de gekregen informatie hebt onthouden.

- Hoe wordt het van je af schrijven van negatieve ervaringen genoemd ook wel genoemd?
- Op welke vlakken kan het van je af schrijven van negatieve ervaringen positief werken? Probeer er zoveel mogelijk te noemen.
- Wat is belangrijk bij het van je af schrijven van je negatieve ervaringen? Probeer zoveel mogelijk punten te noemen.

Vragenlijst 4 Example of how to answer in Figure 1, answer in Figure 2.

Je hebt in dit experiment 5 herinneringen kort beschreven, deze hierna op een schaal van minst naar meest aangrijpend geplaatst. Je hebt hier uiteindelijk 1 van herinnering uitgewerkt. Vul hieronder in welke herinnering je hebt gekozen om uit te werken door bij het gekozen nummer een kruisje te zetten. Ook heb je elke herinnering een getal tussen 1 en 100 gegeven die representeerde hoe erg deze herinnering was. Vul ook in welke getallen je de 5 herinneringen hebt gegeven op de schaal van 1 tot 100. Hieronder staat een voorbeeld van hoe je deze schaal kan invullen.
Als je niet precies meer weet welke nummers je had gegeven of welke je had gekozen, kan je de experimentleider vragen je dit te laten zien (het is mogelijk dit te doen zonder dat de experimentleider je beschrijvingen te zien krijgt). Roep in dit geval nu even de experimentleider erbij.

Vragenlijst 5 Example of how to answer in Figure 3, answer in Figure 4.

De volgende vragen gaan over hoeveel je je verdiept hebt tijdens het beschrijven van je herinnering. Zet een kruisje op het punt wat het beste correspondeert met je mening. Als je mening volledig overeen komt met een van de woorden aan het

Voorbeeld:

	Minst erg				Meest erg
Ordening	**1**	**2**	**3**	**4**	**5**
Zet een kruisje bij welke je uitgewerkt hebt		X			
Hoe erg op een schaal van 1 tot 100.	70	75	84	85	86

Figure 1: Example of how to answer

Vul deze zelf in:

	Minst erg				Meest erg
Ordening	**1**	**2**	**3**	**4**	**5**
Zet een kruisje bij welke je uitgewerkt hebt					
Hoe erg op een schaal van 1 tot 100.					

Figure 2: Answering scale

Voorbeeld:

- Ik vond dit experiment leuk om te doen

Helemaal niet leuk — X — *Heel leuk*

Hier zeg ik dat ik dit experiment grotendeels niet leuk vond, maar ook niet helemaal.

Helemaal niet leuk — X — *Heel leuk*

Hier zeg ik dat ik dit experiment erg leuk vond om te doen.

Helemaal niet leuk — X — *Heel leuk*

Hier zeg ik dat ik het experiment niet leuk vond, maar ook niet niet leuk.

Figure 3: Example of how to answer

Helemaal oneens — Helemaal eens

Figure 4: Answering scale

Figure 5: Answering scale

uiteinde van de lijn, zet je het kruisje aan die kant van de lijn. Zit het precies in het midden, zet je het kruisje in het midden, etc.

- Ik heb echt mijn diepste gevoelens onderzocht en beschreven.
- De herinnering die ik beschreven heb is zeer confronterend.
- Ik heb tijdens het beschrijven de herinnering ook gelinkt aan andere levensgebeurtenissen.
- Het beschrijven van de herinnering heeft de manier waarop ik er nu naar kijk beïnvloed.

Vragenlijst 6 Answered in Figure 5, with free text box for reasons why.

De volgende vragen gaan over de virtuele coach en de informatie die je hebt gekregen. De oefening waar in de vragen naar verwezen wordt is degene waarbij je 1 herinnering 15 minuten lang hebt uitgewerkt in het dagboek.

- De informatie over het beschrijven van je herinneringen voegde iets toe aan deze oefening
- De aanwezigheid van een virtuele coach voegde iets toe aan deze oefening
- De mogelijkheden in het dagboek voegden iets toe aan deze oefening
- Het feit dat de virtuele coach tegen me sprak voegde iets toe aan de oefening

Translated to English version

Questionnaire 1 Answered on 7 pt. Likert scale from 'Do not agree at all' to 'Completely agree'.

The following questionnaires are about the virtual coach and your cooperation during the experimental session. Pick the best answer for every question. The first circle corresponds with 'don't agree at all', the 7th with 'agree fully' . If your opinion is somewhere in-between, select which circle best corresponds with your opinion.

- I felt I was just perceiving questions of a virtual coach.

- The virtual coach seemed real to me.
- I felt that an intelligent being was present as coach.
- I believe what the virtual coach tells me.
- I trust the virtual coach.
- I felt that the virtual coach was not totally honest about his/her feelings toward me.
- I can imagine that the interaction with the virtual coach can happen the same in real life.
- I think the virtual coach knows its business.
- I think the virtual coach knows what he/she's talking about
- I believe the virtual coach liked me
- I think the virtual coach is knowledgeable.
- I felt as if I was interacting with a real virtual coach.
- I believe the virtual coach was genuinely concerned form y welfare.
- The virtual coach seemed trustworthy.
- I felt uncomfortable with the virtual coach.
- My motivation to execute the task set by the virtual coach was very low.
- I felt that the virtual coach appreciated me.
- I wanted to execute all task set by the virtual coach well.
- I trust that the virtual coach wants the best for me
- I think the virtual coach is an expert on writing about your memories.
- I didn't think it important to execute the tasks set by the virtual coach well.
- I did my best to do what the virtual coach asked of me

Questionnaire 2 Answered on 7 pt. Likert scale from 'Do not agree at all' to 'Completely agree'.

The following questions are about the information you got from the virtual coach about writing about negative memories. Select the best answer for every question. The first circle corresponds with 'don't agree at all', the 7th with 'agree fully' . If your opinion is somewhere in-between, select which circle best corresponds with your opinion.

- The information came from an expert.
- It was difficult to keep my attention with the information.
- The information was too long.
- I felt the information was presented in an unpersonal manner.
- The information was unclear.
- I felt the information was presented in a pleasant way.
- The information was boring.
- I felt the information was presented in a distant manner.
- The information retained the attention well.
- I was convinced by the information
- I felt the information was trustworthy.
- The information was too simple.
- The information was presented in a nice way.
- The information was complicated.
- The information was presented in a friendly manner.
- The information was understandable.
- I trust the information.

Questionnaire 3 Answered with free writing in text box

The following questions are meant to test how well you've remembered the information you received.

- How do we call the writing about negative experiences?
- On which things can writing about negative memories have a positive effect? Try to name as much as possible.
- What is important when writing about negative memories? Try to name as much points as possible.

Questionnaire 4 See Figures 1 and 2 in the Dutch version for example and answering scale.

In this experiment you have listed 5 memories, and ordered them on a scale from least to most stressful. In the end you further described 1 memory. Fill in below which memory you've chosen to write about in detail by putting an X at the chosen number. You also indicated how negative this memory was on a scale from 1 to 100. Fill in which numbers you've given the 5 memories on the scale from 1 to 100. Below is an example of how you can fill in this scale.
If you don't remember exactly which numbers you gave or which you've chosen, you can ask the experimenter to show you this (it is possible to do this without the experimenter seeing your descriptions). If you wish to do this, call the experimenter in now.

Questionnaire 5 See Figures 3 and 4 in the Dutch version for example answering scale.

The following questions are about how much you went in detail during the description of your memory. Put an X at the point which best corresponds with your opinion. If your opinion matches fully with one of the words at the edges of the line you put the X on that side of the line. Is it exactly in the middle, you put the X in the middle, etc.

- I truly studied and described my deepest feelings.
- The memory I described is very confronting.
- I have linked my memory to other life events during the description.
- Describing the memory has influenced the way I look at it now.

Questionnaire 6 See Figure 5 in Dutch version for answering scale, also given was a text box for reasons why.

The following questions are about the virtual coach and the information you've received. The exercise which is referred to in the questions is the one where you described one memory in detail for 15 minutes in the diary.

- The information about describing your memories added something to this exercise.
- The presence of the virtual coach added something to this exercise.
- The possibilities of the diary added something to this exercise.
- The fact that the coach spoke to me added something to this exercise.

Appendix III

Dataset database

Category	What?	Subcategory	What?
Empathy	Show empathy, e.g. I'm glad you feel this way.	Negatief	Say you feel bad/sorry, e.g. 'I'm sorry you feel this way'.
Note Work ethic	Note work of patient, e.g. 'You've been working very hard on your sessions'.		
Note PCL	Note scores from PCL, e.g. 'I see your scores have been rising'.	Dropping	e.g. 'Your scores are dropping'.
		Rising	e.g. 'Your scores are rising.
		Stable	e.g. 'There's little change in your scores.
Note trust	Note trust patient, e.g. 'You still have trust in the therapy'.	Low	e.g. 'I see you have little trust in the therapy'.
		High	e.g. 'I see you have trust in a good outcome'.
Future	Refer tot he future. E.g. 'You'll work on your memories coming sessions.	PCL	Refer to scores in the future. E.g. 'In the rest of the therapy your complaints will lessen'.
Question	Question.	Why	Question of 'why' type. E.g. 'How come you feel this way?'
Give perspective	Put current situation into perspective. E.g. 'It's very normal that you feel this way. Many people in therapy feel like this'.		
Motivation	General motivating statements	Compliment	e.g. 'You're doing very well!'
		Go on like this	e.g. 'Keep going like this!'.
		It's going well	e.g. 'You're doing well!'
		Hold on	e.g. 'Hold on!'
		Good luck	e.g. 'Good luck!'
Together	Refer to doing things together. E.g. 'let's tackle this session together'.		
Psychoeducation	Refer to the rationale of treatment. E.g. ' You can always read back the rationale behind this therapy'.		
Tips	Tips, e.g. 'you can try to do a relaxation exercise before you start'.		
Contact	Mention contact with a care giver, e.g. 'You can always contact your therapist if you're not feeling well'.		

Figure 6: Categories and subcategories of statements

	PCL	Dropping			Stable			Rising		
	Trust	High	Med.	Low	High	Med.	Low	High	Med.	Low
Category	Subcategory									
Contact		0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01
Empathy		0.04	0.07	0.04	0.00	0.01	0.10	0.04	0.10	0.32
	Negative	0.00	0.01	0.04	0.00	0.01	0.08	0.04	0.08	0.30
Motivation		0.77	0.66	0.54	0.77	0.66	0.60	0.77	0.60	0.54
	Compliment	0.19	0.01	0.01	0.06	0.00	0.01	0.23	0.03	0.06
	Go on like this	0.31	0.14	0.07	0.07	0.03	0.00	0.00	0.01	0.00
	It's going well	0.06	0.06	0.04	0.13	0.09	0.06	0.02	0.04	0.02
	Hold on	0.03	0.11	0.06	0.08	0.22	0.32	0.11	0.15	0.27
	Good luck	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Psychoeducation		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Give perspective		0.04	0.04	0.30	0.13	0.56	0.80	0.61	0.66	0.66
Together		0.01	0.07	0.05	0.07	0.13	0.13	0.13	0.13	0.13
Status		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tips		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Future		0.09	0.04	0.12	0.12	0.06	0.28	0.09	0.09	0.04
	PCL	0.09	0.00	0.09	0.09	0.06	0.16	0.09	0.09	0.03
Question		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Why	0.01	0.01	0.03	0.01	0.01	0.00	0.00	0.02	0.02
Note work ethic		0.01	0.01	0.00	0.02	0.01	0.01	0.01	0.01	0.02
Note PCL		0.95	0.97	0.97	0.95	0.97	0.95	0.95	0.93	0.89
	Stable	0.00	0.00	0.00	0.94	0.95	0.93	0.00	0.00	0.00
	Dropping	0.89	0.93	0.93	0.08	0.11	0.11	0.05	0.00	0.02
	Rising	0.00	0.00	0.00	0.05	0.02	0.02	0.95	0.95	0.88
Note trust		0.18	0.41	0.48	0.23	0.23	0.23	0.23	0.34	0.18
	High	0.09	0.00	0.00	0.13	0.00	0.00	0.13	0.00	0.00
	Low	0.00	0.42	0.49	0.00	0.24	0.24	0.00	0.36	0.19

Figure 7: Probabilities of a given category in a given situation, defined by PCL score and Trust

PCL:	Dropping			Stable			Rising		
Trust:	High	Med.	Low	High	Med.	Low	High	Med.	Low
Mean Nr. statements	3.08	3.42	3.58	3.31	3.62	3.73	3.84	4.04	3.89

Figure 8: Mean nr. of statements in a given answer per situation, defined by PCL score and Trust

Feedback Statements

Personalized motivational feedback as generated by system.

Dropping PCL – High Trust

- You've achieved a lot in a short time in therapy, your initially substantial complaints are getting less. Compliments on facing the sessions. The symptoms will lessen automatically.
- You can see the tension getting less, and that's very quick. You're doing well! Continue like this!
- You've achieved a lot in a short time in therapy, your initially substantial complaints are getting less. Try to continue like this. You're doing well!

Dropping PCL – Doubts

- The treatment seems to have a good effect on you. You probably doubt if the treatment is useful. You're working hard, continue like this!
- Step by step you see your complaints getting less. Now it's important to stick with it! We're over half way in and you're doing well.
- You're doing better this week than last week. I'm sorry to hear you have little trust in the future. You're doing well.

Dropping PCL – Low trust

- Perhaps you're hesitating whether the therapy is working for you. That's a shame, that you don't have a lot of trust. Your complaints have gotten a lot less. The end is in sight, continue like this!
- I'm sorry to hear that you don't have a lot of trust in the future. Perhaps you're hesitant whether the therapy is working for you. Your complaints started to get a lot less once we started to discuss the bad memories. It's going really well.
- You're probably doubting whether the treatment is useful. That's a shame, that you don't have a lot of trust. You've noticed that your complaints are getting less. You're doing very well.

Stable PCL – High trust

- Because we're now really starting to work on your memories you might feel a bit up-and-down. Try to continue as you are. After this phase we often see the complaints starting to drop.

- It's normal that the complaints don't drop immediately. In your case they're the same as when you started working on the memory. Hold on! You're doing well

- You indicate that you experience little change in your complaints since the start of therapy. That's okay, you're doing well! Now it's important to stick to it!

Stable PCL – Doubts

- Your complaints show little consistency. Sometimes they get less, and then they get worse again. That makes sense, you're working on your past and that's tough. Stick to it now! Let's see together what you need to continue.

- It's difficult to do and you're not noticing a lot yet. You have doubts whether this treatment can help you. It's common that when you work on your past, your complaints get worse. Stick to it!

- Your complaints show little consistency. But that's understandable because we're now working on your memory. It's primarily important to stick to it now! You're still doing well.

Stable PCL – Low trust

- Your complaints show little consistency. Sometimes they get less, and then they get worse again. But the results are also as you can expect when you start intensive treatment. Have hope. You can count on it that the complaints will lessen when you're working on your next memory.

- You notice your complaints going up and down. That's a shame. Your reaction is normal in this stage of therapy. You're doing well, try to stick to it!

- Your complaints are fluctuating, but don't really seem to go down. That makes a lot of sense in a situation like this. You're going in the right direction! Have courage!

Rising PCL – High trust

- Your complaints might have gotten worse. The falling and getting up fits with doing therapy. Confronting new memories will be a challenge, but you're doing well. You're doing something difficult, you're really going for it.

- You probably also notice that your complaints are getting worse. This is a normal reaction in this stage of therapy. It's good to see that you have trust in a good outcome. Try to stick to it.

- You probably also notice that your complaints are getting worse. But it happens often in trauma treatment, and most of the time this means you're doing well and getting closer to the core. You're going in the right direction! Well done!

Rising PCL – Doubts

- You notice that your complaints are getting worse. This can be a part of the treatment and says nothing about the eventual result. You doubt whether you'll be fully recovered at the end. This is the moment where it's important to stick to it!
- You probably also notice that your complaints have gotten worse recently. It sounds strange, but that is a good sign. Stick with it! Let's look together at what you need to continue.
- The last couple of sessions have given rise to a lot of tension. Perhaps you have doubts about whether the therapy works for you. But that is understandable as we're working with the memory now. Stick to it!

Rising PCL – Low trust

- I see you indicate that your complaints have gotten substantially worse. I'm sorry to hear that. However, it is always hard work before we see any results. Stick with it and continue.
- You're feeling a lot from this treatment, aren't you! Perhaps you have doubts whether the therapy works for you. We see this often during trauma treatment. Just a couple more sessions to go, hold on!
- I see your complaints are getting worse over the sessions. Therapy knows several phases and at the beginning of every phase you can see a rise in complaints because it is new. You're still doing well. Stick with it and continue.

General feedback

- Treatment is difficult, but only lasts a couple of weeks. For three-quarter of people the complaints get a lot less due to this treatment. Without treatment the complaints usually last, or even get worse. That would be even more difficult for you. So let's go for this together so that it might be difficult now, but it won't be in the future.
- From scientific research we know that the majority of people benefit from this treatment. We also know it is a difficult treatment, because we ask you to expose yourself to memories and images you'd rather not remember. By exposing yourself to these images and memories instead of fighting them, you'll notice that you're able to bear the fear and tension. And that the images will elicit less tension over time.

- Try to stick with it, even though your tension gets high. Only then can you check if your fear of not being able to handle the tension is true. That's difficult, but you've noticed before that if you stick with it, you can deal with the tension. And that the tension often even gets less after a while. That will bring you a step closer to recovery.
- You've now been working on this treatment for a while. Perhaps you've noticed already that your complaints are slowly getting less. It's important to stick with it now, so we can achieve the largest result.
- Some people experience the treatment for PTSD as difficult. They also notice that the complaints can temporarily get worse. Do you recognize this? Continue, despite the difficulty. You can do it!
- Good that you're actively working on the treatment. Try to stick to it and you'll notice your complaints getting less and less.
- Do you recognize that you sometimes avoid working on the treatment? Avoidance is one of the characteristics of PTSD and one of the reasons your problems stay as they are. Do you catch yourself at this? Try to put yourself to doing the session anyway.
- Hold onto the fact that the memories we're working on are in the past. Stay in the here and now: you're the one who can control the emotions.
- It's completely normal that therapy can elicit difficult emotions. Remember that this will get less during therapy.

No feedback statements

- Thank you for filling in the questionnaires. We'll now continue with the session.
- Thanks for answering these questions. Let's continue with the session.
- Thank you for answering the questions.
- Thanks for filling out these questions. Let's now continue with the session.
- Thanks for answering the questionnaires.
- Thank you for your answer to these questionnaires.
- Well done. Let's continue with the session now.
- Well done answering the questions. We'll now continue with the session.
- Well done filling in the questionnaires.

Questionnaires

PTSD-Check List (PCL) This questionnaire was rated on a 5pt. Likert scale ranging from 'not at all' to 'extremely'

For this questionnaire, please keep in mind the most stressful or traumatic experience you have had. In the past month, how much were you bothered by:

- Repeated, disturbing dreams of the stressful experience?
- Suddenly feeling or acting as if the stressful experience were actually happening again (as if you were actually back there reliving it)?
- Avoiding memories, thoughts, or feelings related to the stressful experience?
- Avoiding external reminders of the stressful experience (for example, people, places, conversations, activities, objects, or situations)?
- Being "superalert" or watchful or on guard?
- Feeling jumpy or easily startled?

Question per situation This question was answered on a slider ranging from 'it decreased a lot', via 'nothing changed' to 'it increased a lot'.

After seeing the reaction by your virtual coach, how did these comments **change** your:

- Motivation to continue
- Trust in a good therapy outcome

Overall questions This question was answered on a slider ranging from 'it decreased a lot', via 'nothing changed' to 'it increased a lot'.

ow did the comments of the virtual coach change your:

- Motivation to continue
- Trust in a good therapy outcome

How much did you feel that:

- The virtual coach really addressed your needs (PSQ1)
- The virtual coach took you seriously
- You're satisfied with the (emotional) support you received from the virtual coach (PSQ1)

- The virtual coach replied appropriately to you
- The virtual coach payed full attention to what you were trying to tell him/her (TPS-5)
- The virtual coach listened to your preferences
- The virtual coach only thought about what is best for you (TPS-8)

Acknowledgements

At the end of this book, I would like to extend my gratitude to everyone who contributed in some, small or larger, way to this work. First, to my supervisors for their support. To Mark, for thinking of me when this project was first starting, and for all the new insights and final critical looks along the way. And to Willem-Paul, for your efforts in making this dissertation the best it could be. For all the meetings, discussions and revisions which shaped this research.

Although this dissertation is my work, it was done within the larger context of the VESP project, and would not have been the same without all the other partners involved. Firstly, thanks to Marieke for all your input on the clinical side of things, and for your continued work on testing our ideas in the field. I wish you all the luck, in finishing your own dissertation work, but also with your new family. Also a big thanks to Ben, for all the work on the 3D editor which really completes 3MR2. To Nexh and Eric, for your input on our ideas. To everyone from CleVR who assisted in the integration of the agent into 3MR, for letting me visit and learn from you. And to all our clinical partners, Jan and Colin in particular, for generously sharing your knowledge with us, many times over.

My thanks also to all of my colleagues in Delft. To Catholijn for the continued efforts to make II a welcoming department for everyone. To Bart and Ruud, for always being open for any technical questions, and keeping us operational. And to Anita, for assistance in so many small things, and for allowing me to regularly come in and adjust your blinds. To Dwi, for your help in the beginning, and saving me when I got stuck working with Vizard. To all of my room mates. Vanessa, for your critical thinking, and to Siska, for all your cheer. To Corine, for welcoming me to II, for sharing frustrations and helping when I got stuck on translations. To Ursula, for your input on my system and testing 3MR. To Ding, for the good food and teaching us about China. To Fransizka, for all the cake, and to Bernd. And to all other present and former colleagues, for lunchtime or after-colloquium chats, for participating in my experiments and for general good times.

A big thank you also to all my friends, as it's definitely also the off-work times that make work possible. And a special thank you to Silvia and Maaike, who started their own PhD projects around the same time I did. For sharing stories, of frustrations and successes, and the little things. And finally, to my family, my parents in particular, for always supporting me in whatever I wanted to pursue. And to Jeroen, for all the support and advice, even if I'm generally too stubborn to listen. This wouldn't have happened without you.

Curriculum Vitæ

Myrthe Lotte TIELMAN

08-12-1989 Born in Maarssen, the Netherlands.

Education

2008-2011 Bachelor Cognitive Artificial Intelligence
Utrecht University, Utrecht

2011-2013 Master Cognitive Artificial Intelligence
Utrecht University, Utrecht

Work experience

2012-2013 Research Internship
TNO Human Factors, Soesterberg

2013-2017 PhD. Interactive Intelligence
Delft University of Technology, Delft
Thesis: A Virtual Agent for Post-Traumatic Stress Disorder Treatment
Promotor: Prof. dr. M.A. Neerincx

2017- Post-doctoral researcher. Interactive Intelligence
Delft University of Technology, Delft

List of Publications

6. **M.L. Tielman, M.A. Neerincx, M. van Meggelen, I. Franken, W.P. Brinkman**, *How should a virtual agent present psychoeducation? Influence of verbal and textual presentation on adherence*, Health & Technology **25**, 6 (2017).

5. **M.L. Tielman, M.A. Neerincx, R. Bidarra, B. Kybartas, W.P. Brinkman**, *A Therapy System for Post-Traumatic Stress Disorder Using a virtual agent and virtual storytelling to reconstruct traumatic memories.*, Journal of Medical Systems **41**, 125 (2017).

4. **M.L Tielman, M.A. Neerincx, W.P. Brinkman**, *Generating Situation-Based Motivational Feedback in a PTSD E-health System*, Int. Conf. on Intelligent Virtual Agents (2017).

3. **M.L. Tielman, M. van Meggelen, M.A. Neerincx, W.P. Brinkman**, *An Ontology-based Question System for a Virtual Coach Assisting in Trauma Recollection*, Int. Conf. on Intelligent Virtual Agents (2015).

2. **M.L. Tielman, W.P. Brinkman, M.A. Neerincx**, *Design Guidelines for a Virtual Coach for Post-Traumatic Stress Disorder patients*, Int. Conf. on Intelligent Virtual Agents (2014).

1. **M.L. Tielman, M.A. Neerincx, J.J. Meyer , R. Looije**, *Adaptive Emotional Expression in Robot-Child Interaction*, Int. Conf. on Human-Robot Interaction (2014).