Conversational Agent-Based Behaviour Change Support Systems: Comparing the Technical Aspects of Recent Applications

Adriaan de Vos, Jasper Nieuwdorp, Ruben Keulemans
Interactive Intelligence Group – Delft University of Technology
{a.d.devos, j.nieuwdorp, r.keulemans}@student.tudelft.nl

Abstract This literature review consists of two parts. The *Narrative* part describes the development of Conversational Agents (CAs) and the techniques they use. The development of textual and embodied CAs, dialog generation algorithms, techniques to make the CAs intelligent and different forms of input and output are described. The *Systematic* part compares 18 recent, specific conversational agent implementations in the health domain used for behaviour change; CA-based Behaviour Change Support Systems (BCSSs). The implementations are classified based on certain features to compare them. Additionally a comparison is made between 11 recent approaches without concrete implementations on technical details and type of addressed issue. The *Narrative* part provides the historical and technical foundation for the classification and analysis in the *Systematic* part. The review concludes with a discussion of our findings and recommendations for future development of CA-based BCSSs.

Keywords Conversational Agent (CA), virtual agent, Behaviour Change Support System (BCSS), health, intervention, dialog system, e-coach, features.

1 Introduction

Health care provided by humans has various disadvantages: it is expensive and inefficient because only certain patients can be treated at a time, and a health care provider is mostly bound to their own knowledge and experience. Furthermore, it is not anonymous. With the advent of computers and other personal digital devices, such as smartphones and tablets, opportunities have arisen to address these issues by means of digital health care systems. Nowadays, different medical systems exist to treat a variety of disorders and injuries. In this review a specific type of system is reviewed: Conversational Agent-based Behaviour Change Support Systems (BCSSs), medical support systems to help users change their behaviour by interacting with a digital agent. For the scope of this report, the following definitions are used:

- *Conversational Agent.* Digital entity being able to have an interactive conversation with a human using natural language.
- *Behaviour Change Support System.* A computer-based system able to change a persons physical or mental state with a predetermined goal.

CA-based BCSSs have different features and implementations and aim to treat many disorders. This review gives an overview of the different possibilities and techniques used in early- and recent approaches. To this end the review is divided into two main parts: The *Narrative* part describes the developments of CAs and the corresponding techniques. The *Systematic* part compares 18 recent CA-based BCSS implementations in detail with regard to features and technical details. Furthermore, 11 novel models without concrete implementations are listed and compared. The *Narrative* part thus provides the historical an technical basis for the *Systematic* part. A conclusion and discussion concludes the report, therein both parts are compared and remarkable observations are stated.

This report answers the following research question: "What are the different techniques that are used for conversational agents for behaviour change therapy?". To this end, two sub-questions were defined. The narrative part will answer "Which mechanisms have been used in conversational agents throughout the years?" For the systematic part the sub-question is "Which modern conversational agent techniques are prevalent for behaviour change therapy?"

2 Research Method

The process for selection for literature of the systematic part is illustrated in the PRISMA diagram in Figure 1. After initial research, the criteria for the literature to be selected were decided. All references included were written in English and published after 2011, as the focus for the systematic part of the paper was modern conversational agents. When the final query was made for the database search, a manual screening of the records was performed. The records were equally divided among researchers for this, without duplication. Based on the title and the abstract of a record it was tested on the following criteria: Mentions Conversational Agents (following the definition in section 1), lies within the health domain for humans and must include a technical approach. If a record's title or abstract did not meet one or more of those criteria the record was excluded in this step. This reduced the number of records to 93, after which they were divided among the researchers in such manner that each record would be reviewed by exactly two researchers. Next, every researcher labeled the records for inclusion (or exclusion) after a full text study. There were four reasons for which records would be excluded from further research: no conversational agent had been used, not applicable to behaviour change for humans, full text not available, a lack of technical details. A full overview of excluded records with their reason can be found in appendix A.2, Table 3. Since the labeling by the researchers took place separately and blind, the correlation between researchers indicates agreement and reproducible results. The average agreement correlation between researchers was 0.84, ranging from 0.77 to 0.90. The remaining papers were included in the systematic part. In the case of disagreement between researches the arguments for exclusion were discussed until agreement was reached. Depending on whether it concerns a specific implementation it was included in either Table 1 or Table 2.

For the narrative part, a less systematic approach was taken. Based on literature discovered in the database search and specific searches in Scopus the appropriate literature was selected. Based on these sources the overview of mechanisms in conversational agents applied in health care was written.

3 Narrative Part – A comparison of the mechanisms used in conversational agents in health care throughout the years

Affective computing is an emerging discipline in the field of computer science and the implementation of active computer support in healthcare has gained momentum in recent years [1]. This has not always been the case as both technology and society had to mature in order to create an environment where conversational agents are ready to assist or even replace healthcare professionals. This section will give an overview of the developments in technology that enabled conversational agents to reach the point they are at today. Furthermore, it will highlight some crucial advancements that stood at the forefront of mechanisms that are vital to BCSSs.

3.1 Pattern Matching

In 1968, pattern matching became popular due to the usage of regular expressions within a text editor [2]. Pattern matching is a very efficient way of programmatically editing a text or string. Originally pattern matching worked by constructing a Finite State Machine (FSM) from the keywords, and then using the FSM to process the text string in a single pass [3]. When the more advanced tree-based pattern matching algorithm was invented, it gave a big boost to the versatility and effective use of pattern matching [4]. This technological progress has been used to quickly and in real-time edit relatively big amounts of text data. Retrieval-based BCCSs, see Section 3.3, make use of pattern matching to understand the response or questions from the patients, and they can be used to edit and personalize the responses to the patient [5].

3.2 Rule-based and AIML

ELIZA by Jospeh Weizenbaum [6] is widely regarded as the first ever conversational agent. It uses pattern matching to mimic human conversation and is rule-based. It has typical Rogerian therapist-like responses since it was designed to be perceived as a real doctor [7]. Eliza works as follows: it has certain predefined rules that are triggered by keywords. If a certain keyword is found, with pattern matching, it will transform the input to a response. Rule-based systems are common and relatively simple to develop because they primarily make use of a finite state machine.

Artificial Intelligence Markup Language (AIML) is an useful tool for creating a rule-based conversational agent with pattern matching. It is an important development, as it provides a foundation for many natural language software agents. It consists of categories that can be considered as the knowledge base of a system, a category also has pattern and template elements. When a pattern is matched on the input provided by the user, the appropriate template is used and transformed as output by the system. AIML formed the basis for A.L.I.C.E. (Artificial Linguistic Internet Computer Entity). A.L.I.C.E. [8] is a conversational agent that has won multiple Loebner prizes (2000, 2001 and 2004)[9]. It is an instantiation of AIML with more that 40.000 categories. A.L.I.C.E. falls in the category of supervised learning, since a "botmaster" can add new categories to change the behaviour of the CA. A.L.I.C.E. also formed the basis for Pandorabots, a server sided AIML interpreter that can be used to deploy conversational agents [8].

3.3 Retrieval-based versus Generative-based

Retrieval-based and generative-based BCCSs use different ways of determining the dialogue response to the patient. Retrieval-based conversational agents require many hand written rules or big sets of hand labeled data (corpus) to become effective. This method is relatively easy to develop because it consists of creating pattern matching rules to form a finite state machine. But this method is also time-consuming because the rules and the corpus need to be hand edited. Most current conversational agents use retrieval-based methods combined with a markup language like AIML[10] or ECML[11]. Partially observable Markov decision processes are often used to learn optimal dialog policies for the dialogue manager [12].

More modern conversational agents use the generative-based method. This approach is more difficult to develop but it promises a better end-result [13]. Deep learning and reinforcement learning are effective learning solutions for creating the dialogue manager of conversational agents, even when there are no hand written rules or hand labeled training data [13].

Reinforcement learning is an useful tool for creating a generative-based conversational agent, more on this can be found in Section 3.6.2.

3.4 Mental Models

While ELIZA was designed to be a therapist, PARRY (1971) [14] simulated a patient (a person suffering from paranoid schizophrenia). While both are rule-based, PARRY has one major advantage compared to ELIZA, it has a mental model, meaning that it has affective variables that simulate a state of mind for the program. This also made it possible for PARRY to have conversational strategies. Even though this is being debated, PARRY was the first CA to pass the Turing test [15].

3.5 Input Recognition

The patient's input to the conversational agent will be processed by a recognizer. Often this will be done through a Natural Language Interface. Natural Language Interfaces are key to creating a conversational agent that mimics a human psychotherapist. Immersion and effectiveness of the therapy will increase if patients can communicate with their voice and facial expressions as opposed to using a text-only interface [16]. The input from the patient will not only consist of the sentences, but will also contain possible emotional features observed through speech or facial expressions. The subsections below describe the various techniques used and their disadvantages and advantages.

3.5.1 Text Input

For some conversational agents text processing is relatively easy because the patient only has a few options to choose from [17]. This allows the conversational agent to fully understand all responses and optimally utilize them to improve the therapy. The big disadvantage with this is that the patient interaction is low. The patient only has a handful of options to choose from, and has no way to express himself with their own sentences. This causes the conversational agent to have limited emotional awareness of the patient.

The more advanced and state-of-the-art conversational agents enable the patients to type their own sentences as a response to the conversational agents questions. Some conversational agents enable a full two-sided conversation by allowing the patient and conversational agent to both ask questions to each other [18]. When the patient is able to formulate their own sentences they can fully express themselves instead of having to choose from a few options. This enables the patient to give more personal and emotional responses. The big challenge with this technique is how the conversational agent is able to understand the sentences from the patient. This is often done through Natural Language processing and is a complicated but promising technique that will be discussed later in this paper.

3.5.2 Speech Input

Speech processing is a natural way to communicate for humans and is thus a very efficient way of enabling the patient to easily and naturally communicate with the conversational agent. Patients are most of the time not confined to certain voice commands but can freely formulate sentences that they want to say. A big advantage with using speech as a medium to communicate is that it is inherently multidimensional. The meaning of the spoken sentence is not only the pronounced words, but also the various emotional features that can be extracted with Artificial Intelligence [19]. Of course, there will be a problem with speech recognition if the patient is slurring his words or has a dialect. Having problems with the speech recognition could cause problems with the immersion of the patient if the conversational agent does not understand the patient.

3.5.3 Facial Expressions

The affective state of a patient can be marginally determined based on speech processing. But human emotion is best visible by observing facial expressions [20]. Facial Expression Recognition is a very promising technique for efficiently and reliably monitoring the affective state of the patient. An important prerequisite for reliable facial expressions is the visibility of the face of the patient. This can often be obstructed by hands, looking the other way or glasses/beards. The processing of the facial expressions is often done with machine learning classification techniques.

3.6 Artificial Intelligence Usage

In this paper, Artificial Intelligence is defined as an algorithm that learns, improves or revises itself with a predetermined goal. Common IF statements or other established rules for decision making that are programmed by the developers do not fall under our definition of Artificial Intelligence. This section will describe and explain the various Artificial Intelligence techniques that have been used for BCCSs in the researched papers. The big advantage of using these techniques is that irregular, multidimensional data can be processed more reliably than with classic programming techniques.

3.6.1 Modeling Affective State

BCCSs show promising results for behaviour change therapy when the conversational agent can model the affective state and thus understand the patient better [16]. Modeling the affective state is done by using a combination of analyzing the input text, spoken voice and facial expressions. This data is hard to model with traditional programming so Artificial Intelligence is a promising technique and the various aspects will be discussed in the next section.

A Random Forest Classifier is a machine learning technique that works especially well with a multidimensional dataset because it uses hyperplanes to classify the various features of the data points. Automatically classifying the facial expression of a human is a promising usage of random forest classifiers to extract the emotion from the geometrical facial features from the patient as shown in [19].

A Dynamic Bayesian Network that uses a Hidden Markov model is a valid way to model the affective state of a patient [19]. This model can be used to gradually evolve from one dialog step to the next one and is able to simulate probabilistic reasoning and uncertainties that often occur within real world data. Dynamic Bayesian Networks can also be effectively used to reason and track hidden user goals [18].

3.6.2 Reinforcement Learning

Reinforcement learning provides robustness against typing errors and speech recognition errors in noisy environments, and it also reduces the cost of creating the hand written, often complex, rule-based dialogue managers. Experiments show that Reinforcement Learning outperforms the handwritten rule-based dialogue approach [21].

Modern conversational agents that use reinforcement learning make use of partially observable Markov decision processes, more information about this in Section 3.6.4.

Development of a reinforcement learning dialogue manager needs a hand-crafted reward function and a large dialogue corpus. This dialogue corpus can be collected in a relatively efficient way by letting a group of people, or a group of conversational agents converse in a so-called 'Wizard-of-Oz' experiment [22]. This is, of course, a disputable way of creating the corpus because the reward function cannot take into account the emotional state of the patient as it is very difficult to create honest reward signals from human users. After the dialogue corpus is created, both the reward function and the dialogue corpus will go in the dialogue manager which uses reinforcement learning to create and optimize a policy to reach the end-goal.

Reinforcement learning will optimize the conversational agent by tweaking the various parameters from the dialogue model used by the dialogue manager [23]. This allows the conversational agent to find the most effective way to function, allowing for better model than rule-based systems.

3.6.3 Natural Language Processing

Natural Language Processing is the science of enabling computers to understand the natural human language. This is traditionally very hard to do because the human language has many of nuances and exceptions to various rules. Throughout history, many researchers tried to manually model the human language with hard coding rules about the grammar and sentence structures [24]. This works well for computer languages but was not effective for natural languages because it is not robust to the variation, nuances and exceptions humans use in their language. The statistical revolution in the 1980s and 1990s has changed the relationship between linguistics and computational languages significantly [25]. This revolution was the start of using probabilistic models and Bayesian techniques for Natural Language Processing.

Probabilistic models and Bayesian techniques are part of Machine Learning. These techniques allows the computer program to reason about the natural language in soft, probabilistic decisions instead of hard values. Hard values allow for a yes or a no. Soft decisions allow for a multitude if different responses with each having their own relative certainty. The disadvantages and advantages are well explained in [4]. The biggest advantage is that machine learning based natural language processing techniques will get more accurate by with more input data as this allows developers to improve the system with a relatively low effort. This technique also works better than hand-crafted rules on unfamiliar input (words that have not yet been seen) and erroneous input (words that are misspelled). There are many different techniques that have been used for natural language processing and the field is still improving at a rapid pace.

3.6.4 Markov Decision Process

In 1960, Ronald A. Howard published "Dynamic Programming and Markov Processes" [26]. It mathematically describes a way for computers to make choices and is used in many fields. One of the properties of a Markov Decision Process (MDP) is that it can be solved by both linear and dynamic programming. For conversational agents it is often used to generate responses in a natural language generator where it can form appropriate responses by deciding on the next word in a sentence [27]. Versions of MDSs such as Fuzzy Markov decision processes (FMDPs), Reinforcement learning and partially observable Markov decision process (POMDP) are also commonly used in conversational agents [21].

4 Systematic Part – Comparing the mechanisms used in state-of-the-art conversational agents

Recent CA-based BCSS applications have different implementations for certain submodules, while their general architecture is similar. These BCSSs are compared to analyze their similarities and differences. The implementations are listed in Table 1 to give a general overview of the different design choices. First, what stands out is described, then specific features are discussed in detail. Recent models without concrete implementations offering various improvements for BCSSs are listed and compared in Table 2.

4.1 Existing Implementations

Document Description			Input			Output					Technical Details				
Author	Reference	Behaviour Change Technique	Text	Speech	Face Expression	Text	Speech	Animated	Embodied	Affective Feedback	Platform	Programming Language	Stateful	Generated / Retrieval	Artificial Intelligence
Al-Taee, 2013	[28]	IMS	❖	✓		✓	✓		✓		R	PHP	✓	R	
Alencar, 2014	[29]	CD	✓				✓	✓	✓	✓	W	PHP	✓	R	
Amini, 2013	[20]	MI	✓		✓	✓	✓	✓	✓	✓	PC		✓	G	
Beun, 2016	[17]	PS	✓			✓				✓	M	ECML	✓	R	
Beun, 2017	[5]	CD	✓			✓				✓	M	ECML	✓	R	
De Carolis, 2017	[19]	CD		✓	✓		✓		✓	✓	R		✓	R	RFC
Fonfara, 2014	[18]			✓			✓		✓		R		✓	G	BN
Hartanto, 2016	[30]	E	❖	✓			✓	✓	✓	✓	VR	C#	✓	R	
Kamphorst, 2014	[31]	COMBI	✓			✓					W, M		✓	G	
Kanaoka, 2015	[32]	MI		✓			✓		✓		R		✓	G	
Kavakli, 2012	[33]	CT		✓			✓	✓	✓		PC		✓	G	
Klaassen, 2013	[34]	MBF	✓			✓					*	BML	✓	R	
Li, 2016	[35]	ES		✓	✓		✓	✓	✓	✓	R	C	✓	R	
Lisetti, 2012	[36]	MI		✓	✓		✓	✓	✓	✓	PC	.NET	✓	G	
Lisetti, 2013	[37]	BMI			✓		✓	✓	✓	✓	PC		✓	G	
Ranjbartabar, 2016	[38]	EFT		✓			✓	✓	✓	✓	PC	C#	✓	R	
Yasavur, 2013	[23]	MI		✓	✓		✓	✓	✓		PC		✓	G	RL
Yasavur, 2014	[39]	BI		✓	✓		✓	✓	✓	✓	PC		✓	G	RL

Table 1: Characteristics of 18 recent CA-based BCSS's

Behaviour Change Technique

IMS Information-**M**otivation-**S**trategy
CD Collaborative **D**ialog
MI Motivational **I**nterviewing
E Experience Based
COMBI Combination of different models
CT Coping **T**heory
MBF Monitored **B**ased **F**eedback
ES Empathizing-**S**ystemizing
BMF Brief **M**otivational **I**nterventions
BI Brief **I**nterventions

Platform

R Robot
W Web
PC Personal **C**omputer
M Mobile Phone Application
VR Virtual **R**eality
***** Cross-platform

Artificial Intelligence

RFC Random **F**orest **C**lassifier
BN Bayesian **N**etworks
RL Reinforcement **L**earning

Input

❖ Utilizes text input, but not for the dialogue.

Columns Clarification

Animated The CA shows it's own emotional state when interacting with the patient.
Embodied The CA has a digitally rendered avatar or real world robot that interacts with the patient.
Affective Feedback The CA will react and take into account the emotional states of the patient to optimize the behaviour therapy.
Stateful Agent maintains information about history used for future answer generation.
Generated Answers are selected from a set of predefined answer sentences.
Retrieved A decision process is used to compose the answer sentences.

4.1.1 General Observations

While analyzing the table, several interesting properties are noteworthy. Three authors are re-occurring, they improved upon an earlier developed system but do not state this explicitly. [17] first focuses on the implementations of persuasive strategies for insomnia in a mobile e-coaching system. Later in [5] the earlier approach was generalized. A user interface paradigm is developed for automated e-coaching that can be used for different coaching processes. [36] developed an embodied avatar and a simple BCSS for motivational interviewing and used this in [37] where a extensive CA-based BCSSs was developed. The system developed in [23] was text-only and mainly focused on implementing a suitable reinforcement learning model to train the agent. [40] improved upon the earlier approach by including a speech-recognizer and spoken system output.

There is a great variety in used behaviour change technique. Most systems are embodied CAs and most systems support affective feedback.

4.1.2 Input Recognition

The way the system gets information from the user, whether it is direct dialogue or secondary information varies per system. From the systems reviewed, it is clear that there are two main approaches, text and speech. There is slight preference for speech input over text input, which can be explained by the type of system. If a system outputs the dialogue as speech, and takes speech as input it mimics a conversation one might have with a therapist [37]. Another factor is that embodied robots all have speech recognition. Systems that do use text input for the dialogue do not focus on mimicking a conversation with a therapist but rather short text exchanges from an e-coach [31, 34]. There are two systems that take both text and speech input, [30] and [28]. It should be noted that the text input they take is not a part of the dialogue, but offers the subject the opportunity to enter data before or after a conversation. A system with a unique approach is described in [37]. The subject is presented with multiple choice options and then the system evaluates the facial expression and the answer to decide on the next step in the treatment. Looking at the combination of text-only input and output results in 4 systems, so a minority of the systems is completely text-based.

4.1.3 Output Generation

There are different ways a CA can return an output to the user. From the table, it becomes clear that embodied agents that respond to users with speech are by far the most popular. Text and speech are almost mutually exclusive. This leads to a categorization with 3 main categories: text-based, embodied, and embodied with affective feedback. All of these require varying ways to generate this output which will be explained in this section.

Text generation

From the selected systems that give feedback to the user via text only 2 out of 9 use generative models. The strong preference for retrieval-based systems can be explained by looking at the advantages of such systems. Retrieval-based systems cannot make mistakes in constructing sentences as all sentences are predefined. Also, domain is important to consider here. Since all systems deal with behaviour change and a specification on top of that, the domain in which the CAs operate is closed and predefined. This makes a retrieval-based system viable. For a generative system to operate in a closed domain, it requires access to a domain-specific knowledge-base [41]. One of the generative-based systems [20] has a dialog module that consists of 3 main components: An utterance planning module that forms sentences that lead to a planned goal (using various sub goals), psychometric instruments, which can be described as a collection of questions and interaction that can fulfill sub-goals, and a score evaluator used for both the motivational interviewing part and checking for the fulfillment of the goals set by the utterance planning module. A more common technique for communicating with the subject is speech. This can be explained because most systems are generating the dialogues text-based and then use Text to Speech (TTS) software to output them as speech [29].

Embodied

While text input/output is enough to be considered a CA, this is not the most effective way to approach behaviour change. In Lisetti [16] key features for designing CAs in a therapy setting are identified. Most of these features cannot be conveyed through text alone. Embodied agents can convey these features to a varying degree. The importance of this is also reflected in Table 1 as all but four systems use a form of embodied output. Many systems take this extra step and create an animated avatar to portray the conversational agent. This underlines the main key feature of a BCSS mentioned by [16]: a human face as interface. Animating the avatar is commonly done by adding animation cues to the speech generator to illustrate the sentiment of a certain text [33]. The quality of this can be improved by synchronizing the lip movements of the avatar to the text, as described in [20]. The platform seems to be an important factor in deciding whether an animation is a suitable choice. Only CAs that are accessed via a PC (or a web interface) have an animated avatar.

Systems that do not incorporate an animated avatar but have an embodied implementation are robots. These are very effective in situations where the need for a more empathetic approach is high [19]. It is worth noting that all the robot-based systems used off-the-shelf robots. A way for animation-based and robot-based systems to improve the empathetic behaviour is to use affective feedback. A small majority of the systems discussed use it, but have very different approaches to it. The workings can be generalized as the system identifying the emotional state of the subject and/or the dialogue and then communicating signals of an appropriate emotional response back to the subject. Especially relevant is [35] as recognizing and responding to the emotional state of the subject is a key factor in the behaviour of the robot. A very detailed example of an implementation is given in [20]. Especially noteworthy is [30], a virtual reality system. This goes further than just an animated avatar as it emerges the subject in a virtual situation where also the surroundings are adjusted. A method of affective feedback here is the Anxiety feedback loop. This feedback loop works by frequently checking the current anxiety levels of the patient and changing the virtual reality environment to control the anxiety levels of the patient. This study had many technical difficulties but the immersion offered by this system, and the compatibility with earlier developed methodologies for BCSS make this a very promising approach.

4.1.4 Technical Details

Although some applications are developed for robots or VR, most are developed to use on common customer hardware like PCs and smartphones to make the support system easily accessible for the test group. This is presumably because, for example, robots and VR technologies are not yet fully developed and ready for mainstream use.

Only half of the papers provide details about the used programming language(s). Among the ones providing this information different programming languages are used. The vast majority uses a high level programming language like C# and PHP. Some approaches use mostly custom written code like [37], while others mainly make use of existing frameworks [29].

All reviewed implementations are stateful, making the CA take into account the dialogue history. Earlier approaches such as [7] did not maintain state information.

Out of the 18 papers, 10 use a generated answer system, the other 8 use a retrieval-based system. There seems to be a slight preference towards generated systems, probably because these systems contribute to a more natural dialogue. Retrieval systems are still popular because they are generally easy to implement and do not need intensive training and parameter tweaking.

Only 4 approaches use AI. As AI usage has gained popularity in various systems the last years, it is remarkable that only a few systems use AI.

4.2 Novel Approaches

Several approaches try to improve CA-based BCSSs. These approaches are mostly models without a concrete implementation or systems not suitable to be placed in Table 1 because of the lack of information. The different approaches are listed in Table 2 with a summary, important technical details, and the papers main contribution. This table is analyzed with regard to the aforementioned criteria.

4.2.1 General Observations

The reviewed papers show various approaches to improve CA-based BCSSs, i.e. they focus on different aspects of the system or they create a model for a specific problem. Some models focus on optimizing the dialog by improving the CAs in multiple ways: increase the credibility, ability to explain, ability to detect and express emotions. Other approaches aim to improve aspects such as increasing the employability of CAs, accurately tracking user progress, determining when to intervene in a learning process. Approaches like [22] generalize the scope by including multi-topic support, while approaches like [42] specify the scope by making custom models around a specific disorder. The majority does not do either and the improved part does not affect the CAs scope.

4.2.2 Technical Details

Some approaches are discrete and deterministic while others are interconnected and probabilistic, generally including more mathematical modelling. Examples of the first type are [43], where a strict predefined grammar is used and [44], which uses predicate logic to reason. Examples of the second type are [45] and [42] both using detailed probabilistic mathematical models.

The models of [22] and [39] can both be visualized as a graphs. [22] models the user and system states as nodes and their interaction as transitions. [39] models nodes as questions and answers as transitions between questions.
Only 3 out of the 11 approaches state the usage of AI to improve the CA.

Author	Ref.	Summary	Technical Details	Main Contribution
Fitrianie, 2015	[11]	Proposal of E-Coach Markup Language (ECML). Using a standard XML specification to define the interaction model between the coach and the patient on a fine-grained level.	Modular XML structure. Low computation overhead so it works well on mobile devices.	Creating a standard for dialogue scripting
Goldberg, 2015	[46]	Military study to model the perception, behaviours and judgment of expert human coaches and to support practical and effective learning that is guided by computer-based agents.	Adaptive training on big data with artificial intelligence and accessible on mobile devices.	Combination of state-of-the-art techniques in a single package for maximum effectiveness
Harbers, 2012	[44]	Studies two theory of mind models to model CAs with a theory of mind. A theory of mind helps the CA to ascribe concepts like knowledge and intentions to others.	Uses 2APL, a language comparable to Prolog, to implement both models.	Improve the agent's believability and ability to explain
Hunter, 2015	[45]	This approach models an asymmetric dialogue, meaning that only the system presents arguments. The system maintains a model of the user, including the users possible arguments believed by the user, to determine which new arguments to present.	Uses custom mathematical model for system, including models for asymmetric dialogues and probabilistic user models.	Improve user model, determine optimal dialogue decisions by doing so
Medina-Medina, 2016	[47]	This model tracks user progress in, possibly related, tasks. Progressions between user states forms the knowledge basis for the CA.	A user state is modeled and has user and a goal objective with certain partial objectives. The progression of a certain partial objective is modeled using a progression line with a progression state based on different state transitions.	Track user progress
Menendez, 2014	[48]	Model of CA with an emotional state, this state changes based on input data. Using a specific text generator the emotional state is expressed in the meaning of the output sentences.	Performs deep linguistic analysis on the input sentences. Uses a combination of Granular Linguistic Model of Phenomena and Fuzzy Finite State Machine models to simulate the agents emotion.	Detect emotions in input, express emotion in output
Mitchell, 2013	[49]	Describes a model to precisely determine the moment when to intervene in a learning process. Students program a small application and are interrupted at certain moments by a textual CA. Student actions of adjusting the code or asking the CA are evaluated in the decision process.	Uses an MDP framework to learn an intervention policy, where current student action, task trajectory and last action are the categories of states.	Determine intervention moment in a learning process
Ophir, 2015	[43]	Model aims to teach a robot how to control people by recognizing their thoughts. Uses cognitive behavioral therapy theory to detect disordered thoughts in natural language.	A Bacchus-Naur Form denoted grammar is defined to specifically generate so-called extreme terms to detect disordered thoughts.	Detecting and classifying disordered thoughts
Procee, 2014	[42]	Presents an agent based model of the procrastination disorder. The model is specifically designed around the various aspects of procrastination, and their positive or negative influence on tasks, personal- and other factors.	Uses a greedy like algorithm to compute the procrastination behaviour of an individual. Contains utility- and priority functions to specifically model procrastination behaviour. Model is implemented in MatLab.	Understanding user behaviour
Ren, 2015	[22]	Modelling both the user and the system as a finite state machine (FSM) results in a two finite state machine design (TFSM). The model aims to improve affective interaction and supports multi-topic conversations.	The Dialog Manager can be visualized as a TFSM, nodes are states in which the system or user makes a decision, edges are actions between states. There is a distinction between inner- and inter-FSM transitions to track the TFSM state.	Improves affective interaction
Yasavur, 2014	[39]	Model to easily input CA-based health assessments. By specifying questions and their relations a spoken CA can take assessments.	Dialog manager can be visualized as a graph, nodes are the questions, edges are the answer keys between questions. Uses a POMDP (reinforcement learning) to make decisions.	Increases employability of CAs

Table 2: Overview of recent approaches to improve CA-based BCSS's

5 Discussion

In general, the improvements of the different implementations and models are fragmented. Most approaches enhance only a single aspect leaving the other parts intact. Approaches combining or using previous work, like [46], are rare, resulting in a variety of single-point improvement approaches. It is remarkable to see that research in this field generally does not attempt to take into account earlier approaches.
A standard for writing textual CA dialogues, AIML, exists, but a standard for affective feedback is missing and is desirable to have. The lack of this standard results in researchers solving the same problem anew instead of investing time in new research. A general format to use different therapies on the same system would be another desirable improvement, [37] makes an attempt to accomplish this.
The possibilities of smartphones, smartwatches and activity trackers are not sufficiently utilized. A CA could use the information from the different sensors available to learn and plan, and these devices can be used to intervene the user when needed.
As the usage of AI techniques gained popularity in recent computer systems, the implemented approaches lack AI usage in general. More AI techniques to train the dialog manager and detect emotions in text and face, would make the CA more intelligent and believable.
As this review studies papers from different fields including, but not limited to: Computer Science, Psychology and Healthcare, certain aspects of the research may be heavily dependent on the background of the researchers. This results in different choices for the systems design. The expertise in certain aspects of the system differs per field.

6 Limitations of this study

Since we had a restricted amount of time and people available for this research, we limited ourselves to only one source for our literature selection. While we could have included sources other than Elsevier's Scopus, we felt that would mostly complicate matters and we already had a good amount of quality results. Furthermore we focused on the technical details of chat bots for behaviour change while literature on that subject focuses on the application, rather than the technical details. Combined with the fact that a conversational agent for behaviour change does not differ that much from a normal conversational agent on an technical level made it difficult to find detailed information on certain systems for our systematic part. Any comparison of systems that is not exhaustive of course has to be taken as is with it limitations. Another limitation of this study is our self-imposed filter criteria of excluding papers from before 2011, as we felt that a (broad) 5 years is reasonable for the speed at which the field of conversational agents moves to label it as modern.

7 Conclusion

In general, standards are missing for the encoding of affective feedback and different therapies. Most approaches are single-point improvements for the system, without taking into account earlier approaches. While single-point improvements are welcome and well-designed, the true strength lies in combining these approaches. For example [46] is a promising paper that combines adaptive training, big data, point of need training and Artificial Intelligence in one single package. This is a paper from the industry and it shows very well what can be accomplished if we combine multiple research papers. Another very promising paper is [19], this paper uses various recent technological advancements for a very positive end result. The paper uses a physical robot, affective feedback based on modeling the state of mind of the user, artificial intelligence for recognizing emotion in speech and facial expressions, and natural language processing. To conclude, BCCSs have shown great improvement over the past few years, but the true strength will lie in introducing industry standards and combining technological advancements to efficiently develop effective systems in the future.

8 Acknowledgements

We would like to express our gratitude to Franziska Burger for her continuous support as our supervisor. The advice and insights she provided was of immense value during the completion of this literary study. Furthermore, the time she spend on reviewing and providing feedback on our performance is greatly appreciated.

References

[1] S. Poria, E. Cambria, R. Bajpai, and A. Hussain. A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion*, 37:98–125, 2017.

[2] Ken Thompson. Programming techniques: Regular expression search algorithm. *Commun. ACM*, 11(6):419–422, June 1968.

[3] A.V. Aho and M.J. Corasick. Efficient string matching: An aid to bibliographic search. *Communications of the ACM*, 18(6):333–340, 1975. cited By 1640.

[4] K. Kukich. Techniques for automatically correcting words in text. *ACM Computing Surveys (CSUR)*, 24(4):377–439, 1992. cited By 558.

[5] Robbert Jan Beun, Siska Fitrianie, Fiemke Griffioen-Both, Sandor Spruit, Corine Horsch, Jaap Lancee, and Willem-Paul Brinkman. Talk and tools: the best of both worlds in mobile user interfaces for e-coaching. *Personal and Ubiquitous Computing*, 21(4):661–674, 2017.

[6] Joseph Weizenbaum. Eliza—a computer program for the study of natural language communication between man and machine. *Commun. ACM*, 9(1):36–45, January 1966.

[7] J Epstein and WD Klinkenberg. From eliza to internet: A brief history of computerized assessment. *Computers in Human Behavior*, 17(3):295–314, 2001.

[8] Inc. A.L.I.C.E. Artificial Intelligence Foundation. Alicebot, 2007. [Online; accessed 28-November-2017].

[9] H. G. Loebner. Home page of the loebner prize, 2015. [Online; accessed 28-November-2017].

[10] Richard Wallace. The elements of aiml style. *Alice AI Foundation*, 2003.

[11] S. Fitrianie, F. Griffioen-Both, S. Spruit, J. Lancee, and R. J. Beun. Automated dialogue generation for behavior intervention on mobile devices. In E. M. Shakshuki, editor, *6th International Conference on Emerging Ubiquitous Systems and Pervasive Networks, EUSPN 2015*, volume 63, pages 236–243. Elsevier, 2015.

[12] D. Griol and J.M. Molina. Context-aware conversational agents using pomdps and agenda-based simulation. *Advances in Intelligent and Soft Computing*, 94:29–36, 2011. cited By 0.

[13] Y. Sun, Y. Zhang, Y. Chen, and R. Jin. Conversational recommendation system with unsupervised learning. In *RecSys 2016 - Proceedings of the 10th ACM Conference on Recommender Systems*, pages 397–398, 2016. cited By 0.

[14] Kenneth Mark Colby, Sylvia Weber, and Franklin Dennis Hilf. Artificial paranoia. *Artificial Intelligence*, 2(1):1 – 25, 1971.

[15] Pinar Saygin Ayse, Cicekli Ilyas, and Akman Varol. Turing test: 50 years later. *Minds and Machines*, 10(4):463–518, 2000.

[16] C. L. Lisetti. Embodied conversational agents for psychotherapy. In *Proceedings of the Conference Workshop on Technology in Mental Health (CHI'08)*, volume 1, 2008.

[17] R. J. Beun, W. P. Brinkman, S. Fitrianie, F. Griffioen-Both, C. Horsch, J. Lancee, and S. Spruit. Improving adherence in automated e-coaching: A case from insomnia therapy. In M. Tscheligi, B. De Ruyter, A. Meschtscherjakov, V. Fuchsberger, M. Murer, and M. Tscheligi, editors, *11th International Conference on Persuasive Technology, PERSUASIVE 2016*, volume 9638, pages 276–287. Springer Verlag, 2016.

[18] J. Fonfara, S. Hellbach, and H. J. Böhme. Imitating dialog strategies under uncertainty. In A. Chellali, M. Mallem, and P. Horain, editors, *The 6th international conference on Intelligent Human Computer Interaction, IHCI 2014*, volume 39, pages 131–138. Elsevier, 2014.

[19] B. De Carolis, S. Ferilli, and G. Palestra. Simulating empathic behavior in a social assistive robot. *Multimedia Tools and Applications*, 76(4):5073–5094, 2017.

[20] R. Amini, C. Lisetti, U. Yasavur, and N. Rishe. On-demand virtual health counselor for delivering behavior-change health interventions. In *2013 1st IEEE International Conference on Healthcare Informatics, ICHI 2013*, pages 46–55, 2013.

[21] S. Young, M. Gašić, B. Thomson, and J.D. Williams. Pomdp-based statistical spoken dialog systems: A review. *Proceedings of the IEEE*, 101(5):1160–1179, 2013. cited By 100.

[22] F. Ren, Y. Wang, and C. Quan. Tfsm-based dialogue management model framework for affective dialogue systems. *IEEJ Transactions on Electrical and Electronic Engineering*, 10(4):404–410, 2015.

[23] U. Yasavur, C. Lisetti, and N. Rishe. Modeling brief alcohol intervention dialogue with mdps for delivery by ecas. In *13th International Conference on Intelligent Virtual Agents, IVA 2013*, volume 8108 LNAI, pages 92–105. Springer, Berlin, Heidelberg, Edinburgh, 2013.

[24] S.R. Petrick. On natural language based computer systems. *IBM Journal of Research and Development*, 20(4):314–325, 1976. cited By 33.

[25] Mark Johnson. How the statistical revolution changes (computational) linguistics. *Proceedings of the EACL 2009 Workshop on the Interaction between Linguistics and Computational Linguistics*, pages 3–11, 2009.

[26] R.A. Howard. Dynamic processes and markov processes. *MIT Press, Cambridge, MA*, 1960.

[27] T. Dodson, N. Mattei, and J. Goldsmith. A natural language argumentation interface for explanation generation in markov decision processes. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6992 LNAI:42–55, 2011. cited By 4.

[28] M. A. Al-Taee, A. H. Sungoor, S. N. Abood, and N. Y. Philip. Web-of-things inspired e-health platform for integrated diabetes care management. In *2013 2nd IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies, AEECT 2013*. IEEE Computer Society, 2013.

[29] M. Alencar and J. F. Netto. Tutor collaborator using multi-agent system. In T. Yuizono, H. Ogata, T. Yuizono, G. Zurita, N. Baloian, H. Ogata, G. Zurita, N. Baloian, T. Inoue, and T. Inoue, editors, *7th International Conference on Collaboration Technologies and Social Computing, CollabTech 2014*, volume 460, pages 153–159. Springer Verlag, 2014.

[30] D. Hartanto, W. P. Brinkman, I. L. Kampmann, N. Morina, P. G. M. Emmelkamp, and M. A. Neerincx. Home-based virtual reality exposure therapy with virtual health agent support. In D. Giakoumis, G. Lopez, A. Matic, S. Serino, and P. Cipresso, editors, *5th International Conference on Pervasive Computing Paradigms for Mental Health, MindCare 2015*, volume 604, pages 85–98. Springer Verlag, 2016.

[31] B. A. Kamphorst, M. C. A. Klein, and A. Van Wissen. Autonomous e-coaching in the wild: Empirical validation of a model-based reasoning system. In *13th International Conference*

on Autonomous Agents and Multiagent Systems, AAMAS 2014, volume 1, pages 725–732. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2014.

[32] T. Kanaoka and B. Mutlu. Designing a motivational agent for behavior change in physical activity. In *33rd Annual CHI Conference on Human Factors in Computing Systems, CHI EA 2015*, volume 18, pages 1445–1450. Association for Computing Machinery, 2015.

[33] M. Kavakli, M. Li, and T. Rudra. Towards the design of a virtual sociologist on aborigines substance abuse: A coping-theory perspective. In *2012 1st IEEE International Conference on Computational Intelligence and Cybernetics, CyberneticsCom 2012*, pages 50–54, 2012.

[34] R. Klaassen, R. Op Den Akker, and H. Op Den Akker. Feedback presentation for mobile personalised digital physical activity coaching platforms. In *6th ACM International Conference on PErvasive Technologies Related to Assistive Environments, PETRA 2013*, 2013.

[35] C. Li, Q. Jia, and Y. Feng. Human-robot interactoin design for robot-assisted intervention for children with autism based on e-s theory. In *8th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2016*, volume 2, pages 320–324. Institute of Electrical and Electronics Engineers Inc., 2016.

[36] C. Lisetti, U. Yasavur, C. De Leon, R. Amini, N. Rishe, and U. Visser. Building an on-demand avatar-based health intervention for behavior change. In *25th International Florida Artificial Intelligence Research Society Conference, FLAIRS-25*, pages 175–180, 2012.

[37] C. Lisetti, R. Amini, U. Yasavur, and N. Rishe. I can help you change! an empathic virtual agent delivers behavior change health interventions. *ACM Transactions on Management Information Systems*, 4(4), 2013.

[38] H. Ranjbartabar and D. Richards. A virtual emotional freedom therapy practitioner (demonstration). In *15th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2016*, pages 1471–1473. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2016.

[39] U. Yasavur, C. Lisetti, and N. Rishe. Let's talk! speaking virtual counselor offers you a brief intervention. *Journal on Multimodal User Interfaces*, 8(4):381–398, 2014.

[40] U. Yasavur, C. Lisetti, and N. Rishe. Statistical dialog manager design tool for health screening and assessments. In *14th International Conference on Intelligent Virtual Agents, IVA 2014*, volume 8637 LNAI, pages 492–495. Springer Verlag, Boston, MA, 2014.

[41] Z. Babar, A. Lapouchnian, and E. Yu. Chatbot design - reasoning about design options using iq^* and process architecture. In *10th International iQ^* Workshop, iStar 2017 - co-located with the 29th International Conference on Advanced Information Systems Engineering, CAiSE 2017*, volume 1829, pages 73–78, 2017. cited By 0.

[42] R. Procee, B. A. Kamphorst, A. Van Wissen, and J. J. Meyer. An agent-based model of procrastination. In T. Schaub, G. Friedrich, and B. O'Sullivan, editors, *21st European Conference on Artificial Intelligence, ECAI 2014*, volume 263, pages 747–752. IOS Press, 2014.

[43] D. Ophir and J. Zheng. A sociable robot as a psychotherapist. In *10th International Workshop on Robot Motion and Control, RoMoCo 2015*, pages 199–206. Institute of Electrical and Electronics Engineers Inc., 2015.

[44] M. Harbers, K. Van Den Bosch, and J. J. Meyer. Modeling agents with a theory of mind: Theory-theory versus simulation theory. *Web Intelligence and Agent Systems*, 10(3):331–343, 2012.

[45] A. Hunter. Modelling the persuadee in asymmetric argumentation dialogues for persuasion. In M. Wooldridge and Q. Yang, editors, *24th International Joint Conference on Artificial Intelligence, IJCAI 2015*, volume 2015-January, pages 3055–3061. International Joint Conferences on Artificial Intelligence, 2015.

[46] B. Goldberg, R. Sottilare, J. Moss, and A. Sinatra. Dimensions of instructional management research in support of a generalized intelligent framework for tutoring (gift). In F. Longo, R. Sottilare, and A. G. Bruzzone, editors, *5th International Defense and Homeland Security Simulation Workshop, DHSS 2015*, pages 33–41. Dime University of Genoa, 2015.

[47] N. Medina-Medina, Z. Callejas, K. Benghazi, and M. Noguera. User progress modelling in counselling systems: An application to an adaptive virtual coach. In I. Comyn-Wattiau, I. Y. Song, S. Yamamoto, M. Saeki, and K. Tanaka, editors, *35th International Conference on Conceptual Modelling, ER 2016 held in conjunction with Workshops on AHA, MoBiD, MORE-BI, MReBA, QMMQ, SCME and WM2SP, 2016*, volume 9974 LNCS, pages 479–487. Springer Verlag, 2016.

[48] C. Menendez, L. Eciolaza, and G. Trivino. Generating advices with emotional content for promoting efficient consumption of energy. *International Journal of Uncertainty, Fuzziness and Knowlege-Based Systems*, 22(5):677–697, 2014.

[49] C. M. Mitchell, K. E. Boyer, and J. C. Lester. A markov decision process model of tutorial intervention in task-oriented dialogue. In *16th International Conference on Artificial Intelligence in Education, AIED 2013*, volume 7926 LNAI, pages 828–831. Springer-Verlag Berlin Heidelberg, Memphis, TN, 2013.

Appendices

A - Research Method

A.1 PRISMA statement

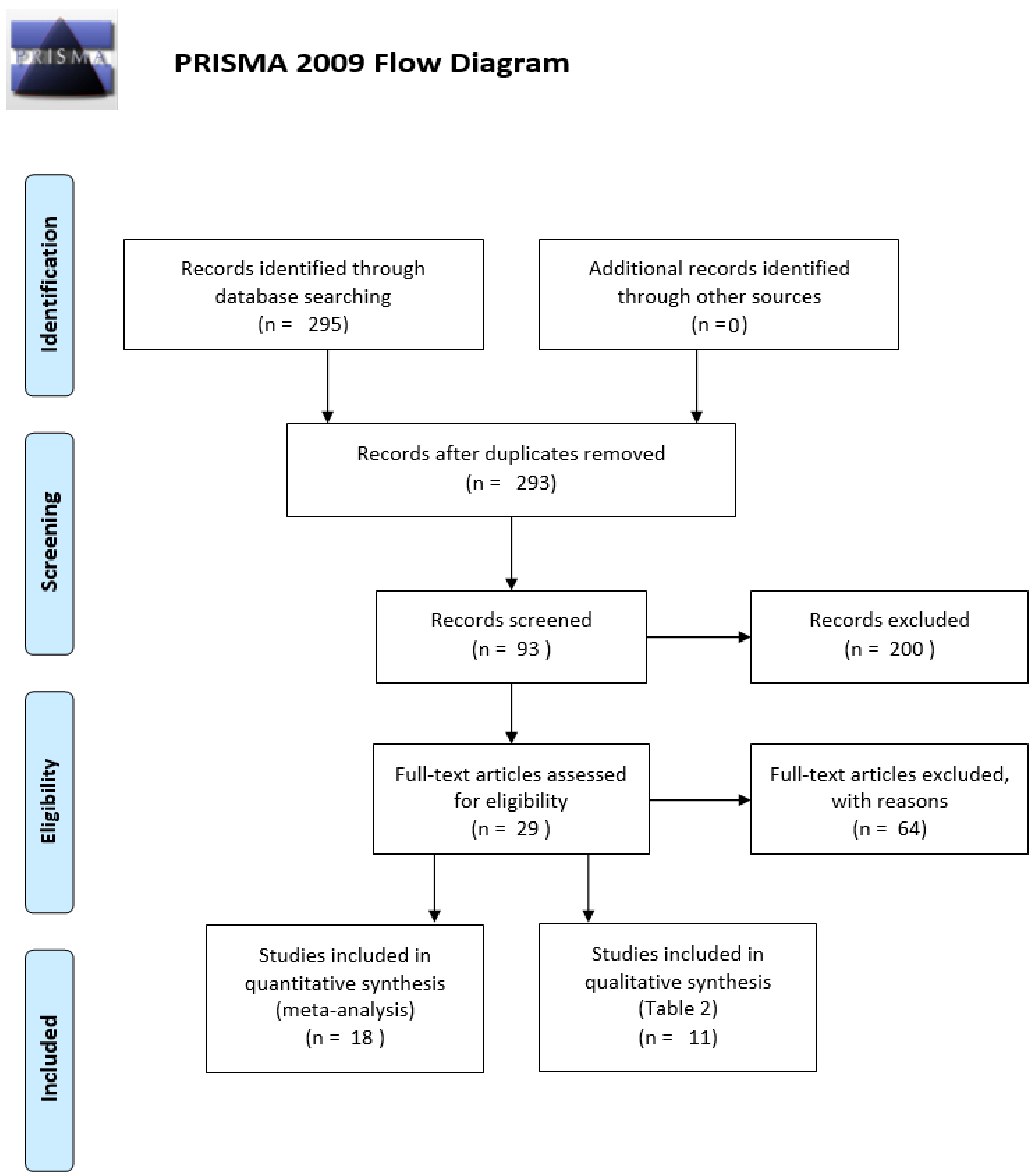

Figure 1: Prisma Statement

A.2 Excluded literature

Author, year	Ref.	Reason
Garcia, 2017	[50]	1
Ghayoumi, 2016	[51]	1
Manuvinakurike, 2013	[52]	1
Manuvinakurike, 2014	[53]	1
Formolo, 2017	[54]	2
Gulenko, 2014	[55]	2
Jones, 2014	[56]	2
Leu, 2014	[57]	2
Litman, 2014	[58]	2
Lorenz, 2016	[59]	2
Ochs, 2012	[60]	2
Rieser, 2014	[61]	2
Thomason, 2013	[62]	2
Tikka, 2017	[63]	2
Valstar, 2016	[64]	2
Zhang, 2016	[65]	2
Beun, 2012	[66]	3
Birnbaum, 2016	[67]	3
Boratto, 2017	[68]	3
Celikyilmaz, 2014	[69]	3
Creed, 2012	[70]	3
Ebert, 2016	[71]	3
Barakova, 2015	[72]	3
reczek, 2013	[73]	3
Kamphorst, 2017	[74]	3
Kennedy, 2012	[75]	3
Kreps, 2013	[76]	3
Kulyk, 2014	[77]	3
Hartanto 2015a	[78]	3
Henkel, 2016	[79]	3
Lehto, 2013	[80]	3
LeRouge, 2016	[81]	3

Author, year	Ref.	Reason	Comment
Nooitgedagt, 2017	[82]	3	
Ong, 2016	[83]	3	
Pantic, 2016	[84]	3	
Riccardi, 2014	[85]	3	
Ring, 2015	[86]	3	
Rutjes, 2016	[87]	3	
Scholten, 2016	[88]	3	
Simons, 2012	[89]	3	
Berkel, 2013	[90]	3	
Zwaan, 2012	[91]	3	
Whalon, 2015	[92]	3	
Yoshii, 2014	[93]	3	
Gnjatović, 2014	[94]	4	
Han, 2015	[95]	4	
Joo, 2016	[96]	4	
Shamekhi, 2017	[97]	4	
Wang, 2016	[98]	4	
Pisica, 2015	[99]	1,2	
Spanakis, 2017	[100]	1,2	
Spaulding, 2016	[101]	1,2	
Wils, 2016	[102]	1,2	
Yasavur, 2012	[103]	1,2	
Bevil, 2016b	[104]	1,3	
Burton, 2016	[105]	1,3	
D'Alfonso, 2017	[106]	1,3	
Zuckerman, 2015	[107]	1,3	
Carneiro, 2015	[108]	2,3	
Künzel, 2012	[109]	2,3	Text in Portuguese
Ren, 2016	[110]	2,3	
Ren, 2014	[111]	2,3	
Aranyi, 2016	[112]	1,2,3	
Bevil, 2016a	[113]	1,2,3	

Table 3: Reasons for exclusion of rejected papers

<u>Reason Clasiffication</u>

1 The record does not utilize a CA (that matches our definition).
2 The record does not operate in the domain of human behaviour change.
3 The record contains an insufficient amount of technical detail.
4 The full text of this record could not be obtained

References of excluded literature

[50] O. Garcia, R. S. Alonso, F. Guevara, J. F. De Paz, G. Villarrubia, J. M. Corchado, O. Abrishambaf, and Z. Vale. Use of context-aware social computing to improve energy efficiency in public buildings. In *2016 IEEE Symposium Series on Computational Intelligence, SSCI 2016*. Institute of Electrical and Electronics Engineers Inc., 2017.

[51] M. Ghayoumi, M. Thafa, and A. K. Bansal. Towards formal multimodal analysis of emotions for affective computing. In *22nd International Conference on Distributed Multimedia Systems, DMS 2016*, pages 48–54. Knowledge Systems Institute Graduate School, 2016.

[52] R. Manuvinakurike, B. Barry, and T. Bickmore. Indexing stories for conversational health interventions. In *2013 AAAI Spring Symposium*, volume SS-13-03, pages 27–32, Palo Alto, CA, 2013.

[53] R. Manuvinakurike, W. F. Velicer, and T. W. Bickmore. Automated indexing of internet stories for health behavior change: Weight loss attitude pilot study. *Journal of Medical Internet Research*, 16(12), 2014.

[54] D. Formolo and T. Bosse. A conversational agent that reacts to vocal signals, 2017.

[55] I. Gulenko. Chatbot for it security training: Using motivational interviewing to improve security behaviour. In A. Panchenko, R. Yavorsky, A. Panchenko, N. Konstantinova, D. Ustalov, D. I. Ignatov, and M. Yu Khachay, editors, *3rd International Conference on Analysis of Images, Social Networks and Texts, AIST 2014*, volume 1197, pages 7–16. CEUR-WS, 2014.

[56] H. Jones, M. Chollet, M. Ochs, N. Sabouret, and C. Pelachaud. Expressing social attitudes in virtual agents for social coaching. In *13th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2014*, volume 2, pages 1409–1410. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2014.

[57] F. Y. Leu, J. C. Liu, Y. T. Hsu, and Y. L. Huang. The simulation of an emotional robot implemented with fuzzy logic. *Soft Computing*, 18(9):1729–1743, 2014.

[58] D. Litman and K. Forbes-Riley. Evaluating a spoken dialogue system that detects and adapts to user affective states. In *15th Annual Meeting of the Special Interest Group on Discourse and Dialogue, SIGDIAL 2014*, pages 181–185. Association for Computational Linguistics (ACL), 2014.

[59] T. Lorenz, A. Weiss, and S. Hirche. Synchrony and reciprocity: Key mechanisms for social companion robots in therapy and care. *International Journal of Social Robotics*, 8(1):125–143, 2016.

[60] M. Ochs, D. Sadek, and C. Pelachaud. A formal model of emotions for an empathic rational dialog agent. *Autonomous Agents and Multi-Agent Systems*, 24(3):410–440, 2012.

[61] V. Rieser, O. Lemon, and S. Keizer. Natural language generation as incremental planning under uncertainty: Adaptive information presentation for statistical dialogue systems. *IEEE Transactions on Audio, Speech and Language Processing*, 22(5):979–994, 2014.

[62] J. Thomason and D. Litman. Differences in user responses to awizard-of-oz versus automated system. In *2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL HLT 2013*, pages 796–801. Association for Computational Linguistics (ACL), 2013.

[63] P. Tikka and H. Oinas-Kukkonen. Contributing or receiving-the role of social interaction styles in persuasion over a social networking platform. *Personal and Ubiquitous Computing*, 21(4):705–721, 2017.

[64] M. Valstar, T. Baur, A. Cafaro, A. Ghitulescu, B. Potard, and J. Wagner. Ask alice: An artificial retrieval of information agent. In C. Pelachaud, Y. I. Nakano, T. Nishida, C. Busso, L. P. Morency, and E. Andre, editors, *18th ACM International Conference on Multimodal Interaction, ICMI 2016*, pages 419–420. Association for Computing Machinery, Inc, 2016.

[65] J. Zhang, N. M. Thalmann, and J. Zheng. Combining memory and emotion with dialog on social companion: A review. In *29th International Conference on Computer Animation and Social Agents, CASA 2016*, volume 23-25-May-2016, pages 1–9. Association for Computing Machinery, 2016.

[66] Robbert Jan Beun. Persuasive strategies in mobile insomnia therapy: alignment, adaptation, and motivational support. *Personal and Ubiquitous Computing*, 17(6):1187–1195, 2012.

[67] G. E. Birnbaum, M. Mizrahi, G. Hoffman, H. T. Reis, E. J. Finkel, and O. Sass. Machines as a source of consolation: Robot responsiveness increases human approach behavior and desire for companionship. In *11th Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI 2016*, volume 2016-April, pages 165–171. IEEE Computer Society, 2016.

[68] L. Boratto, S. Carta, F. Mulas, and P. Pilloni. An e-coaching ecosystem: design and effectiveness analysis of the engagement of remote coaching on athletes. *Personal and Ubiquitous Computing*, 21(4):689–704, 2017.

[69] A. Celikyilmaz, Z. Feizollahi, D. Hakkani-Tur, and R. Sarikaya. Resolving referring expressions in conversational dialogs for natural user interfaces. In *2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014*, pages 2094–2104. Association for Computational Linguistics (ACL), 2014.

[70] C. Creed and R. Beale. User interactions with an affective nutritional coach. *Interacting with Computers*, 24(5):339–350, 2012.

[71] D. D. Ebert, D. Lehr, E. Heber, H. Riper, P. Cuijpers, and M. Berking. Internet- and mobile-based stress management for employees with adherence-focused guidance: Efficacy and mechanism of change. *Scandinavian Journal of Work, Environment and Health*, 42(5):382–394, 2016.

[72] Emilia I. Barakova, Prina Bajracharya, Marije Willemsen, Tino Lourens, and Bibi Huskens. Long-term lego therapy with humanoid robot for children with asd. *Expert Systems*, 32(6):698–709, 2015.

[73] J. Greczek, A. Atrash, and M. Matarić. A computational model of graded cueing: Robots encouraging behavior change, 2013.

[74] B. A. Kamphorst. E-coaching systems: What they are, and what they aren't. *Personal and Ubiquitous Computing*, 21(4):625–632, 2017.

[75] C. M. Kennedy, J. Powell, T. H. Payne, J. Ainsworth, A. Boyd, and I. Buchan. Active assistance technology for health-related behavior change: An interdisciplinary review. *Journal of Medical Internet Research*, 14(3), 2012.

[76] G. L. Kreps and L. Neuhauser. Artificial intelligence and immediacy: Designing health communication to personally engage consumers and providers. *Patient Education and Counseling*, 92(2):205–210, 2013.

[77] O. Kulyk, R. Op den Akkery, R. Klaasseny, and L. van Gemert-Pijnen. Personalized virtual coaching for lifestyle support: Principles for design and evaluation. *International Journal on Advances in Life Sciences*, 6(3-4):300–309, 2014.

[78] D. Hartanto, W. P. Brinkman, I. L. Kampmann, N. Morina, P. G. Emmelkamp, and M. A. Neerincx. Design and implementation of home-based virtual reality exposure therapy system with a virtual ecoach, 2015.

[79] Z. Henkel and C. L. Bethel. Increasing psychological well-being through human-robot interaction. In *11th Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI 2016*, volume 2016-April, pages 617–618. IEEE Computer Society, 2016.

[80] T. Lehto, H. Oinas-Kukkonen, T. Pätiälä, and O. Saarelma. Virtual health coaching for consumers: A persuasive systems design perspective. *International Journal of Networking and Virtual Organisations*, 13(1):24–41, 2013.

[81] C. LeRouge, K. Dickhut, C. Lisetti, S. Sangameswaran, and T. Malasanos. Engaging adolescents in a computer-based weight management program: Avatars and virtual coaches could help. *Journal of the American Medical Informatics Association*, 23(1):19–28, 2016.

[82] A. Nooitgedagt, R. J. Beun, and F. Dignum. e-coaching for intensive cardiac rehabilitation a requirement analysis, 2017.

[83] E. C. J. Ong and Z. C. Soriano. A conversational agent to shift students‘ affect state, 2016.

[84] M. Pantic, V. Evers, M. Deisenroth, L. Merino, and B. Schuller. Social and affective robotics tutorial. In *24th ACM Multimedia Conference, MM 2016*, pages 1477–1478. Association for Computing Machinery, Inc, 2016.

[85] G. Riccardi. Towards healthcare personal agents. In *2014 ACM Roadmapping the Future of Multimodal Interaction Research Including Business Opportunities and Challenges, RFMIR 2014*, pages 53–56. Association for Computing Machinery, Inc, 2014.

[86] L. Ring, L. Shi, K. Totzke, and T. Bickmore. Social support agents for older adults: longitudinal affective computing in the home. *Journal on Multimodal User Interfaces*, 9(1):79–88, 2015.

[87] H. Rutjes, M. C. Willemsen, and W. A. Ijsselsteijn. Understanding effective coaching on healthy lifestyle by combining theory- and data-driven approaches. In R. Orji, M. Busch, A. Dijkstra, M. Tscheligi, M. Tscheligi, M. Reisinger, and A. Stibe, editors, *International Workshop on Personalization in Persuasive Technology, PPT 2016*, volume 1582, pages 26–29. CEUR-WS, 2016.

[88] M. Scholten, S. Kelders, and L. Van Gemert-Pijnen. How persuasive is a virtual coach in terms of promoting web-based intervention user engagement? exploring a potential match between user support needs and a virtual coach's motivational capabilities. In L. Siemons, O. Kulyk, H. Oinas-Kukkonen, and L. van Gemert-Pijnen, editors, *4th International Workshop on Behavior Change Support Systems, BCSS 2016*, volume 1573, pages 66–73. CEUR-WS, 2016.

[89] L. P. A. Simons and J. F. Hampe. Mobile application supported lifestyle intervention to lower prostate cancer risk. In *HEALTHINF 2012 - Proceedings of the International Conference on Health Informatics*, pages 443–448, Vilamoura, Algarve, 2012.

[90] J. Van Berkel, C. R. L. Boot, K. I. Proper, P. M. Bongers, and A. J. Van Der Beek. Process evaluation of a workplace health promotion intervention aimed at improving work engagement and energy balance. *Journal of Occupational and Environmental Medicine*, 55(1):19–26, 2013.

[91] J. M. van der Zwaan, V. Dignum, and C. M. Jonker. Corpus-based validation of a dialogue model for social support. In *24th Benelux Conference on Artificial Intelligence, BNAIC 2012*, Maastricht, 2012.

[92] K. Whalon, J. R. Martinez, D. Shannon, C. Butcher, and M. F. Hanline. The impact of reading to engage children with autism in language and learning (recall). *Topics in Early Childhood Special Education*, 35(2):102–115, 2015.

[93] A. Yoshii and T. Nakajima. Design implications of a conversational agent in an implicitly-persuasive application. *International Journal of Software Engineering and its Applications*, 8(10):203–212, 2014.

[94] M. Gnjatović and V. Delić. End-user design of emotion-adaptive dialogue strategies for therapeutic purposes, 2014.

[95] Z. Han, F. Ren, and D. Miao. A synthetic and computational language model for interactive dialogue system. In A. Palma dos Reis, J. Roth, P. Isaias, A. P. Abraham, P. Kommers, H. F. Betancort, and L. Rodrigues, editors, *European Conference on Data Mining 2015, ECDM 2015 and International Conferences on Intelligent Systems and Agents 2015, ISA 2015 and Theory and Practice in Modern Computing 2015, TPMC 2015*, pages 73–80. IADIS, 2015.

[96] Y. K. Joo and R. J. Lee-Won. An agent-based intervention to assist drivers under stereotype threat: Effects of in-vehicle agents' attributional error feedback. *Cyberpsychology, Behavior, and Social Networking*, 19(10):615–620, 2016.

[97] A. Shamekhi, T. Bickmore, A. Lestoquoy, and P. Gardiner. Augmenting group medical visits with conversational agents for stress management behavior change, 2017.

[98] Y. Wang. Design and implementation of a friendly environment for long-lasting stay in space. In *67th International Astronautical Congress, IAC 2016*. International Astronautical Federation, IAF, 2016.

[99] I. Pisica, C. J. Axon, P. R. Hobson, G. A. Taylor, and D. C. H. Wallom. A multi-agent model for assessing electricity tariffs. In *2014 IEEE PES Innovative Smart Grid Technologies Conference Europe, ISGT-Europe 2014*, volume 2015-January. IEEE Computer Society, 2015.

[100] G. Spanakis, G. Weiss, B. Boh, L. Lemmens, and A. Roefs. Machine learning techniques in eating behavior e-coaching: Balancing between generalization and personalization. *Personal and Ubiquitous Computing*, 21(4):645–659, 2017.

[101] S. Spaulding, G. Gordon, and C. Breazeal. Affect-aware student models for robot tutors. In *15th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2016*, pages 864–872. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2016.

[102] P. Wills, P. Baxter, J. Kennedy, E. Senft, and T. Belpaeme. Socially contingent humanoid robot head behaviour results in increased charity donations. In *11th Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI 2016*, volume 2016-April, pages 533–534. IEEE Computer Society, 2016.

[103] U. Yasavur, R. Amini, and C. Lisetti. User modeling for pervasive alcohol intervention systems. In *Joint 1st Int. Workshop on Recommendation Technologies for Lifesytle Change,*

LIFESTYLE 2012 and the 1st Int. Workshop on Interfaces for Recommender Systems, InterfaceRS 2012 - Workshops at the 6th ACM Conf. on Recommender Systems, RECSYS 2012, volume 891, pages 29–34, Dublin, 2012.

[104] R. Bevill, C. H. Park, H. J. Kim, J. W. Lee, A. Rennie, M. Jeon, and A. M. Howard. Interactive robotic framework for multi-sensory therapy for children with autism spectrum disorder. In *11th Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI 2016*, volume 2016-April, pages 421–422. IEEE Computer Society, 2016.

[105] C. Burton, A. Szentagotai Tatar, B. McKinstry, C. Matheson, S. Matu, R. Moldovan, M. Macnab, E. Farrow, D. David, C. Pagliari, A. Serrano Blanco, and M. Wolters. Pilot randomised controlled trial of help4mood, an embodied virtual agent-based system to support treatment of depression. *Journal of Telemedicine and Telecare*, 22(6):348–355, 2016.

[106] S. D'Alfonso, O. Santesteban-Echarri, S. Rice, G. Wadley, R. Lederman, C. Miles, J. Gleeson, and M. Alvarez-Jimenez. Artificial intelligence-assisted online social therapy for youth mental health. *Frontiers in Psychology*, 8(JUN), 2017.

[107] O. Zuckerman and G. Hoffman. Empathy objects: Robotic devices as conversation companions. In *9th International Conference on Tangible, Embedded, and Embodied Interaction, TEI 2015*, pages 593–598. Association for Computing Machinery, Inc, 2015.

[108] J. Carneiro, D. Martinho, G. Marreiros, and P. Novais. A general template to configure multi-criteria problems in ubiquitous gdss. *International Journal of Software Engineering and its Applications*, 9(11):193–206, 2015.

[109] M. F. Künzel, A. K. D. Silva, R. Frozza, D. Bagatini, and B. Lux. Dimi 3d - companion agent in a learning virtual environment. In *5th Workshop and School of Agent Systems, their Environment and Applications, WESAAC 2011*, pages 107–109, Curitiba, Parana, 2012.

[110] F. Ren, Y. Wang, and C. Quan. A novel factored pomdp model for affective dialogue management. *Journal of Intelligent and Fuzzy Systems*, 31(1):127–136, 2016.

[111] J. Ren, B. Jack, D. Schulman, and T. Bickmore. Supporting longitudinal change in many health behaviors. In *32nd Annual ACM Conference on Human Factors in Computing Systems, CHI EA 2014*, pages 1657–1662, Toronto, ON, 2014. Association for Computing Machinery.

[112] G. Aranyi, F. Pecune, F. Charles, C. Pelachaud, and M. Cavazza. Affective interaction with a virtual character through an fnirs brain-computer interface. *Frontiers in Computational Neuroscience*, 10(JULY), 2016.

[113] R. Bevill, P. Azzi, M. Spadafora, C. H. Park, H. J. Kim, J. W. Lee, K. Raihan, M. Jeon, and A. M. Howard. Multisensory robotic therapy to promote natural emotional interaction

for children with asd. In *11th Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI 2016*, volume 2016-April, page 571. IEEE Computer Society, 2016.