

Please quote as: Winkler, R. & Söllner, M. (2018): Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis. In: Academy of Management Annual Meeting (AOM). Chicago, USA.

Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis

ABSTRACT

Chatbots are becoming a ubiquitous trend in many fields such as medicine, product and service industry, and education. Chatbots are computer programs used to conduct auditory or textual conversations. A growing body of evidence suggests that these programs have the potential to change the way students learn and search for information. Especially in large-scale learning scenarios with more than 100 students per lecturer, chatbots are able to solve the problem of individual student support. However, until now, there has been no systematic, structured overview of their use in education. The aim of this paper is therefore to conduct a systematic literature review based on a multi-perspective framework, from which we have derived initial search questions, synthesized past research, and highlighted future research directions. We reviewed titles and abstracts of 1405 articles drawn from management, education, information systems, and psychology literature before examining and individually coding a relevant subset of 80 articles. The results show that chatbots are in the very beginning of entering education. Few studies suggest the potential of chatbots for improving learning processes and outcomes. Nevertheless, past research has revealed that the effectiveness of chatbots in education is complex and depends on a variety of factors. With our literature review, we make two principal contributions: first, we structure and synthesize past research by using an input-process-output framework, and secondly, we use the framework to highlight research gaps for guiding future research in that area.

Keywords: chatbot, learning process, technology-mediated learning, higher education

INTRODUCTION

In recent decades, the number of students per lecturer has constantly risen (Nicol & Macfarlane-Dick, 2006). Large-scale lectures at universities with more than 100 students per lecturer and massive open online courses (MOOCs) are increasingly becoming the default learning scenario. Consequently, individualized support provided by lecturers is nearly impossible and students are unable to engage in effective learning (Brinton, Rill, Ha, Chiang, Smith, & Ju, 2015). Several studies have revealed that this lack of individualized support leads to weak learning outcomes, high drop out rates and dissatisfaction (Brinton et al., 2015; Eom, Wen, & Ashill, 2006; Hone & El Said, 2016). The best solution would be to have one teacher per student. Obviously, this is not possible due to financial and organizational restrictions (Oeste, Lehmann, Janson, Söllner, & Leimeister, 2015).

Chatbots have the potential to solve this problem using the examples of other sectors. Chatbots have a growing presence in modern society, becoming integral parts of everything from personal assistants on mobile devices to technical support help over telephone lines, and even being used for health interventions (Serban et al., 2017). In 2015, the size of the chatbot market comprised 113 million U.S. dollar and is projected to be 994.5 million U.S. dollar in 2024 (McKinsey & Company, 2016; Statista, 2017). The usage of messaging apps in the past years has also increased exponentially. In 2016, it is estimated that about 75% of all smartphone users used some sort of messaging apps (Maruti Techlabs, 2017). Moreover, analysts predict that by 2020, 30% of all web browsing sessions will be done without a screen, 50% of all searches will be by voice commands, and customers will manage 85% of their enterprise relationships without interacting with a human being (Gartner, 2016; Olson, 2016).

This rapid increase of chatbots comes with four main advantages. Firstly, the implementation of chatbots saves customer service costs by replacing nearly all human assistants. For example,

Juniper Research reported that by 2022 chatbots will save companies worldwide about eight billion dollars annually in customer supporting costs. On average, a chatbot interaction with a customer saves them a cost of 0.70 \$ per interaction compared to traditional support conversations (Juniper Research, 2017). Secondly, chatbots increase user satisfaction by speeding up response times and being available twenty-four hours a day. For instance, the famous hotel chain Marriott International started using a chatbot for hotel reservations in 2017. Since then, 44 percent of all registered members using Facebook received successful assistance related to their stay or reservation (Nguyen, 2017). Thirdly, In contrast to websites and apps, chatbots can anticipate customer questions and thus interact proactively with their users and display exactly the information users are searching for. Without a chatbot, customers and other users would need to trigger a conversation with human assistants themselves or search the required information on websites. This leads to inconvenient and inefficient, as well as easily avoidable, search processes. Chatbots can initiate communication according to the user's specific context, such as location or clickstreams, and therefore makes the customer feel personally addressed (Howlett, 2017). Fourthly, chatbots are also a great business analysis tools, since conversations between users and chatbots can automatically be analyzed to better understand customer requirements, and therefore, improve products and services. Despite all these advantages, chatbots are still at the beginning stages of their expansion and 2018 is promised to be a year where chatbots spread into more fields, such as education (Chatbot's Magazine, 2017).

Until now, the implementation of chatbots in education has been rather scarce. Instead, studies have often tried to implement pedagogical agents and traditional intelligent tutoring systems in learning scenarios (Baker, 2016; Goos, Hartmanis, van Leeuwen, Goettl, Halff, Redfield, & Shute, 1998; Govindasamy, 2014). Pedagogical agents are human-like interfaces between the learner and the content in an educational environment and traditional intelligent

tutoring systems are computer systems that aim to provide immediate and customized instruction or feedback to the learner (Goos et al., 1998; Kim, Baylor, & Shen, 2007). In contrast to these learning technologies, chatbots interact with students in a synchronous way making it possible to react on individual intents. This allows students to remain in the driver seat and actively control their learning process, which is the major success factor for effective learning according to the predominant constructivistic learning theory (Glaserfeld, 1987).

Chatbots in education promise to have a significant positive impact on learning success and student satisfaction. A small number of studies have already shown successfully implemented chatbots in learning scenarios (Dutta, 2017; Huang, Lee, Kwon, & Kim, 2017; Kerly, Hall, & Bull, 2007). For example, the University of Georgia created a chatbot based on IBM's Watson platform called 'Jill Watson', which was developed specifically to handle forum posts by students enrolled in a computer science course (Ashok, Brian, Mithun, Shanu, Abhinaya, & Bryan, 2015). As a result, students were more engaged in the course, and wished that they had the same opportunity in other courses as well (Lip ko, 2016). Especially in large-scale learning scenarios at universities or in massive open online courses (MOOCs), chatbots have the potential to compensate the insufficient individual support of lecturers which is one of the main factors leading, for example, to retention rates less than 10% for MOOCs (Hone & El Said, 2016). This means that chatbots can help providing individual learning support with limited investment of financial and organizational resources.

Despite the mentioned benefits of chatbots in general education, chatbots can play a major role in management education as well. Several managerial competency frameworks suggest that making judgments and decisions, providing and receiving feedback, analytical thinking and technological awareness are crucial competencies for future managers (OECD, 2014; Ruth, 2006; United Nations, 2017). Chatbots can help to develop each of these skills. Firstly, chatbots can

deliver future managers the right information at the right time to make right judgements and decisions. Secondly, chatbots are promising tools to provide continuing feedback to lecturers and students. Thirdly, using chatbots as learning partners improve the ability of students to collect and analyze relevant data by empowering students to analyze a problem first and then getting the information quickly. Last but not least, future managers get trained to work hand in hand with digital assistants, which becomes standard in future management activities.

The goal of this systematic literature review is to set the ground for future research regarding chatbots in education. We therefore structured and synthesized past research through a multi-perspective framework consisting of an input, process and output perspective. This framework should help future researchers to fill important research gaps in order to bring chatbots into (management) education. This paper is structured as follows. In section two and three, we formalize the three main concepts “chatbot”, “technology-mediated learning” and “learning outcome”, and present the development of our theoretical framework. Section four describes the applied method of our review, section five the findings, section six the discussion, section seven describes the implications for future research and practice by showing future research directions and section eight completes with the conclusion.

MAIN CONCEPTS

In this section, we describe the three core concepts “chatbot”, “technology-mediated learning” and “learning outcome” forming the conceptual basis of our framework.

Chatbots

Chatbots are computer programs which attempt to simulate conversations of human beings via text or voice interactions (Rouse, 2017). According to this definition, chatbots comprise all kinds of software enabling humans to make a conversation with a computer. This includes talkbots,

chatterbots, conversational agents, artificial conversational entities and also virtual assistants such as Amazon's Alexa or Google's Home. Chatbots are typically used in dialog systems for different reasons including customer service, information acquisition, etc. (Serban et al., 2017).

Chatbots have a long history going back to 1964, when the first chatbot ELIZA was created. ELIZA analyzed input sentences and created its response based on reassembly rules associated with a decomposition of the input (Weizenbaum, 1966). More than thirty years later, the chatbot A.L.I.C.E. was developed using Artificial Intelligence Markup Language, which is based on categories containing a stimulus, or pattern, and a template in the response. Category patterns are then matched to find the most appropriate response to a user input (Wallace, 2007). In 2008, Cleverbot was launched and, unlike other chatbots, its responses were not pre-programmed. Instead, it learned directly from human input, where a user would type in a comment or question and Cleverbot would find all keywords or an exact phrase matching the input (Gehl, 2014). Nowadays, chatbots are able to capture a wide range of user cases and are able to steer the user in a desired direction.

The remarkable increase of chatbots in the last years indicates three dimensions to classify them. Dimension one, "building approaches", distinguishes between retrieval-based models and generative models. Retrieval-based models use a repository of predefined responses and some kind of algorithm to pick an appropriate response based on the input and/or context. This algorithm can be very simple, such as rule-based expression match, or more complex, such as machine learning classifiers. Generative models do not reply with pre-defined responses, instead, they generate responses out of the input with the help of machine learning techniques. Both approaches have advantages and disadvantages. On the one hand, retrieval-based methods are easier to build and do not make grammatical mistakes, on the other hand, they are unable to handle unseen cases where no predefined response exists. Generative models are able to refer back to contextual entity

information mentioned earlier in the conversation. This let the chatbot feel more human and capable of longer dialogs (Liu, Lowe, Serban, Noseworthy, Charlin, & Pineau, 2016). The second dimension is related to the input mode of chatbots. Newly developed chatbots tend to prefer speech over text as input for a conversation. Both ways seem to have advantages. Speech input enables the user to have a more natural and convenient conversation with chatbots compared to text input, which enables the user to trigger conversations in places where sound might not be social accepted, such as schools, offices, etc. (Abu Shawar & Atwell, 2007). The third dimension addresses the inclusion of contextual information. In human-to-human conversations, people keep track of what the dialog partner has said or done in order to use that information for further conversations. For example, humans can ask “how has it been going?” when the dialog partner had an important presentation the same day. Humans automatically know that this question refers to the presentation whereas chatbots struggle with that. Therefore, chatbot developer try to embed contextual information such as linguistic (e.g. something mentioned earlier in the text) and physical context (f.e. date/time, location, user information) in order to select the right responses (Britz, 2016).

Technology-Mediated Learning

Technology-mediated learning (TML) is described as “*an environment in which the learner’s interactions with learning materials (readings, assignments, exercises, etc.), peers, and/or instructors are mediated through advanced information technologies*” (Alavi & Leidner, 2001). TML services have been in the focus of research for decades coming from disciplines such as psychology, education, business or information systems (Gupta, Bostrom, & Huber, 2010). In practice, TML works in many forms and combines different learning styles and methods (Gupta & Bostrom, 2009; Söllner, Bitzer, Janson, & Leimeister, 2017):

- Web or computer-based
- Asynchronous or synchronous

- Instructor-led or self-paced
- Individual-based or team-based (collaborative learning)

The theoretical model of TML described by Gupta and Bostom (2009) states that learning outcomes in TML are impacted by procedural and structural factors. Structural factors consist mainly of the learning method and procedural factors deals with students' actions while interacting with the structures in their learning processes. Both factors need to be analyzed to fully understand the impact of technology on learning outcomes (Söllner et al., 2017).

Chatbot-mediated learning (CML) can be considered as one way of TML. It is mostly web-based, synchronous, self-paced, and concentrates on an individual student. CML provides a new individual learning experience where students can proactively use chatbots in their learning process in order to increase their learning process quality and learning outcome.

Learning outcomes

Learning outcome is a multi-faceted construct. How to classify and measure learning outcomes resulting out of a learning process is highly discussed in past research. Cognitive psychologists and neuroscientists both provide evidence supporting a distinction between declarative and procedural knowledge. Declarative knowledge refers to factual knowledge such as knowing: "World War II started in 1938". In contrast, procedural knowledge refers to process-related knowledge such as knowing how and therefore being able to demonstrate how such as "knowing how to structure one's own learning process". These two types of knowledge are not observable by a third person. Knowledge, which is observable by a third party, can be summarized under psychomotor skills. Apart these types of knowledge, Goleman (2013) pointed out that there exists a third category named affective knowledge. Affective knowledge means managing own emotions and the inner potentials for positive relationships. These inner potentials further motivates

us to engage more in some activities than in others. In summary, human learning outcomes can be divided into cognitive, affective and psychomotor.

Anderson et al. (2001) tied on these findings and tried to classify learning outcomes with a matrix consisting of a knowledge and a cognitive process dimension. The knowledge dimension consists of factual, conceptual, procedural and metacognitive knowledge. Apart from already described factual and procedural knowledge, they added conceptual and metacognitive knowledge. Conceptual knowledge shows the interrelationships among basic concepts and metacognitive knowledge is the knowledge of cognition in general (see table 1).

INSERT TABLE 1 ABOUT HERE

The second dimension describes cognitive processes. It consists of remembering, understanding, applying, analyzing, evaluating and creating. This dimension is considered to be a hierarchy in the sense that the six major categories are believed to differ in their complexity, with remembering being less complex than understanding and so on. Chatbots can help to achieve more complex categories in all types of knowledge. Despite these cognitive learning outcomes, learner satisfaction is often used for measuring affective learning outcomes (Arbaugh, 2001; Gupta et al., 2010).

THEORETICAL FRAMEWORK

In this section we describe how we developed our theoretical framework used to derive initial research questions, integrate past research and highlight future research directions. We followed the theoretical model of technology-mediated learning (TML) described by Gupta and Bostrom (2009) in order to consider chatbot research in education out of three different

perspectives: structure (input), learning process (process), and learning outcome (output, see figure 1).

INSERT FIGURE 1 ABOUT HERE

The input perspective of chatbot-mediated learning (CML) consists of individual predispositions of learners such as computer experience (Arbaugh, 2001), self-efficacy (Lim, Lee, & Nam, 2007), learning styles (Ozkan & Koseler, 2009), and learning motivation (Klein, Noe, & Wang, 2006). These individual differences can have an influence on CML process quality and CML outcomes. The other part of the input perspective consists of chatbot's quality and design which is another key success factor helping students to increase learning outcomes. The second perspective addresses the learning process (Siau, Sheng, & Nah, 2006). Söllner et al. (2017) argued that the process perspective is crucial for a full understanding of how technology impacts learning. The learning process is a complex phenomenon including cognitive processes and interactions based on learning methods, individual differences between learners and other elements of the teaching/learning scenarios that influence TML outcomes (Gupta et al., 2010; Gupta & Bostrom, 2013). For instance, if chatbots does not meet the requirements of students such as easy use and access, the chatbot adds little value and is not used any more. Thus, the chatbot-mediated learning process quality can have a remarkable impact on learning outcomes. Last but not least, the outcome perspective considers how input and process factors influence different types of learning outcomes (Gupta & Bostrom, 2013).

METHODS

We conducted a systematic literature review based on Webster and Watson (2002), and vom Brocke (2015), according to which, a review should be concept-centric and follow basic principles, such as defining the search scope and justifying all decisions made during the search process. First of all, we defined the search scope by using Cooper's Taxonomy (1988, see figure 2).

INSERT FIGURE 2 ABOUT HERE

We focused on outcomes of chatbot research studies and their underlying theories and practical applications, because the major goal of the review is to show the potential of chatbots for increasing learning outcomes. Further, we integrate and synthesize past literature by bridging the gap between management, information systems, education and psychology research with a common framework. By this means, we want to identify issues central guiding future research directions. We review chatbot literature out of a neutral perspective since our focus is to synthesize literature from different research fields and make an exhaustive literature review with selective citation. This means that we tried to cover nearly the entire literature on chatbots in educational settings and included a selected, representative sample of works in our findings. Our literature review addresses specialized scholars in the field of (management) education as well as teachers and educational institutions.

We tried to justify all our decisions and made the process transparent by following a five-stage framework suggested by Arksey and O'Malley (2005). Thus, the five stages of our chatbot literature review consists of a) identifying initial research questions b) identifying relevant studies c) study selection d) charting the data and e) collating, summarizing, and reporting the results.

Step 1: Identifying the initial research questions

We analyzed and structured past literature out of three perspectives: input, process, and output. Hence, we use our developed framework to derive the initial research questions (see figure 3).

INSERT FIGURE 3 ABOUT HERE

First of all, we occupied an input perspective on chatbots in education consisting of individual student differences and different types of chatbots. Thus, we derived the following questions starting with the input perspective:

1. How do individual differences of students affect chatbot-mediated learning (CML) processes and CML outcomes?
2. In which educational settings are chatbots applied?
3. What approaches are being used to build and design a chatbot in learning settings and how does that influence CML processes and CML outcomes?

Question one refers to student characteristics, which can have a significant impact on technology-mediated learning (Gupta & Bostrom, 2009). Question two tries to capture all fields of learning settings, where chatbots are already successfully implemented and question three highlights the importance of building and designing chatbots in order to have an impact on learning.

Second, we derived a question out of the process perspective:

4. What are determinants of CML process quality and how do they relate to chatbot design and CML learning outcomes?

Question four is a core question of our review since even a well-developed chatbot might be useless if it does not produce additional value to students during the learning process.

Last but not least, we derived an initial research question from the output perspective:

5. How can different kinds of learning outcomes be influenced by chatbot design, individual differences of students and CML process quality?

To set the ground for future research and practical implementation of chatbots it is important to outline which kind of learning outcomes can be achieved.

Step 2: Identifying relevant studies

We applied a wide definition of chatbots in order to capture available literature from different research fields such as management, education, information systems, and psychology. First of all, we made an informal literature search in order to identify keywords for the actual open database search resulting in the search terms depicted below:

(*“conversational agent” OR “chat bot” OR “chatbot” OR “pedagogical agent” OR “intelligent tutor system” OR “dialogue system” OR “smart personal assistant” OR “smart assistant”*)

After reviewing literature of database *science direct*, we added “dialogue system” to our search string, because we considered this key word to be a central concept in capturing most of the chatbot literature.

Before reviewing the literature, we developed inclusion and exclusion criteria. The last ten to fifteen years have seen a rapid improvement and spreading of chatbots due to new technologies such as natural language processing, neural networks and machine learning (Pelk, 2016). Therefore, we considered the time period 2005 to 2018 to be appropriate. We further included articles related to the fields of management, education, information systems, and psychology. Moreover, we included studies on chatbots trying to mediate learning processes. This means that they helped to create any kind of learning outcomes. We therefore excluded chatbots without intended learning outcome (e.g. customer support or just for fun chatbots), bots without chatting

function, and recommendation and feedback systems with one-way communication. Databases used for investigation are *Science Direct*, *EBSCO Business Source Premium*, *IEEE Xplore*, *Pro Quest*, *Aisle*, and *ACM DL*. In addition, we undertook a forward and backward search of central articles and used Google Scholar to identify other primary sources within grey literature. We conducted the review over two months, November and December 2017.

Step 3: Study selection

Overall, we identified 1405 articles. A review of the abstracts revealed large numbers of articles that were irrelevant. For example, educational technologies that did not intend to communicate with students and only provided one way recommendations and adaptations (Im & Hars, 2007; Sonwalkar, 2008). Further, our literature storing system removed some articles because they were duplicates of other databases. Guided by the defined inclusion and exclusion criteria, we identified 236 relevant articles. In the next step, we thoroughly read the full texts, resulting in a representative and manageable amount of 80 articles.

Step 4: Data charting and collation

In the next step, we mapped the concepts of our selected articles to the input-process-output framework and our initial research questions. For example, the article “Automated interventions for multiple health behaviors using conversational agents” from Bickmore et al. (2010) discusses the implementation and effect of chatbots in medicine addressing question two of our review. The next section discusses the findings for our review.

FINDINGS

The systematic review includes 80 articles. In general, it is worth noting that fifteen articles were published in 2017 and twelve articles in 2016. This reflects the trend of chatbots already

described in the introduction section. In the next sections we structure and discuss our findings based on the initial research questions.

Question 1: How do individual differences of students affect chatbot-mediated learning (CML) processes and CML outcomes?

Several individual differences of students influence CML process quality and the related learning outcome through a complex interaction of students with learning methods and structures (Gupta et al., 2010). Firstly, the attitude and trust of students towards technology plays a major role. Students with a positive attitude towards the value of chatbots feel more content in their CML processes (Söllner et al., 2017). Secondly, learning characteristics such as trait emotions and personality traits have a significant effect on agent-directed emotions during the learning process. Students with trait emotions during learning, such as anger, anxiety or joy, and stable personality traits, such as agreeableness, conscientiousness, and neuroticism, have significant influence on chatbot-directed emotions during CML processes (Harley, Carter, Papaionnou, Bouchet, Landis, Azevedo, & Karabachian, 2016). Thirdly, educational background and social and technological skills influence the way how students interact with chatbots. For example, students with a background in humanities tend to have a more natural dialogue with human-like sentences during CML processes, whereas computer science students tend to be more formal trying to challenge chatbots to their limits of ‘intelligence’ (Novielli, Rosis, & Mazzotta, 2010). Furthermore, students with better technological skills and a stronger need for interaction benefit more from chatbots (Ben Mimoun, Mohammed Slim, & Poncin, 2015). Finally, self-efficacy and self-regulated skills of students influence CML processes and the related learning outcomes. High self-efficacy and well developed self-regulated skills of students have a positive effect on CML process quality and CML

outcomes (Söllner et al., 2017). To sum it up, we see that CML process quality and related learning outcomes depends significantly on individual learner characteristics.

Question 2: In which educational settings are chatbots applied?

Health and well-being interventions

Several studies under review investigated chatbots used as medical agents supporting students in their medical education or patients with their therapy. In medical education, chatbots are often used on mobile devices in order to make learning independent of place and time. These chatbots are mostly integrated into a learning platform, thus helping students through tests and instructional material (Alepis & Virvou, 2011). In patient therapy, chatbots provide information and counseling to hospital patients at the time of hospital discharge and react on patient questions. The aim of these interventions is to provide individual support to patients helping them to follow their therapy. For example, one study showed that patients with depressive symptoms prefer to receive support from chatbots rather than doctors or nurses and showed that they have a significantly higher therapeutic relationship compared to patients without a chatbot. With the help of a chatbot, patients learn how to help themselves with a degree of independence of human assistance (Bickmore et al., 2010). In the well-being sector, chatbots help students to change their unhealthy behavior by triggering conversations with them. For example, daily conversations with a chatbot can lead to greater physical activity and promote fruit and vegetable consumption (Bickmore, Schulman, & Sidner, 2013).

Language learning

Chatbots enable students to receive individual support in a mistake-friendly environment through giving them the chance to apply their language skills anytime and anywhere. By providing

students with daily conversation contexts, students get the chance to use their knowledge in a simulated real-life situation. This reduces anxiety and increases the willingness to communicate in a foreign language (Ayedoun, Hayashi, & Seta, 2015). Nevertheless, this effect seems to decline rapidly over time as students lose interest in chatbots as language partners compared to human learning partners. This could happen because of a simple novelty effect or simply the weaker value of chatbots compared to human assistants (Fryer, Ainley, Thompson, Gibson, & Sherlock, 2017b).

Feedback and metacognitive thinking

Chatbots in education can also be used to provide feedback to students about their performance and can trigger metacognitive thinking processes by discussing student learning behaviors in a synchronous way. Today, one grade at the end of semester is often the only feedback students get since lecturers have nearly no time for individual formative feedback during the courses. However, formative feedback during learning processes is, according to Hattie (2012), one of the most important factors in increasing students' performance and motivation. Chatbots used as formative feedback tools have an additional positive effect on engagement indicators and task completion since students can follow-up with questions on their assessments (Lundqvist, Pursey, & Williams, 2013). In addition, chatbots can also be used to help students improving their metacognitive thinking considered as key competences in a rapid changing world where lifelong learning gets more and more important. For example, the use of a chatbot during a course showed that the discrepancy between computer assessment and student assessment sank significantly (Kerly, Ellis, & Bull, 2008). The major advantage of chatbots in comparison to asynchronous ways of communication is that students reflect their beliefs during discussion.

Motivation and self-efficacy

Chatbots can be used to increase student motivation in learning and strengthen their self-perception. Motivation and self-efficacy are one of the most important factors of learning success (Schunk, 1991). For example, chatbots can be used to create curiosity by asking students challenging questions which is a main driver of intrinsic motivation (Oudeyer, Gottlieb, & Lopes, 2016). Further, chatbots can strengthen the self-belief of students by giving them more control of their learning process (van der Meij, van der Meij, & Harmsen, 2015).

All in all, the review shows that chatbots are implemented for a variety of purposes. To sum it up, chatbots they try to offer individual support in order to give people more control over their learning process. This leads to a higher CML process quality and higher learning outcomes.

Question 3: What approaches are being used to build and design a chatbot in learning settings and how does that influence CML processes and CML outcomes?

Flow Chatbots

Flow chatbots are tree-based using a previously defined path defined by the chatbot's developer. A large number of articles in our review have used this kind of chatbot. When a user input is entered, the chatbot tries to match it with the existing database. If a match is found, the answer is generated, if not, the chatbot tries to gain additional information or provides the user with buttons to choose from. This kind of chatbot requires a large dataset in order to keep the frustration level of students caused by wrong responses low (Allison, 2012).

Artificially Intelligent Chatbots

Artificially intelligent chatbots allow the user to engage in a much freer way and are therefore more similar to a real human-to-human conversation. Basically, past literature

differentiates between one-way and two-way artificially intelligent chatbots. For one-way artificially intelligent chatbots, the chatbot is trying to understand what the user is saying by using machine learning techniques (Dutta, 2017). Two-way artificially intelligent chatbots use on top of that artificial intelligence to feed the information back to the user. Instead of picking a response from a pool of pre-made answers, two-way artificially intelligent construct the most accurate answer word-by-word before delivering it to the user. These chatbots are able to learn over time, based on real inputs from the users in terms of right and wrong answers of chatbots. For example, the famous chatbot jabberwacky makes conversation by reflecting user input learned in similar contexts with previous conversations (Angeli & Brahnam, 2008). Most recent studies under review used one-way artificially intelligent chatbots helping to understand student intents better and therefore increase CML process quality and learning outcome. Furthermore, we did not found any two-way artificially intelligent chatbots in educational settings.

Chatbots with integrated speech recognition

Recent articles in our review investigated the benefits of speech input during learning sessions (Abdul-Kader et al., 2015; Griol, Molina, & Callejas, 2017a; Saadatzi, Pennington, Welch, Graham, & Scott, 2017). This trend reflects practical implementations such as Amazon's Alexa or Google's Home. Speech recognition within a chatbot converts user inputs by recognizing the vibrations in the air while the user is speaking. The chatbot is then translating this analog data into digital data that is understandable by a computer (Graves & Jailty, 2014). Despite recognizing the content of speech messages, emotional states of users can also serve as input for responses (Alepis & Virvou, 2011). This is a great chance to recognize affective behavior of students such as boredom, anxiety, excitement, etc. in order to create responses on that (Griol, Molina, & Callejas, 2017b).

Chatbots with integrated context-data

More and more chatbots try to include contextual information to increase productivity and personality of chatbots. Articles under review showed mainly two different ways of including context information in chatbot conversations. One way is to capture physical context data of students by analyzing their behavior on learning platforms. This enables chatbots to anticipate students' questions or adapt responses to students' learning style leading to a higher CML process quality (Crockett, Latham, & Whitton, 2017; Latham, Crockett, McLean, & Edmonds, 2012). Another way of including context data into chatbot responses is by capturing the emotional status of students. Emotions such as anger, frustration, fear and joy can be attached by speech-based emotion recognition as well as sentiment analysis of linguistic transcripts leading to a more individualized learning support (Griol et al., 2017b; Pérez, Cerezo, Serón, & Rodríguez, 2016).

Our review has shown that different types of chatbots have different influences on CML processes and the related learning outcomes. It strongly depends on the specific use case of chatbots in the learning process in order to identify which kind of chatbot is most useful.

Question 4: What are determinants of CML process quality and how do they relate to chatbot design and CML learning outcomes?

Comparison between human-to-chatbot and human-to-human communication

The way of communication between human-to-chatbot and human-to-human differs significantly. Several studies in our review revealed that students interacting with chatbots during their learning process are less agreeable, less open, less extroverted, communicate for longer durations with shorter messages, lack much of the richness of vocabulary, and exhibit greater profanity (Corti & Gillespie, 2016; Hill, Randolph Ford, & Farreras, 2015; Mou & Xu, 2017). On the one hand, shorter messages and the lack of richness of vocabulary makes it harder for chatbots

to understand student' intents. On the other hand, people tend to adapt their language to the language of chatbots similar to when they are speaking with children or using search engines such as google (Hill et al., 2015). This can increase the effectiveness of chatbots. Moreover, the greater profanity of people when interacting with self-learning chatbots can lead to undesirable answers. One example might be Microsoft's tay chatbot on twitter, which was manipulated to provide racist and profane responses within twenty-four hours (Vincent & James, 2016). Chatbot developers have to understand this different way of communication in order to improve CML processes and learning outcomes.

Comparison between spoken and text input

One influencing factor of CML process quality is the input mode of chatbots. A few articles investigated the comparison between text and spoken inputs. They revealed that spoken input produces a richer language, a friendlier conversation and longer dialogues (Cremonesi, Ricci, Berkovsky, & Tuzhilin, 2017; Novielli et al., 2010). Furthermore, speech can be considered as a more natural and convenient way of communication leading to increased user satisfaction. Nevertheless, speech input is not superior at all. There might be learning locations such as during lectures or in the bus where speech input is not possible (Abdul-Kader et al., 2015). Thus, the combination of speech and input modes might lead to the best CML process quality.

Structural embeddedness in learning scenario

Recent research has analyzed that the structural embeddedness of technology during the learning process is a key success factor for learning process quality (Bitzer & Janson 2014; Gupta & Bostrom, 2013). When students does not recognize the higher value of chatbots during learning immediately and the instructions given on chatbots are not clear, the CML process quality is low.

For instance, if chatbots are not directly accessible where students questions arise or if lecturers does not explain how to best use them, chatbots might stay unsuccessful (Söllner et al., 2017).

Despite different student characteristics and chatbot types classified as input factors, the way how students implement chatbots during their learning process is crucial to understand its impact on learning outcomes.

Question 5: How can different kinds of learning outcomes be influenced by chatbot design, individual differences of students and CML process quality?

Visualization of Chatbots

The optic of chatbots is a key factor for achieving learning success of students. Articles under review showed controversial conclusions. One article revealed that chatbots should be embodied, show emotions and have a personality (Berry, Butler, & Rosis, 2005). Another article was able to show via eye-tracking that an embodiment of chatbots is not needed since students does not pay attention to it (Ben Mimoun et al., 2015). Moreover, the gender of chatbots does not play a role regarding famous chatbot measures such as usage, spoken topics, etc. (Brahnam & Angeli, 2012).

Context-Awareness

Chatbots should include context information in order to increase cognitive and affective learning outcomes. This means that a chatbot should consider the cognitive as well as emotional status of students in order to help them to reach their learning goals. Especially, detecting knowledge gaps of students is essential to increase cognitive learning outcomes of students (Pérez et al., 2016). Moreover, detecting the emotional status and including relational behavior such as caring and empathy helps students to improve their affective learning outcomes (Foster &

Oberlander, 2010). This is also the reason why chatbots should include small talk (Bickmore et al., 2013; Kerly et al., 2007).

Proactive guided by students

Chatbots can help students to actively control and regulate learning processes in order to measure metacognitive learning outcomes. The predominant constructivistic learning theory states that this is the only way how learning occurs (Glaserfeld, 1987). Chatbots are therefore more successful when they empower students to self-regulate their own learning process leading to better metacognitive learning outcomes (Duffy & Azevedo, 2015).

Integration in existing learning and instant messaging tools

Chatbots can only be successful if they are easily accessible and respond fast. Thus, they should be implemented in existing learning and messaging systems, where students directly conduct their tasks or communicate with each other. Studies in our review showed that one of the success factors of a chatbot is that students can receive the required information fast and they do not have to use too many applications during the learning process. The chatbots should therefore be implemented directly where students conduct tasks, such as learning management systems, or where they communicate outside of the learning platforms, such as facebook and whatsapp (Pereira, 2016).

The review shows that chatbots are able to improve affective, cognitive and metacognitive learning outcomes. It is thereby essential that chatbots are easily accessible, respond fast, empower students to control their learning process and consider contextual information of students such as emotional status.

DISCUSSION

In this section, we quickly summarize the findings, highlight and comment important issues and discuss implications structured by our research questions.

Question 1: How do individual differences of students affect the chatbot-mediated learning processes and CML outcomes?

Several articles in our review have shown that a number of individual differences of students such as personality traits, educational background, and self-regulated skills of students have an impact on CML process quality. It can therefore be assumed that chatbots should be tailored to the individual differences of students. However, these findings are rather difficult to interpret because no study analyzed whether personalized chatbots have a significant impact on learning outcomes. Future studies, which take these variables into account, will need to be undertaken. On top of this, there are still many unanswered questions about how chatbots can be designed to adapt to individual differences, such as personality traits and self-regulation skills.

Question 2: In which learning settings are chatbots applied?

Prior studies have noted the importance of chatbots for a variety of purposes including medical education and therapy, language learning, receiving feedback, and strengthen motivation and self-efficacy. Chatbots across all fields showed that students appreciate using chatbots due to a mistake-friendlier learning environment compared to talking to human authorities. One interesting was made by Bickmore et al. (2010), where patients prefer talking to chatbots over doctors and nurses. However, Fryer et al. (2017a) states that most of these effects occur because of a novelty effect and decline over time. This finding is consistent with technological implementations in other fields such as information systems (Wells, Campbell, Valacich, & Featherman, 2010). Thus, further work is required to establish the long term effects of chatbots on learning processes and outcomes.

Furthermore, very little was found in the literature on the question of systematically evaluating chatbots. This may be explained by the fact that learning processes and outcomes are complex and influenced by a lot of factors. It is therefore difficult to preclude all other factors than chatbots. However, there are a variety of methods to evaluate the usefulness of chatbots in the articles under review. In therapy, studies used self-report ratings containing measures such as satisfaction and usability. In language learning, studies used metrics like the amount of conversations (Ayedoun et al., 2015). In studies about receiving feedback and strengthen motivation and self-efficacy, they used dialogue time as important metric (van der Meij et al., 2015). To develop a full picture of the potential of chatbots in education, additional studies regarding a systematic method for evaluating chatbots in education are necessary.

Question 3: What approaches are being used to build and design a chatbot in learning settings and how does that influence CML processes and CML outcomes?

Our review showed that chatbot building approaches and chatbot design can be distinguished mainly in three different dimensions. First, whether they are retrieval based or generative, second, whether they use speech or/and text as input mode, and third, whether they integrate context-data. Most of the studies under review before 2016 showed very basic, rule-based chatbots. With upcoming technological improvements regarding natural language processing and machine learning techniques, more and more generative, artificially intelligent chatbots arised. The most interesting finding was that of Dutta et al. (2017). He developed an intelligent Chatbot tool to assist high school students for learning general knowledge subjects using natural language processing and machine learning algorithms for recognizing the correct intent. The review also showed that for closed domains with a fixed set of possible questions rule-based chatbots might be more than sufficient. In all other cases, chatbots with artificial intelligence techniques lead to better CML

process quality and CML outcomes. Thus, further research should concentrate on the development and design of artificially intelligent chatbots.

Question 4: How do students interact with chatbots during their learning process and how is that related to learning outcomes?

Our review has revealed that the way how student interact with chatbots is fundamentally different to interactions with human teaching staff. Studies showed that students apply simpler sentences, lack much rich vocabulary and get distracted more easily. These findings suggest that for chatbots it is even more difficult to understand students' intents. The challenge here is to provide follow-up questions to gather the right intentions and to guide students back whenever they get distracted. This is an important issue for future research. Moreover, our review indicates that studies are not about replacing lecturers with chatbots. It is more about creating a co-dependent and intelligent relationship between teacher and chatbot utilizing both for their strengths and deliver the best student experience (Hughes, 2017). An important research question for the future is thus how and where human or chatbot assistance is preferable.

Apart from that, recent studies under review recognized the success of voice assistants such as Amazon's Alexa and Google's Home and therefore included speech input in their educational chatbots. One interesting finding was that of Griol et al. (2017b) who combined speech-base and linguistic classification in order to successfully detect emotions. This finding support the work of other studies in this area trying to include emotions in chatbot systems. Despite these promising results, questions remain on how speech input changes CML learning processes as well as the resulting learning outcomes.

Question 5: How can different kinds of learning outcomes be influenced by chatbot design, individual differences of students and CML process quality?

Our findings reveal that chatbots can be used to improve different kinds of learning outcomes such as affective, cognitive and metacognitive learning outcomes. However, we found contradictory results on how chatbots can achieve these learning outcomes. For example, studies showed controversial insights regarding the visualization of chatbots. One part of the studies preferred to use embodiment and building up a personality of the chatbot (Berry et al., 2005), whereas others came to the opposite conclusion (Ben Mimoun et al., 2015). Moreover, our findings suggest that chatbots should be used to empower students to actively control and regulate the learning process. Further studies should therefore analyze how chatbots can help students to actively learn new content and reflect on their learning process, and investigate which visualization method works best.

IMPLICATIONS FOR RESEARCH AND PRACTICE

In this section we summarize and elaborate on unresolved issues with chatbots in education by using our input-process-output framework (see table 2). Furthermore, we will take a closer look at the potential of chatbots in management education and outline practical implications.

INSERT TABLE 2 ABOUT HERE

The review has identified ten important research gaps in literature that need to be addressed in order to show the value of chatbots in education. The input perspective deals with different student characteristics and chatbot types influencing CML processes and CML outcomes. First of all, there are many unanswered questions on how chatbots can successfully be tailored to individual differences of students and how these personalized chatbots influence CML processes and CML outcomes. Moreover, existing studies under review showed different ways of evaluating the

usefulness of chatbots. Further investigations are needed to develop systematic evaluation methods for chatbots in education in order to ensure that they improve learning outcomes. Furthermore, practical implications such as Amazon's Alexa or Google's Home indicate how to successfully design and develop artificially intelligent chatbots in education. More research on this topic is required to provide greater insights into the design and development of educational chatbots.

Out of a process perspective, further research should be undertaken to investigate the additional value of conversations during the CML process in comparison to one-directional ways of learning such as learning material or forums on a learning platform. This is an essential step for highlighting the importance of chatbots over other learning technologies. Moreover, the inclusion of learning theories to identify parts of a learning process that can be supported by chatbots is crucial. Another identified research gap addresses the important question at which point in the learning process human or chatbot assistance leads to better learning outcomes.

Out of an output perspective, more empirical evidence of CML processes on CML outcomes have to be collected. Until now, most studies did not choose research designs trying to measure significant increases on learning outcomes. Moreover, further work should be done to investigate how self-efficacy and self-regulation skills can be increased by chatbots. Furthermore, existing literature does not address specific educational areas such as management education. Chatbots can help future managers to improve their lifelong learning abilities and managerial competencies. Further research is required to explore which competencies can be supported by chatbots.

This literature review shows practitioners the future potential and benefits of chatbots. It helps to identify possible chatbot use cases for instructors in their own lectures in order to offer additional, individual support to students.

LIMITATIONS

Our review exhibits some limitations. We excluded promising fields of application of chatbots such as enterprise-customer interactions or just for fun chatbots, where learning is not a primary function. Nevertheless, we think that chatbots in education should learn from these fields. Moreover, we know that our multi-perspective view on chatbots in education is only one way of collecting and synthesizing existing literature on chatbots. Consequently, the suggested research gaps cannot be considered to be complete.

CONCLUSION

Aim of this review was to synthesize, integrate and structure existing chatbot literature in education. The review has considered chatbots in education out of three different perspectives identifying a lot of factors influencing the effectiveness of chatbots in education. The effectiveness of chatbots in education depends on individual student differences, the ways of building chatbots, and the chatbot mediated learning process quality. Another aim of this review was to provide a framework for future research directions on chatbots in education helping to bring chatbots into education. With this review, it is becoming evident that chatbots play an increasing role in future education. Chatbots have the potential to create individual learning experiences for students and therefore increase learning outcomes and support lecturers and their teaching staff. This is especially true for large-scale lecturers and massive open online courses with more than hundred students per lecture where individual support is due to financial and organizational restrictions not possible.

FIGURE 1

Theoretical framework for literature review

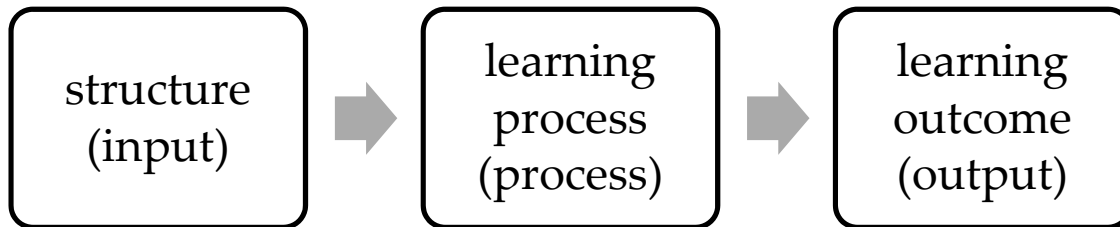


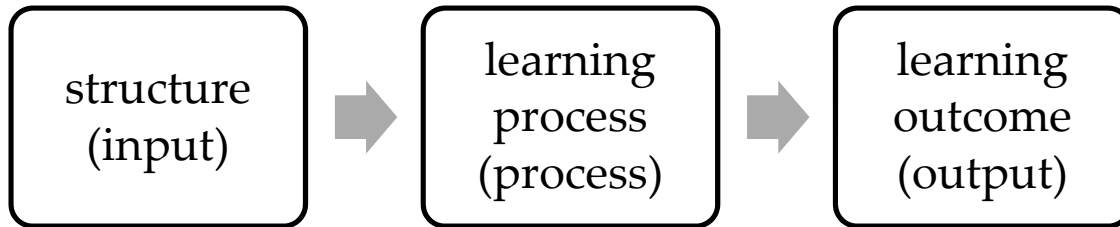
FIGURE 2

Search scope of our review

Characteristics	Categories			
Focus	Research Outcome	Research Methods	Theories	Practices or Applications
Goal	Integration	Criticism	Identification of central issues	
Perspective	Neutral Representation		Espousal of Position	
Coverage	Exhaustive	Exhaustive with Selective Citation	Representative	Central or Pivotal
Organization	Historical	Conceptual		Methodological
Audience	Specialized Scholars	General Scholars	Practitioners or Policy Makers	General Public

FIGURE 3

Initial search questions derived from our framework



1. How do individual differences of students affect CML processes and CML outcomes?
2. In which kinds of educational settings are chatbots applied?
3. What approaches are being used to build and design a chatbot in education and how does that influence CML processes and CML outcomes?

4. What are determinants of CML process quality and how do they relate to chatbot design and CML learning outcomes?

5. How can different kinds of learning outcomes be influenced by chatbot design, individual differences of students and CML process quality?

TABLE 1

Taxonomy of learning outcomes described by Anderson and Krathwohl (2001).

	remember	understand	apply	analyze	evaluate	create
Factual						
Conceptual						
Procedural						
metacognitive						

TABLE 2

Research gaps structured by our input-process-output framework

	Structures (input)	Learning process (process)	Learning outcome (output)
Question 1	<ol style="list-style-type: none"> 1. Adaptation of chatbots to individual student differences 2. Influence of personalized chatbots on learning process and outcome 		
Question 2	<ol style="list-style-type: none"> 3. Evaluation methods of chatbots in education 		
Question 3	<ol style="list-style-type: none"> 4. Design and development of artificially intelligent chatbots 		
Question 4		<ol style="list-style-type: none"> 5. Influence of conversational learning settings on constructing new knowledge 6. Integration of chatbots in different steps of the learning process with the help of learning theories 7. Human vs. chatbot assistance during the learning process 	
Question 5			<ol style="list-style-type: none"> 8. Empirical evidence on the influence of chatbot-mediated learning processes on learning outcomes 9. Influence of chatbots on self-efficacy and self-regulation skills 10. Influence of chatbots on managerial competencies

REFERENCES

- Abdul-Kader, Sameera A., & Woods, J. 2015. Survey on chatbot design techniques in speech conversation systems. *Int. Journal on Advanced Computer Science Applied*, 6(7): 72–80.
- Abu Shawar, B., & Atwell, E. (Eds.) 2007. *Chatbots: Are they really useful?*
- Alavi, M., & Leidner, D. E. 2001. Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS quarterly*, 25(1): 107.
- Alepis, E., & Virvou, M. 2011. Automatic generation of emotions in tutoring agents for affective e-learning in medical education. *Expert Systems with Applications*, 38(8): 9840–9847.
- Allison, D. 2012. Chatbots in the library: Is it time? *Library Hi Tech*, 30(1): 95–107.
- Anderson, L. W., & Krathwohl, D. R. 2001. A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives. *New York. Longman.*, 9(2): 137–175.
- Angeli, A. de, & Brahnam, S. 2008. I hate you! Disinhibition with virtual partners. *Interacting with Computers*, 20(3): 302–310.
- Arbaugh, J. B. 2001. How instructor immediacy behaviors affect student satisfaction and learning in web-based courses. *Business Communication Quarterly*, 64(4): 42–54.
- Arksey, H., & O'Malley, L. 2005. Scoping studies: Towards a methodological framework. *International journal of social research methodology*, 8(1): 19–32.
- Ashok, G., Brian, C., Mithun, K., Shanu, S., Abhinaya, S., & Bryan, W. 2015. Using Watson for Enhancing Human-Computer Co-Creativity. *AAAI Symposium*: 22–29.
- Ayedoun, E., Hayashi, Y., & Seta, K. 2015. A Conversational Agent to Encourage Willingness to Communicate in the Context of English as a Foreign Language. *Procedia Computer Science*, 60(1): 1433–1442.
- Baker, R. S. 2016. Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education*, 26(2): 600–614.
- Ben Mimoun, Mohammed Slim, & Poncin, I. 2015. A valued agent: How ECAs affect website customers' satisfaction and behaviors. *Journal of Retailing and Consumer Services*, 26: 70–82.
- Berry, D. C., Butler, L. T., & Rosis, F. de 2005. Evaluating a realistic agent in an advice-giving task. *International Journal of Human-Computer Studies*, 63(3): 304–327.
- Bickmore, T. W., Mitchell, S. E., Jack, B. W., Paasche-Orlow, M. K., Pfeifer, L. M., & O'Donnell, J. 2010. Response to a relational agent by hospital patients with depressive symptoms. *Interacting with Computers*, 22(4): 289–298.

- Bickmore, T. W., Schulman, D., & Sidner, C. 2013. Automated interventions for multiple health behaviors using conversational agents. *Patient Education and Counseling*, 92(2): 142–148.
- Bitzer, P., & Janson, A. Towards a Holistic Understanding of Technology Mediated Learning Sciences, *European Conference on Information Systems 2014 (ECIS)*. Tel Aviv, Israel.
- Brahnam, S., & Angeli, A. de 2012. Gender affordances of conversational agents. *Interacting with Computers*, 24(3): 139–153.
- Brinton, C. G., Rill, R., Ha, S., Chiang, M., Smith, R., & Ju, W. 2015. Individualization for Education at Scale: MIIC Design and Preliminary Evaluation. *IEEE Transactions on Learning Technologies*, 8(1): 136–148.
- Britz, D. Deep Learning for Chatbots, Part 1 – Introduction; <http://www.wildml.com/2016/04/deep-learning-for-chatbots-part-1-introduction/>, 5 Jan 2018, 05 Jan 2018.
- Chatbot's Magazine. 6 Ways Artificial Intelligence and Chatbots Are Changing Education; <https://chatbotsmagazine.com/six-ways-a-i-and-chatbots-are-changing-education-c22e2d319bbf>, 05 Jan 2018.
- Cooper, H. M. 1988. Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge, Technology & Policy*, 1(1): 104–126.
- Corti, K., & Gillespie, A. 2016. Co-constructing intersubjectivity with artificial conversational agents: People are more likely to initiate repairs of misunderstandings with agents represented as human. *Computers in Human Behavior*, 58: 431–442.
- Cremonesi, P., Ricci, F., Berkovsky, S., & Tuzhilin, A. (Eds.) 2017. *RecSys '17: Proceedings of the Eleventh ACM Conference on Recommender Systems, August 27-31, 2017, Como, Italy*. New York: Association for Computing Machinery.
- Crockett, K., Latham, A., & Whitton, N. 2017. On predicting learning styles in conversational intelligent tutoring systems using fuzzy decision trees. *International Journal of Human-Computer Studies*, 97: 98–115.
- Duffy, M. C., & Azevedo, R. 2015. Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior*, 52: 338–348.
- Dutta, D. 2017. *Developing an Intelligent Chat-bot Tool to assist high school students for learning general knowledge subjects*. Georgia Institute of Technology. Atlanta.
- Eom, S. B., Wen, H. J., & Ashill, N. 2006. The determinants of students' perceived learning outcomes and satisfaction in university online education: An empirical investigation. *Decision Sciences Journal of Innovative Education*, 4(2): 215–235.

- Foster, M. E., & Oberlander, J. 2010. User preferences can drive facial expressions: Evaluating an embodied conversational agent in a recommender dialogue system. *User modeling and user-adapted interaction*, 20(4): 341–381.
- Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. 2017a. Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior*, 75: 461–468.
- Gartner. Gartner Predicts a Virtual World of Exponential Change - Smarter With Gartner; <https://www.gartner.com/smarterwithgartner/gartner-predicts-a-virtual-world-of-exponential-change/>, 04 Jan 2018.
- Gehl, R. W. 2014. Teaching to the Turing Test with Cleverbot. *Transformations: The Journal of Inclusive Scholarship and Pedagogy*, 24(1-2): 56–66.
- Glasersfeld, E. v. 1987. Constructivism. *The concise Corsini encyclopedia of psychology and behavioral science*, 6: 19–21.
- Goleman, D., Boyatzis, R. E., & McKee, A. 2013. *Primal leadership: Unleashing the power of emotional intelligence*: Harvard Business Press.
- Goos, G., Hartmanis, J., van Leeuwen, J., Goettl, B. P., Halff, H. M., Redfield, C. L., & Shute, V. J. (Eds.) 1998. *Intelligent Tutoring Systems*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Govindasamy, M. K. 2014. Animated Pedagogical Agents: A Review of Agent Technology Software in Electronic Learning Environments. *Journal of Educational Multimedia and Hypermedia*, 23(2): 163–188.
- Graves, A., & Jaitly, N. (Eds.) 2014. *Towards end-to-end speech recognition with recurrent neural networks*.
- Griol, D., Molina, J. M., & Callejas, Z. 2017a. Combining speech-based and linguistic classifiers to recognize emotion in user spoken utterances. *Neurocomputing*.
- Griol, D., Molina, J. M., & Callejas, Z. 2017b. Incorporating android conversational agents in m-learning apps. *Expert Systems*, 34(4).
- Gupta, S., & Bostrom, R. 2013. Research Note—An Investigation of the Appropriation of Technology-Mediated Training Methods Incorporating Enactive and Collaborative Learning. *Information Systems Research*, 24(2): 454–469.
- Gupta, S., & Bostrom, R. P. 2009. Technology-mediated learning: A comprehensive theoretical model. *Journal of the Association for Information Systems*, 10(9): 686.
- Gupta, S., Bostrom, R. P., & Huber, M. 2010. End-user training methods: What we know, need to know. *ACM SIGMIS Database*, 41(4): 9–39.
- Harley, J. M., Carter, C. K., Papaionnou, N., Bouchet, F., Landis, R. S., Azevedo, R., & Karabachian, L. 2016. Examining the predictive relationship between personality and emotion

- traits and students' agent-directed emotions: Towards emotionally-adaptive agent-based learning environments. *User Modeling and User-Adapted Interaction*, 26(2-3): 177–219.
- Hattie, J. 2012. *Visible learning for teachers: Maximizing impact on learning*: Routledge.
- Hill, J., Randolph Ford, W., & Farreras, I. G. 2015. Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in Human Behavior*, 49: 245–250.
- Hone, K. S., & El Said, G. R. 2016a. Exploring the factors affecting MOOC retention: A survey study. *Computers & Education*, 98: 157–168.
- Howlett, N. 2017. How machine learning is developing to get more insight from complex voice-of-customer data. *applied marketing analytics*, 3(3).
- Huang, J.-X., Lee, K.-S., Kwon, O.-W., & Kim, Y.-K. 2017. A chatbot for a dialogue-based second language learning system. *CALL in a climate of change: adapting to turbulent global conditions*: 151.
- Hughes, C. 2018: The Year of the Chatbot – Chatbots Magazine;
<https://chatbotsmagazine.com/2018-the-year-of-the-chatbot-fc5a5f780a31>, 08 Jan 2018.
- Im, I., & Hars, A. 2007. Does a one-size recommendation system fit all? The effectiveness of collaborative filtering based recommendation systems across different domains and search modes. *ACM Transactions on Information Systems (TOIS)*, 26(1): 4.
- Juniper Research. Chatbots Infographic - Key Statistics 2017;
<https://www.juniperresearch.com/resources/infographics/chatbots-infographic-key-statistics-2017>, 05 Jan 2018.
- Kerly, A., Ellis, R., & Bull, S. 2008. CALMsystem: A Conversational Agent for Learner Modelling. *Knowledge-Based Systems*, 21(3): 238–246.
- Kerly, A., Hall, P., & Bull, S. 2007. Bringing chatbots into education: Towards natural language negotiation of open learner models. *Knowledge-Based Systems*, 20(2): 177–185.
- Kim, Y., Baylor, A. L., & Shen, E. 2007. Pedagogical Agents as Learning Companions: The Impact of Agent Emotion and Gender. *Journal of Computer Assisted Learning*, 23(3): 220–234.
- Klein, H. J., Noe, R. A., & Wang, C. 2006. Motivation to learn and course outcomes: The impact of delivery mode, learning goal orientation, and perceived barriers and enablers. *Personnel psychology*, 59(3): 665–702.
- Latham, A., Crockett, K., McLean, D., & Edmonds, B. 2012. A conversational intelligent tutoring system to automatically predict learning styles. *Computers & Education*, 59(1): 95–109.

- Lim, H., Lee, S.-G., & Nam, K. 2007. Validating E-learning factors affecting training effectiveness. *International Journal of Information Management*, 27(1): 22–35.
- Lip ko, H. Meet Jill Watson: Georgia Tech's first AI teaching assistant; <https://pe.gatech.edu/blog/meet-jill-watson-georgia-techs-first-ai-teaching-assistant>, 5 Jan 2018, 05 Jan 2018.
- Liu, C.-W., Lowe, R., Serban, I. V., Noseworthy, M., Charlin, L., & Pineau, J. 2016. How NOT to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation. *arXiv preprint arXiv:1603.08023*.
- Lundqvist, K. O., Pursey, G., & Williams, S. (Eds.) 2013. *Design and implementation of conversational agents for harvesting feedback in eLearning systems*: Springer.
- Maruti Techlabs. Why can chatbots replace Mobile Apps immediately? - Maruti Techlabs; <https://www.marutitech.com/why-can-chatbots-replace-mobile-apps-immediately/>, 04 Jan 2018.
- McKinsey & Company. Winning the expectations game in customer care; <https://www.mckinsey.com/business-functions/operations/our-insights/winning-the-expectations-game-in-customer-care>.
- Mou, Y., & Xu, K. 2017. The media inequality: Comparing the initial human-human and human-AI social interactions. *Computers in Human Behavior*, 72: 432–440.
- Nguyen, M.-H. From Fortune 500s to small businesses, real businesses are already using chatbots to improve their service; <http://www.businessinsider.com/examples-brands-companies-chatbots-business-2017-10>, 11 Oct 2017, 05 Jan 2018.
- Nicol, D. J., & Macfarlane-Dick, D. 2006. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education*, 31(2): 199–218.
- Novielli, N., Rosis, F. de, & Mazzotta, I. 2010. User attitude towards an embodied conversational agent: Effects of the interaction mode. *Journal of Pragmatics*, 42(9): 2385–2397.
- OECD. OECD Competency Framework; <http://www.oecd.org/careers/oecdcorecompetencies.htm>, 05 Jan 2018.
- Oeste, S., Lehmann, K., Janson, A., Söllner, M., & Leimeister, J. M. 2015. Redesigning University Large Scale Lectures: How To Activate The Learner. *Academy of Management Proceedings*, 2015(1): 14650.
- Olson, C. Just say it: The future of search is voice and personal digital assistants; <https://www.campaignlive.co.uk/article/just-say-it-future-search-voice-personal-digital-assistants/1392459>, 05 Jan 2018.
- Oudeyer, P.-Y., Gottlieb, J., & Lopes, M. 2016. Chapter 11 - Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. In B. S. a. S. Knecht (Ed.),

Progress in Brain Research : Motivation Theory, Neurobiology and Applications, Volume 229: 257–284: Elsevier.

Ozkan, S., & Koseler, R. 2009. Multi-dimensional students' evaluation of e-learning systems in the higher education context: An empirical investigation. *Computers & Education*, 53(4): 1285–1296.

Pelk, H. Machine learning, neural networks and algorithms – ITNEXT; <https://itnext.io/machine-learning-neural-networks-and-algorithms-4cd27a105b89>, 06 Jan 2018.

Pereira, J. 2016. Leveraging chatbots to improve self-guided learning through conversational quizzes. In F. J. García-Peñalvo (Ed.), *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality - TEEM '16*: 911–918. New York, New York, USA: ACM Press.

Pérez, J., Cerezo, E., Serón, F. J., & Rodríguez, L.-F. 2016. A cognitive-affective architecture for ECAs. *Biologically Inspired Cognitive Architectures*, 18: 33–40.

Rouse, M. What is chatbot?; <http://searchcrm.techtarget.com/definition/chatbot>, 05 Jan 2018.

Ruth, D. 2006. Frameworks of managerial competence: Limits, problems and suggestions. *Journal of European Industrial Training*, 30(3): 206–226.

Saadatzi, M. N., Pennington, R. C., Welch, K. C., Graham, J. H., & Scott, R. E. 2017. The Use of an Autonomous Pedagogical Agent and Automatic Speech Recognition for Teaching Sight Words to Students with Autism Spectrum Disorder. *Journal of Special Education Technology*, 32(3): 173–183.

Schunk, D. H. 1991. Self-efficacy and academic motivation. *Educational Psychologist*, 26(3-4): 207–231.

Serban, I. V., Sankar, C., Germain, M., Zhang, S., Lin, Z., Subramanian, S., Kim, T., Pieper, M., Chandar, S., & Ke, N. R. 2017. A deep reinforcement learning chatbot. *arXiv preprint arXiv:1709.02349*.

Siau, K., Sheng, H., & Nah, F.-H. 2006. Use of a classroom response system to enhance classroom interactivity. *IEEE Transactions on Education*, 49(3): 398–403.

Söllner, M., Bitzer, P., Janson, A., & Leimeister, J. M. 2017. Process is king: Evaluating the performance of technology-mediated learning in vocational software training. *Journal of Information Technology*, 18(2): 159.

Sonwalkar, N. 2008. Adaptive individualization: The next generation of online education. *On the horizon*, 16(1): 44–47.

Statista. Chatbot Market size 2015-2024; <https://www.statista.com/statistics/656596/worldwide-chatbot-market/>.

- United Nations. United Nations Competencies for the future; <https://hr.un.org/materials/united-nations-competencies-future>, 02 Jan 2018.
- van der Meij, H., van der Meij, J., & Harmsen, R. 2015. Animated pedagogical agents effects on enhancing student motivation and learning in a science inquiry learning environment. *Educational Technology Research and Development*, 63(3): 381–403.
- Vincent, & James. Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day; <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>, 06 Jan 2018.
- Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. 2015. Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. *CAIS*, 37: 9.
- Wallace, R. S. 2007. *Chapter 00. The Anatomy of ALICE*: Retrieved.
- Webster, J., & Watson, R. T. 2002. Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*: 13–23.
- Weizenbaum, J. 1966. ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1): 36–45.
- Wells, J. D., Campbell, D. E., Valacich, J. S., & Featherman, M. 2010. The effect of perceived novelty on the adoption of information technology innovations: A risk/reward perspective. *Decision Sciences*, 41(4): 813–843.