Understanding Artificial Agents as Facilitators of Learning

Understanding Artificial Agents as Facilitators of Learning

Sofia Fountoukidou

This work was financially supported by project MAMEM that has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement number: 644780.

A catalogue record is available from the Eindhoven University of Technology Library.

ISBN: 978-94-6423-419-0
NUR: 964

Keywords: Artificial agent, learning outcomes, behavioral modelling

Cover design: Veronica Galati
Lay out design: Veronica Galati

Printed by: ProefschriftMaken || www.proefschriftmaken.nl

Understanding Artificial Agents as Facilitators of Learning

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr.ir. F.P.T. Baaijens,
voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op vrijdag 1 oktober 2021 om 13:30 uur.

door

Sofia Fountoukidou

geboren te Thessaloniki, Griekenland.

Dit proefschrift van het proefontwerp is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter: prof.dr.ir G.J. van Houtum

1e promotor: prof.dr. C.J.H. Midden

2e promotor: dr. J.R.C. Ham

copromotor(en): dr. U. Matzat

leden: prof.dr. E.O Postma (Universiteit Tilburg)

prof.dr. E. Christiansen (Aalborg University)

prof.dr. P. Markopoulos

Het onderzoek of ontwerp dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

Contents

Chapter 1

General Introduction

Imagine a situation where you are forced to stay and work from home for a long period of time with minimal social interaction because of a viral pandemic. Now, imagine that in order to prevent negative effects of this social isolation you decide to look for a new hobby. Thus, you subscribe to an online piano tutorial, but the moment you log in, you find yourself in a somewhat strange situation; instead of a human teacher, you are presented with an artificial one that waves at you and says: *"Ciao, I am Marco, and I will be your piano teacher"*. Before you know it, Marco starts explaining all the necessary first steps that will make you the next Mozart: how to read musical notes and their corresponding piano keys.

Independent of what the first reaction to your artificial teacher was, one question really matters for you: *is an artificial agent effective for attaining your learning goals?* This question was the starting point of the current thesis.

The existing literature reports mixed findings with regard to the overall effect of artificial pedagogical agents on learning (Castro-Alonso, Wong, Adesope & Paas, 2021; Heidig & Clarebout, 2011; Martha, & Santoso, 2019; Schroeder, Adesope, & Gilbert, 2013). In the introduction of this thesis, I will argue that such mixed effects about artificial agents' effectiveness are found because the question is too broad to receive a simple answer. Therefore, based on this argument, the overall goal of the current thesis is to examine under which conditions and in which ways an artificial agent can facilitate learning. Our fundamental claim is that an artificial agent that acts as a model can successfully do so.

In this chapter, I will start by discussing teaching and learning in both their traditional and modern form, then, I will describe the emergence of artificial pedagogical agents and present important gaps of knowledge in the research field of artificial pedagogical agents. Lastly, I will present the research goals of the current thesis followed by an overview of the next four chapters.

1.1 Teaching and learning

The term education has usually a broad meaning and it refers to both intentional (e.g., formal educational settings) and unintentional (e.g., watching television) teaching. For the current thesis, it is important to consider what kind of endeavour is intentional teaching. At first glance, the answer seems simple enough; one in which knowledge and skills are transmitted to one or more students by a more educated person we use to call teacher. A limitation of such an answer is that it does not take into consideration the concept of learning. In fact, the instructional process always has two parties who are directly involved: the teacher and the student. Within this process, the activity of a teacher is called teaching, and the activity of students is called learning. It is this reciprocal relationship between these two agents that is of paramount pedagogical importance and, despite the fact that they have different roles, it is the activity they perform together that is of interest in teaching (Kansanen, 1999). Indeed, these concepts are represented in a well-accepted definition of teaching by Andersons and Burns (1989, pp. 8), who define teaching as "…an interpersonal, interactive activity, typically involving verbal communication, which is undertaken for the purpose of helping one or more students learn or change the ways in which they can or will behave".

The recognition that the student is an active creator of knowledge induced fundamental changes in the role of the teacher. Traditionally, the teacher was considered the fountain of all knowledge (see e.g., Sequeira, 2012). This suggests a picture of students sitting in rows in front of the teacher who is talking and passing information to students with the aid of a chalkboard, while the students either listen passively or take their own notes. However, in recent times, teaching has been reconceived in light of modern ambitions and more insight into the benefits of student-centered instruction (Cohen, 2011; Keiler, 2018). Thus, the function of the teacher changed from a disseminator of information to that of a facilitator of learning. That is, the teacher's role has become to assist students to learn for themselves (Moustafa, Ben-Zvi-Assaraf & Eshach, 2013).

1.2 The role of the teacher as a facilitator

Given the transformation of the teacher's role from disseminator of information to facilitator of learning, new questions inevitably arose within the educational community. Two important questions, which are also the focal point of the current thesis, pertain to the teacher's instructional method and behavior and their effect on learning.

To start with, one question that arose is which teaching practices, other than the traditional method of learning through books and lectures, can promote learning? Although this is a broad question, important for the current thesis is that earlier research shows that learning is generally more effective when it is based on experience (Kolb, 2014; Raja, 2018). That is, when a learner can experience a concept it is more likely to be learned. Experiencing a concept can happen either directly (e.g., experiencing something yourself) or indirectly (observing someone else experience something).

Observational learning (often called modeling) as an instructional method has received increased attention over the recent decades. Modeling, which is an important element of Bandura's social cognitive theory (Bandura, 1986), refers to the fact that much of our learning derives from vicarious experience. Thus, the concept of modeling refers to learning by observing another person's behavior (the so-called model). One of the principal mechanisms by which modeling influences learning is by stimulating self-efficacy (Bandura, 1977). Self-efficacy is defined as one's beliefs about his/her ability to perform specific tasks. Observing someone else performing a target behavior raises individuals' beliefs about their own ability to perform it successfully (Bandura, 1986). A plethora of studies has confirmed the positive impact of self-efficacy on learning outcomes, such as task performance (i.e., Agarwal, Sambamurthy, & Stair, 2000; Bouffard-Bouchard, 1990; Moos & Azevedo, 2009).

In the educational context, most research on observational learning has involved modeling of psychomotor tasks (i.e., tasks that focus on the development of physical skills, such as performing arts and sports; see Bloom, 1994). This type of modeling has been called behavioral modeling. Nonetheless, a lot of educational research has demonstrated the effectiveness of observational learning for purely cognitive tasks as well (i.e., tasks that focus on the development of cognitive skills, such as mathematics; see Bloom,1994). This alternative form of modeling is called cognitive modeling. Cognitive modeling is based on a process of attending (or "listening") to one's thoughts as one performs an activity and utilizes self-instructional thoughts to guide performance (Wouters, Pass & Merrienboer, 2008).

Another question that arose as a consequence of the transformation of the teacher's role, relates to whether and how teachers' communication skills affect learning (Prozesky, 2000). Communication can be both verbal (involving words and sentences) and nonverbal (e.g., through facial expression, spatial behavior gesture, and nonverbal vocalization). In the current thesis, I will focus on the teachers' nonverbal communication, as it has been found to play a more important role for learning than verbal communication (Mehrabian, 1981; Jones, 2017). Specifically, some forms of teachers' nonverbal behavior have been found to increase "nonverbal immediacy" (Andersen, 1979). The nonverbal immediacy concept refers to the ability of teachers to create psychological closeness to their students through nonverbal communication (Mehrabian, 1981). This concept is grounded in approach-avoidance theory, which asserts that people "are drawn toward the person and things they like, evaluate highly, and prefer; and they avoid or move away from things they dislike, evaluate negatively, or do not prefer" (Mehrabian, 1981, p. 1). Several forms of nonverbal behavior of teachers have been found to play a crucial role in student's learning, such as proximity, eye gaze, gestures, body position, facial and vocal expressiveness (Witt & Wheeless, 2001). Cumulative evidence has revealed that human teachers' nonverbal immediacy behavior promotes affective learning as also cognitive learning (i.e., perceptions of learning and recall) (Ellis, Carmon & Pike, 2016; Witt, Wheeless & Allen, 2004).

1.3 Technology as a medium

One important factor that helped the transformation of the teacher's role as facilitator of learning is the recent advancement of educational technology (Jan, 2017). Due to the increased access to information and educational opportunities (i.e., distance learning) that technology has enabled, in many classrooms today we see the teacher's role shifting to the "guide on the side" as students take more responsibility for their own learning. Schools and universities across the globe are beginning to redesign learning spaces to enable this new model of education, foster more interaction and small group work, and use technology as an enabler.

1.4 An artificial agent as a teacher

Given the transformation of the teacher into the role of a facilitator of learning and the increased availability and sophistication of educational technology in recent decades, researchers started wondering: Can technology become the teacher itself (Mousavinasab et al., 2021)? Such a question led to a new line of research, which examined, amongst other, the use of pedagogical artificial agents in multimedia learning environments. Pedagogical agents are defined as anthropomorph virtual characters to serve various instructional functions (Valetsianos & Miller, 2008). To date, pedagogical artificial agents are becoming more common as facilitators to training in educational settings, private institutions, and the military.

The development of electronic pedagogical agents can be traced back to the 1970s' Intelligent Tutoring Systems (ITS). An ITS exhibits characteristics similar to a human tutor such that it may be able to answer student questions, detect misconceptions, and provide feedback. While the original ITS were abstract entities that focused on tutoring (i.e., SCOLAR tutor system; Woolf, 2010), the next three decades saw advances in agent representation (i.e., visual embodiment) and interactive capabilities (Clarebout, Elen, Johnson & Shaw, 2002). Over the years, ITS evolved into modern virtual characters that encompass complex visual forms, are able to interact with learners using multiple channels of communication (e.g., text, speech, and deictic gestures), and are able to exhibit social skills and intelligence by communicating with users on a broad range of issues that include not just the tutoring topic, but also topics of broader interest.

The vision and role of agents in the learning ecology has also shifted during the last decades. While ITS were initially seen as abstract intelligent systems able to assist learners cognitively (e.g., by posing or answering questions relevant to student tasks), more recently, agents are seen as inherently social artifacts (for a review, see, Heidig & Clarebout, 2011). In addition, the field has expanded its scope in terms of roles that pedagogical agents might play in learning environments. Such roles include tutors, coaches, and actors (Payr, 2003), experts, motivators and mentors (Baylor & Kim, 2005), learning companions (Kim, Baylor, & Shen, 2007); change agents (Kim & Baylor, 2008), and lifelong learning partners (Chou, Chan, & Lin, 2003).

There are multiple reasons to employ pedagogical agents in multimedia learning environments. One important reason is that a pedagogical agent may supplement face-to-face learning by providing a low-cost accessible alternative to a human lecturer. Some of the largest costs and time commitments have gone into video strategies used for Massive Open Online Courses such as those featured on platforms like Coursera (Li, Kizilcec, Bailenson & Ju, 2016). Development costs for such courses vary from $38,000 to $325,000 (Hollands & Tirthali, 2014, p.12) with the largest expenses being videography and the hiring of teaching assistants (Lewin, 2013).

Another reason for the inclusion of a pedagogical agent in online learning environments is to improve consistency with which instruction is delivered. Such consistency might be attained by keeping important agent characteristics consistent over teaching instances, and using pedagogical agents makes this consistency possible. There is an array of such characteristics that can impact the teaching process, and which might be different between human instructors. Examples of these characteristics are instructor experience, confidence, perceived credibility and his/her interaction with the learners and/or learning environment (Swanson & Falkman, 1997). Pedagogical agents' content delivery is predetermined and programmed, making it well suited to address consistency concerns.

1.5 Knowledge gaps: What we need to know about artificial pedagogical agents

Despite artificial agents' vast potential as educational tools, findings regarding their effectiveness for learning are mixed. Meta-analyses revealed that artificial agents were associated with a small but positive effect on learning (Castro-Alonso, Wong, Adesope & Paas, 2021; Schroeder, Adesope, & Gilbert, 2013), while systematic reviews showed that the majority of studies found nonsignificant effects (Heidig & Clarebout, 2011; Martha, & Santoso, 2019). Empirical research on artificial agents generally falls into one of three categories: (1) studies that focus on the simple presence of an agent; (2) studies that focus on appearance or visible features; and (3) studies that focus on their behavior (Lane, 2016). Despite voices echoing that artificial agents may be effective due to their pedagogy rather than merely their appearance (e.g., Moreno, 2005), research on pedagogical behavior is far less common than research on, for instance, the agent's visible features. This is mainly because, pedagogical behavior does not technically require an embodied agent (Lane, 2016). As a consequence, there is a knowledge gap when it comes to the instructional method that an artificial agent applies in the learning environment. Given that human teachers' instructional method has been found to have tremendous impact on students' learning (i.e., Beas & Salanova, 2006), I argue that more research is warranted on artificial agents' instructional method. I further claim that examining the instructional method employed by an artificial agent can help clarifying why many studies found no effects while, to the contrary, several other studies present positive effects.

As discussed above, artificial agents' pedagogical behavior, such as the instructional method they employ, does not technically require the visual presence of an artificial agent. Thus, a point of confusion concerns the debate about artificial agents' visibility in multimedia leaning settings. This debate mainly relates to whether it would be more effective if the instructional design were presented by simpler means of communication rather than by an embodied character (i.e., using simple arrows or color coding of key points rather than having an artificial agent "point" to text or parts of instructional graphics) (i.e., see Choi & Clark, 2006).

More, specifically, two competing perspectives exist in the literature on whether the visibility of an artificial agent in multimedia settings hinders or augments learning. These have been labelled as "agents-as-complements" versus "agents-as-distractors" (Frechette & Moreno, 2010). On the one hand, theories supporting the agents-as complements perspective, like social presence theory (i.e., Hoyt, Blascovich & Swinth, 2003) and social agency theory (i.e., Moreno, Mayer, Spires & Lester, 2001), argue in short that an agent's visual presence increases student motivation, which in turn leads to greater invested effort during learning and more well-formed mental models of the taught concepts.

On the other hand, theories supporting the agents-as distractors perspective, like seductive details (Mayer, 2001) and cognitive load theory (Sweller 2004; Sweller, Ayres, & Kalyuga, 2011), hold that the inclusion of an agent might hinder rather than foster learning. Moreno et al. (2001) termed this "interference" and reasoned that the presence of the agent can hamper learning, because "any additional material that is not necessary to make the lesson interesting reduces effective working-memory capacity and, thereby, interferes with the core material" (p. 186). Overall, these theories predict two types of adverse effects of the visual presence of the agent: cognitive distractions (i.e., inability to pay attention and comprehend learning content) and affective distractions (i.e., disruptive feelings leading to impediment of learning goals) (Frechette & Moreno, 2010). Therefore, according to these theories, an instructional design will be more successful when unnecessary or distracting elements (i.e., artificial agents) are removed from the presentation, thus freeing the learner's cognitive resources to process the content that is most central to learning (Moreno & Mayer, 2000).

Neither of these two sets of theories have been univocally supported by past research. Rather, studies of the effect of an artificial agent on learners' motivation (Carlotto & Jaques, 2016; Chen & Chou, 2015; Dinçer & Doğanay, 2017; Lin, Ginns, Wang & Zhang, 2020; Park, 2015; van der Meij, van der Meij & Harmsen, 2015) as also studies on the effect of an artificial agent on learners' cognitive load (Dinçer & Doğanay, 2017; Frechette & Moreno, 2010; Moreno et al., 2001) show mixed and inconsistent findings.

Overall, the existence of these opposing theories and principles, as well as the contradictory empirical findings, demonstrate another crucial inconsistency that the field suffers from and that restricts insights about the effectiveness of pedagogical agents' ability to facilitate learning in computer-based environments. We still need to find out whether pedagogical agents' visibility is an asset or a limitation for learning. I propose that for solving this second inconsistency we need to have a closer look at the conditions under which an artificial agent's visual presence is relevant and, therefore, essential for goal achievement, or irrelevant and, thus, unnecessary additive. As discussed below, I claim that we need to have a closer look at Bloom's taxonomy (1994) and distinguish between psychomotor tasks (i.e., that focus on psychomotor learning) and cognitive tasks (i.e., that focus on cognitive learning). Overall, I argue that this distinction helps explaining under which conditions pedagogical agents' visibility constitutes a barrier to or a facilitator of learning.

A third gap of knowledge relates to the conditions that increase the effectiveness of an artificial agent for learning. For this, a closer inspection of the behavior of human teachers could be taken into consideration. As sketched above (Section 1.2), in traditional classroom settings with human teachers, various nonverbal forms of teacher behavior have been found to increase immediacy and, subsequently, learning (Ellis, Carmon & Pike, 2016; Witt, Wheeless & Allen, 2004). Visual nonverbal forms of behavior of artificial agents, such as the use of gestures and facial expressions, have received increasing attention over the last years (Baylor & Kim, 2009). However, there is limited evidence on whether and, more importantly, how artificial agents' vocal nonverbal behavior (i.e., vocal expressiveness) can influence learning outcomes. Given that it has been shown that it is mainly the artificial agent's voice that is responsible for increased learning gains rather than its visual presence (Atkinson, 2002; Bente et al., 2008; Krämer & Bente, 2010), I claim that we need to pay more attention to the effect of an artificial agent's vocal expressiveness on learning outcomes and to the underlying mechanisms that explain these effects.

1.6 Research goals

For more than two decades now, researchers could not ascertain whether pedagogical agents can facilitate learning. As described above, findings of past studies often are contradicting. I argue that in light of the great variety of artificial agents used in past studies, as also the specific educational functions of these agents, this issue is too broad to receive a simple answer.

Therefore, it is important to formulate more specific questions. In this spirit, a more fruitful approach would be to ask under which conditions artificial agents can facilitate learning. A crucial condition that been neglected by earlier research is the instructional method an artificial agent applies in the multimedia learning environment (Heidig & Clarebout, 2011; Schroeder & Gotch, 2015).

In more detail, past research suggested that pedagogical agents could make use of different instructional roles, such as modeling/demonstrating, coaching/scaffolding and acting as an information source (see Table 1 for a definition of these methods) (Schroeder & Gotch, 2015). Indeed, Schroeder and Gotch (2015), in their review, found that 36% of the agents provided coaching or scaffolding, while 64% of agents acted as an information delivery vehicle. Despite the fact that modeling has been found to be a very effective instructional method with human models (i.e., Compeau, & Higgins, 1995a, 1995b), no studies were identified using artificial agents as models (Schroeder & Gotch, 2015; Veletsianos & Russell, 2014). To the best of my knowledge, there is still no empirical evidence on whether an artificial agent as a model could facilitate learning.

The current research contributes to answering the question whether an artificial agent could be effective for learning by examining the conditions for such effectiveness. Thus, the overall goal of the current thesis is to examine under which conditions and in which ways an artificial agent could facilitate learning.

Table 1 Agent's instructional methods in a learning environment (taken from Schroeder & Gotch, 2015).

Agent's instructional method	**Description**
Modeling/demonstrating	The agent physically accomplished some task, thereby showing the learner how to successfully complete it (signaling the learner's attention through gaze or gesture is not considered demonstrating or modeling).
Coaching/scaffolding	The agent may have some level of artificial intelligence in order to provide individualized feedback or provides tips that are not part of the instructional material but assist the student in completing the task.
Information source	The agent provides the learning material to the students (i.e., narration).

In order to attain our overall goal, we broke it down into three sub-goals. The **first sub-goal** of the current thesis is to answer the fundamental question of whether modeling by an artificial agent is effective for learning. That is, the type of instructional method of the artificial agent is one fundamental condition to consider.

Further, as discussed above, two competing perspectives exist in the literature on whether the visibility of an artificial agent in multimedia settings hinders or augments learning. ("agents-as-complements" versus "agents-as-distractors"). I argue that it is crucial to investigate this issue further, because in contrast to modeling taking place in classrooms, modeling taking place in multimedia settings can, indeed, occur without the actual visual presence of a model. Therefore, the **second sub-goal** of the current thesis is to examine the conditions under which the visual presence of the artificial model is beneficial for learning. I claim that such a critical condition is the type of learning task being demonstrated by an artificial model. Modeling by human teachers has been found an effective instructional method for both psychomotor tasks and also for purely cognitive tasks (i.e., Gist, 1989; Gorrell & Capron,1990). Nonetheless, I claim that this is not the case for modeling by an artificial agent. Specifically, I hypothesize that an artificial agent's visual presence is beneficial when the learning task to be modelled is psychomotor (i.e., behavioral modeling), because it provides a prototype. However, I further argue that when it comes to modeling of a purely cognitive task (i.e., cognitive modeling) this will be less the case. That is, purely cognitive tasks entail actions that are not readily observable, and therefore I claim that the visual presence of the artificial model is decorational (when modeling such tasks) and therefore less beneficial.

After investigating the fundamental question of whether modeling by an artificial model is effective, and whether the type of learning task is a decisive condition for the inclusion of an artificial agent's visual presence, a **third sub-goal** of the current thesis is to examine the conditions that increase the effectiveness of an artificial agent as a model for learning. Specifically, I argue that the effectiveness of an artificial agent as a model depends on the nonverbal behavior that it appears to exhibit. That is, whether, similar to human teachers, specific forms of nonverbal behavior of an artificial model can create immediacy and, subsequently, increase learning. Furthermore, I examine whether the underlying mechanisms of motivation and attention explain the effect of immediacy on learning. Overall, the third sub-goal focuses on artificial agent's vocal expressiveness (i.e., pitch tone, pitch variation and speech rate) as a powerful form of nonverbal behavior that could strengthen students' learning outcomes.

1.7 Overview of the chapters

The current thesis analyses under what conditions and in which ways artificial agents contribute to enhancing learning outcomes. Overall, six studies organized in three empirical chapters were conducted to shed light on the topic of artificial agents as educational tools in multimedia environments. These studies rest on the idea that artificial agents can act as models in order to improve learning.

Specifically, the research question examined in **Chapter 2** is whether modeling by an artificial agent is an effective instructional method for learning. To answer this question, we conducted two experimental studies, in which modeling by an artificial agent was compared to other commonly used non-modeling instructional methods: a) agent-delivered instructional narration (=agent as a source of information), b) no agent, text-only instruction, and c) no agent, voice-only instructional narration. I expect that modeling by an artificial agent is more effective than the other non-modeling methods in enhancing learning-related beliefs (i.e., self-efficacy and technology ease of-use) as also learning (declarative knowledge and task performance).

Next, the research question examined in **Chapter 3** is whether the positive effects of modeling by an artificial agent on learning depends on the visual presence of the artificial agent. Specifically, the first experimental study of **Chapter 3** aims to examine effects of the interaction between the on-screen visibility of an artificial model (presence vs. absence) and type of task (psychomotor vs. cognitive) on learning outcomes (recall, affective beliefs, and task performance). Further, in the second experimental study of Chapter 3, I aim to extend these findings, by examining whether learners' perceived cognitive load changes depending on the match between the visibility of the artificial model and the type of task. Overall, I posit that the visual presence of the artificial model facilitates learning outcomes and reduces cognitive load for a psychomotor task, but it does not provide any additional learning benefit for a cognitive task.

Then, the research question examined in **Chapter 4** is whether and how an artificial models' nonverbal behavior can increase learning outcomes (affective and cognitive learning). To answer this question, I conducted two experimental studies, in which an artificial model showing strong vocal expressiveness (i.e., higher pitch tone, more pitch variation, higher speech rate) was compared to an artificial model that shows weak vocal expressiveness (i.e., lower pitch tone, less pitch variation, lower speech rate). Overall, I posit that vocal expressiveness of an artificial agent is related to learning outcomes because it promotes immediacy. Furthermore, I expect that immediacy, in turn, increases students' motivation and attention and thereby subsequently facilitates learning.

Lastly, in **Chapter 5,** I discuss the main findings and contribution of the studies presented in this thesis and identify general limitations and directions for future research. In addition, I briefly address ethical considerations of using artificial agents as educational tools. Finally, I present the General Conclusion of the thesis.

Chapter 2

Effects of an artificial agent as a behavioral model on motivational and learning outcomes[1]

[1]This chapter is based on:
Fountoukidou, S., Ham, J., Matzat, U., & Midden, C. (2019). Effects of an artificial agent as a behavioral model on motivational and learning outcomes. *Computers in Human Behavior*, 97, pp.84-93.
Fountoukidou, S., Ham, J., Matzat, U., & Midden, C. (2018). Using an artificial agent as a behavior model to promote assistive technology acceptance. In *Persuasive Technology - 13th International Conference, PERSUASIVE 2018, Proceedings* (pp. 285-296). (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Vol. 10809 LNCS). Springer.

2.1 General Introduction

In recent times, technology has been increasingly used in the pursuit of persuading people towards various beneficial behaviors. Hence, persuasive technologies are defined as the class of technologies, which are intentionally designed to alter or reinforce individuals' psychological attributes, (i.e., attitudes or motivations), which are further presumed to affect behavior, through the use of various persuasive strategies (i.e., persuasive messages and reminders) (Fogg, 2003). Such technologies are regularly used in various contexts, such as marketing, education, and health-related contexts. Their overall potential in influencing psychological attributes and behaviors has been shown excessively in earlier research (Hamari, Koivisto, & Pakkanen, 2014).

Though persuasive technology can take on different functional roles (i.e., as tools and media), it can exert a strong influence when it takes the form of a social actor (Ham & Midden, 2014). Artificial social agents -often on-screen animated characters-have been the target of increasing interest due to their ability to simulate social interaction. Artificial agents have been employed in various settings (Dehn & Van Mulken, 2000), with their anticipated positive impact to be attributed to their ability to provide social cues (Reeves & Nass, 1996).

Recently, a more specific line of research in the domain of education and learning drew attention to the role of embedded artificial agents (so-called pedagogical agents) in multimedia learning. Their inclusion represents an attempt to introduce more instructional support and persuasive elements in learning settings (Clark & Choi, 2005). However, contrary to the purpose they are designed to serve, the general motivating or learning facilitating effect of the embedded agents has been questioned. A recent literature review drew a discouraging picture concerning the overall advantage of artificial agents for learning (Heidig & Clarebout, 2011). This is, the majority of the reviewed experiments (39 in total) yielded non-significant results on learning out-comes when comparing the scores of participants in agent and no-agent groups. Motivational outcomes (i.e., self-efficacy) likewise revealed non-significant differences (i.e., Baylor & Ryu, 2003; Domagk, 2010). Hence, it has been proposed that further research is needed to examine conditions under which an artificial agent could facilitate learning, such as its instructional method of teaching. I claim that examining the agent's instructional method helps to clarify under which conditions agents yield better learning outcomes.

Undoubtedly, a human teacher's instructional method has a serious impact on students' motivation and learning outcomes (Beas & Salanova, 2006). Behavioral modeling has been found to be a powerful method of education across a diverse range of behavioral domains. Behavioral modeling, originated from Bandura's social cognitive theory, posits that much of our learning derives from vicarious experience and advocates the concept of modeling or learning by observing another person's behavior (the so-called model) (Bandura, 1986). Using agents for behavioral modeling means the employment of an agent that verbally explains and physically accomplishes a task, thereby showing the learner how to successfully complete it.

It has been proposed that artificial agents could serve as an alternative source for persuasive human behavioral models, because of their capability to show complex tasks, for example by means of gestures and expressions (Baylor & Plant, 2005; Johnson, Rickel, & Lester, 2000; Kim, Baylor, & Shen, 2007; Krämer & Bente, 2010; Plant, Baylor, Doerr, & Rosenberg-Kima, 2009). In fact, evidence suggests that gestures performed by an agent in multimedia learning environments influence learning outcomes (i.e., transfer, retention) (Davis, 2018; Johnson, Ozogul, & Reisslein, 2015). In addition to this, Wang, Li, Mayer, and Liu (2018) showed that a gesturing agent positively affects learners' cognitive processing during learning, as measured with a number of eye-tracking measures (i.e., number of fixations). Moreover, related work provided support for the signaling hypothesis, that the presence of an agent facilitates learning only when used to signal relevant on-screen visual information (Johnson, Ozogul, Moreno, & Reisslein, 2013; Moreno, Reislein, & Ozogul, 2010). Despite such encouraging findings, the effect of an artificial agent in the role of a behavioral model on learning has not yet been studied. Therefore, the main objective of the current work is to investigate the effects of an artificial agent employed as a behavioral model on individuals' motivational (i.e., self-efficacy) and learning outcomes (knowledge, task performance).

One of the principal mechanisms by which human behavioral modeling influences learning is stimulating self-efficacy (Bandura, 1977). Self-efficacy is defined as one's beliefs about his/her ability to perform specific tasks. Observing someone else performing the target behavior raises individuals' beliefs about their own ability to perform it successfully (Bandura, 1986). A plethora of studies have confirmed the positive impact of self-efficacy on learning outcomes, such as task performance (i.e., Agarwal, Sambamurthy, & Stair, 2000; Bouffard-Bouchard, 1990; Moos & Azevedo, 2009).

In the domain of technological innovation adoption, the focal do-main of this research, computer self-efficacy has been defined as one's belief about own ability to perform a specific computer activity (Compeau, & Higgins, 1995a, 1995b). Research on behavioral modeling in computer training highlighted the key role of computer self-efficacy in determining computer skill acquisition (Gist, 1989). Furthermore, these studies demonstrated that behavioral modeling yields higher scores of computer self-efficacy compared to other commonly used non-modeling methods (i.e., lecture-based instruction, self-manual).

Computer self-efficacy has also been found to be an important predictor of individuals' subjective evaluation of a system. The Technology Acceptance Model 3 (TAM 3), one of the widely used theoretical models examining individual reactions to computing technology, posits that computer self-efficacy is one of the determinants of system-specific perceived ease of use (Venkatesh & Bala, 2008). According to TAM, system's perceived ease of use, defined as the degree to which the prospective user expects the target system to be free of effort, is one of the two drivers of individuals' intention to use a system (with the other being perceived usefulness). Venkatesh and Davis (1996) found that general computer self-efficacy had a direct effect on ease of use perceptions, both before and after hands-on experience with the software (Venkatesh & Davis, 1996). Thus, training has been suggested as an intervention for increasing users' system perceptions of ease use by influencing its determinants (i.e., computer self-efficacy).

What is more, an essential learning construct is declarative knowledge. Declarative knowledge is defined by Anderson (1985, p199) as "knowledge about facts and things". Yi and Davis (2003) developed a model showing, amongst other, that behavioral modeling significantly influences declarative knowledge via the four component processes (attention, retention, production, motivation). Similarly, Szymanski (1988) suggested that since knowledge organization in long-term memory depends greatly on knowledge-related experiences, verbal instructions (i.e., lectures) are not likely to be effective (Szymanski, 1988). Instead, the author proposed behavioral modeling as an approach to developing accurate declarative knowledge structures.

Nonetheless, the most essential objective of learning is the actual development of skills. Yi and Davis (2003) showed that self-efficacy and declarative knowledge are two distinct causal pathways by which behavioral modeling influences behavioral performance (i.e., procedural knowledge). Earlier research provided some evidence of a correlation between declarative knowledge and task performance. Concerning the impact of computer self-efficacy on task performance, numerous studies have reported significant empirical relationships (i.e., Johnson & Marakas, 2000; Martocchio & Judge, 1997).

2.1.1 The current work

Earlier research provides inconsistent evidence for effects of pedagogical agents on learning outcomes (i.e., Heidig & Clarebout, 2011). We suggest that examining the agent's instructional method helps to clarify under which conditions agents influence motivational and learning outcomes. In Chapter 2, we report about two studies. The purpose of Study 1 was to investigate the effect of an agent as a behavioral model, compared to two common non-modeling instructional methods (agent-delivered instructional narration and no-agent, text-only instruction), on learners' beliefs of their computer-self efficacy and system's perceived ease of use. Building on the results of Study 1, the purpose of Study 2 was to extend the insights into the effects of agent-delivered modeling, as compared to other non-modeling methods, by focusing on learners' cognition (i.e., declarative knowledge) and behavior (i.e., task performance). Further, in Study 2, we substituted the no-agent, text-only condition with a no-agent, voice-only narration condition. Moreover, we replicated effects of agent-delivered modeling on individuals' self-efficacy and perceived ease of use, so as to further strengthen the findings of Study 1. Thus, the combination of both Study 1 and Study 2 has an additional replication value.

2.2 Study 1

One of the principal mechanisms by which human behavioral modeling influences learning is stimulating self-efficacy, which, in turn, has been found to be a predictor of individuals' system's perceived ease of use (Venkatesh & Bala, 2008). In Study 1, we predicted that agent-delivered modeling as an instructional method will positively influence users' computer self-efficacy (H1) and perceived ease of use of a novel technology (H2). To test these hypotheses, we compared agent-delivered modeling to two non-modeling instructional methods (i.e., agent-delivered instructional narration and no-agent, text-only instruction). Taking into account the findings and recommendations of earlier studies (Agarwal et al., 2000; Bandura, 1997; Marakas, Yi, & Johnson, 1998), we examined the impact of agent-delivered modeling on specific computer-self efficacy (perceptions of ability to perform specific computer-related tasks), as opposed to general computer self-efficacy (judgment of efficacy across multiple computer application domains).

2.2.1 Method

Participants and design

A total of 197 individuals participated in the study. The participants were recruited using a local participant database, and most of them were students from Eindhoven University of Technology. Of these participants, 122 (61.9%) were males and 74 (37.6%) of them were females (one person did not answer the question about gender). The age of the sample ranged from 19 to 29, with a mean age of 23 (SD = 2.44). One-hundred fifteen participants were educated to undergraduate level or higher, and 77 had completed high school (5 persons did not state their educational background). The vast majority of the participants (95.5%) reported using computers on a daily basis, with a computer use frequency for more than 12 h per week (82.5%). The average general computer self-efficacy of the population was high (M = 5.51 on a scale from 1 to 7, SD = 0.74), which is concordant with the participants' stated extensive computer use. In addition, more than half of the participants (63.5%) reported no previous experience with using assistive computer technologies (i.e., software and/or hardware).

The study employed a between-subjects design, with the participants being randomly assigned to one of three experimental conditions: agent-delivered modeling, agent-delivered instructional narration, and no-agent, text-only instruction. We pre-tested the success of our manipulation, by randomly allocating 10 participants to one of these conditions and afterwards asking them about their perceptions of the type of the instructional method they received (i.e., demonstration, narration, or textual instruction). Results indicated that our manipulations worked as expected. The study's dependent variables were specific computer self-efficacy and perceived ease of use. Inclusion criteria were participants' fluency in English. Overall, the duration of the study was approximately 20 min, for which participants received 5€ as compensation for their participation.

Apparatus

The study's instructional material pertained to an eye-tracking software, called GazeTheWeb

(GTW). GTW, illustrated in Figure 1, is a web-browser, developed to be controlled solely with the eyes, using eye-tracking hardware (for more information see Menges, Kumar, Müller, & Sengupta, 2017). Participants were unfamiliar with the system.

The 3D animated artificial agent used in this study was created using the CrazyTalk 8 software. The agent was designed to resemble participants' characteristics in terms of appearance, according to the guidelines derived from the earlier literature (Baylor & Plant, 2005; Plant et al., 2009; Rosenberg-Kima, Baylor, Plant, & Doerr, 2008). Since the participants of this study were young individuals, the agent was designed to be young (~25 years), attractive (as manipulated by the agent's facial features) and "cool" (as manipulated by the agent's clothing and hairstyle).

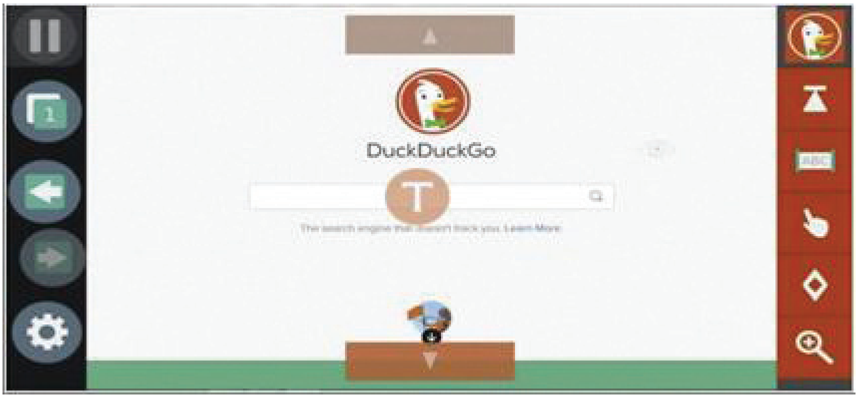


Figure 1 The homepage of the GazeTheWeb interface for performing the required web search.

Measures

Specific computer self-efficacy was assessed by asking participants to answer five questions regarding their perceived ability to perform the necessary steps of the instructed computer task, using GTW. These questions were self-constructed (see Appendix A.1 for the items on this scale). Specifically, to develop measures for specific computer self-efficacy, recommendations for earlier work on question construction for this construct were closely followed (Bandura, 1997; Marakas et al., 1988). Participants answered these five questions on a 7-point rating scale, ranging from 1 (strongly disagree) to 7 (strongly agree). We obtained a reliable measure (Cronbach's $\alpha = 0.80$) of specific computer self-efficacy by averaging participants' answers to this set of questions.

General computer self-efficacy was assessed by asking participants to answer eight questions regarding their perceived ability to use un-familiar computer technologies in general. This 8-item scale was originally created by Compeau and Higgins (1995b) (see Appendix A.1 for the items on this scale). Participants answered these questions on a 7-point rating scale, ranging from 1 (strongly disagree) to 7 (strongly agree). We obtained a reliable measure of general computer self-efficacy (Cronbach's $\alpha = 0.75$) by averaging participants' answers to this set of questions.

System's perceived ease of use was assessed by asking participants to answer four questions regarding their personal evaluation of the mental effort that is needed to use GTW. This 4-item scale was originally created by Davis (1989, 1993) (see Appendix A.1 for the items on this scale). Participants answered these questions on a 7-point rating scale, ranging from 1 (strongly disagree) to 7 (strongly agree). We obtained a reliable measure of perceived ease of use (Cronbach's $\alpha = 0.81$) by averaging participants' answers to this set of questions.

Lastly, demographic questions of age, gender, education, and level of computer use were asked.

Procedure

Participants were asked to read and sign an informed consent form, stating the general purpose of the research and their willingness to participate in this study. Then, participants were randomly assigned to one of the three outlined experimental conditions and they were asked to watch an instructional video (split into two screens) on how to perform a web search with their eyes using the GTW browser. It was while the participants watched the video that the manipulation of agent-delivered modeling took place. Figure 2 shows sample screenshots for each condition.

In more detail, the video in the agent-delivered modeling condition was split into the following two screens: on the right-hand side, an artificial agent appeared to use the GTW system by moving its eyes and head, so as to conduct a web search, while verbally explaining the task-related features of the system; the left-hand side of the screen contained a display of the system, exposing participants to the progressive effects of the agent's actions (i.e., produced by its gaze) in real time. Overall, the agent did not appear to perform any other movements (i.e., hand gestures), other than using its eyes and head to perform the demonstration of the eye-tracking software.

The video in the agent-delivered instructional narration condition was split into the following two screens: on the right-hand side, the (same) artificial agent appeared to provide (the same) verbal instructions on how to conduct a web search using GTW, explaining the task-related features of the system; while the left-hand side of the screen contained a display of the system, exposing participants to progressive screenshots of the system with labels, highlighting the commands the verbal explanation was referring each time. The agent appeared to be still (i.e., no head movements or hand gestures), although 'its human-like gaze behavior remained (i.e., blinking).

Lastly, in the no-agent, text-only instruction condition, the right-hand side of the screen contained a textbox displaying written instructions. Thus, participants in this condition were provided with the same system instructions, but they could not see or listen to the agent. The left-hand side of the screen was identical to the agent-delivered instructional narration condition (i.e. labels highlighting the system's commands).

After the end of the instructional videos, participants were re-quested to answer an online questionnaire. Lastly, they were debriefed, paid and thanked for their contribution.

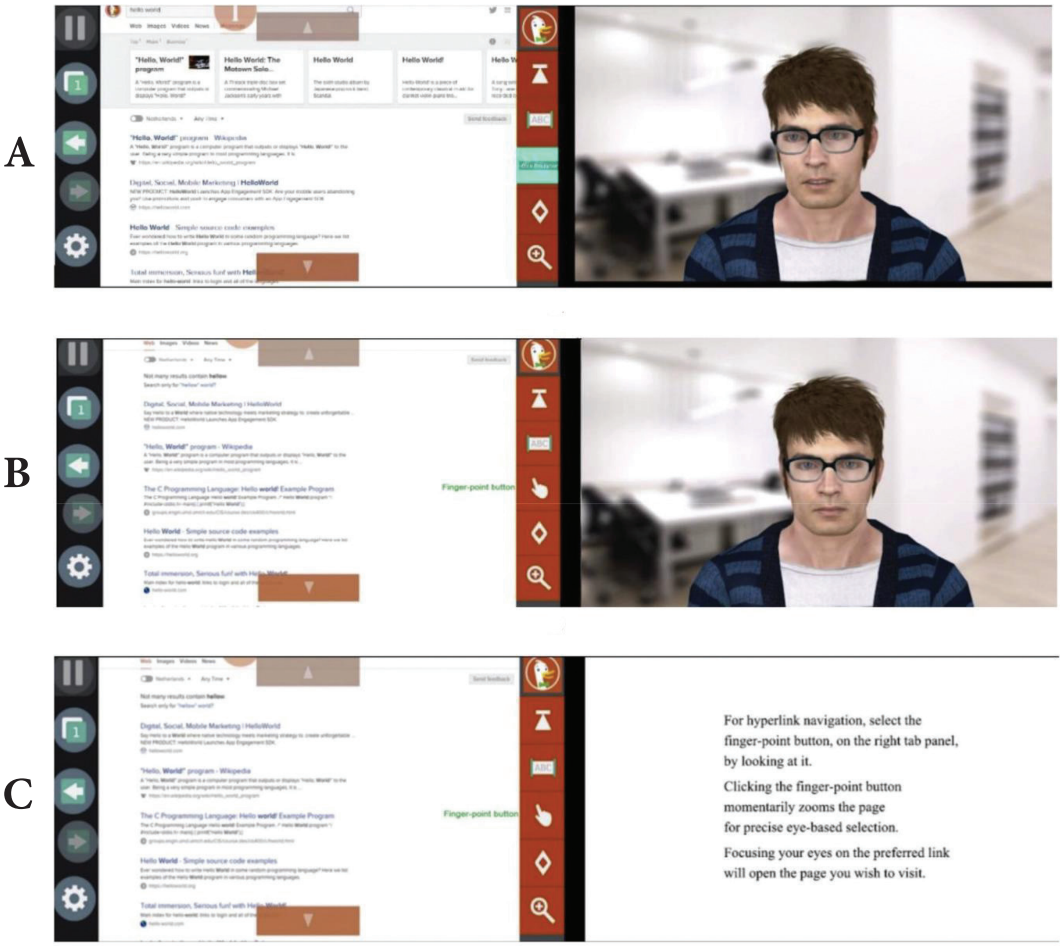


Figure 2 Different types of instructional methods: (a) Agent-delivered modeling; the agent tilts the head to focus its gaze to the system feature, which, as a result of this action, becomes activated (blue button on the left-hand side) (b) Agent- delivered instructional narration; the agent is motionless while explaining the system feature, which is high-lighted in the left-hand side screenshot (c) No-agent, text-only instruction; the agent has been substituted by a text-box, which provides instructions of the function of the system feature, highlighted in the left-hand side screenshot.

2.2.2 Results

Specific computer self-efficacy: A one-way analysis of covariance (ANCOVA) was conducted to determine the effect of the type of instructional method on participants' specific computer self-efficacy, after controlling for their general computer self-efficacy[2] .

After controlling for general computer self-efficacy, the significant main effect of the type of instruction on specific computer self-efficacy remained, $F(2,193) = 6.832$, $p < .01$, $\eta p^2 = 0.066$. Planned contrasts revealed that specific self-efficacy was significantly higher for the participants in the agent-delivered modeling condition ($N = 66$, $M = 6.1$, $SD = 0.80$), as compared to the participants in the agent-de-livered instructional narration condition ($N = 66$, $M = 5.6$, $SD = 0.92$), $t(193) = -3.48$, $p < .01$, and as compared to participants in the text-only instruction condition ($N = 65$, $M = 5.7$, $SD = 0.93$), $t(193) = -2.82$, $p < .01$. No significant difference was found between participants in the two non-modeling conditions after controlling for general self-efficacy.

Perceived ease of use: A one-way ANCOVA was conducted to determine the effect of the type of instructional method on participants' perceived ease of use, after controlling for their general computer self-efficacy[3]. Results demonstrated a significant relationship between the covariate general computer self-efficacy and perceived ease of use, $F(1,193) = 27.203$, $p < .01$, $\eta p^2 = 0.124$. More important, the findings revealed a marginally significant main effect of the type of instruction on perceived ease of use after controlling for the general computer self-efficacy, $F(2, 193) = 2.882$, $p = .058$, $\eta p^2 = 0.029$. Planned contrasts revealed that perceived ease of use was significantly higher for participants in the agent-delivered modeling condition ($N = 66$, $M = 4.78$, $SD = 1.04$), as compared to participants in the agent-delivered instructional narration condition ($N = 66$, $M = 4.36$, $SD = 0.99$), $t(193) = -2.34$, $p < .05$. Nonetheless, results showed no evidence for a significant difference in perceived ease of use, between participants in the agent-delivered modeling condition and participants in the text-only instruction condition ($N = 65$, $M = 4.54$, $SD = 1.14$), $t(193) = -1.34$, $p > .05$. Similarly, no significant difference was found between participants in the two non-modeling conditions after controlling for general self-efficacy.

2.2.3 Discussion

The results of the current study supported our first hypothesis showing that participants in the agent-delivered modeling condition reported higher computer self-efficacy, as compared to participants in the two non-modeling conditions. The results, by and large, support our claim that the agent's instructional method is key for finding out under what con-

[2] When we did not include the general self-efficacy covariate in the analysis (ANOVA), the results were comparable and in line with our first hypothesis, $F(2,194) = 5.10$, $p = .007$, $\eta^2 = 0.05$.

[3] When we did not include the general self-efficacy covariate in the analysis (i.e., ANOVA), the results were compa rable and partially supported our second hypothesis, $F(2,194) = 2.23$, $p = .11$.

ditions pedagogical agents yield better self-efficacy beliefs. Also, the results are in line with past research on the effect of behavioral modeling (conducted by a human agent) on users' computer self-efficacy, as compared to other non-modeling methods (i.e., lecture training and self-manual) (Compeau & Higgins, 1995a, 1995b; Gist, 1989). These findings suggest that an artificial agent functioning as a behavioral model can be implemented in a digital learning setting as an alternative solution to a human agent, enhancing learners' belief in their own capabilities to perform a modeled task.

Additionally, we found that participants in the agent-delivered modeling condition had higher perceptions of ease of use of the system, compared to participants in the agent-delivered instructional narration condition, also when controlling for their general computer self-efficacy. However, contrary to our hypothesis, no differences on perceived ease of use were found between participants in the agent-delivered modeling and the no-agent, text-only condition. Therefore, our second hypothesis was only partially supported.

This finding could be explained when taking into consideration the role of direct experience with a system in the formation of individuals' perceptions of ease of use. According to Venkatesh (2000), the more concrete a person's behavioral experience with a system is, the more accurate his/her judgments are regarding the ease of use of the system. However, prior to any system experience, users anchor their system perceptions of ease of use to their more abstract beliefs they have about technologies (Venkatesh & Davis, 1996). Behavioral modeling is a form of vicarious experience that can serve as an anchor point for such abstract beliefs, leading to the formation of more accurate perceptions of system-specific ease of use. Thus, it might be that, since participants in the text-only instruction condition were provided with less concrete system experience, they mostly relied on their general schema about technologies, which, judging from their reported extensive experience with technologies, was positive. Although in the current study we controlled for participants' general computer self-efficacy beliefs, other known anchors of perceived ease of use were not measured. (i.e., computer anxiety, computer playfulness).

Besides the advantage of behavioral modeling over other instructional methods (i.e., textual instruction) to provide vicarious learning experience, an additional factor that might have contributed to the possible difference in system experience between participants the agent-delivered modeling condition and the text-only instruction condition pertain to the methodological design, and specifically to the different modalities used (oral vs. written instructions). The split-attention principle states that separately presenting mutually referring written text and pictures (i.e., such as in the text-only instruction condition) requires learners to split their attention between both the information sources, as well as to mentally integrate them (Ayres & Sweller, 2005; Sweller, Ayres, & Kalyuga, 2011). Therefore, participants who received written instruction might have faced with more challenges in processing and integration of information, resulting in acquiring less system experience, and relieving even more to their abstract technology-related beliefs when evaluating the specific system's ease of use.

2.3 Study 2

In Study 1, we provided evidence for the effect of the artificial agent functioning as a behavioral model on users' beliefs of their computer self-efficacy and the system's perceived ease of use. However, learning is a multidimensional process that also encompasses cognitive (i.e., declarative knowledge), and skill-based (i.e., performance) constructs (Kanfer & Ackerman, 1989; Kraiger, Ford, & Salas, 1993). Thus, as a next step, we designed a follow-up study in order to extend our insights into the effect of an artificial agent functioning as a behavioral model on those learning outcomes. We argue that an artificial agent that takes on the instructional role of a behavioral model can improve users' learning, both at knowledge and at performance level, as compared to other non-modeling instructional methods.

In more detail, in Study 2 we predicted that agent-delivered behavioral modeling would enhance users' declarative knowledge (H1) and also their task performance (H2) when compared to two non-modeling instructional methods (agent-delivered instructional narration and no-agent, voice-only instructional narration). We also re-tested effects of agent-delivered modeling on individuals' self-efficacy (H3) and perceived ease of use (H4), so as to further strengthen findings of Study 1.

In Study 2, we substituted the written instructions (no-agent, text-only instruction condition) that we used in Study 1 with spoken instructions (i.e., no-agent, voice-only instructional narration). This is, to avoid any effect of the type of presentation of instruction (written vs oral) on participants' learning.

Overall, the present study aimed to add to the existing research on artificial agents by distinguishing the independent effects of agent-de-livered modeling on multidimensional outcomes of learning. Particularly, the study's goal is to contribute to the existing empirical evidence, by showing the facilitating effect of an agent not only on individuals' motivational constructs (self-efficacy), but also on cognition and, most importantly, on behavior.

2.3.1 Method

Participants and design

The experiment was divided into two parts. A total of 99 participants completed the first part (instructional videos), while 7 participants were withdrawn from the second part (task performance with the system) by the experimenter, due to technical difficulties they faced in using the system (i.e., system calibration failure). Nonetheless, participant withdrawal did not induce major changes to the initial sample's demographic characteristics. The participants were recruited from a local participant database, and most of them were students from Eindhoven University of Technology. Sixty-two of the participants were males and 37 of them were females. The age of the sample ranged from 19 to 55, with a mean age of 25.4 (*SD* = 6.29). Sixty-three participants were educated to undergraduate level or higher, and 35 have completed high school (one case did not report educational background). The vast majority of the participants reported using computers on a daily basis (98%), with a computer use frequency for more than 12 h per week (76.8%). The average general computer self-efficacy of the population was high (M = 5.51, SD = 0.73), which is concordant with the participants' extensive computer use. Nonetheless, more than half of the participants (58.6%) reported no previous experience with using assistive computer technologies (i.e., software and/or hardware).

The study employed a between-subjects design, with the participants being randomly assigned to one of the three experimental conditions: agent-delivered modeling, agent-delivered instructional narration, and no-agent, voice-only instructional narration. The success of the manipulation was checked with a single item measure: "During the instructional video, I was able to observe a trainer using GazeTheWeb to perform a web search step by step". Participants could answer this question on a 7-point rating scale, ranging from 1 (strongly disagree) to 7 (strongly agree). Results provided support for our manipulation, confirming that participants in the agent-delivered modeling condition agreed significantly more receiving an agent-delivered task demonstration (N = 34, M = 6.44, SD = 1.1), as compared to participants in the agent-delivered instructional narration condition (N = 34, M = 5.62, SD = 1.5), $t(96) = 2.341$, $p = .02$, and as compared to participants in the voice-only, instructional narration condition (N = 31, M = 5.13, SD = 1.6), $t(96) = 3.643$, $p < .001$. The study's dependent variables were declarative knowledge, task performance, specific computer self-efficacy beliefs, and ease of use perceptions. Inclusion criteria were fluency in English and exclusion criteria pertained to participation in Study 1. Overall, the duration of the study was approximately 30 min, for which participants received 5€ as compensation for their participation.

Apparatus

The software system used in the current study was the GazeTheWeb software (GTW). The eye-tracking hardware used was the Tobii EyeX.
A "user booklet" was created in a paper form and was given to the participants just before the second part of the study (actual system use). This booklet included: 1) information about the task assignment (i.e., "perform the same task as the one you saw in the instructional videos"), as well as task instructions (i.e., "perform the task as fast and as accurately as possible") (see Appendix A.2).

Measures

The dependent variable of the first hypothesis was declarative knowledge and it was assessed by asking participants to answer a series of multiple-choice questions, the most commonly used method for knowledge assessment (i.e., Kraiger et al., 1993; Yi & Davis, 2003). The multiple-choice questions for declarative knowledge were defined as items that assess 'pure recall' of specific isolated pieces of knowledge, such as facts, definitions, terminologies, and concepts (Abu-Zaid & Khan, 2013). The multiple-choice questions of the current study were constructed to measure participants' memory of the GTW system's commands, icons, and layout, as described in the instructional videos. Participants were requested to answer 15 multiple-choice questions with four choices each, of which one answer was correct (see Appendix A.2 for the multiple-choice questions). The internal consistency of the multiple-choice items was acceptable (α (KR-20) = 0.70). We constructed a declarative knowledge measure by counting participants' number of correct answers to these 15 questions.

The dependent variable of the second hypothesis was participants' task performance using GTW. Participants had to perform four sub-tasks with GTW (see Appendix A.2 for details on the task performance assignment). Task performance was assessed using three performance indicators: 1) overall task completeness; 2) task completion time (i.e. speed); and 3) task accuracy (i.e., the number of errors). These performance measurements were in agreement with performance measurements suggested by earlier literature (Förster, Higgins, & Bianco, 2003; Sweeney, Maguire, & Shackel, 1993).

To ensure the calculation accuracy of the performance measurements, use cases, descriptions of all the required interactions of users with the system, were generated prior to conducting the experiment. Thus, following the guidelines from earlier literature (Schneider & Winters, 2001), we created use cases that defined: 1) the main paths (cases where a user successfully performs a task, according to the basic course of action presented in the instructional videos); 2) alternative paths (cases different from the basic paths, which still lead to a task completeness); 3) exception paths (unintended paths through the GTW system, due to either participant's missing information or system us-ability problems). To increase the precision of our use case methodology, the overall computer task was divided into four subtasks, with this task segmentation to be based on the sequential content of the instructional videos. Thus, separate use cases were generated for each of the four subtasks.

Overall task completeness was calculated as a percentage, after summing up the number of the subtasks that were successfully completed by a participant (with 100% success being the completion of all four sub-tasks). Task completion time (measured in milliseconds), was calculated by subtracting the start-time from the end-time of each of the four subtasks. We constructed a measure of task completion time by averaging a participant's completion time for these four subtasks. Task accuracy was calculated by adding the number of errors made during each subtask (i.e., pre-defined in the exception paths). All types of errors were given the same weight (i.e., 1). We constructed a measure of task accuracy by averaging a participant's number of errors for the subtasks. The performance scores were calculated by two researchers independently, who were both blinded to the experimental conditions. There was a 100% agreement on the performance measurements between the two raters.

The two dependent variables of the third and fourth hypothesis were specific computer-self-efficacy and perceived ease of use. We used the same measures for these variables as in Study 1 (see Appendix A.1 for the items on these scales). We obtained a reliable measure of specific computer self-efficacy (Cronbach's α = 0.86), and perceived ease of use (Cronbach's α = 0.84), by averaging participants' answers to each set of questions. General computer self-efficacy was also assessed in the same manner as in Study 1 (see Appendix A.1 for the items on this scale). We obtained a reliable measure of general computer self-efficacy (Cronbach's $\alpha = 0.79$) by averaging participants' answers to this set of questions.

Lastly, participants were also requested to answer a series of questions about their gender, age, education level, and frequency of computer use and any previous AT use.

Procedure

Participants were asked to read and sign an informed consent form, stating the general purpose of the research and their willingness to participate in this study. The experiment was divided into two parts. In the first part, participants were randomly assigned to one of the three experimental conditions and they were asked to watch an instructional video (split into two screens) on how to perform a web search with their eyes using the GTW browser. The first part of the current study was identical to that of our earlier research (see 2.2.1), with the only diference to be in one of the three experimental conditions. This is, the text-only narration condition was substituted with the voice-only, instructional narration condition. In more detail, the voice-only instructional narration condition was identical to the agent-delivered instructional narration condition, with the only difference being that the right-hand side of the screen depicted a neutral background of an office, instead of the artificial agent (see Figure 3). Thus, participants in this condition were provided with the same verbal instructions as in the other two conditions, but the physical appearance of the agent on-screen was removed.

After the end of the instructional videos, participants were requested to answer an online questionnaire and to complete the multiple-choice test.

Next, the experiment proceeded to the second part, where participants were requestedto perform a web search using the GTW system. Instructions were provided to all participants through a user booklet before they started the task. This is, each participant was requested to search for the "hello world" Wikipedia page, with an overall task mission to complete an unfinished sentence on the user booklet by copying the two words, found in the last paragraph of the correspondent Wikipedia web page. Overall, this computer task required participants to perform four sequential computer subtasks, namely, 1) web search initiation 2) typing of search term 3) hyperlink navigation, 4) page navigation. Next, the experiment leader assisted each participant with the eye-tracker calibration (i.e., the default calibration process of the specific eye-tracker model has been followed). To ensure the accuracy of the calibration and to avoid initial task performance errors due to participants' unfamiliarity with eye-tracking technologies, participants were asked to play an eye-tracking game for a brief amount of time, with the experiment leader being present. Afterwards, participants were asked to perform the task as fast and as accurate as possible, but without having a specified time limit and in absence of the experiment leader. After completing the task, participants were debriefed, paid and thanked for their contribution.

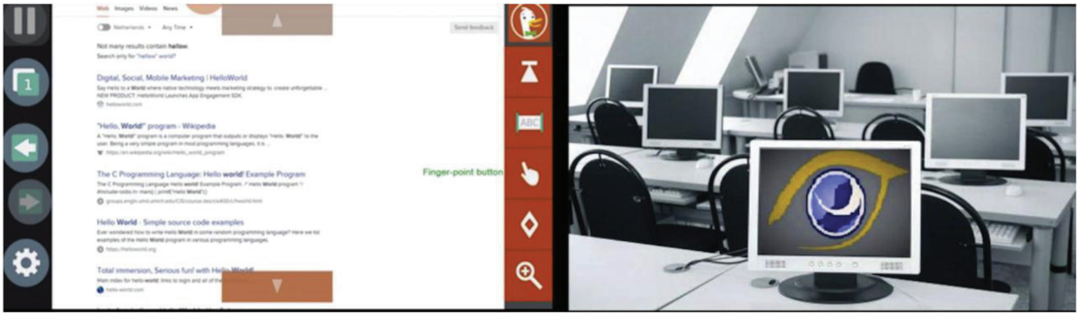


Figure 3 Voice-only instructional narration: On the left side, labels highlight a specific command (i.e., finger-point button), while on the right side, a neutral background was presented. A voice provided participants with information related to the shown feature.

2.3.2 Results

Declarative knowledge: A one-way ANOVA was conducted to investigate the effect of the type of instructional method on participants' declarative knowledge. We found a significant main effect of the type of instructional method on participants' declarative knowledge, F(2,96 = 3.025, p = .05, $\eta_p^2 = 0.060$. Planned contrast analysis only partially supported our first hypothesis, indicating that participants in the agent-delivered modeling condition (N = 34, M = .83, SD = 0.33) scored significantly higher compared to participants in the agent-delivered instructional narration condition (N = 34, M = 0.73, SD = 0.28), $t(96) = -2.458$, p = .016, r = 0.24. To the contrary, results provided no evidence for a significant difference on participants' declarative knowledge between the agent-delivered modeling condition and the voice-only, instructional narration condition, (N = 31, M = 0.77, SD = 0.14), $t(96) = 1.263$, p = .21, r = 0.13.

Task performance: A one-way multivariate analysis of variance (MANOVA) was conducted to investigate the effect of the type of instructional method on participants' task performance with the system. Three measures of participants' task performance were analyzed: overall task completeness, mean completion time and mean number of errors. Due to the fact that all participants included in the analysis had successfully completed the overall task, only the other two performance measures were analyzed (completion time and number of errors). Using Pillai's trace, we found a statistically significant MANOVA effect of the type of instructional method on the combined dependent variables, V = 0.11, $F(4,178) = 2.767$, p = .029, $\eta_p^2 = 0.059$. Follow-up univariate ANOVAs showed that both participants' mean completion time ($F(2, 89) = 5.889$, $p = .004$, $\eta_p^2 = 0.117$) and mean number of errors ($F(2,89) = 3.896$, $p = .024$, $\eta_p^2 = 0.080$) were significantly different between the three experimental conditions. In line with our second hypothesis, simple contrasts revealed that participants in the agent-delivered modeling condition ($N = 32$, $M = 18.6$, $SD = 6.90$) completed the task significantly faster, compared to participants in both, the agent-delivered instructional narration condition ($N = 31$, $M = 35, 6$, $SD = 26.7$), $p = .002$, and the voice-only, instructional narration condition ($N = 29$, $M = 32.4$, $SD = 24.1$), $p = .01$ Similarly, participants in the agent-delivered modeling condition ($N = 32$, $M = 0.375$, $SD = 0.39$) performed the task with significantly fewer errors, as compared to participants, in both, the agent-delivered instructional narration condition ($N = 31$, $M = 0.831$, $SD = 0.90$), p = .01, and the voice-only, instructional narration condition, ($N = 29$, $M = 0.759$, $SD = 0.70$), p = .03. Results provided no evidence for a significant difference in the two measures of task performance between the participants in the two non-modeling conditions.

Specific computer self-efficacy. Further, a one-way between-subjects analysis of covariance (ANCOVA) was conducted to investigate the effect of the type of instructional method on participants' specific computer self-efficacy, after controlling for their general computer self-efficacy[4]. One extreme outlier was found in the data, as assessed by inspection of a boxplot. Therefore, we reported our analysis without that one participant. Results showed a significant relationship between the covariate general computer self-efficacy and specific computer self-efficacy, $F(1,94) = 15.496$, $p < .001$, $\eta_p^2 = 0.142$. After controlling for general computer self-efficacy, the significant main effect of the type of instructional method on specific computer self-efficacy remained, $F(2,94) = 5.538$, $p = .005$, $\eta_p^2 = 0.105$. Planned contrast analysis revealed that the participants in the agent-delivered modeling condition ($N = 33$, $M = 6.38$, $SD = 0.49$) exhibited significantly higher specific computer self-efficacy, as compared to participants in the agent-delivered instructional narration condition ($N = 34$, $M = 5.82$, $SD = 0.88$), $t(94) = -2.521$, $p = .002$, and, as compared to the participants in the voice-only, instructional narration condition ($N = 31$, $M = 5.78$, $SD = 0.92$), $t(94) = -3.136$, $p = .013$. Results provided no evidence for a significant difference in participants' computer self-efficacy between the two non-modeling conditions also when controlling for their general computer self-efficacy.

Perceived Ease of use. A one-way ANCOVA was conducted to compare the effect of the type of instructional method on participants' perceived ease of use, after controlling for their general computer self-efficacy[5]. Results showed a non-significant relationship between the covariate general computer self-efficacy and perceived ease of use, $F(1,95) = 1.736$, $p = .19$, η_p^2

[4]When we did not include the general self-efficacy covariate in the analysis (i.e., ANOVA), the results were comparable and in line with our first hypothesis, $F(2, 95) = 5.970$, $p = .004$, $\eta^2 = 0.063$.
[5]When we did not include the general self-efficacy covariate in the analysis (i.e., ANOVA), the results were comparable and not in line with our second hypothesis, $F(2,96) = 1.034$, $p = .36$, $\eta^2 = 0.021$.

= 0.018. Likewise, results provided no evidence for a significant main effect of the type of instructional method on participants' perceived ease of use, also when controlling for their general computer self-efficacy, $F(2,95) = 0.777$, $p = .46$, $\eta_p^2 = 0.016$.

2.3.3 Discussion

Results of Study 2 showed that participants in the agent-delivered modeling condition scored significantly higher on a declarative knowledge assessment compared to participants in the agent-delivered instructional narration. This finding indicates that using an artificial agent as a behavioral model is a more effective approach to enhance learners' declarative knowledge acquisition compared to having an agent only narrating. This result is in line with suggestions from the earlier literature regarding the superiority of human behavioral modeling over lecturing in order to develop accurate declarative knowledge structures (Szymanski, 1988). However, the results showed that this advantage was not present when agent-delivered modeling was com-pared to voice-only instructional narration. We argue that this is be-cause of the study's design-related characteristics rather than because of elements inherent to modeling. Specifically, the screen of the instructional videos was split into two parts, but, contrary to the two agent conditions, in the voice-only instructional narration condition, only one side of the screen contained visual information (the other side only contained a static photo). Thus, according to cognitive load theory (Sweller et al., 2011), in the absence of an on-screen artificial agent (i.e., less visual information), participants in the voice-only condition, might have experienced less cognitive load, resulting in better verbal information processing and acquisition, as compared to the agent-de-livered instructional narration.

In line with our prediction and earlier research in the field of computer training (i.e., Compeau & Higgins, 1995a; Compeau & Higgins, 1995b), participants in the agent-delivered modeling condition showed better task performance when using the system, compared to participants in the two non-modeling conditions. Findings revealed that these participants were significantly faster and made fewer errors when completing the task, compared to those in the two non-modeling conditions. We found no significant differences between individuals in the two non-modeling conditions. Thus, our third hypothesis was sup-ported, providing evidence that it is the instructional approach of an artificial agent (i.e., modeling) that can positively influence learners' behavior (i.e., task performance) rather than the agent's mere presence.

Lastly, similar to the findings of Study 1, results of Study 2 showed that participants in the agent-delivered modeling condition showed higher computer self-efficacy, as compared to participants in the two non-modeling conditions. To the contrary, Study 2 did not provide evidence for a significant difference in perceived ease of use between the agent-delivered modeling and the two non-modeling conditions. The lack of success in reproducing the effect of agent-delivered modeling on participants' perceptions of ease of use, as compared to the agent-delivered instructional narration, could be attributed to the smaller sample size as compared to that of Study 1. More power was required to detect such an effect on users' perceived ease of use.

2.4 General discussion

Artificial agents have been recently employed in multimedia learning environments, in order to introduce more instructional support and persuasive elements (Clark & Choi, 2005). However, contrary to the purpose they are designed to serve, the literature draws a discouraging picture concerning the overall general impact of artificial agents on motivational and learning outcomes. Based on past work (i.e., Heidig & Clarebout, 2011), we argued that a more fruitful approach is to ask under what conditions artificial agents might facilitate learning. While most pedagogical research focused on the agents' design, other conditions of their use, such as the agents' instructional role, have been neglected.

The current work examined behavioral modeling as a facilitating instructional role that an artificial agent can take in a multimedia learning environment. Past work employing human models (i.e., Compeau & Higgins, 1995a; Compeau & Higgins, 1995b) revealed that behavioral modeling, yields higher scores of computer self-efficacy and better task performance, compared to other commonly used non-modeling instructional methods. Findings of Study 1 and Study 2 showed that, similar to a human model, an artificial model can positively in-fluence users' motivational (computer self-efficacy and perceptions of ease of use of the system), cognitive (declarative knowledge) and, most importantly, skill-based (i.e., task performance) constructs of learning, as compared to other popular non-modeling instructional methods. Such results further suggest that it is the agent's instructional approach -provision of a behavioral model for social learning- rather than its mere physical presence that has a positive impact on learning.

2.4.1 Limitations, future research, and practical recommendations

Like every study, this research is not without limitations. The mixed results regarding the system's perceived ease of use could be attributed to the different design-related characteristics among conditions that could have an impact on participants' overall cognitive capacity. This is, the text-only instruction condition of Study 1 delivered only visual in-formation to the users, while the other two agent-delivered conditions provided information through, both, visual and auditory modalities. To the contrary, in the voice-only instructional narration method of Study 2, individuals were required to watch and process less information in their visual working memory.

In fact, past research revealed that on-screen agents can sometimes be a source of distraction for individuals' learning leading to a negative effect on retention and transfer (Veletsianos, Miller, & Doering, 2009). Though one could claim that the physical appearance of a model is a prerequisite for modeling, this might not always be the case in online multimedia learning environments. Although the study's findings provide evidence that behavioral modeling performed by an agent is an effective instructional method, the study's design does not allow to make inferences about the sole effect of the physical presence of the artificial model on learning. Future research might examine the mere effect of an artificial agent's physical presence as a model on people's learning gain, taking into consideration conditions under which conditions the artificial model's physical appearance on-screen facilitates learning (i.e., based on the type of task to be modeled). Additionally, future studies could consider whether various characteristics of the artificial agent as a behavioral model (i.e., the non-verbal behavior of the artificial model) could further augment learning.

What is more, although the effectiveness of agent-delivered modeling was tested with people who were computer literate, as shown by their high general computer self-efficacy and frequency of use, we are convinced that these findings can be generalized to individuals with lacking computer skills. In fact, modeling has been shown to be a more effective instructional method for people with lower general computer self-efficacy, as compared to other non-modeling methods (Gist, Schwoerer, & Rosen, 1989). Future research might examine the effectiveness of modeling to a population with different individual characteristics (i.e., with low computer literacy). Additionally, concerning individual characteristics, despite random assignment across treatment conditions, as in every experimental design, it is possible that the groups obtained had pre-existing differences in a quality that systematically altered the response of one group to the treatment as compared with the others.

Besides characteristics of the population, other opportunities for future work involves the nature of the computer task to be demonstrated. Earlier research found that the success of modeling might differ depending on the system that is modeled (Compeau & Higgins, 1995a). It has been proposed that familiarity with the system reduces the effectiveness of modeling. In our experiments, participants were un-familiar with the study's specific system (i.e., eye-tracking browser), although they were familiar with the general task (i.e.., Wikipedia search). Nonetheless, the study's focus was on the demonstration of a novel eye-tracking system (i.e., how to use a browser using eyes) and not Wikipedia search per se.

Lastly, we compared agent-delivered modeling with other popular non-modeling methods. However, the potential of agent technologies for education is vast. For example, recent findings suggest that students' learning can be enhanced through the act of teaching agents (i.e., learning by teaching) (Biswas, Leelawong, Schwartz, Vye, & The Teachable Agents Group at Vanderbilt, 2005). Future research might examine learning by being taught versus learning by teaching, using modeling as an instructional method (i.e., students are provided with agent-delivered modeling vs. students teach an agent by modeling a behavior). Future research might examine learning by being taught (i.e., students are provided with agent-delivered modeling) versus learning by teaching (i.e., students teach an agent by modeling a behavior) using modeling as a teaching method.

The current work provides further support for the effectiveness of persuasive technologies that take the form of a social actor. It also augments work within educational research, by identifying artificial agents as effective behavioral models that have the potential to successfully replace human models in multimedia environments. This is important given the widespread use of online education and distance learning, and therefore, the compelling need to ameliorate distance education course quality. Overall, artificial agents embedded in digital settings as models can provide an important technology for the improvement of distance education.

Chapter 3

The effect of the visual presence of an artificial model and type of learning task on cognitive load and learning outcomes[6]

[6]This chapter is based on:
Fountoukidou, S., Matzat, U., Ham, J., & Midden, C. (2021). *The effect of the visual presence of an artificial model and type of learning task on cognitive load and learning outcomes*. Manuscript submitted for publication.

3.1 General Introduction

Innovative educational technology tools have a great premise for improving learning, yet they are often not used to their full potential. Attempts to utilize more of a technology's capabilities have led researchers to investigate instructional software tools such as artificial pedagogical agents -on-screen, animated characters. Such agents are designed to facilitate learning by providing instructional support and motivation in multimedia learning environments (Clark & Choi, 2005). Nonetheless, findings regarding their effectiveness for learning are mixed. A meta-analysis revealed that agents were associated with a small but positive effect on learning (Schroeder, Adesope, & Gilbert, 2013), while systematic reviews showed that the majority of studies found nonsignificant effects (Heidig & Clarebout, 2011; Martha, & Santoso, 2019). Nonetheless, in light of the great variety of artificial agents used in past studies, as also the specific functions they may execute, the question of whether they can generally facilitate learning might be too broad. A more fruitful approach is to ask under which conditions pedagogical agents can facilitate learning. While the majority of pedagogical research mainly focused on the impact of an agent's design characteristics on learning (i.e., appearance), other conditions of their use, such as the artificial agents' instructional method, have been neglected (Heidig & Clarebout, 2011).

Undeniably, the instructional method used by a teacher – either human or artificial- has an important role to play in student learning. Modeling has been found to be an effective instructional method in enhancing learning. This method entails a model that physically accomplishes a task thereby demonstrating to learners how to successfully complete it (Bandura, 1986). However, the effectiveness of modeling by an artificial agent has not been examined until recently (Schroeder & Gotch 2015). In their study, Fountoukidou, Ham, Matzat and Midden (2019) examined whether modeling by an artificial agent is effective for learning. Findings revealed that an artificial agent as a model is more effective than the other non-modeling methods (i.e., agent-instructional narration, voice-only-instructional narration, text-only-instruction) in enhancing motivational (i.e., self-efficacy), cognitive (declarative knowledge) and behavioral (task performance) learning outcomes.

Although these findings provide support in favor of artificial modeling over other instructional methods, the design of the study of Fountoukidou et al. (2019) does not allow to infer whether the positive effects of an artificial agent as a model on learning depends on the visual presence of the agent. Therefore, these findings cannot counter those arguing that it is the instructional method (i.e., modeling) that is responsible for the learning gains, even when no medium is visually present (i.e., the artificial model) (Clark & Choi, 2007).

In fact, two competing perspectives exist in the literature on whether the visual presence of artificial agents in multimedia settings hinders or augments learning. These have been labelled as "agents-as-complements" versus "agents-as-distractors" (Frechette & Moreno, 2010). Theories supporting the agents-as complements perspective, like social presence theory (i.e., Hoyt, Blascovich & Swinth, 2003) and social agency theory (i.e., Moreno, Mayer, Spires & Lester, 2001), argue in short that an agent's visual presence increases student motivation, which in turn leads to greater invested effort during learning and more well-formed mental models of the taught concepts. In terms of the evidence for a motivational effect of the artificial agent's visual presence, existing literature is inconclusive. That is, some studies have provided support of the agents-as-complements perspective (e.g., Chen & Chou, 2015; Dinçer & Doğanay, 2017; Lin, Ginns, Wang & Zhang, 2020; Park, 2015), while some other studies did not (Carlotto & Jaques, 2016; van der Meij, van der Meij & Harmsen, 2015).

On the other hand, theories supporting the agents-as distractors perspective, like seductive details (Mayer, 2001) and cognitive load theory (Sweller 2004; Sweller, Ayres, & Kalyuga, 2011), hold that the inclusion of an agent might hinder rather than foster learning. Moreno et al. (2001) termed this "interference" and reasoned that the presence of the agent can hamper learning, because "any additional material that is not necessary to make the lesson interesting reduces effective working-memory capacity and, thereby, interferes with the core material" (p. 186). Overall, these theories predict two types of adverse effects of the visual presence of the agent: cognitive distractions (i.e., inability to pay attention and comprehend learning content) and affective distractions (i.e., disruptive feelings leading to impediment of learning goals) (Frechette & Moreno, 2010). Therefore, according to these theories, an instructional design will be more successful when unnecessary or distracting elements (i.e. artificial agents) are removed from the presentation, thus freeing the learner's cognitive resources to process the content that is most central to learning (Moreno & Mayer, 2000). Results on the effects of an artificial agent on cognitive load are mixed. That is, some studies found that artificial agents increase cognitive load (Frechette & Moreno, 2010), some other studies found no difference in cognitive load between agent and no-agent conditions (e.g., Moreno et al., 2001) and some other studies have found that artificial agents reduce cognitive load (e.g., Dinçer & Doğanay, 2017).

Though one could claim that the presence of a model is a prerequisite for modeling in traditional learning environments, the visual presence of the model is not the necessary in multimedia learning settings. For instance, one may listen to verbal instructions on how to perform a task, while the effects of the model's actions are being demonstrated on the computer screen, but without the model being visible. One question, then, arises: what is the sole effect of the artificial model's visual presence on learning outcomes? Since this question is very broad, the current research's aim is to study *under which conditions does the visual presence of an artificial model facilitate learning?*

Modeling involves the visual observation of the behavior of a model when performing a task. According to Bandura (1986), a model is only effective when it is relevant to the modeled behavior. However, the modeled behavior is determined every time by the learning task. This work posits that the type of learning task could define whether the artificial model's visual presence augments the attainment of learning outcomes. In other words, we argue that the type of learning task could be a decisive factor of whether an artificial model would be considered as relevant and, therefore, supplemental to the instruction, or as irrelevant and, thus, unnecessary additive.

According to Bloom's taxonomy (1994), learning tasks are categorized into three domains, that is, psychomotor (i.e., skill-based domain), cognitive (i.e., knowledge-based domain), and affective (i.e., attitudinal-based domain). Consequently, modeling of cognitive tasks (called cognitive modeling) has been proposed as another form of modeling of psychomotor tasks (called behavioral modeling) (Collins, Brown & Newman, 1989). Specifically, cognitive modeling pertains to the observation of a model's performance of cognitive skills and processes (i.e., solving of problems), which requires the explication of thoughts and reasons that underlie the performance of actions or choices (Wouters, Pass & Merrienboer, 2008). Overall, past research has shown that similar to behavioral modeling, cognitive modeling (via human teachers) is more effective in enhancing learners' performance, as compared to other teaching methods (i.e., lecture) (i.e., Gist, 1989; Gorrell & Capron,1990).

In the present work, we argue that an artificial model's visual presence is particularly relevant for behavioral modeling, which pertains to the demonstration of psychomotor learning tasks (i.e., learning of appropriate muscle movements to perform a task). Specifically, we argue that for behavioral modeling, the visual presence of an artificial model is highly required, as it facilitates the construction of a mental model of the psychomotor task taught, by providing a prototype. However, this might be less the case for cognitive modeling that focuses on the demonstration of purely cognitive tasks (e.g., mathematics). We postulate that this might be because purely cognitive tasks entail actions that are not readily observable. Thus, these mental actions need to be inferred either from physical actions that follow from them (i.e., writing down mathematical equations), or they need to be made explicit in order to be observed (i.e., talking out loud). Thus, we argue that for cognitive modeling, the agent's visual presence is decorational (and therefore not required), because only the task demonstration (with verbal instructions included) contains the core material to be learned.

3.1.1 The current work

The current work maintains that artificial agents can enhance learning, under certain conditions, in which their visual presence facilitate learners' cognitive processes. Hence, more research is needed that takes into consideration the conditions under which agents enhance learning outcomes. In this research, we hold that the type of modeling (behavioral or cognitive), which is based on the learning task at hand (psychomotor or cognitive), is a crucial factor that determines whether the visual presence of an agent enhances or hinders learning outcomes.

In more detail, in this chapter, we report about two studies. The primary purpose of Study 1 is to examine effects of the interaction between the on-screen visibility of an artificial model (presence vs. absence) and type of task (psychomotor vs. cognitive) on learning outcomes. Such an interaction, to the best of our knowledge, has not been investigated before. Since, learning is a multidimensional process that also consists of cognitive and skill-based and affective constructs, the learning outcomes examined are knowledge (i.e., recall), physical skills (i.e., task performance) and motivational and affective beliefs (i.e., self-efficacy, affect towards the instructional material, affect towards the artificial instructor).

What is more, any type of learning task, whether psychomotor or cognitive, may consist of various levels of complexity. In Study 1, we further argue that as the level of complexity of a psychomotor task increases, the visual information provided by a model becomes more important for the construction of a more accurate mental model of such a task. We further argue that a learner's more accurate mental model of the task will manifest itself as better task performance.

In Study 1 it is implicitly assumed that the visual presence of the artificial agent as a model has a different effect on learners' cognitive load depending on the type of task it models (i.e., reduction of cognitive load for psychomotor tasks). However, this argument is not tested explicitly. Thus, building on the results of Study 1, the purpose of Study 2 is firstly to replicate part of the findings (i.e., the interaction between the on-screen visibility of an artificial model and type of task on task performance and recall) and secondly to extent these findings, by examining: 1) effects of the interaction between the on-screen visibility of an artificial model (presence vs. absence) and type of task (psychomotor vs. cognitive) on learners' performance-related cognitive load; and, 2) the effect of the visibility of the artificial model (presence vs. absence) on learners' recall-related cognitive load for the psychomotor task.

3.2 Study 1

In the current research, we argue that the type of learning task (psychomotor versus cognitive) could define whether the artificial model's visual presence augments the attainment of learning outcomes. In Study 1 we predict that the positive effect of the visual presence of the artificial model on individuals' a) task performance b) self-efficacy, c) recall and d) affective beliefs is larger for a psychomotor task than for a cognitive task (H1). We further predict that the effect of the visual presence of the artificial model on individuals' task performance is larger for the difficult level of a psychomotor task, as opposed to the easy level of psychomotor task. In addition, we hypothesize that this will not be the case for purely cognitive tasks since the very essence of such tasks lies mostly in the provision of audio rather than visual information (H2).

3.2.1 Method

Participants and design

Overall, 138[7] individuals participated in this study, with most of them being students at Eindhoven University of Technology. Specifically, 82 participants had an undergraduate education or higher and 54 participants had finished high school (two participants did not state their educational level). Of these participants, 65 (47.1%) were males and 73 (52.9%)

[7] Data on task performance was missing for 6 participants due to technical difficulties with the recordings. Nonetheless, such participant exclusion did not induce major changes to the initial sample's demographic characteristics.

of them were females. The age of the sample ranged from 18 to 44, with a mean age of 23 (SD = 3.71).

The study employed a 2 x 2 x 2 factorial design with on-screen visibility of the artificial model (visual presence vs. absence) as a between-subject factor, type of task (psychomotor vs. cognitive) and level of task complexity (easy vs. difficult) as two within-subject factors. The study's dependent variables were a) task performance b) self-efficacy, c) recall and d) affective beliefs (towards the instructional material and the artificial instructor). The study's inclusion criteria included English language fluency. Overall, the study lasted for approximately 30 minutes, and participants were compensated for their participation (5 euros).

The Tetris game as the study's learning material

The study's overall learning material pertained to a well-known computer game, called Tetris. Since one factor of the study was type of task (psychomotor and cognitive), we created two different variations of the Tetris method of play. These two variations of the original Tetris game were generated in such a way, to reflect the two types of learning task, while maintaining the fundamental rules the original Tetris game (i.e., how to move and rotate game pieces)[8].

In more detail, the cognitive task was created as a form of cognitive activity, while the psychomotor task was designed to be a form of a motor activity. Both psychomotor and cognitive learning tasks were devised to be equivalent with respect to duration and they were both based upon the same fundamental game rules (i.e., moving and rotating game pieces) even though the learning objectives differ (i.e., cognitive vs. psychomotor).

In addition, since one factor of the study pertained to the level of task complexity, we developed two levels of task complexity per learning task; this is, the easy level and the difficult level.

Finally, before the full study took place, a pilot study was conducted, in order to test these self-constructed psychomotor and cognitive learning tasks. Specifically, this work was directed towards a threefold goal: firstly, to verify that the two levels of complexity of both tasks (easy and difficult) worked as intended (without ceiling and floor effects). Secondly, to ensure that both the cognitive and psychomotor task, although different in nature, were comparable in terms of complexity (complexity was measured subjectively as perceived task complexity, and, objectively as task performance). Lastly, a third motivation was to acquire insights about the task performance of participants, to help develop a more accurate scoring system and increase the reliability of the measurements. Results of the pilot indicated that all three goals were achieved as intended.

[8]Tetris is a puzzle game, where a player has to move and rotate the game pieces, which fall down the playing field, with the goal to create horizontal lines without gaps. When such a line is created, it gets destroyed and any block above this deleted line will fall. If the game pieces land above the top of the playing field, the game is over

Cognitive learning task

Cognitive modeling was designed to be a demonstration of a purely cognitive task. Cognitive tasks deal with how a student acquires, processes and utilizes knowledge. It is the "thinking" domain (Kasilingam, Ramalingam & Chinnavan, 2014). Mathematics are often used in experiments as an example of cognitive learning. Thus, we developed a version of Tetris game where a player had to follow a set of (self-constructed) game rules, in order to construct and, then, solve (simple) math operations. Constructing and solving math operations was a prerequisite, to accomplish certain game actions (i.e., to move and rotate the game pieces in the desired position within the playing field).

Different game rules for constructing math operations were developed, for the two levels of complexity (easy and difficult level). In order to construct these game rules, we performed a cognitive task analysis. In short, our cognitive task analysis included: 1) identification of the cognitive skills to be utilized in the performance of tasks, (this was based on Bloom's (1994) taxonomy of the cognitive domain); 2) creation of levels of complexity per each cognitive skill required; 3) and, finally, requirement extraction for the cognitive learning task and performance assessment.

Overall, the level of complexity for this learning task was chosen to be based on the level of element interactivity in solving equations, according to the cognitive load theory (i.e., example of low element interactivity is $x = 2+3$; example of high element interactivity: $x + 3 = 5$) (Sweller & Chandler 1994; Chandler & Sweller, 1996).

Psychomotor learning task

Behavioral modeling was designed to be a demonstration of a psychomotor task. Psychomotor tasks deal with learning objectives, which most often relate to some muscular or motor skill acquisition. Therefore, a version of Tetris was self-constructed, where participants had to perform specific hand movements. Thus, we developed a version of Tetris game where a player had to follow a set of (self-constructed) game rules on how to perform specific hand movements. Performing these set of hand movements was a prerequisite, in order to accomplish certain game actions (i.e., to move and rotate the game pieces in the desired position within the playing field).

Different sets of hand movements were developed, for the two levels of complexity (easy and difficult level). In order to construct these game rules, we performed a psychomotor task analysis. In short, our psychomotor task analysis included: 1) identification of the motor skills to be utilized in the performance of tasks, (this was based on Bloom's (1994) taxonomy of the psychomotor domain); 2) creation of levels of complexity per each motor skill included; 3) and, finally, requirement extraction for the psychomotor learning task and psychomotor performance assessment.

Overall, the level of complexity for this learning task was chosen to be based on the level of familiarity with the psychomotor movements required (i.e., whether participants had a pre-constructed mental model of a specific movement prior to this game).

Artificial modeling

Both cognitive and behavioral modeling were presented to participants in the form of instructional videos. Since the on-screen visibility of the artificial model was one of the study's between subject factors, we designed the same instructional videos with and without the visual presence of the artificial agent.

The instructional video, in which the artificial model was visually present was split into the following two screens: The right-hand side of the screen consists of an artificial agent demonstrating, while providing verbal instructions on how to play Tetris in two different ways of interaction: how to move and rotate the game pieces to the desired position, 1) by performing hand movements (i.e., behavioral modeling); and, 2) by solving specific math operations (i.e., cognitive modeling). On, the left-hand side of the screen, participants observed the effects of the model's "real-time actions" on the computer system (i.e., a game piece that has been rotated clockwise as a result of the agent's physical or cognitive activity). Figures 4 shows screenshots of the videos for the cognitive and behavioral modeling were the agent is visible on-screen (for details see artificial agent's modeling of cognitive task and psychomotor task).

The other set of instructional videos was identical with the only difference being that there was no artificial agent visible on-screen. Thus, participants in this condition were provided with the same verbal instructions (i.e., how to move and rotate game pieces by either performing hand movements or solving math operations), and effects of the model's "real-time actions" on the computer system (i.e., the effects of the agent's "real-time actions" on the game). However, participants could not observe the artificial model on-screen demonstrating the task. Figure 5 shows screenshots of the videos for the cognitive and behavioral modeling respectively were the agent was not visible on-screen (for details see no-agent modeling of cognitive task and psychomotor task).

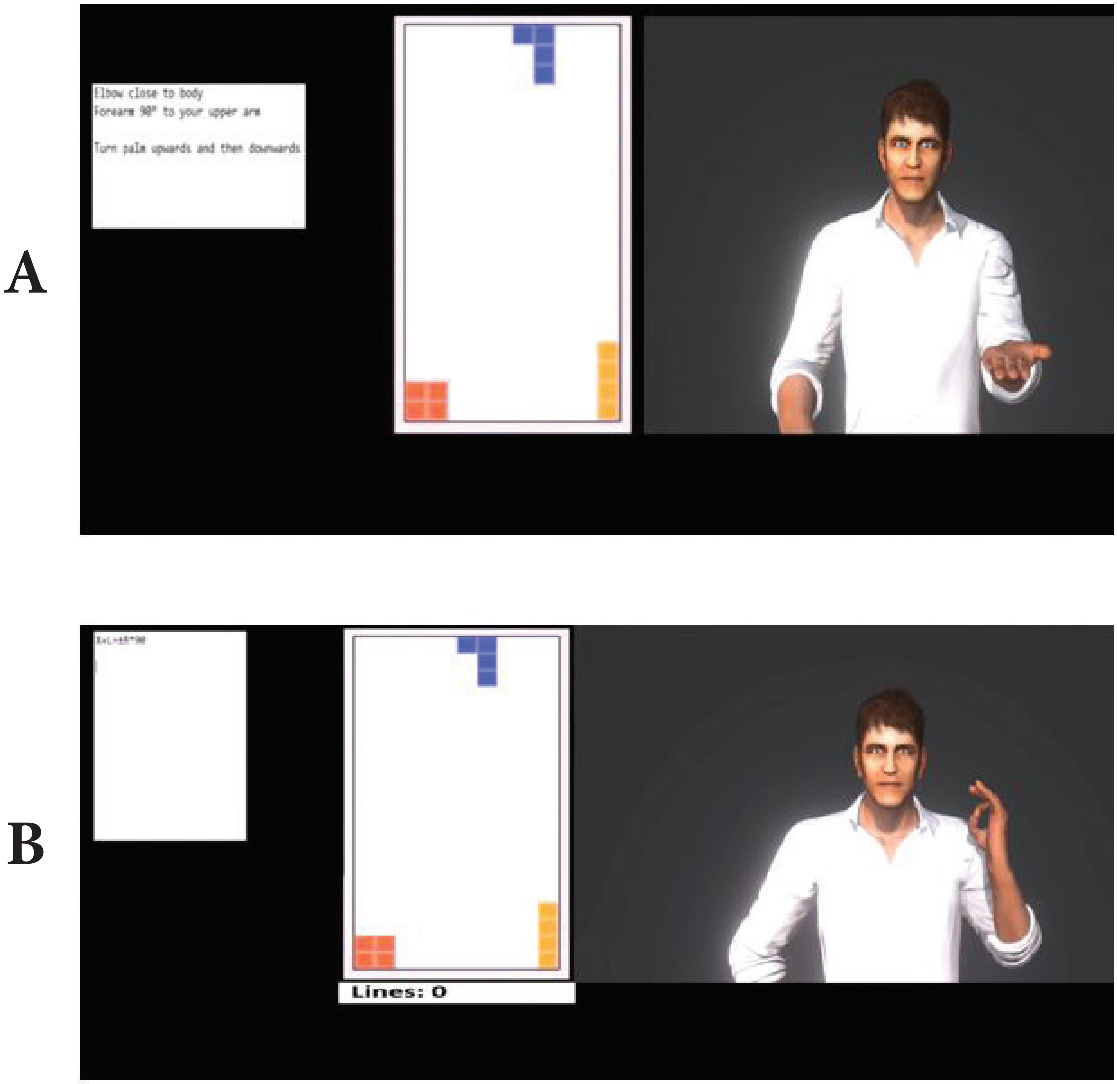


Figure 4 Artificial agent's modeling: a) psychomotor task: the agent demonstrates how a block is rotated by turning the palm upwards and then downwards; b) cognitive task: the agent demonstrates how a block is rotated by solving math equations (in this figure the agent explains that the line score is 0).

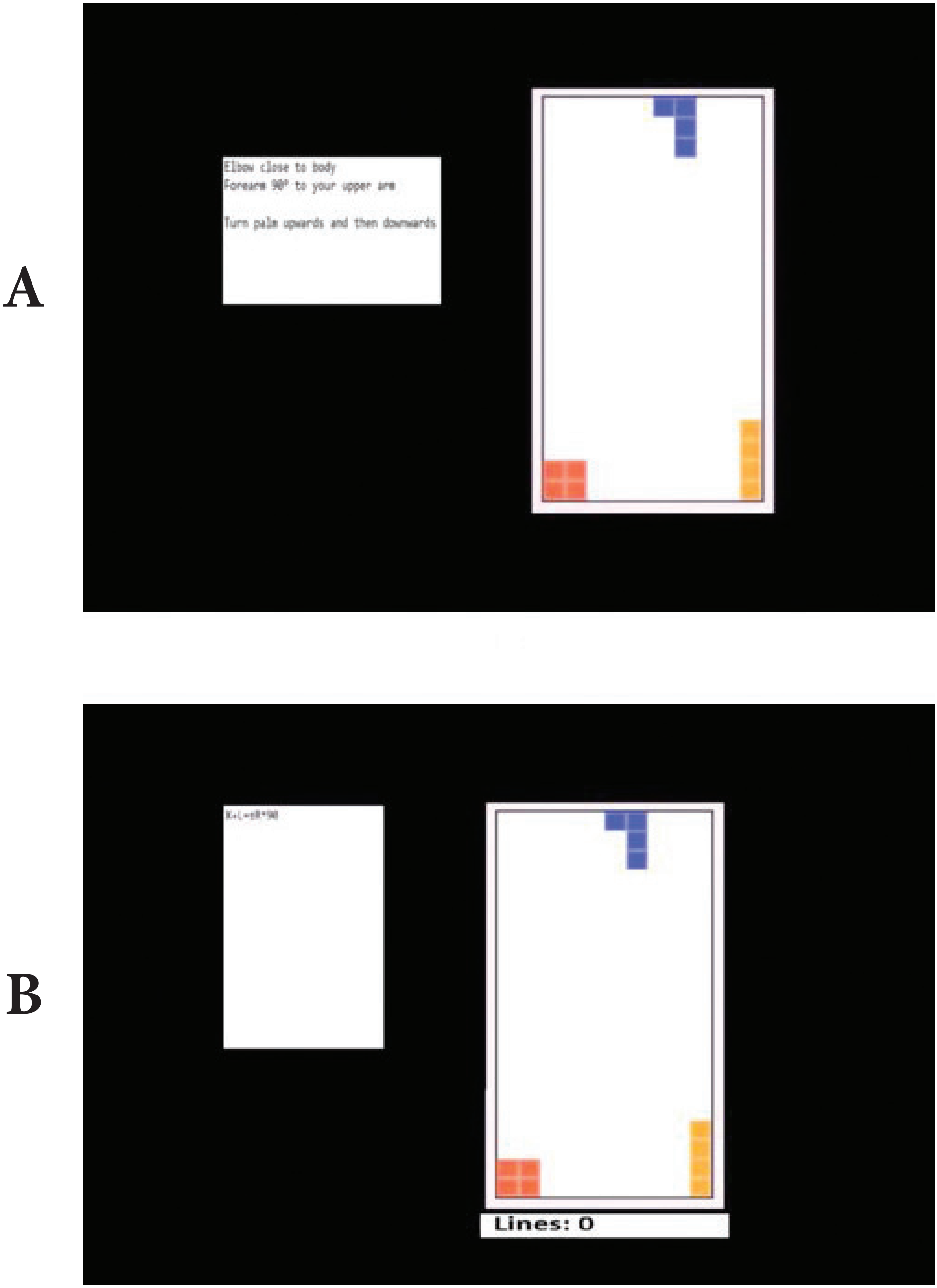


Figure 5 No-agent modeling: a) psychomotor task: a non-visible agent demonstrates a block is rotated by turning the palm upwards and then downwards; b) cognitive task: demonstrates how a block is rotated by solving math equations.

Apparatus

The current study employed a 3D animated artificial agent, which was created using the iClone 7 software. According to guidelines from past literature, the artificial agent was designed to resemble participants' various characteristics (Plant Baylor, Doerr & Rosenberg-Kima, 2009; Rosenberg-Kima, Baylor, Plant & Doerr, 2008). As participants of this study were mainly young individuals, the agent was designed to look young (~25 years), attractive (i.e., agent's facial features) and "cool" (i.e., agent's clothing and hairstyle).

In addition, web cameras were used to record the participants' task performance of the psychomotor learning task.

Measures

Task performance: Participants skill-based learning was assessed with a self-constructed performance assignment. This performance assignment made use of screenshots of a Tetris game in action (for an example of a screenshot see Figure 6). Specifically, there were four performance assignments per learning task (cognitive and psychomotor), which equalled to two assignments for each of the two levels of complexity (easy, difficult). In total, there were eight performance assessments per participant (for details see cognitive performance assignment and psychomotor performance assignment).

These assignments provided participants with a specific game mission every time (i.e., "*move this Tetris block 3 times to the right*"), which was different from the demonstration they watched. For both types of tasks, the performance assignments and given missions were the same (though their expected answers were obviously different). Participants were requested to follow the video instructions in order to accomplish it (i.e., by performing the correct hand movements for the task and by performing the right math operations for the cognitive task). The hand movements of participants in the performance assignment were captured by a web camera, while in the case of the cognitive performance assignment, participants were asked to write down the math operation process and the final result.

Task performance of these assignments was assessed using task accuracy (i.e., the number of errors) as a performance indicator. In more details a scoring system was developed for both tasks by breaking each task down to its comprised steps. Therefore, if all the identified steps were performed exactly as instructed, the performance was scored as correct (re-coded as 2); if one or more of these steps were either incorrectly performed or not performed at all, the performance was scored as wrong (re-coded as 0). Nonetheless, according to the scoring system that we developed, there were some cases that have been assessed as "half correct" (and thus recorded as 1). The specific rules that allowed a performance to be scored as half-correct was based on the idea that one had understood the essence of the task at hand (psychomotor and cognitive) and mistakes were only trivial and not core to the task. Participants' final score (for both psychomotor task and cognitive task) was the sum of the two given assignments for each level of complexity.

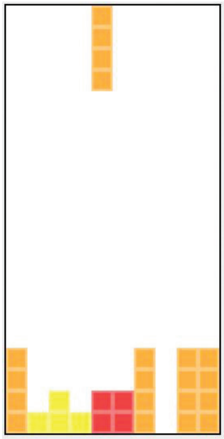


Figure 6 Example of a task performance assignment screenshot. Participants were asked to demonstrate how they would make this piece move for both types of task (e.g., 3 times to the right).

Self-efficacy: participants' self-efficacy per each type of task (psychomotor, cognitive) and each level of complexity (easy and difficult) was assessed by asking participants to answer one question regarding their perceived ability to perform the task (psychomotor or cognitive). Participants could choose an option on a 7-point rating scale, ranging from 1 to 7 (i.e., easy level: "*I believe I have the ability to move a Tetris block left and right using the math equations presented in the instructional video*" for the cognitive task; "*I believe I have the ability to rotate a Tetris block clockwise and countercloscwise using the hand gestures presented in the instructional video*" for the psychomotor task).

Recall: Participants' recall of the instructions was assessed with a recall test consisted of four self-constructed multiple-choice questions for both cognitive (i.e., *What is the math operation for moving a block to the right?*) and psychomotor task (i.e., *What is the hand gesture for moving a block to the right?*). Participants could select one correct answer out 5 options, with one being "*I do not know*" (see Appendix B.1 for the recall test for both cognitive and psychomotor task). The two recall tests were constructed to be as comparable as possible in terms of questions; however due to the different nature of the cognitive and psychomotor domain from which the tasks were derived the answers to most of the questions differ. Participants' answers were either correct (re-coded as 1), or wrong (re-coded as 0). Participants' final score (for both psychomotor task and cognitive task) was the sum of their correct answers for each level of complexity.

Affective beliefs: Participants' affective beliefs were assessed by asking participants to answer two sets of items for each task (psychomotor, cognitive), measuring: 1) their affect towards the instructional material; 2) their affect towards the (artificial) instructor. Both these components of affective learning contained three questions each and were administered through a 7-point semantic differential scale (Andersen, 1979) (constructed (see Appendix B.1 for the items on these scales). We constructed reliable measures of participants' affect towards the instructional material (Cronbach's $a = 0.86$) and towards the (artificial) instructor (Cronbach's $\alpha = 0.88$) for the psychomotor task by averaging participants' answers to each set of questions. Similarly, we constructed reliable measures of participants' affect towards the instructional material (Cronbach's $a = 0.86$) and towards the (artificial) instructor (Cronbach's $a = 0.88$) for the cognitive task by averaging participants' answers to each set of questions.

Procedure

After being welcomed in the main hall of the lab building, participants were asked to read and sign the study's informed consent form. Then, participants were randomly assigned to one of the two experimental conditions (visual presence of the artificial model vs. absence of the artificial model).

All participants watched instructional videos on how to play the Tetris game in two different ways (hand movements for the psychomotor task and math operations for of a cognitive task), which were randomly presented in a successive fashion. Given that there were two types of tasks, psychomotor and cognitive) of a different level of complexity (easy and difficult), each participant was requested to watch four instructional videos in total (psychomotor-easy, psychomotor-difficult, cognitive-easy, and cognitive-difficult).

After the end of each video, participants were given two performance assignments per complexity level and they were requested to show what they have learned in practice. After the end of each task, participants were required to fill in a questionnaire and then to answer to a recall test. Figure 7 illustrates the experimental procedure. Lastly, they were debriefed and compensated for their participation.

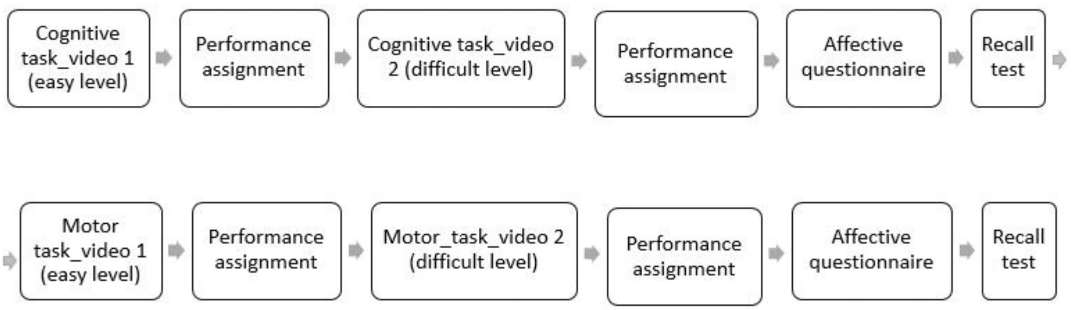


Figure 7 Study's experimental procedure with the two types of task (cognitive and psychomotor) and the two levels of complexity (easy and difficult).

3.2.2 Results

First we tested H1, pertaining to whether the effect of the visual presence of the artificial model on individuals' a) task performance b) self-efficacy, c) recall and d) affective beliefs is larger for the psychomotor task (i.e., behavioral modeling) than for the cognitive task (i.e., cognitive modeling).

Next, we continued with H2, examining whether the effect of the visual presence of the artificial model on individuals' task performance is larger for the difficult level of psychomotor task than for the easy level of psychomotor task, and whether this difference disappears for the cognitive task.

Due to the fact that we created two levels of complexity per learning task, for task comparability reasons, we included task complexity as a within-subject factor to all our analyses (with the exception of the analysis of affective beliefs, as they were measured for each learning task as a whole (see Figure 7).

Testing the interaction between the artificial model's on-screen visibility and type of task on learning (H1).

To test H1, our experimental design consisted of two, within subject factors (type of task and level of task complexity) and one, between subject factor (visibility of the artificial model), thus, a 3-way mixed ANOVA analysis was conducted for the first three dependent variables. However, since two variables of affective beliefs were measured (teacher liking and content liking), a 2-way mixed MANOVA analysis was conducted for our fourth dependent variable. Below we discuss the results of our analyses.

a) *Task performance*: In line with hypothesis 1a, we found a significant interaction effect between visibility of the artificial model and type of task for task for performance, $F(1, 130) = 35.155$, $p < .001$, $\eta_p^2 = .213$. This indicates that the agent's presence resulted in different changes of task performance scores, depending on the type of task (= type of modeling).

Specifically, as expected, separate independent t-tests revealed a significant difference in performance scores for the psychomotor task when the artificial model was visually present ($N = 67$, $M = 1.8$, $SD = .25$) as compared to when it was absent ($N = 65$, $M = 1.2$, $SD = .61$), $t(136) = 8.120$, $p < .001$. However, there was no significant difference found under the condition of the cognitive task where the artificial model was visually present ($N = 67$, $M = 1.1$, $SD= .68$) and when it was absent ($N = 65$, $M =1.2$, $SD = .71$), $t(136) = -.697$, $p = .48$.

b) *Self-efficacy*[9]: In line with our hypothesis 1b, we found a significant interaction effect

[9]There was no significant interaction effect between visibility of the artificial model, type of task and level of complexity for self-efficacy, $F(1, 136) = .694$, $p = .40$.

between visibility of the artificial model and type of task for self-efficacy, $F(1, 136) = 36.05$, $p < .01$, $\eta_p^2 = .067$. This indicates that the agent's visual presence provoked different changes in self-efficacy responses depending on the type of task (psychomotor or cognitive).

Specifically, as expected, separate independent t-test analyses revealed a significant difference in self-efficacy for psychomotor task were the artificial model was visually present ($N = 69$, $M = 6.5$, $SD = .91$), as compared to when it was absent ($N = 69$, $M = 5.8$, $SD = .94$), $t(136) = 4.248$, $p < .001$. Moreover, in line with H1b, results provided no evidence for a significant difference in self-efficacy for the cognitive task when the artificial model was visually present ($N = 69$, $M = 5.9$, $SD = 1.0$) versus when it was not visually present ($N = 69$, $M =5.7$, $SD = 1.17$), $t(136) = .377$, $p = .70$.

c) *Recall*[10]: Our hypothesis 1c was not supported for recall, as we did not find evidence for a significant interaction effect between visibility of the artificial model and type of task for recall, $F(1, 136) = 2.954$, $p= .09$. This indicates that the visual presence did not provoke different recall responses regardless of the type of task.

Main effects: There was a significant main effect of type of task $F(1,136) = 51.746$, $p < .001$, $\eta_p^2 = .276$. This is, participants showed higher recall after receiving cognitive modeling ($M =.951$, $SE = .011$) than after receiving behavioral modeling ($M = .746$, $SE = .026$). There was no main effect of the level of complexity found, $F(1, 136) = 2.024$, $p = .15$. What is more, there was also a significant main effect of visibility of the artificial model $F(1,136) = 5.507$, $p = .02$, $\eta_p^2 = .039$. This is, regardless of type of task, participants scored higher when the artificial model was not visually present ($M = .882$, $SE = .02$) rather than when it was visible ($M = .815$, $SE = .02$).

d) *Affective beliefs*: In line with our hypothesis 1d, we found a significant interaction effect between visibility of the artificial model and type of task for the two dependent variables combined, Wilk's $\Lambda = .771$, $F(2,135) = 19.992$, $p < .001$, $\eta_p^2 = .229$. Univariate tests revealed that there is a significant interaction effect on both teacher liking, $F(1, 136) = 39.994$, $p < .001$, $\eta_p^2 = .227$, and content liking, $F(1, 136) = 21.030$, $p < .001$, $\eta_p^2 = .134$. This indicates that the agent's visual presence provoked different changes in affective responses depending on the type of task. Separate univariate ANOVAs revealed a significant difference between the two conditions on teacher liking, $F(1, 136) = 12.583$, $p = .001$, $\eta_p^2 = .085$, and content liking, $F(1, 136) = 25.776$, $p < .001$, $\eta_p^2 = .159$.

Specifically, as expected by H1b, independent t-test analyses revealed a significant difference in teacher liking for behavioral modeling were the artificial model was visually present ($N = 69$, $M = 5.4$, $SD = 1.1$) as compared to when it was not visually present ($N = 69$, $M = 4.3$, $SD = 1.1$), $t(136) = 5.696$, $p < .001$. In line with H1b, we did not find a

[10]There was no significant interaction effect between visibility of the artificial model, type of task and level of complexity for recall, $F(1, 136) = .681$, $p = 41$.

significant difference in teacher liking for the cognitive task between the agent's visual presence condition (N = 69, *M* = 5.2, SD= .92) and no-presence condition (*N* = 69, *M* = 5.2, SD = .95), $t(136) = -.091$, $p = .92$. Similarly, we found a significant difference in content liking for the psychomotor task between the agent modeling condition (*N* = 69, *M* = 5.5, *SD* = 1.1) and no-agent modeling condition (N = 69, *M* = 4.3, *SD* = 1.2), $t(136) = 6.176$, $p < .001$. In line with H1b, results showed a non-significant difference in content liking for the cognitive task where the artificial model was visually present (*N* = 69, *M* = 5.4, *SD*= .97) and when it was absent (*N* = 69, *M* =5.2, *SD* = .91), $t(136) = 1.650$, $p = .1$.

Overall, in line with H1, our results revealed that learners provided with a psychomotor task (behavioral modeling) from an on-screen artificial agent had better task performance, and they further reported higher self-efficacy and affective beliefs, as compared to learners who received the same instructions but without the artificial model being visible to them. To the contrary, as expected, under the condition of a cognitive task (cognitive modeling), the visual presence of the artificial model was not found to influence learners' task performance, self-efficacy and affective beliefs. Regarding recall, contrary to H1, current findings suggested that that regardless of the type of task, learners scored higher when the artificial model was not visually present.

Testing the interaction between the artificial model's on-screen visibility, type of task and task complexity on task performance (H2).

To test H2, our experimental design consisted of two, within subject factors (type of task and level of task complexity) and one between subject factor (visibility of the artificial model), thus, a 3-way mixed ANOVA analysis was conducted for task performance.

In line with our second hypothesis, we found a significant interaction effect between visibility of the artificial model, type of task and level of complexity for task performance, $F(1, 130) = 17.41$, $p < .001$. This finding indicates that participants scored differently for the two levels of complexity depending on whether the model was visually present or not and depending on type of task (psychomotor or cognitive).

Following, as expected, we found a significant interaction effect between visibility of the artificial model and level of complexity for the psychomotor task on task performance, $F(1, 130) = 38.93$, $p < .001$. To the contrary, but as expected, we did not find a significant interaction effect between visibility of the artificial model and level of complexity for the cognitive task for task performance, $F(1, 130) = .48$, $p = .49$. This result indicates that only for the psychomotor task, participants performed differently for the two levels of complexity depending on whether the model was visually present or not. Further, for the psychomotor task, separate independent t-test analyses for the easy level revealed a significant difference in performance scores when the artificial model was visually present (*N* = 68, *M* = 3.7, *SD* = .71) as compared to when it was absent (*N* = 66, *M* = 3.2, *SD* = 1.25), $t(132) = 2.459$, $p = .015$. Similarly, we found a significant difference in performance scores for the difficult level of the psychomotor task between the agent modeling condition (N = 69, M = 3.8, SD = .60) and no-agent modeling condition (*N* = 68, *M* = 1.6, 2, *SD* = 1.87), $t(135) = 9.124$, $p < .001$.

3.2.3 Discussion

Taking the stance that an artificial agent can become a facilitator of learning underspecific conditions, in Study 1 we argued that the benefit of an artificial model's visual presence for learning is conditional on the type of learning task to be modelled. This is, the type of learning task (i.e., cognitive or psychomotor) is a critical factor of whether an artificial model is considered relevant and, therefore, supplemental to the instruction, or irrelevant to it and, thus, a distractive or unnecessary additive.

Consequently, our first hypothesis (H1) was that the effect of the visual presence of the artificial model on individuals' learning-related outcomes (task performance, self-efficacy, recall and affect) would be larger for psychomotor tasks (behavioral modeling) than for cognitive tasks (cognitive modeling). Results of Study 1 supported our first hypothesis for all learning outcomes, except for recall. In other words, findings revealed that when it comes to the demonstration of a psychomotor task (behavioral modeling), the visual presence of the artificial model enhanced learners' task performance, self-efficacy and affective beliefs, as compared to those who received the same instructions but without the artificial model being visible to them. Furthermore, and in line with our reasoning leading to H1, under the condition of a cognitive task demonstration (cognitive modeling), the visual presence of the artificial model was not found to influence individuals' learning outcomes. These findings support the study's argument that the additional value of the presence of the artificial model depends on the learning task to be modelled. We anticipated that when it comes to behavioral modeling the visual presence of an artificial model would be helpful, as it facilitates the construction of a mental model of the specific psychomotor task by providing a prototype. However, for cognitive modeling, which pertains to purely cognitive tasks where actions are not readily observable, the artificial model's visual presence is decorational and, thus, unnecessary. What is more, the level of task complexity was found to be a crucial factor regarding learning through artificial behavioral modeling. In line with our second hypothesis (H2), our findings revealed that, for the psychomotor task, the effect of the artificial model's visual presence was larger for the difficult level than for the easy level. Thus, the findings confirm our initial argument. That is, as the level of complexity of a psychomotor task increases, the visual information provided by the artificial model becomes more important for learners' construction of a more accurate mental model of the task, and, consequently, for better task performance.

One of the main concerns in the literature is that the inclusion of an artificial agent may create more extraneous cognitive load (i.e., the type of load resulted from the way the learning material is presented), as it requires the participant to process additional information. The current study's findings help reduce such concerns by showing that an artificial agent is beneficial in enhancing learners' task performance, motivational and affective beliefs, but only under specific conditions (i.e., the type of learning task that it models), in which their visual presence facilitates learners' cognitive processes (i.e., construction of a more accurate mental model of a given task). What is more, the study's results provide evidence that under certain conditions (i.e., increased difficulty of a psychomotor task being modelled), the visual presence of an artificial agent becomes even more important for improving learners' physical skills (i.e. task performance). However, in this study we did not examine the effect of the visibility of the artificial model and type of task on cognitive load. Future research might expand the findings by taking into consideration the interplay of the visibility of the artificial model and type of task on learners' cognitive load.

Nonetheless, contrary to our hypothesis, findings showed that, regardless of type of modeling received, participants had better recall of the instructions when the artificial model was not visually present, rather than when it was. This surprising result may be explained in light of cognitive load theory and its concept of the redundancy effect (Sweller et al., 2011). This effect may occur when multiple sources of information can be understood separately without the need for mental integration. Hence, it might be that under the condition of behavioral modeling, the artificial model's visual demonstration was redundant for recalling task instructions and it might have caused an unnecessary increase of extraneous cognitive load. Therefore, its visual presence might have inhibited participants' cognitive processing of the auditory narration of the task. To the contrary, this was not the case for individuals' task performance of the psychomotor task. We argue that in this case, optimum task performance was based on the successful integration of the two types of information provided: the visual demonstration and the auditory narration. Thus, no or little extraneous cognitive load was created. Future work might further examine whether a redundancy effect could explain the study's unexpected finding.

Other opportunities for further research include the study's learning material. In the current study we created two variations of the Tetris game to reflect the two types of learning task to be demonstrated (psychomotor and cognitive). Although participants were not familiar with these two versions, they were all familiar with the fundamental rules of the original Tetris game (i.e., how to move and rotate game pieces). Future research might examine the impact of the visual presence of an artificial model on learning using other learning material (i.e., different or unfamiliar to the participants), which might further enhance the learning effect of a visually present artificial model. Furthermore, the current study's population mainly comprised young students with a high level of computer literacy. Future research might examine the effect of the artificial model's visibility and type of task on learning to a population with other individual characteristics (i.e., with low computer literacy). Next, the study only tested such an effect on learner' short-term learning outcomes. Further assessment of this effect on long-term learning outcomes is suggested.

3.3 Study 2

In Study 1, we tested the type of instructional task (cognitive and psychomotor) in conjunction with the presence or the absence of an artificial model for learning. Findings showed that for psychomotor tasks, unlike for cognitive tasks, the visual presence of an artificial model was beneficial, increasing learners' task performance, self-efficacy and affective beliefs. The findings were in accordance with our hypothesis, which was based on the notion that behavioral modeling helps the learner to construct a mental model of the specific psychomotor task by providing a prototype. To the contrary, and as expected, this was not the case for cognitive modeling. This is because, as we claim, for cognitive modeling the relevant performed actions are of a cognitive nature and thus for a third party unobservable, making the artificial model's visual presence unnecessary.

Thus, in Study 1, it was implicitly assumed that the visual presence of the artificial agent has a different effect on learners' cognitive processes depending on the type of task it models. However, this argument has not been tested explicitly. This omission led to the current, follow-up study that aims not only to replicate but also to extent findings, by examining whether learners' performance-related cognitive load does change depending on the match between the visibility of the artificial model and the type of task.

According to cognitive load theory, cognitive load can either be intrinsic, that is, caused by the difficulty of the material itself, or extraneous, that is, due to its presentation (Kalyuga, 2011). The main concern in the literature is that the inclusion of an artificial agent may create more extraneous cognitive load, as it requires the participant to process additional information. It has been suggested that an artificial agent's properties (i.e., appearance, facial expressions, gestures, voice) could create an information rich display that could overwhelm a learner's working memory and decrease learning (Clark and Choi, 2007; Mayer & Moreno, 1998). On the other hand, it is argued that the features of artificial agents improve learner motivation and interest by creating a social agency environment (Atkinson, Mayer, & Merrill, 2005; Mayer, Sabko, & Mautone, 2003). A recent systematic review by Schroeder and Adesope (2014) found no clear direction among studies on cognitive load and artificial agents. The authors concluded that further research needs to be conducted to understand the impact that artificial agents have on cognitive load. Overall, in this Study 2, we posit that the visual presence of the artificial model helps learners construct an adequate cognitive representation of a psychomotor task, but it does not provide any additional benefit for a cognitive task.

What is more, in Study 1, contrary to our hypothesis, results showed that, regardless of the type of task, learners' recall of the instructions was increased when the artificial model was not visually present. Cognitive load theory and its concept of redundancy effect (Sweller et al., 2011) was suggested to explain this unexpected finding. That is, it might be that under the condition of behavioral modeling, the artificial model's demonstration was redundant for recalling the verbal instructions and it might have caused an unnecessary increase of extraneous cognitive load (see Study 1). To the contrary, we further argue that this is not the case for the task performance of the psychomotor task. That is because, as we claim, optimal task performance is based on the successful integration of both types of information provided by the artificial model: the visual demonstration and verbal instructions. Consequently, no or little extraneous cognitive load is created. In Study 2 we seek to explore whether the redundancy effect explains this unexpected finding on recall.

Taking into consideration all of the above, the goals in the present Study 2 are threefold: firstly, we will replicate parts of the Study 1 and test the interaction effect between the visibility of the artificial agent and the type of task on task performance and recall. Our first hypothesis is that the positive effect of the visual presence of the artificial model on individuals' task performance is larger for a psychomotor task than for a cognitive task (H1a). We further expect that, regardless of the type of task, learners' recall is better when the artificial model is not visually present. (H1b). Next, the study's second goal is to examine the interaction effect between the visibility of the artificial model and type of task on learners' perceived cognitive load of their task performance. Our second hypothesis is that the visual presence of the artificial agent reduces leaners' cognitive load for the psychomotor task performance but not the cognitive task performance (H2). The study's third goal is testing the effect of the visibility of the artificial model on perceived cognitive load of the recall test for the psychomotor task. We expect that the visual presence of the agent increases learners cognitive load related to recall in the psychomotor task (H3).

Lastly, in the literature, the role between cognitive load and test performance remains unclear. According to Kirschner's (2002) proposition, test performance is determined by cognitive load. This appears to be in accordance with the very essence of the cognitive load theory itself, which claims that increased cognitive load may negatively impact performance and vice versa (Sweller, Ayres & Kalyuga, 2011). Nonetheless, to the best of our knowledge, the implied mediating role of cognitive load in the relation between instruction and performance has not been empirically tested. Therefore, for exploratory reasons, we examine whether cognitive load mediates the effect of the visibility of the artificial model on task performance and recall for the psychomotor task.

3.3.1 Method

Participants and design

A total of 97 individuals participated in the study. Participants were recruited using a local participant database, and most of them were students from Eindhoven University of Technology. Of these participants, 51 were males and 46 were females. The age of the sample ranged from 18 to 33, with a mean age of 23 (*SD* = 2.9). Most of the participants were students at a Bachelor and Master level. The study employed a 2 x 2 factorial design with on-screen visibility of the artificial model (visual presence vs. absence) as a between-subject factor, type of task (psychomotor vs. cognitive) as a within-subject factor. The study's dependent variables were task performance, recall, and cognitive load. Inclusion criteria were participants' fluency in English and familiarity with the Tetris game. The duration of the study was approximately 30 minutes, for which participants received €5 as compensation for their participation.

Materials

Artificial modeling

In the same way as in Study 1, participants were presented with an artificial agent modeling both a psychomotor and a cognitive task. Both cognitive and behavioral modeling were presented to participants in the form of instructional videos. Since the on-screen visibility of the artificial model was one of the study's between subject factors, the same instructional videos were designed with and without the visual presence of the artificial agent (for more details on the construction of the instructional videos and the artificial agent see Study 1).

The instructional video in which the artificial agent was visually present was split into the following two screens: The right-hand side of the screen consists of an artificial model demonstrating, while providing verbal instructions on how to play the Tetris game in two different ways of interaction: how to rotate the game pieces to the desired position, 1) by performing hand movements (i.e., behavioral task); and, 2) by solving specific math operations (i.e., cognitive task). On, the left-hand side of the screen, participants observed the effects of the model's "real-time actions" on the computer system (i.e., a game piece that has been rotated clockwise as a result of the agent's physical or cognitive activity).

The other instructional video was identical with the only difference being that there was no artificial agent visible on-screen. Thus, participants in this condition were provided with the same verbal instructions (i.e., how to rotate game pieces by either performing hand movements or solving math operations), and effects of the model's "real-time actions" on the computer system (i.e., the effects of the agent's "real-time actions" on the game). However, in this video, participants could not observe the artificial model on-screen demonstrating the task.

Measures

Task Performance test: A participant's performance for both psychomotor and cognitive task was measured with the use of a performance test developed in Study 1. Both task performance tests contained two exercises each, asking participants to rotate Tetris game pieces in a specific way, according to the artificial model's instructions. Task accuracy (i.e., the number of errors) was used as a performance indicator. Participants' final score (for both psychomotor task and cognitive task) was the sum of the two given exercises. Performance was scored as correct (re-coded as 2); as wrong (re-coded; and as ed as "half correct" (re-corded as 1) (for more information on the task performance scoring system, see Study 1).

Recall test: A participant's recall of the video instructions was measured with a use of a recall test. The recall test consisted of five self-constructed gap filling questions (see Appendix B.2) and two self-constructed multiple-choice questions for both cognitive ("What is the math operation for rotating a block clockwise /anticlockwise?") and psychomotor task ("What is the hand gesture for rotating a block clockwise anticlockwise?") (the two multiple-choice questions were derived from Study 1). Participants could select one correct answer out five options, with one being “I do not know”. The two recall tests were constructed to be as comparable as possible in terms of questions; however due to the different nature of the cognitive and psychomotor domain from which the tasks were derived the answers to most of the questions differ. Participants’ final score (for both psychomotor task and cognitive task) was the sum of the gap filling and multiple-choice questions. Participants’ answers were either correct (re-coded as 1), or wrong (re-coded as 0).

Cognitive load: A participant’s cognitive load related to the task performance for both types of task and related to the recall test for psychomotor task, was measured in the same way. This is with the use of the subjective measures of perceived mental effort, task difficulty and extraneous cognitive load.

First of all, we measured participants’ subjective mental effort by using three of the six subscales of the NASA-Task Load (NASA-TL) Index (Hart & Staveland, 1988) that measure factors associated with completing a task: (1) mental demands 2) effort and 3) frustration level. The resulting three-item scale ranged from 1 (very little) to 7 (very much). Secondly, we measured subjective task difficulty by Kalyuga, Chandler and Sweller’s (1999). This is a three-item scale ranging from 1 (not at all the case) to 10 (completely the case). These two scales of mental effort and task difficulty are the most often applied measure for assessing cognitive load (Sweller, Ayres and Kalyuga, 2011). Thirdly, we administered the three-item extraneous cognitive load scale, ranging from 0 (not at all the case) to 10 (completely the case), developed by Leppink et al., (2013). We use these three scales in the following ways: for cognitive load related to task performance, we constructed reliable measures of participants’ mental effort (Cronbach’s $a = 0.75$), task difficulty (Cronbach’s $a = 0.81$) and extraneous cognitive load (Cronbach’s $a = 0.8$) by averaging participants’ answers to each set of questions. Overall, due to the fact that the items of the three scales were highly correlated, we constructed an overall scale comprised of the three subscales measuring cognitive load for performance (Cronbach’s $a = 0.88$) (see Appendix B.2 for the items on this scale). Similarly, for cognitive load related to recall of the psychomotor task, we were able to construct reliable measures of participants’ mental effort (Cronbach’s $a = 0.8$), task difficulty (Cronbach’s $a = 0.79$) and extraneous cognitive load (Cronbach’s $a = 0.8$) by averaging participants’ answers to each set of questions. Overall, due to the fact that the items of the three scales were highly correlated, we constructed an overall scale comprised of the three subscales measuring cognitive load for recall in the psychomotor task (Cronbach’s $a = 0.88$) (see Appendix B.2 for the items on this scale).

Procedure

Participants were invited per email and welcomed in the central hall of the lab building. After arrival, participants were first asked to read and sign an informed consent form, stating the general purpose of the research and their willingness to participate in the study. Then, they were randomly assigned to one of the two experimental conditions (visual presence of the artificial model vs. absence of the artificial model).

Next, participants watched the two instructional videos on how to play the Tetris game in two different ways (hand movements for the psychomotor task and math operations for of a cognitive task) that were randomly presented in a successive fashion. After each video, participants completed a task performance test, and answered a set of items about the cognitive load they experienced in completing the task performance test. Next, they filled in a recall test and lastly, used for the psychomotor task, they completed a set of items about the cognitive load they experienced in completing the recall test (cognitive load in completing the recall test for the cognitive task was measured as well, but is not reported here since we have no hypothesis about it in this study). Figure 8 illustrates the experimental procedure. Finally, participants were debriefed, paid and thanked for their contribution.

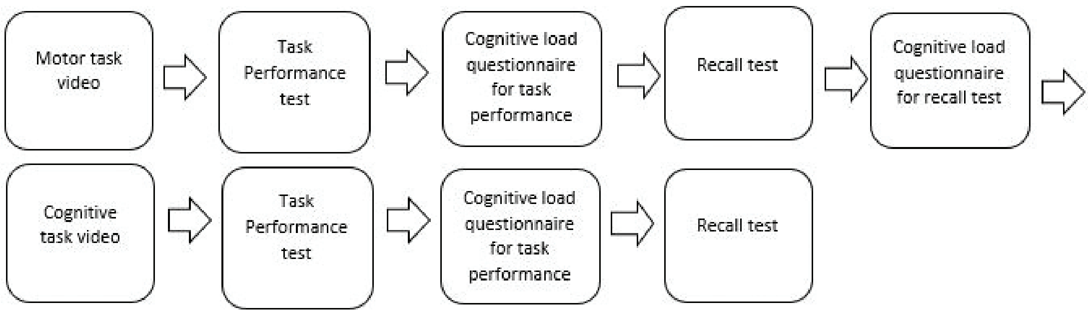


Figure 8 Study's experimental procedure (psychomotor and cognitive tasks were presented in random order).

3.3.2 Results

Our experimental design consisted of one within subject factor (type of task) and one between-subject factor (visibility of the artificial model). First, we tested H1(a), pertaining to whether the effect of the visual presence of the artificial model on individuals' task performance is larger for the psychomotor task (i.e., behavioral modeling) than for the cognitive task (i.e., cognitive modeling). Additionally, we investigated whether learners' recall is better when the artificial model is not visually present. (H1b). Next, we continued with H2, testing whether the effect of the visual presence of the artificial model on individuals' cognitive load is larger for the psychomotor task performance than for the cognitive task performance. Lastly, the effect of the visibility of the artificial model on perceived cognitive load related to the recall of the psychomotor task was examined (H3).

To test H1(a), a 2-way mixed ANOVA analysis was conducted for task performance. In line with hypothesis 1a, we found a significant interaction effect between visibility of the artificial model and type of task on task performance, $F\,(1, 93) = 26.13$, $p < .001$, $\eta_p^2 = .219$. This indicates that the agent's presence resulted in different changes of task performance scores, depending on the type of task (= type of modeling). Specifically, as expected, separate independent t-tests revealed a significant difference in performance scores for psychomotor task when the artificial model was visually present ($N = 48$, $M = 3.6$, $SD = 1.34$) as compared to when it was absent ($N = 47$, $M = 1.6$, $SD = 63$), $t(93) = -10.42$, $p < .001$. However, there was no significant difference found under the condition of the cognitive task where the artificial model was visually present ($N = 47$, $M = 2.3$, $SD = 1.55$) and when it was absent ($N = 48$, $M = 2.1$, $SD = 1.68$), $t(95) = -.952$, $p = .34$.

To test H1(b), a 2-way mixed ANOVA analysis was conducted for the recall test[11]. Contrary to our hypothesis, we did not find a significant main effect of visibility of the artificial model on recall $F\ (1,95) = 1.133$, p = .29. This is, the visual presence of the artificial model did not have any impact, (neither positive nor negative) on participants' recall. Following, we found a significant main effect of type of task on recall $F(1,95) = 55.95$, $p < .001$, $\eta_p^2 = .217$. That is, participants showed higher recall after receiving cognitive modeling ($M = 5.49$, $SD = 1.46$) than after receiving behavioral modeling ($M = 4.42$, $SD = 1.75$). Lastly, in accordance with the findings of our Study 1, we did not find a significant interaction effect between visibility of the artificial model and type of task on recall, $F(1, 95) = .880$, $p = .35$. This indicates that the visual presence did not provoke different recall responses for the two types of task.

To test H2, a 2-way mixed ANOVA analysis was conducted for cognitive load related to task performance[12]. In line with our second hypothesis, we found a significant interaction effect between visibility of the artificial model and type of task on cognitive load related to task performance, $F(1, 95) = 12.39$, $p < .01$, $\eta_p^2 = .115$. This indicates that the agent's presence resulted in different changes of cognitive load perceptions, depending on the type of task participants performed. Specifically, as expected, separate independent t-tests revealed a significant difference in cognitive load perceptions for psychomotor task when the artificial model was visually present ($N = 48$, $M = 1.8$, $SD = .99$) as compared to when it was absent ($N = 49$, M = 3.1, SD = 1.5), $t(95) = 4.545$, $p < .001$. However, there was no significant difference found on cognitive load of task performance under the condition of the cognitive task where the artificial model was visually present ($N = 48$, $M = 3.0$, $SD = 1.17$) and when it was absent ($N = 49$, $M =3.1$, $SD = 1.26$), $t(95) = .38$, $p = .70$.

To test H3, an independent-samples t-test analysis was conducted for cognitive load related to recall test of the psychomotor task. Contrary to our hypothesis, there was no significant difference found on cognitive load of recall of the psychomotor task where the artificial model was visually present ($N = 48$, $M = 3.4$, $SD = 1.30$) and when it was absent ($N = 49$, M

[11]Results were similar when an analysis was conducted on recall, measured with either only multiple-choice questions or only the gap-filling questions.

[12]Results were similar when an analysis was conducted on cognitive load, measured as either mental effort, task difficulty or extraneous cognitive load.

=3.5, *SD* = 1.68), *t*(95) = .22, *p* = .82. Thus, the visual presence of the artificial model was not found to have any impact, (neither positive nor negative) on participants' cognitive load experienced in completing the recall test of the psychomotor task.

Mediation analysis

Our aim was to explore whether cognitive load could explain part of the anticipated effect of the visibility of the artificial model on task performance and recall of the psychomotor motor task. Since we found no significant difference of the visibility of the artificial model on recall, we conducted a mediation analysis only for psychomotor task performance.

A mediation analysis was conducted, using dummy coding-artificial model's visual presence and absence. The analysis was performed using the PROCESS custom dialog for SPSS, as developed by Hayes (2018). The results are reported in Figure 9. Below we provide a summary of the main findings.

The analysis showed the visibility of the artificial model was a significant predictor of the psychomotor task performance R^2 = .74, as well as of cognitive load, R^2 = .42. Participants' cognitive load was not found to be significantly associated with psychomotor task performance. Similarly, cognitive load[13] was not found to mediate (part of) the effect of the visibility of the artificial model on psychomotor task performance.

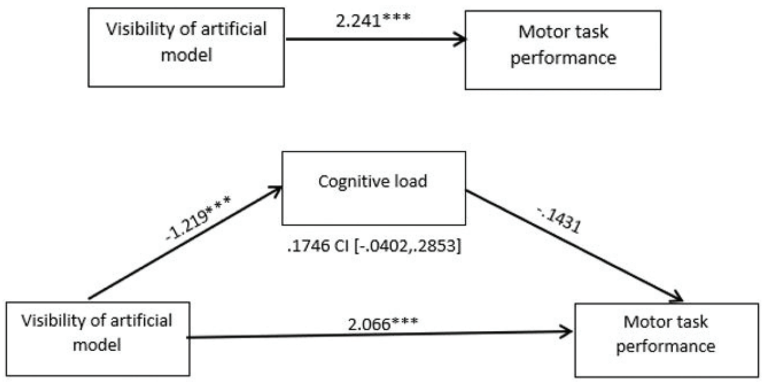


Figure 9 Mediation analysis of the difference in learners' psychomotor task performance towards the visibility of the artificial model (visually present vs. absent). All estimates are in unstandardized units. Based on 1000 bootstrap samples (bias corrected), *** *p* < .001.

[13]When conducting mediation analysis measuring cognitive load with the three questionnaires separately, only perceived task difficulty was found to partially mediate the effect of the visibility of the artificial model on task performance, explaining an additional 5% of the variance in task performance.

3.3.3 Discussion

Although the use of artificial agents in multimedia learning environments is an attempt to enhance learning, research has produced mixed evidence for enhanced learning (Schroeder, Adesope & Gilbert, 2013). Hence, there is much disagreement among researchers on whether the visual presence of an artificial agent is actually needed (Davis & Antonenko, 2017).

In Study 1, we argued that an artificial agent can be beneficial to the learning process under specific conditions. Given the potential of artificial agents to replace human models (Fountoukidou et al., 2019), Study 1 tested the type of instructional task (cognitive and psychomotor) in conjunction with the presence or the absence of an artificial model for learning. Findings showed that for psychomotor tasks, unlike for cognitive tasks, the visual presence of an artificial model was beneficial, increasing learners' self-efficacy, affective beliefs and task performance. Nonetheless, contrary to our hypothesis, results showed that, regardless of the type of task, learners' recall was increased when the artificial model was not visually present.

The first goal of the current, follow-up study was to replicate the effect of the interaction between visibility of the artificial model and the type of task on task performance and recall. Results partially supported our first hypothesis, showing that that when it comes to the demonstration of a psychomotor task (behavioral modeling), the visual presence of the artificial model enhanced learners' task performance, as compared to those who received the same psychomotor task instructions, but without the artificial model being visually present. Furthermore, and in line with findings of Study 1, under the condition of a cognitive task demonstration (cognitive modeling), the visual presence of the artificial model was not found to influence individuals' task performance. These results support the current work's argument that the additional value of the visual presence of the artificial model depends on the learning task to be modelled. We anticipated that when it comes to behavioral modeling the visual presence of an artificial model would increase task performance, as it facilitates the construction of a mental model of the specific psychomotor task by providing a prototype. However, for cognitive modeling, which pertains to purely cognitive tasks where actions are not readily observable, the artificial model's visual presence is unnecessary.

Nonetheless, contrary to our first hypothesis and findings of Study 1, the visual presence of the artificial model did not have a negative impact on learners' recall. We attribute this to the slight difference on how recall was measured in the two studies. In the first study recall was measured with four multiple choice questions, since participants were presented with four instructional videos per task. In the current follow-up study though, only the two instructional videos were presented to participants (since we focused only on one level of difficulty), and, therefore recall had to be measured with two multiple choice questions (that are the same as in the first study), and with five additional gap filling questions[14]. Consequently, results do not provide evidence that the visual presence of the artificial model increases learners' cognitive load related to recall (H3). Future work could make use of all four multiple-choice questions used in Study 1 to measure recall and then retest whether the visual presence of an artificial model negatively affect learners' recall and whether it also increases perceived cognitive load related to recall.

Next, the current study's second goal was to expand the findings of Study 1, by examining the effect of the interaction between visibility of the artificial model and type of task on learners' cognitive load perceptions related to their task performance. Results supported our second hypothesis, showing that that when it comes to the demonstration of a psychomotor task (behavioral modeling), the visual presence of the artificial model decreased learners' perceived cognitive load for the task performance, as compared to those who received the same psychomotor task instructions, but without the artificial model being visually present. Furthermore, and in line with findings of Study 1, under the condition of a cognitive task demonstration (cognitive modeling), the visual presence of the artificial model was not found to influence individuals' perceived cognitive load with respect to the task performance. Overall, the results showed that agents can be beneficial in decreasing cognitive load perceptions regarding their task performance under specific conditions, like the type of task being modelled by an artificial agent. Such results are essential given the lack of clarity of the direction of the artificial agents' effect on cognitive load (Schroeder & Adesope, 2014). Nonetheless, the study does not provide evidence that artificial agents can also be a source of increased cognitive load related to task performance for the cognitive task. We argue that this might be because the level of complexity of the performance test we constructed for the cognitive task (solving math operations), was not high enough for the study's participants, the majority of whom were students at a technical university. Thus, it might be that participants were able to ignore any distracting information in order to decrease cognitive load (i.e., De Jong, 2010, Moreno, 2005). Therefore, future studies could increase the level of complexity of the cognitive task to examine whether the visual presence of an artificial model can increase cognitive load of task performance.

Lastly, in Study 2, we explored whether cognitive load mediates the effect of the visual presence of the artificial model on psychomotor task performance. The study does not provide strong evidence that cognitive load explains the effect of the visibility of an artificial model on task performance. We argue that this might be due to conceptual or methodological issues regarding the cognitive load construct. Concerning the former, it might be that, as Paas and van Merrienboer (1993a) proposed, task performance is an assessment factor reflecting part of cognitive load. Regarding the latter, it might be that the scales used do not adequately capture the essence of the cognitive load construct. It has been admitted by Paas, Tuovinen, Tabbers, and Van Gerven (2003) that "the question of how to measure the multidimensional

[14]Similar to the findings of the current study, in Study 1 we do not find a significant difference of the visibility of the artificial model on recall, when recall is measured with only the two (out of four) multiple choice questions.

construct of cognitive load has proven difficult for researchers" (p. 66). This study is among a few that tested the mediating role of cognitive load between instruction and performance. More research is required to shed light into the relationship between test performance and cognitive load.

Overall, the contribution of the current follow-up study is threefold: Firstly, it replicates the interaction effect between visibility of an artificial model and type of learning task on learn ers' task performance found in Study 1. Thus, the study increases the confidence that when it comes to motor task (behavioral modeling), the visual presence of the artificial model enhances learners' task performance, as compared to the provision of the same instructions but without the artificial model being visible on-screen. To the contrary, when it comes to a cognitive task (cognitive modeling), this positive effect of the visual presence of an artificial model on task performance disappears. Secondly, this research extends the findings of Study 1 by revealing that the visual presence of the artificial model has a different effect on leaners perceived cognitive load depending on the type of task they perform. That is, the visual presence of the artificial model reduced leaners' perceived cognitive load for their motor task performance but not for the cognitive task performance. Thirdly, cognitive load was not clearly found to explain the effect of the visibility of the artificial model on task performance. Therefore, these findings stress the importance of conducting more research, so as to help solving any conceptual and/or methodological challenges regarding the cognitive load construct and its relationship to test performance.

3.4 General discussion

Artificial agents have been recently employed in multimedia learning environments, in order to provide more instructional support and motivational elements (Clark & Choi, 2005). However, existing literature reported mixed results concerning the artificial agent's impact on learning outcomes. (i.e., Heidig & Clarebout, 2011; Martha, & Santoso, 2019; Schroeder, Adesope, & Gilbert, 2013). In this Chapter, we argue that a more fruitful approach is to ask under what conditions artificial agents might facilitate learning. While the majority of pedagogical research mainly focused on the agents' design, other conditions of their use, such as the agents' instructional role, have been neglected.

Earlier work showed that modeling can be an effective instructional role that an artificial agent can take in order to enhance learning outcomes in digital settings (Fountoukidou, Ham, Matzat & Midden, 2019). In fact, there are two types of modeling based on the type of learning task to be modeled: behavioral modeling pertaining to the demonstration of psychomotor tasks and cognitive modeling concerning the demonstration of (purely) cognitive tasks (Collins et al., 1989; Wouters et al., 2008). Nonetheless, given that the visual presence of a model is not a prerequisite in a digital environment, the sole effect of an artificial model's visual presence on learning remains unclear. Thus, the question that arises is whether its visual presence facilitates learning. The existing literature contains contradictory theories concerning the overall impact of an artificial agent's on-screen presence on learning. Specifically, theories such as social presence theory and social agency theory argue that the artificial agent's physical presence leads to well-formed mental models of concepts taught and better learning due to an increased motivation (i.e., Hoyt et al., 2003; Moreno et al., 2001). Nonetheless, findings of recent studies are inconclusive in terms of the motivational effect of the artificial agent's visual presence (i.e., Chen & Chou, 2015; Dinçer & Doğanay, 2017; Lin et al., 2020; Park, 2015). On the other hand, theories, such as cognitive load theory (Sweller, 1988; Sweller 2004), hold that that such on-screen presence can impose cognitive and affective distractions and, thus, hamper learning. However, studies on the effects of an artificial agent on cognitive load reported opposing results (i.e., Dinçer & Doğanay, 2017; Frechette & Moreno, 2010; Moreno et al., 2001).

We conducted two studies with the primary goal to examine whether the benefit of the visual presence of the artificial model on learning is dependent on the type of learning task to be modeled (psychomotor or cognitive). Findings of Study 1 revealed that for a psychomotor task (behavioral modeling), the visual presence of the artificial model enhanced learners' task performance, self-efficacy and affective beliefs, as compared to those who received the same instructions but without the artificial model being visible to them. To the contrary, and as expected, in a cognitive task (cognitive modeling), this positive effect of the visual presence of an artificial model on learning was not found. Further, it was shown that the level of task complexity is another factor to be considered, as the effect of the visual presence of the artificial model on task performance was larger for the difficult level than for the easy level of the psychomotor task. However, unexpectedly, results of Study 1 showed that independent from type of the learning task, recall was hampered by the visual presence of the artificial model.

In Study 2, we replicated the interaction effect between visibility of an artificial model and type of learning task on learners' task performance found in Study 1. Thus, the two studies combined increase confidence on the essential role of the visual presence of an artificial agent for learning during a psychomotor task (behavioral modeling). Furthermore, Study 2 extends the findings of Study 1 by revealing that the visual presence of the artificial model has a different effect on leaners' perceived cognitive load depending on the type of task they perform. That is, the visual presence of the artificial model reduced leaners' perceived cognitive load for their psychomotor task performance but not for the cognitive task performance. However, perceived cognitive load was not clearly found to mediate the effect of the visibility of the artificial model on task performance. Finally, in Study 2, we were not able to replicate findings of Study 1 regarding the visual presence of an artificial agent having a negative impact on learners' recall for both types of tasks.

Overall, findings of the two studies provide strong evidence that the visual presence of the artificial model enhanced learners' self-efficacy, affective beliefs and task performance, as it also minimized cognitive load associated with task performance, for psychomotor tasks (behavioral modeling), but less so for cognitive tasks (cognitive modeling).

3.4.1 Implications

The current work has both theoretical and practical implications. Regarding theoretical implications, our findings support the argument that the question on whether artificial agents facilitate learning can only be answered by taking into consideration the specific conditions of their use. Findings of the current work confirms that the type of learning task that an artificial agent models is an important condition.

Furthermore, our two studies combined add to the scientific debate on whether artificial agents augment or hinder learning. Specifically, our results do not provide evidence that the sole visual presence of an artificial agent (as in the case of cognitive modeling) has a positive impact on learning (i.e., task performance, self-efficacy and affective beliefs) as claimed by earlier theories (i.e., social presence theory, Hoyt, Blascovich & Swinth, 2003; and social agency theory, Moreno et al., 2001). More importantly, according to our findings, the sole visual presence of an artificial agent can even have aversive effects on certain learning outcomes (i.e., recall of the instructions). Thus, we argue that theories supporting the instructional value of the sole visual presence of an artificial agent, like social presence theory and social agency theory, are incomplete. Future research could identify and examine conditions under which the inclusion of an artificial agent's sole visual presence can facilitate learning. We hypothesize that a framework for such conditions could be related to learners' feelings of loneliness and isolation as a consequence of reduced social presence and psychological immediacy in online learning environments, compared to in-person instruction (Jeste, Lee & Cacioppo, 2020). In fact, earlier research has shown that social exclusion increases both attentiveness to social cues (Pickett, Gardner & Knowles, 2004) and attributions of human-likeness to artificial agents (Epley, Waytz, Akalis & Cacioppo, 2008). Further, it has been found that socially excluded people are more easily persuaded by an artificial agent to change their behavior (Ruijten, Midden & Ham, 2015).

Similarly, our findings do not provide support for the set of theories that claim that artificial agents are distractors of learning, such as cognitive load theory (Sweller 2004; Sweller, Ayres, & Kalyuga, 2011). That is, the current research, provides evidence against cognitive load theory, when the type of task being modeled is considered. More specifically, our findings reveal the effectiveness of modeling by an artificial in increasing learners' self-efficacy, affective beliefs, and task performance, while minimizing performance-related cognitive load for psychomotor tasks as opposed to purely cognitive tasks. However, some support of this set of theories was found for cognitive learning. That is, the current research provided some evidence that an artificial model can negatively affect learners' recall regardless of the type of learning task it demonstrates. However, we could not to replicate this finding, as recall in Study 2 was measured in a different way from Study 1 (i.e., in Study 2 there were less multiple-choice questions and additional use of gap-filling questions). Therefore, we argue that future research could further examine the impact of the visibility of an artificial agent as a model on cognitive learning, by having as reference how recall was measured in Study 1. Furthermore, we propose that theories opposing the instructional benefit of an artificial agent, like cognitive load theory could benefit by specifying conditions under which the coexistence of visualizations and verbal explanations (as in the case of modeling by ana artificial agent) can facilitate or hamper learning. Our findings suggest that such a condition might be the type of learning outcome (cognitive, psychomotor).

The current work has also practical implications. In more details, our studies provide practical knowledge on how to optimally design artificial agents as models to achieve positive learning outcomes based on the type of task at hand. Specifically, the visual presence of an artificial model is recommended for psychomotor tasks as it positively influences task performance, self-efficacy, and affective learning, as it also minimized task performance-related cognitive load. To the contrary the inclusion of an artificial model is unnecessary when the tasks being designed are cognitive in nature. Lastly, the inclusion of an artificial model to the learning process becomes even more important as the level of complexity of the psychomotor task increases.

Overall, our two studies are important as they draw attention to and increase our understanding on the conditions under which an artificial agent can facilitate learning. These studies contribute to the development of artificial agents as powerful educational tools that one day can help improve education and lifelong learning.

Chapter 4

The Effect of an artificial agent's vocal expressiveness on immediacy and learning outcomes[15]

[15]This chapter is based on:
Fountoukidou, S., Matzat, U., Ham, J., & Midden, C. (2021). *The effect of an artificial agent's vocal expressiveness on immediacy and learning outcomes.* Manuscript submitted for publication.
Fountoukidou, S., Matzat, U., Ham, J., & Midden, C. (2019). Effects of a virtual model's pitch and speech rate on affective and cognitive learning. In E. Karapanos, E. Kyza, H. Oinas-Kukkonen, P. Karppinen, & K. T. Win (Eds.), *Persuasive Technology: Development of Persuasive and Behavior Change Support Systems - 14th International Conf erence, PERSUASIVE 2019, Proceedings* (pp. 16-27). (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Vol. 11433 LNCS). Springer

4.1 General Introduction

Pedagogical artificial agents are animated characters that are embedded in virtual learning environments. They are seen as potential tools to create a social presence that primes learners to deeply process the learning material (Kim & Baylor, 2015). Artificial teachers are expected to play a crucial role in the years to come, as the educational landscape is changing and reshaping by artificial intelligence (Chassignola, Khoroshavinb, Klimovac & Bilyatdinovac, 2018). Nonetheless, earlier literature draws a discouraging picture regarding the effects of artificial agents on learning (i.e., Schroeder, Adesope & Gilbert, 2013; Heidig & Clarebout, 2011). One of the reasons for the insignificant effects of artificial agents on learning might be that their potential for subtle nonverbal behaviors other than visual cues that we know from human-human interaction has received little attention (Krämer & Bente, 2010). In fact, it has been argued that nonverbal cues of artificial teachers could increase their social presence leading to learning gains (Baylor & Kim, 2009). However, there are many unanswered questions on how artificial agents' nonverbal cues should be implemented to accomplish such aims in multimedia settings.

A closer inspection of the behavior of human teachers could be taken into consideration in order to inform the behavioral design of pedagogical artificial agents. In traditional classroom settings, teachers' nonverbal behavior plays a crucial role in student learning. Specifically, some forms of teachers' nonverbal behavior are found to increase "nonverbal immediacy" (Andersen, 1979). The nonverbal immediacy concept refers to the ability of teachers to create psychological closeness with their students through nonverbal communication (Mehrabian, 1981). This concept is grounded in approach-avoidance theory, which asserts that people "are drawn toward the person and things they like, evaluate highly, and prefer; and they avoid or move away from things they dislike, evaluate negatively, or do not prefer" (Mehrabian, 1981, p. 1). Several nonverbal cues of teachers have been found to play a crucial role in student's learning, such as proximity, eye gaze, gestures, body position, facial and vocal expressiveness (Witt & Wheeless, 2001). Cumulative evidence has revealed that human teachers' nonverbal immediacy behavior promotes affective learning (student beliefs and motivations) and cognitive learning (immediate recall and perceived learning) (Ellis, Carmon & Pike, 2016; Witt, Wheeless & Allen, 2004).

Two major explanations have been proposed for explaining the effect of immediacy on learning: motivational theory (Christopher, 1990) and arousal–attention theory (Kelly & Gorham, 1988). Motivational theory suggests that some forms of teacher behavior may increase student (state) motivation by stimulating students, directing their efforts and, in turn, influence affective learning. Nonetheless, only few empirical studies have investigated this, providing support for motivational theory (i.e., Christopher, 1990; McCroskey, Richmond & Benett, 2006). Arousal-attention theory posits that that immediacy is associated with increased arousal, and if increased arousal focuses attention, increases the intensity of information processing, and improves memory (immediate recall and especially delayed recall) (Phaf & Wolters, 1986). Again, only few studies empirically tested the viability of arousal-attention theory (Comstock et al, 1995; Kelly & Gorham, 1988). Further examination of motivation and attention as mediators of path from immediacy to affective and/or cognitive change has been deemed necessary (Witt, Wheeless & Allen, 2004).

Visual nonverbal cues of artificial pedagogical agents, such as the use of gestures and facial expressions, have received increasing attention over the last years (Baylor & Kim, 2009). Nonetheless, it has been excessively shown that it is mainly the artificial agent's voice that is responsible for increased learning rather than its visual presence (Atkinson, 2002; Bente et al., 2008; Krämer & Bente, 2010). Previous work found that some of the characteristics of a speaker's voice that can affect learning (i.e., transfer and social perception) in multimedia settings are mechanization (human vs machine-synthesized voice) (Mayer et al., 2003; Atkinson et al., 2005), accent (native vs foreign accent) (Mayer et al., 2003), gender (male vs female voice) (Linek et al., 2010), dialect (regional dialect vs standard speech) (Rey & Steib, 2013) and slang (youth slang vs standard speech) (Schneider et al., 2015). However, there is limited evidence on whether and, more importantly, how artificial teachers' vocal nonverbal cues can influence multimedia learning outcomes.

In fact, only few studies have provided evidence that indicates that vocal expressiveness of an artificial agent can benefit learning. Liew and colleagues (2020) found that an enthusiastic voice of a virtual speaker led to higher transfer performance and social ratings when compared to a calm voice. However, the study featured an invisible narrator that had no visual cues, such as face and body. Thus, it is not known whether the positive effect of vocal expressiveness can also manifest in a multimedia environment presented by an on-screen artificial agent that inevitably involves visual cues. Further, two studies examined vocal nonverbal cues of a robot and found an effect of an expressive voice, as compared to a flat voice, on affective and cognitive learning (Westlund et al., 2017; Kennedy, Baxter & Balpaeme, 2017). However, it is not known whether results of studies that used physically embodied robots also apply to artificial agents that are not physically present. To the best of our knowledge, only one study investigated effects of a vocally expressive artificial agent (i.e., use of additional pauses, louder voice and better enunciated words) on learning, as compared to a non-vocally expressive artificial agent (Valetsianos 2009). Results provided evidence of a benefit of vocal expressiveness on affective and cognitive learning. However, since the study employed a quasi-experimental design, one limitation pertains to its low internal validity (Grabbe, 2015). What is more, further delineations of verbal expressiveness such as pitch tone, pitch variation, and speech rate require further exploration (Valetsianos, 2009). Thus, the aim of the current work is to examine whether vocal expressiveness (operationalized as pitch tone, pitch variation and speech rate) of an artificial agent can create immediacy and enhance learning, as it is proposed by communication literature (i.e., Ellis, Carmon, & Pike, 2016; Witt, Wheeless & Allen, 2004). What is more, this research goes one steps further and test whether the proposed underlying mechanisms of attention and motivation can explain the anticipated effect of vocal expressiveness on learning outcomes.

More recently, it has been argued that the inclusion of nonverbal cues of artificial agents is not a panacea, and caution is needed when constructing such cues, as they can be detrimental to learning (Dehn & Van Mulken, 2000; Clark & Choi, 2006; Woo, 2009; Frechette & Moreno 2010). This is because they could impose an additional processing burden, which is known as extraneous cognitive load, on working memory, because learners have to attend to nonverbal cues by an expressive artificial agent (Sweller 2004). In the present work, we employed an artificial agent that adopted the role of a model (Fountoukidou, Ham, Matzat & Midden,2019). We argue that vocal expressiveness of an artificial model, facilitates, rather than hinders, learning. Our claim is based on the cognitive theory of multimedia learning, according to which there are two separate channels (auditory and visual) for processing information that both have a limited processing capacity (Mayer, 2002). Due to the fact that modeling requires a substantial amount of processing to take place in the visual channel (i.e., demonstration), we argue that the inclusion of vocal nonverbal cues, balances the processing demands, since they are being processed in the auditory channel and not in the visual channel.

4.1.1 The current work

The current work maintains that vocal expressiveness of an artificial agent can play a crucial role in the learning process leading to an increase in affective and cognitive learning. In more detail, in the current research we report about two studies. The aims of Study 1 are threefold: 1) to examine the effect of an artificial agent's vocal expressiveness on immediacy; 2) to test whether an artificial agent showing strong vocal expressiveness (i.e., higher pitch tone, more pitch variation, higher speech rate) will enhance affective and cognitive learning (immediate recall and perceived cognitive learning), as compared to an artificial agent that shows weak vocal expressiveness (i.e., lower pitch tone, less pitch variation, lower speech rate); 3) to examine whether the underlying mechanisms of motivation and attention explain the effect of immediacy (and thereby also of vocal expressiveness) on the two learning outcomes.

It has been argued that when it comes to cognitive learning, immediacy mainly impacts delayed recall (Phaf & Wolters, 1986). However, only few studies examined the effects of immediacy on delayed recall, and their findings are mixed and inconclusive (Comstock et al., 1995; Titsworth, 2001). Building on the results of Study 1, the purpose of Study 2 was to extent these findings, by examining: 1) the effect of vocal expressiveness (strong versus weak) of an artificial agent on immediacy and delayed recall; 2) whether immediacy mediates the effect of vocal expressiveness on delayed recall.

4.2 Study 1

In this work it is argued that vocal expressiveness of an artificial agent (strong versus weak) plays a crucial role in increasing learning outcomes. In Study 1 we predict that an artificial agent that shows strong vocal expressiveness will increase perceptions of immediacy, as compared to an artificial model that shows weak vocal expressiveness (H1). We further predict that strong vocal expressiveness improves affective learning (H2) and cognitive learning (H3), when compared to weak vocal expressiveness. Lastly, we hypothesize motivation to mediate (part of) the effect of immediacy on affective learning (H4) and attention to mediate (part of) the effect of immediacy on cognitive learning (H5).

4.2.1 Method

Participants

One-hundred-forty-four participants participated in this study. (38% females and 62% males). The majority of the population were students from Eindhoven University of Technology. One-hundred-forty-four individuals participated in the study. Most of the participants were students from Eindhoven University of Technology. Specifically, 92 participants (63%) were educated to undergraduate level or higher, 45 participants (31.2%) had completed high school and seven participants (4.8%) chose not to disclose their education.

The study used a between-participants design, with the participants being randomly assigned to one of the two experimental conditions: artificial modeling with strong expressiveness and artificial modeling with weak vocal expressiveness. The dependent variables of the study were immediacy, affective learning, and cognitive learning (perceived and actual). Inclusion criteria were participants' English fluency. The experiment lasted for approximately 30 minutes, and participants were compensated for their participation (5 euros).

Materials

The 3D animated artificial agent, employed in this research, was designed using the CrazyTalk 8 software (see Subsection 2.2.1)

The study's instructional script discusses an eye-tracking software, called GazeTheWeb (GTW). GTW is a gaze-controlled web-browser, which work with an eye-tracking hardware (Menges et al., 2017). (see Subsection 2.2.1 for more details).

The vocal parameters that used to distinguish the strong vocal expressiveness from the weak vocal expressiveness were pitch (pitch tone and pitch variation) and speech rate. The actor's voice was recorded using Audacity software. Pitch analysis of these audio recordings was performed with the use of Praat software.

Artificial agent

The artificial agent used in the current study was developed in such a way so as to share some common characteristics with participants in terms of their appearance. These characteristics were derived from earlier literature (Rosenberg-Kima, Baylor, Plant & Doerr, 2008). Therefore, since the majority of the participants were students at a Dutch University, the agent was designed to be young (<30 years old), attractive (in terms of the artificial agent's facial characteristics) and "cool" (in terms of the artificial agent's clothing and hairstyle). What is more, the artificial agent intentionally lacked strong facial expressions.

Concerning its educational role, the artificial agent appeared as a model, demonstrating the GTW system's functionalities by moving his head and eyes, while providing verbal explanations at the same time (see Figure 10).

The design of the artificial agent, both in terms of appearance and educational role was identical in both experimental conditions, and the only difference was the level of the artificial gent's vocal expressiveness (explained in the subsection 2.2.3).

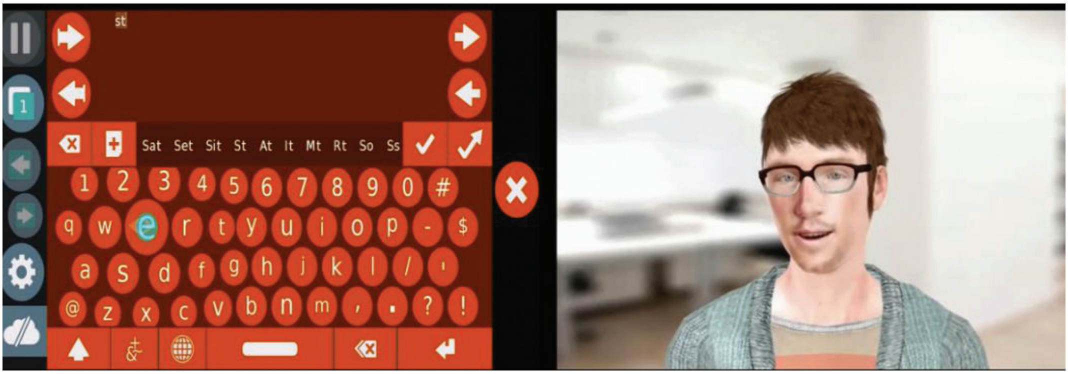


Figure 10 Artificial modeling: on the right side, the agent appeared to demonstrate an action (i.e., typing), while providing verbal explanations; on the left side, the light blue highlights the effect of the artificial agent 's action.

Instructional script

Initially, an instructional script was created, which familiarised participants to the use of a novel, gaze-controlled web browser (GTW), that was unknown to the study's population. Then, two versions of this script were developed. The only difference between these two versions was the level of the artificial agent's vocal expressiveness (strong vs. weak). That is, the two versions differed in terms of average pitch tone, pitch variation, and speech rate (see below for more details). A male actor, whose voice was recorded an and borrowed by the artificial agent, performed both instructional versions. The selection of the voice actor was based on two requirements: clear English pronunciation and good voice acting skills.

Vocal expressiveness

The study concurrently manipulated the vocal parameters of a) pitch tone, b) pitch variation, and c) speech rate, so as to create two different levels of vocal expressiveness (strong vs. weak). The decision of the aforementioned vocal parameters was based on earlier results, suggesting that the combination of a speaker's temporal (i.e., speech rate) and expressive (i.e., pitch) vocal features has the greatest impact on both emotions and cognition (Breitenstein, Lancker & Daum, 2001).

Pitch is defined as the degree of highness or lowness of a tone, determined by the vibration of the vocal folds (i.e., the faster the vibration per second (Hz), the higher the pitch). It is generally measured as the fundamental frequency of the sound wave. There is no globally optimal pitch tone, and it is determined by factors such as culture and context (Gudykunst, Ting-Toomey & Chua, 1988). Nevertheless, according to general guidelines, the average fundamental frequency of a male adult's speech is 120 Hz (Hollien, & Shipp, 1972; Hsiao et al., 1994; Mizuno & Nakajima, 1998). Thus, in this study, the pitch tone boundary so as to differentiate the two levels of vocal expressiveness (strong and weak) was set at ~120 Hz. All in all, according to our calculations, the average pitch tone of the strong vocal expressiveness condition was 260Hz, while the average pitch tone of the weak vocal expressiveness condition was 115 Hz.

In addition to the pitch tone, pitch variation (i.e., intonation) was also manipulated. That is, there was more pitch variation (i.e., voice rises and then falls before it rises again), in the strong vocal expressiveness condition than in the weak vocal expressiveness condition. Hence, the weak vocal expressiveness condition, apart from its lower pitch tone, was also constructed to be "flat" in terms of pitch variation. Concerning the strong vocal expressiveness condition, the artificial agent's voice raising was congruent with important information that participants needed to recall. The development of pitch tone and variation in the strong vocal expressiveness condition was intentionally prepared so as to emphasize, both, affective nonverbal communication (i.e., speaker's feeling and attitude conveyance) and cognitive nonverbal communication (i.e., help in the encoding of a new information) (Frechette & Moreno, 2010).

Speech rate is defined as the speed at which one speaks. It's calculated by the number of words spoken in a minute. An average number of words per minute (wpm) can vary hugely. This is because speech rate is inextricably bound to the speaker's culture, geographical location, subject matter, gender, emotional state, fluency, profession or audience. Nonetheless, according to some general guidelines, that conversational speech generally falls between 125 wpm at the slow end, to 150 wpm in the fast range (Simonds et al., 2006). Overall, the speech rate in the strong vocal expressiveness condition was 133 wpm, as opposed to 119 wpm in the weak vocal expressiveness condition. Inevitably, this manipulation lead to a relatively small difference in the video duration between the two conditions (9, 5 minutes in the strong vocal expressiveness condition as opposed to 10 minutes in the week vocal expressiveness condition).

Measures

Regarding the manipulation check, participants were asked to evaluate the artificial agent's

vocal expressiveness via a self-constructed scale, assessing the vocal parameters of pitch tone, pitch variation and speech rate. This scale consisted of three items and it was administered through a 7-point semantic differential scale (1) use of high vs. low tone of voice; 2) use of vocal variety vs. flat voice; 3) use of fast vs. slow speech rate). We constructed an acceptable measure of perceived vocal expressiveness (Cronbach's α = .68), by averaging participants' answers to this set of questions.
Nonverbal immediacy (H1), was assessed using a scale consisted of six items and it was administered through a 7-point semantic differential scale (i.e., pleasant vs. unpleasant voice, enthusiastic vs. boring voice etc.). This scale was adapted from earlier versions measuring not only vocal but a variety of other nonverbal cues (i.e., facial expressiveness) (Mehrabian, 1981; Richmond, Gorham & McCroskey, 1987; Richmond, McCroskey & Johnson, 2003; Servilha & Costa, 2015) (see Appendix C.1 for the items on this scale). We constructed a reliable measure of nonverbal immediacy (Cronbach's α = .87), by averaging participants' answers to this set of questions.

Affective learning (H2) was assessed by asking participants to estimate three components of their affective perceptions towards the instructional material, towards the artificial instructor, and the likelihood of following the same artificial instructor for other instructional videos. These components of affective learning were administered through a 7-point semantic differential scale (Andersen, 1979; Scott & Wheeless, 1975) (see Appendix C.1 for the items on these scales). We constructed reliable measures of affective perceptions towards the instructional material (Cronbach's a = 0.86), towards the artificial instructor (Cronbach's a = 0.88) and likelihood of following the same artificial instructor for other instructional videos (Cronbach's a = 0.90), by averaging participants answers to each set of questions.

Cognitive learning (H3) was assessed both objectively (recall test) as well as subjetively (perceived cognitive learning). Specifically, recall of the content of the instructional video was measured as an index of cognitive learning, and it was measured through a self-constructed recall test. This recall test contained two methods of knowledge assessment: 1) A "fill-in-the-blanks" test consisted of nine recall items, and, 2) a multiple-choice test consisted of 18 questions (see Appendix C.1 for the recall test). For the gap filling test, participants were asked to recall keywords (exact words or synonyms) spoken by the artificial agent during the video, and to fill in the blanks of a written transcript. For the multiple-choice test, participants were asked to answer a series of questions by selecting the correct answer amongst four optional answers. We constructed two measures of cognitive learning by counting participants' number of correct answers to each test separately. The participants' performance scores were calculated by two researchers independently. There was a 100% agreement on the performance scores between the two researchers.

Furthermore, perceived cognitive learning was assessed by asking participants' to answer two, 7-point scale, questions (Richmond, Gorham & McCroskey, 1987). A "learning loss" score was, then, computed by subtracting the score on the first question (i.e., How much did you learn during the video lesson?) from the score of the second question (i.e., How much do you think you could have learned from this video had you had this ideal instructor?), indicating a learner's overall perceived cognitive learning score. Reliability using this measure in previous research was reported at .94 (Gorham, 1988). Overall, learning loss score has been widely used in communication research as an index of cognitive learning (e.g., Chesebro & McCroskey, 2000).

The main dependent variable of the fourth hypothesis was (state) motivation. Motivation was assessed by asking participants to answer nine questions on how they felt about the instructional video they watched. The questionnaire was administered in a 7-point semantic differential scale taken from Christophel (1990) (we selected the nine out of the 12 questions that were relevant to our study) (see Appendix C.1 for the items on this scale). We constructed a reliable measure of motivation (Cronbach's α = .91), by averaging participants' answers to each set of questions.

The main dependent variable of the fifth hypothesis was attention. Attention was assessed by asking participants to answer four questions, administered in a 7-point scale, about the level of their attention to the instructional video (Yi & Davis, 2003) (see Appendix C.1 for the items on this scale). We constructed a reliable measure of motivation (Cronbach's α = .86), by averaging participants' answers to each question.

Lastly, we explored whether vocal expressiveness would influence how learners perceive the likeability of an artificial agent. Furthermore, and more importantly, likeability of the artificial agent was taken into consideration so as to test whether the influence of vocal expressiveness on affective learning is still mediated by immediacy even when another possible mediating path is considered (i.e., likeability of the artificial agent). The agents' likeability was measured with a subscale of the "Godspeed" scale, developed to assess key concepts of Human-Computer interaction (Bartneck, Croft & Kulic, 2008) (see Appendix C.1 for the items on this scale). The scale was formatted in a 7-point semantic differential, scale. We constructed a reliable scale of likeability (Cronbach's α = .93) by averaging participants' answers to each set of questions.

Procedure

Participants were welcomed in the main hall of the lab building. Each participant was required to read and sign an informed consent form, explaining the general aim of the study and their willingness to participate. Next, they were randomly assigned to one of the two experimental conditions and they were requested to watch an instructional video on a computer monitor regarding the use of GTW browser. The video screen was split into two sides: on the right-hand side, an artificial agent appeared to use the GTW system by moving the head and eyes, while explaining the system functionalities being demonstrated; on the left-hand side, the actual system was displayed, showing participants the effects of the artificial agent's actions on the system in real time (i.e., Figure 10). Participants in both conditions were provided with an identical instructional video, with the only difference being the level of vocal expressiveness in terms of pitch tone, pitch variation and rate.

As a next step, participants were invited to answer an online survey and a recall test. Lastly, they were debriefed, paid and thanked for their participation.

4.2.2 Results

Manipulation check: An independent sample t-test analysis was conducted to check the study's manipulation of vocal expressiveness (i.e., perceptions of vocal parameters of pitch tone, pitch variety and speech rate) between the strong vocal expressiveness and weak vocal expressiveness condition. As expected, the results revealed a statistically significant effect on vocal expressiveness, $t(142) = 9.211$, $p < .001$, with participants in the strong vocal expressiveness condition ($N = 78$, $M = 3.7$, $SD = .9$) to report stronger perceptions of vocal expressiveness as compared to participants in the weak vocal expressiveness condition ($N = 66$, $M = 2.3$, $SD = .8$).

Immediacy: To test H1, an independent sample t-test analysis was conducted to examine the effect of the level of the artificial model's vocal expressiveness on individuals' perceptions of immediacy. Results, supported our hypothesis, revealing a statistically significant effect, $t(142) = 6.873$, $p < .001$, with participants in the strong vocal expressiveness condition ($N = 78$, $M = 4.1$, $SD = 1.3$) to report higher perceptions of immediacy as compared to participants in the weak vocal expressiveness condition ($N = 66$, $M = 2.8$, $SD = 1.0$).

Affective learning: To test H2, a one-way multivariate analysis of variance (MANOVA) was conducted to examine the effect of the level of the artificial model's vocal expressiveness on the three affective dependent variables (affective perceptions towards the instructional material, towards the artificial instructor, and perceived likelihood to follow the same instructor on other instructional material). The results revealed a statistically significant effect of the level of the artificial model's vocal expressiveness on the three dependent variables combined, Wilk's $\Lambda = .85$, $F(3,140) = 7.88$, $p < .001$, $\eta p^2 = .14$.

In line with our hypothesis, separate univariate ANOVAs on the outcome variables revealed a significant treatment effect on: 1) affective perceptions towards instructional material, $F(1, 142) = 7.23$, $p < .01$, $\eta p^2 = .48$, with participants' affect towards the instructional material to be more positive in the strong vocal expressiveness condition ($N = 78$, $M = 5.6$, $SD = .9$), as compared to participants in the weak vocal expressiveness condition ($N = 66$, $M = 5.1$, $SD = 1.0$); 2) affective perceptions towards the artificial instructor, $F(1, 142) = 21.39$, $p < .001$, $\eta p^2 = .13$, with participants' affect towards the artificial instructor to be more positive in the strong vocal expressiveness condition ($N = 78$, $M = 5.4$, $SD = 1.1$), as compared to participants in the weak vocal expressiveness condition ($N = 66$, $M = 4.5$, $SD = 1.1$); 3) the likelihood of following the same artificial instructor for other instructional material, $F(1,142) = 17.82$, $p < .001$, $\eta p^2 = .11$, with participants' perceived likelihood to follow the same instructor on other instructional material to be more positive in the strong vocal expressiveness condition ($N = 78$, $M = 4.4$, $SD = 1.5$), as compared to participants in the weak vocal expressiveness condition ($N = 66$, $M = 3.3$, $SD = 1.4$).

Cognitive learning: To test H3, a one-way multivariate analysis of variance (MANOVA) was conducted to examine the effect of the level of the artificial model's vocal expressiveness on individuals' recall. Recall was measured with a fill-in-the-blanks test and a multiple-choice test. The results revealed a statistically significant effect of the level of the artificial model's vocal expressiveness on the two dependent variables combined, Wilk's $\Lambda = .94$, $F(2, 141) = 4.33$, $p = .01$, $\eta p^2 = .6$.

In line with our hypothesis, separate univariate ANOVAs on the outcome variables revealed a significant treatment effect on fill-in the-blanks test, $F(1, 142) = 5.25$, $p = .02$, $\eta p^2 = .36$, with participants' recall performance to be better in the strong vocal expressiveness condition ($N = 78$, $M = 7.9$, $SD = 2.7$), as compared to participants in the weak vocal expressiveness condition ($N = 66$, $M = 6.9$, $SD = 2.2$). Results showed a non-significant treatment effect on the multiple-choice test, $F\,(1, 142) = .31$, $p > .05$, between participants in the strong vocal expressiveness condition ($N = 78$, $M = 10.4$, $SD = 2.8$) and participants in the weak vocal expressiveness condition ($N = 66$, $M = 10.6$, $SD = 2.9$).

Perceived cognitive learning (learning loss): An independent sample t-test analysis was conducted to examine the effect of the level of the artificial model's vocal expressiveness on individuals' perceptions of learning. As expected, the results revealed a statistically significant effect, $t(142) = -2.36$, $p = .02$, $r = .20$, with participants in the strong vocal expressiveness condition ($N = 78$, $M = .41$, $SE = .1$) to report less learning loss (therefore more perceived cognitive learning), as compared to participants in the weak vocal expressiveness condition ($N = 66$, $M = .83$, $SE = .13$).

Artificial agent's likeability: An independent sample t-test analysis was conducted to examine the effect of the level of the artificial model's vocal expressiveness on individuals' judgments about the artificial agent's likeability. The results revealed a statistically significant effect, $t(142) = 4.44$, $p < .001$, with participants' judgments on the artificial agent's likeability to be more positive in the strong vocal expressiveness condition ($N = 78$, $M = 5.4$, $SD = 1.0$) as compared to participants' judgments in the weak vocal expressiveness condition ($N = 66$, $M = 4.6$, $SD = 1.0$).

Path analyses

Our final aim was to test whether motivation and attention could explain parts of the anticipated effect of immediacy on affective learning (affective perceptions towards the instructional material, towards the artificial instructor and the likelihood of following the same artificial instructor for other instructional material) (H4) and cognitive learning (fill-the-blanks recall test and perceived cognitive learning) (H5). Vocal expressiveness, our manipulated factor, was included as predictor of immediacy. Hence, path analyses were conducted in STATA 14 to test the model (see Figure 11). This type of analyses provides a comprehensive picture of the nature of the associations between the predictor and dependent variables of interest. The overall fit of the models was assessed by the chi-square goodness of fit (X2), comparative fit index (CFI), and root mean square error of approximation (RMSEA). Detailed results of the path analysis for each affective and cognitive learning variable are reported in the following subsections.

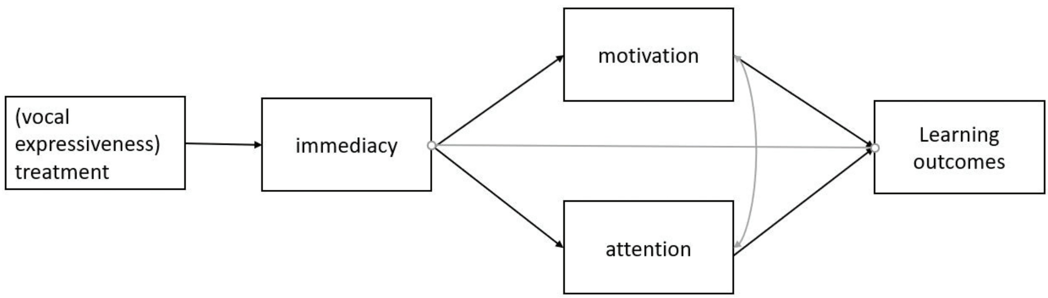


Figure 11 Hypothesized path analysis model.

Path analyses of affective learning outcomes

Concerning affective learning, three path analyses models were tested for the three ffective learning outcomes (affect towards the artificial teacher, likelihood of following the same artificial instructor for other instructional material and affect towards the instructional material). Results are presented in Figures 12, 13 and 14. In more detail, as seen in Figure 12, vocal expressiveness was a positive significant predictor of immediacy. Further, immediacy[16] was found to be a positive significant predictor of affect towards the artificial teacher. Motivation was a significant positive predictor of affect towards the artificial teacher. To the opposite, the path coefficient from attention to affect towards the artificial teacher was positive but non-significant. These findings suggest that motivation partially mediate[17] the effect of immediacy on affect towards the artificial teacher. Attention, however, does not show any mediating effect. Overall, the model explained 43,8% of affect towards the artificial teacher. Similar results were found for participants' likelihood of following the

[16]For further clarification, a mediation analysis was conducted. Immediacy remained a mediator of the effect of vocal expressiveness on all three affective outcomes even when likeability of artificial agent was included as a second mediator. Furthermore, likeability also mediated the effect of vocal expressiveness on all three forms of affective learning.

[17]For further clarification, a mediation analysis was conducted. The total effect of immediacy on affect towards the artificial teacher was found to be significant ($\beta = .57$, $p < .001$). The indirect effect through motivation was also significant ($\beta = .166$, 95% *CI* [.048, .338]). The indirect effect through attention was non-significant ($\beta = .047$, 95% *CI* [-.020, .118]).

same artificial instructor for other instructional material, as seen in Figure 13. Again, these results suggest that only motivation partially mediates[18] the effect of immediacy on likelihood of following the same artificial instructor for other instructional material. Overall, the model explained 54,6% of likelihood of following the same artificial instructor for other instructional material. However, as Figure 14 illustrates, both motivation and attention appear to be significant positive predictors of affect towards the instructional material, while the path coefficient from immediacy to affect towards the instructional material is positive but non-significant. We can conclude that affect towards the instructional material is fully mediated[19] by both attention and motivation. Overall, the model explained 36% of affect towards the instructional material. To summarize, the results of the path analyses supported our hypothesis 4, showing that motivation explain (part of) the effect of immediacy on affective learning outcomes.

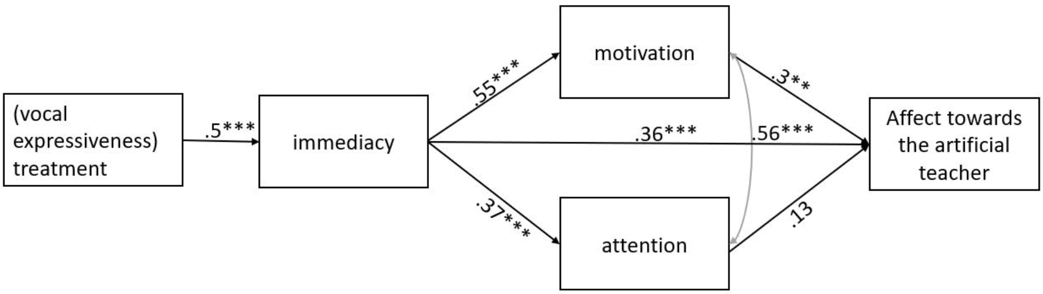


Figure 12 Results of the path analysis for affect towards the artificial teacher. Standardized coefficients are presented. Grey lines indicate non-hypothesized relationships, **p* <. 05, ***p* < .01, ****p* < .001. Overall model fit statistics: $X^2(3) = 2.017$, $p > .05$; *RMSEA* < .001; *CFI* > .95.

[18]For further clarification, a mediation analysis was conducted. The total effect of immediacy on likelihood of following the same artificial instructor for other instructional material was found to be significant ($\beta = .58$, $p < .001$). The indirect effect through motivation was also significant ($\beta = .241$, 95% *CI* [.138, .348]). The indirect effect through attention was non-significant ($\beta = .05$, 95% *CI* [-.001, .104]).

[19]For further clarification, a mediation analysis was conducted. The total effect of immediacy on affect towards the instructional material was found to be significant ($\beta = .40$, $p < .001$). The indirect effect through motivation was also significant ($\beta = .172$, 95% *CI* [.053, .291]). Similarly, the indirect effect through attention was also significant ($\beta = .095$, 95% *CI* [.019, .176]).

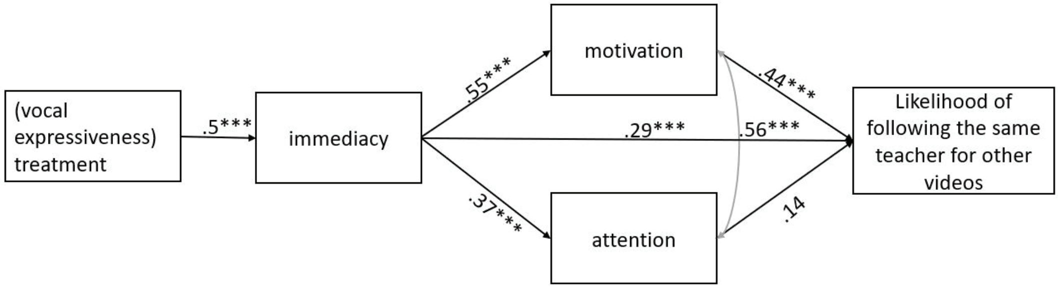


Figure 13 Results of the path analysis for likelihood of following the same artificial instructor for other instructional material. Standardized coefficients are presented. Grey lines indicate non-hypothesized relationships, *$p < .05$, **$p < .01$, ***$p < .001$. Overall model fit statistics: $X^2(3) = 0.721$, $p > .05$; $RMSEA < .001$; $CFI > .95$.

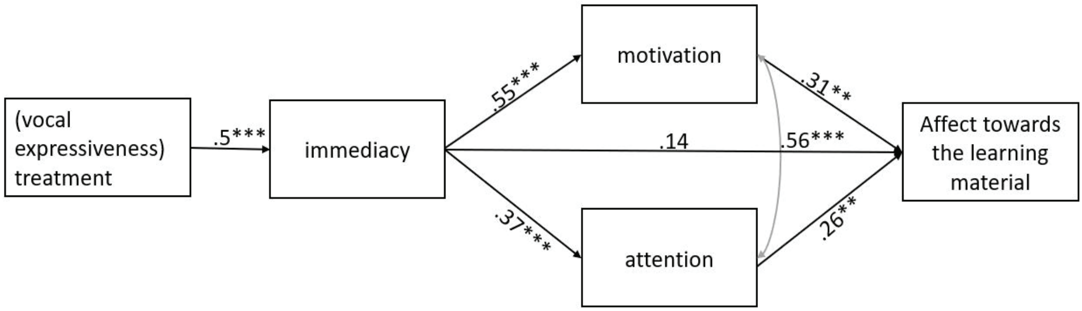


Figure 14 Results of the path analysis for affect towards the artificial teacher. Standardized coefficients are presented. Grey lines indicate non-hypothesized relationships, *$p < .05$, **$p < .01$, ***$p < .001$. Overall model fit statistics: $X^2(3) = 0.138$, $p > .05$; $RMSEA < .001$; $CFI > .95$.

Path analyses of cognitive learning outcomes

Concerning cognitive learning, two path analyses were conducted for the two cognitive learning outcomes (perceived cognitive learning and recall scores). Results are presented in Figures 15 and 16. In more details, as seen in Figure 15, immediacy was found to be a negative significant predictor of perceived cognitive learning (it is negative as it has been measured as learning loss). Nonetheless the path coefficients from both motivation and attention to perceived cognitive learning were non-significant. These findings suggest that neither motivation nor attention mediated the effect of immediacy on perceived cognitive learning. Lastly, as illustrated in Figure 16, neither immediacy nor motivation and attention were found to be significant predictors of participants' recall. The results of the path analyses did not provide support for our hypothesis 5, as attention was not found to be a mediator.

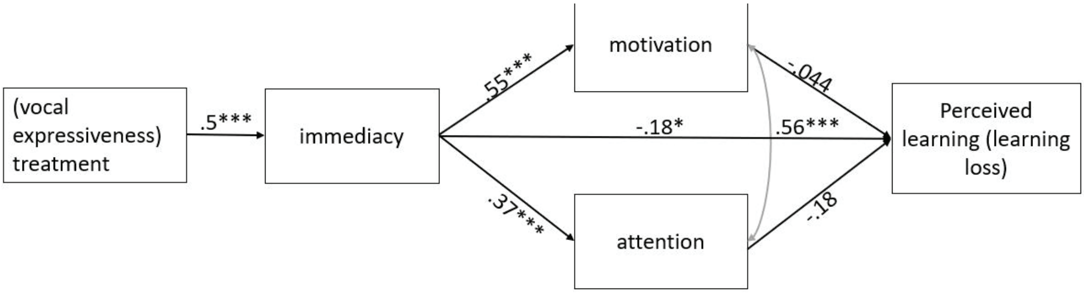


Figure 15 Results of the path analysis for perceived cognitive learning (learning loss). Standardized coefficients are presented. Grey lines indicate non-hypothesized relationships, $^*p < .05$, $^{**}p < .01$, $^{***}p < .001$. Overall model fit statistics: $X^2(3) = 0.754$, $p > .05$; $RMSEA < .001$; $CFI > .95$.

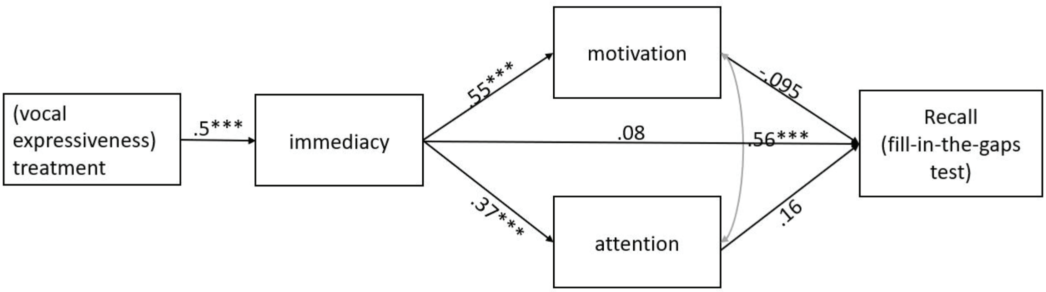


Figure 16 Results of the path analysis for recall scores. Standardized coefficients are presented. Grey lines indicate non-hypothesized relationships, $^*p < .05$, $^{**}p < .01$, $^{***}p < .001$. Overall model fit statistics: $X^2(3) = 4.325$, $p > .05$; $RMSEA < .08$; $CFI > .95$.

4.2.3 Discussion

The current research investigated the influence of an artificial model with strong vocal expressiveness, as compared to the (same) artificial model with weak vocal expressiveness, as a means to increase nonverbal immediacy and subsequently to enhance individuals' affective and cognitive learning. What is more, we tested the proposed underlying mechanisms of motivation and attention to explain the anticipated effect of artificial model's immediacy on affective and cognitive learning respectively.

Our results supported our first hypothesis, showing that an artificial model that shows strong vocal expressiveness can increase perceptions of immediacy, as compared to an artificial model that shows weak vocal expressiveness. This is in accordance with the vast body of literature on human teachers' nonverbal immediacy (i.e., Mehrabian, 1981; Witt, Wheeless & Allen, 2004; Ellis, Carmon & Pike, 2016). Therefore, the study provides evidence that, similar to a human teacher, strong nonverbal cues, such as vocal expressiveness, can influence learners' perceptions of psychological closeness (i.e., immediacy) with an artificial teacher.

Furthermore, the study's results supported our second hypothesis, showing that an artificial model with strong vocal expressiveness can enhance individuals' affective learning, as compared to an artificial model with weak vocal expressiveness. Specifically, according to the current study's findings, participants in the strong vocal expressiveness condition indicated increased affective perceptions towards the instructional material, the artificial teacher as well as an increased likelihood of following the same artificial instructor for other instructional videos. Undeniably, students' affective experiences are important as they have been found to be the central mediator linking teaching behaviors to student reports of learning and other important classroom outcomes (Bolkan, 2015).

Next, results provided partial support for our third hypothesis, showing that an artificial model that shows strong vocal expressiveness can impact learners' recall when assessed with the fill-in-the-blanks test as compared to an artificial model that shows weak vocal expressiveness. To the contrary, no evidence for a difference between the two levels of vocal expressiveness was found when recall was assessed with a multiple-choice test. The reason why we found an effect on only one cognitive test might be due to the difference between the gap filling and multiple-choice tests as methods of knowledge assessment. Our results are in accordance with earlier studies, which found a significant difference in learners' scores of the two types of tests, with learners' multiple-choice scores to be significantly better than their gap filling scores (Medawela et al., 2018; Utari, 2013). These studies did not test why this is the case, thus more exploration of such findings required.

Additionally, we found evidence that an artificial model with strong vocal expressiveness also affects perceived cognitive learning. Specifically, the study's findings provide evidence that strong vocal expressiveness has a positive influence on perceptions of learning as compared to weak vocal expressiveness. Despite the fact that perceived cognitive learning is not as strong as measuring actual cognitive learning, earlier research has found a moderately strong validity coefficient between students' performance on a recall test and reports of how much they believed they learned during a lecture (Chesebro & McCroskey, 2000).

Further, confirming our fourth hypothesis (concerning mediating psychological processes of motivation on affective learning), findings of this study suggest that learners' motivation explains part of the effect of immediacy (and, thus, of vocal expressiveness) on affective learning outcomes. Thus, the current study provides support for motivation theory (Christophel, 1990), which argues what this study's findings revealed; strong nonverbal cues, such as vocal expressiveness, increase perceptions of immediacy, which has a positive effect on motivation leading to enhanced affective learning.

However, contrary to our fifth hypothesis (concerning mediating psychological processes of attention on cognitive learning) neither immediacy nor attention appear to explain the effect of vocal expressiveness on immediate recall. Further though immediacy was found to be predictor of perceived cognitive learning, attention was not. Thus, results of the current Study 1 do not provide evidence in favour of arousal–attention theory, which posits that immediacy stimulates arousal, which, thereby, affects attention and memory leading to greater cognitive learning (Kelly & Gorham, 1988). However, other psychological studies demonstrated the important role of arousal in altering both attention and consolidation of memories (Christianson & Loftus, 1991; Eysenck, 1976; Heuer & Reisberg, 1992; Revelle & Loftus, 1992). Such studies suggest that if arousal acts specifically on memory consolidation, its influence magnifies following a delay, as consolidation is a process that occurs over time. Thus, future research might examine whether attention mediates the effect of immediacy on delayed recall (i.e., one-week past treatment). How the effect of vocal expressiveness on immediate recall can be explained, is still unclear.

Collectively, the findings of Study 1 are in line with past work that emphasised the vital role of human teachers' nonverbal cues in increasing students' affective, cognitive, and perceived cognitive learning in traditional classroom settings (Witt, Wheeless & Allen, 2004). Similarly, the study showed that a teacher's nonverbal cues are related to learning outcomes because they promote immediacy. Nonetheless, earlier studies examined a plethora of nonverbal cues together (i.e., facial cues, posture) as it is difficult to disentangle various cues from each other when human teachers are employed. In addition, the majority of past studies utilized a survey research design, which has been argued to be of limited usefulness when it comes to making conclusions related to students' learning (Comstock et al., 1995; Hess & Smythe, 2001; Witt & Wheeless, 2001). Therefore, an advantage of the current Study 1 is that by employing artificial agents, it was able to experimentally show the single effects of nonverbal cues, such as vocal expressiveness, on immediacy and learning outcomes. All in all, the fact that the same learning mechanisms were found between human and artificial teachers (i.e., alignment of these findings with human teachers' past research) suggests that the study's results are pertinent for human teachers too.

In addition, the current findings are in accordance with the few studies that have provided evidence of the positive effect of vocal expressiveness of artificial agents (Valetsianos, 2009) and robots (Westlund et al., 2017; Kennedy, Baxter & Balpaeme, 2017). However, the current Study 1 goes beyond these earlier studies, by examining the underlying mechanisms of the effect of vocal expressiveness on both affective and cognitive outcomes. This is, Study 1 revealed that vocal expressiveness is related to learning outcomes, because it reduces psychological distance, thus, promoting immediacy. What is more, this research further examined motivation and attention as mediators of the path from immediacy to affective and cognitive learning. These mechanisms were reported as potential explanations in earlier studies, though they were not empirically tested (Valetsianos, 2009). Lastly, showing that the effect of the combination of pitch and speech rate on learning has its own importance, as it can help designers' choice of vocal parameters when constructing vocal expressiveness of artificial teachers.

Despite the study's aforementioned advantages, caution is needed in generalizing the results beyond the study's population characteristics. This is because nonverbal cues vary culturally and contextually (Gudykunst, Ting-Toomey & Chua, 1988). Future research could explore whether different contexts (i.e., geographical location) would produce different results on learning. Furthermore, the artifcial model's vocal expressiveness consisted of both pitch (tone and variation) and speech rate. Future research could examine the single effect of each vocal parameter on learning outcomes. Another limitation of the study pertains to the short duration of multimedia learning (~10 min). Though artificial agent's strong vocal expressiveness was shown to increase affective and cognitive learning, the effects of repeated and prolonged exposure to nonverbal cues are not known.

4.3 Study 2

In Study 1 we found evidence indicating that an artificial agent with strong vocal expressiveness increased affective learning and perceived cognitive learning as compared to an artificial agent with weak vocal expressiveness. However, in Study 1, we found mixed effects regarding the effects of vocal expressiveness on actual cognitive learning (measured as immediate recall). Further, we examined the underlying process of the influence of artificial agent's vocal expressiveness on affective and cognitive learning. Specifically, regarding affective learning, we found support for motivation theory (Christopher, 1990), as findings revealed that vocal expressiveness enhanced immediacy, which in turn increased motivation and, then, increased affective learning. Regarding actual cognitive learning, we did not find evidence supporting attention-arousal theory (Kelly & Gorham, 1988), as results showed that neither immediacy nor attention had an effect on immediate recall.

Though only few studies empirically tested the viability of arousal-attention theory before (Comstock et al, 1995; Kelly & Gorham, 1988), other psychological studies, outside the educational research area, demonstrated the important role of arousal in altering both attention and consolidation of memories (Christianson & Loftus, 1991; Eysenck, 1976; Heuer & Reisberg, 1992; Revelle & Loftus, 1992). Such studies suggest that if arousal acts specifically on memory consolidation, its influence increases following a somewhat longer period, as consolidation is a process that occurs over time. Thus, in the current Study 2 we argue that a plausible reason for the mixed effects of vocal expressiveness on actual cognitive learning, as well as, the lack of evidence of immediacy as a mediator of vocal expressiveness on actual cognitive learning we found in Study 1, might be the fact that we examined immediate recall (i.e., measured immediately after treatment) rather than delayed recall (i.e. several days after treatment).

Unfortunately, little guidance exists in the literature concerning delayed effects of immediacy on learning. In the domain of educational research, a metanalytical review on the relationship between immediacy and learning outcomes with human teachers (Witt, Wheeless & Allen, 2004) revealed that only two studies measured delayed recall (Comstock et al., 1995; Titsworth, 2001), and their results were mixed and inconclusive. Specifically, in the study of Titsworth (2001) videotaped lectures manipulated a plethora of immediacy cues displayed by a human teacher (i.e., “we” versus “I” statements and enhanced versus minimal vocal expressiveness, facial expressions, gestures, and body movement). Findings showed that immediacy cues had an effect on delayed recall (i.e., one-week past treatment) when measured with the “detail test” (assessment of student ability to recall and describe specific facts and details learned from a lecture). However, no significant effects were found when delayed recall was measured with a “concept test” (assessment of student ability to abstract concepts and principles discussed in the lecture by answering multiple-choice questions). To the contrary, the study of Comstock et al. (1995) found that immediacy cues did not affect delayed recall (measured with true-false statements). Unfortunately details of their methodology are not reported (i.e., manipulation of immediacy, time of post-test). These contradictory findings point to the importance of continued research in order to understand how cognitive performance over time is associated with immediacy.

Therefore, in this follow-up study, Study 2, we aimed to investigate whether vocal expressiveness of an artificial agent has an effect on immediacy and delayed recall. Accordingly, we expected strong vocal expressiveness to increase perceptions of immediacy as compared to weak vocal expressiveness (H1). In line with Titsworth (2001), we further predicted that strong vocal expressiveness of an artificial agent would increase delayed recall. as compared to weak vocal expressiveness (H2). Lastly, we predicted that immediacy would mediate the effect of vocal expressiveness on delayed recall (H3).

4.3.1 Method

Participants

The participants of this study included 139 individuals (57 females and 82 males) and the majority of the population were students from Eindhoven University of Technology. Unfortunately, 35 of the participants did not respond to the delayed recall test they received by mail. This resulted in a total sample size of n=105 for the analysis of delayed recall. Of them, 46 were female (44%) and 59 were male (56%). Sixty-one participants (58%) were educated to undergraduate level or higher, 39 participants (37%) had only completed high school.

In the same ways as Study 1, Study 2 used a between-participants design, with the participants being randomly assigned to one of the two experimental conditions: artificial modeling with strong expressiveness and artificial modeling with weak vocal expressiveness. The dependent variables of the study were immediacy and delayed recall. Only participants who were fluent in English were allowed to take part in this study. The experiment lasted for approximately 30 minutes, with a follow-up questi onnaire after 12 days, for which participants received five euros as compensation for their participation.

Materials

The 3D animated artificial agent, employed in this research, was designed using the CrazyTalk 8 software.

In the same way as in Study 1, the instructional script used in Study 2 discusses an eye-tracking software, called GazeTheWeb (Menges et al., 2017).

In the same way as in Study 1, the vocal parameters that were concurrently used in Study 2 to distinguish the strong vocal expressiveness from the weak vocal expressiveness were pitch (pitch tone and pitch variation) and speech rate. The actor's voice was recorded using Audacity software. Pitch analysis of these audio recordings was performed with the use of Praat software.

For more details regarding the construction of the study's artificial agent, instructional script and vocal expressiveness, see the subsection 4.2.1.

Measures

Regarding the manipulation check, in line with Study 1, participants were asked to evaluate the artificial agent's vocal expressiveness via a self-constructed scale, assessing the vocal parameters of pitch tone, pitch variation and speech rate. This scale consisted of three items and it was administered through a 7-point semantic differential scale (1) use of high vs. low tone of voice; 2) use of vocal variety vs. flat voice; 3) use of fast vs. slow speech rate). We constructed an acceptable measure of perceived vocal expressiveness (Cronbach's $a = .64$), by averaging participants' answers to this set of questions.

In the same way as in Study 1, nonverbal immediacy was assessed using a scale consisting of six items, which was administered through a 7-point semantic differential scale (i.e., pleasant vs. unpleasant voice, enthusiastic vs. boring voice etc.) (see Appendix C.1 for the items on this scale). This scale was adapted from earlier versions measuring not only vocal but a variety of other nonverbal cues (i.e., facial expressiveness) (Mehrabian, 1981; Richmond, Gorham & McCroskey, 1987; Richmond, McCroskey & Johnson, 2003; Servilha & Costa, 2015). We constructed a reliable measure of nonverbal immediacy (Cronbach's $a = .87$), by averaging participants' answers to this set of questions.

Delayed recall was measured as an index of cognitive learning and it was assessed 12 days after treatment. Delayed recall was assessed with the use of the recall test developed in Study 1. That is, overall, delayed recall was assessed with a fill-in-the-blanks test consisting of nine recall items and a multiple-choice test of 18 questions (see Appendix C.1 for the recall test).

This recall test was administered via email and students completed it online. The main reason that delayed recall was measured online was to avoid participant attrition between the first session and the delayed post-test session. We constructed two measures of delayed recall by counting participants' number of correct answers to each type of recall tests separately (gap filling test and multiple-choice test). These two measures were found to be moderately positively correlated, $r(104) = .5$, $p < .001$. The participants' performance scores were scored by two researchers independently. There was a 100% agreement on the performance scores between the two researchers.

Procedure

Participants were welcomed in the main hall of the lab building. Each participant was required to read and sign an informed consent form, explaining the general aim of the study and their willingness to participate. Next, they were randomly assigned to one of the two experimental conditions and they were requested to watch an instructional video on a computer monitor regarding the use of GTW browser. The video screen was split into two sides: on the right-hand side, an artificial agent appeared to use the GTW system by moving the head and eyes, while explaining the system functionalities being demonstrated; on the left-hand side, the actual system was displayed, showing participants the effects of the artificial agent's actions on the system in real time. Participants in both conditions were provided with an identical instructional video, with the only difference being the level of vocal expressiveness in terms of pitch tone, pitch variation and rate (for more details regarding the study's instructional videos see Study 1).

After the end of the instructional video, the participants were asked to answer a small survey. Afterwards, participants were paid, thanked for their participation, and asked to respond to the second survey they would receive 12 days after the experiment. Twelve days after the experiment in the lab, participants were sent via email a second survey containing a recall test. Upon completion of the online recall test a debriefing mail was sent to each participant.

4.3.2 Results

Manipulation check: An independent sample t-test analysis was conducted to check the study's manipulation of vocal expressiveness (i.e., perceptions of vocal parameters of pitch tone, pitch variety and speech rate) between the strong vocal expressiveness and weak vocal expressiveness condition. As expected, the results that participants in the strong vocal expressiveness condition (N = 67, M = 3.28, SD = .99) reported stronger perceptions of vocal expressiveness as compared to participants in the weak vocal expressiveness condition (N = 72, M = 2.25, SD = .88), $t(142) = 9.211$, $p < .001$.

Immediacy: To test H1, an independent sample t-test analysis was conducted to examine the effect of the level of the artificial model's vocal expressiveness on individuals' perceptions of immediacy. The results supported our first hypothesis, revealing that participants in the strong vocal expressiveness condition (N = 67, M = 4.24, SD = 1.0) reported higher perceptions of immediacy as compared to participants in the weak vocal expressiveness condition (N = 72, M = 3.23, SD = .89), $t(137) = -6.198$, $p < .001$.

Delayed recall: To test H2, a one-way multivariate analysis of variance (MANOVA) was conducted to examine the effect of the level of the artificial model's vocal expressiveness on both scores measuring individuals' delayed recall. The results did not provide evidence for a statistically significant effect of the level of the artificial model's vocal expressiveness on the two measures of delayed recall combined, Wilk's $\Lambda = .99$, $F(2, 103) = ,249$, $p > .05$. Similarly, there was no significant effect found for any single score, both times $p > .05$.

Given the non-significant effect of vocal expressiveness on delayed recall we could not test our third hypothesis on the mediating role of immediacy. However, immediacy was not found to be significantly correlated with any of the two measures of cognitive learning, $p > .05$.

4.3.3 Discussion

Findings on the effect of pedagogical artificial agents on learning outcomes are mixed (Heidig & Clarebout, 2011; Martha, & Santoso, 2019; Schroeder, Adesope, & Gilbert, 2013). One of the proposed reasons for such inconclusive findings is the artificial agent's untapped potential for subtle nonverbal behavior that we know from human-human interaction.

In Study 1, we found that an artificial agent with strong vocal expressiveness increased immediacy, which in turn, enhanced affective learning and perceived cognitive learning, as compared to an artificial agent with weak vocal expressiveness. What is more, findings of support motivation theory, as motivation was found to explain part of the effect of immediacy (and, thus, of vocal expressiveness) on affective learning outcomes. However, concerning actual cognitive learning (i.e., immediate recall), results were mixed. Additionally, testing the underlying process of the effect of vocal expressiveness on recall, we failed to find evidence for attention-arousal theory, as results showed that neither immediacy nor attention had an effect on immediate recall.

Since other psychological studies outside the educational research area have suggested a larger influence of arousal on memory consolidation over time (Christianson & Loftus, 1991; Eysenck, 1976; Heuer & Reisberg, 1992; Revelle & Loftus, 1992), and given the lack of studies on the effect of immediacy on learners' retention over time (Witt, Wheeless & Allen, 2004), we conducted a follow-up study to focus on the measurement of delayed recall.

Hence, the current Study 2 investigated the effect of an artificial model with strong vocal expressiveness, as compared to the (same) artificial model with weak vocal expressiveness on nonverbal immediacy (H1) and delayed recall (12 days past treatment) (H2). Lastly, we predicted that immediacy would mediate the effect of vocal expressiveness on delayed recall (H3).

Confirming to our first hypothesis, results showed that an artificial agent with strong vocal expressiveness can increase perceptions of immediacy, as compared to an artificial agent with weak vocal expressiveness. This finding is in accordance with our findings of Study 1 and the literature on human teachers' nonverbal immediacy (i.e., Mehrabian, 1981; Witt, Wheeless & Allen, 2004; Ellis, Carmon & Pike, 2016). Therefore, this study strengthens the conclusion that similar to a human teacher, artificial agents with strong nonverbal cues, such as vocal expressiveness, can enhance learners' perceptions of psychological closeness (i.e., immediacy) with an artificial teacher.

Next, not providing support for our second hypothesis, results did not provide evidence for an effect of an artificial agent's vocal expressiveness (strong vs. weak) on delayed recall. As it seems, our results do not support the notion that a delay period of 12 days could consolidate learning effects. On the surface, it appears that our findings of Study 2 are in line with the research of Comstock et al. (1995), in which the authors failed to find an effect of immediacy of human teachers on delayed recall, and in contrast with the study of Titsworth (2001), where such effect was found (for only one type of knowledge assessment). Unfortunately, the article of Comstock et al. (1995) does not provide details of their experimental methodology (e.g.., immediacy manipulation and time of post-test), thus a thorough comparison of the studies (the current study and the two earlier ones) cannot be performed. Nonetheless, one crucial difference between Titsworth (2001) and Comstock et al. (1995) is the way delayed recall was assessed. While Titsworth (2001) measured delayed recall at a detailed (i.e., detail test) and abstract (concept test) level, as it seems Comstock et al. (1995) only measured it at an abstract level with the use of true-false items (there is no detailed information mentioned in their article). The results of Titsworth (2001) are comparable with our findings of Study 1, where we found that vocal expressiveness increased immediate recall when recall was measured in a detailed level (i.e.., gap-filling test), but not when it was measured in a more abstract level (i.e., multiple-choice test).

However, contrary to Titsworth (2001), the present Study 2 does not provide evidence for an effect of immediacy on delayed recall for both types of assessment. One reason for these contradictory findings between these two studies might be the time difference in assessing delayed recall. That is, while Titsworth (2001) measured delayed recall one week after treatment, in Study 2, we measured delayed recall approximately after two weeks' time. A large body of literature highlights the crucial role of the time interval on the consolidation process. However, the time interval that sustains complete consolidation has not yet been clarified; thus, consolidation might encompass a period ranging from weeks to years (Manoli et al., 2018). Given the unclarity on the interval of memory consolidation, our choice of delayed recall assessment after several days was made, to make our method consistent with other studies that included measures of retention over time (1-2 weeks). Nonetheless, given that the task was not very relevant to the participants, it might be that any effects of vocal expressiveness on delayed recall might have disappeared after 12 days.

In addition, to the type of knowledge assessment and time of post-test, the study of Titsworth (2001) manipulated a large number of both verbal and nonverbal immediacy cues next to teacher's vocal expressiveness (i.e., "we" versus "I" statements, facial expressions, gestures and body movement). It might be that, if arousal is the underlying process as proposed, the combination of various immediacy cues, or the single effect of an immediacy cue, other than vocal expressiveness, activated the underlying mechanism of arousal-attention, leading to an increase in delayed recall.

Therefore, although our current findings appear to be in line with Comstock et al. (1995), we postulate that arousal-attention theory might still be valid when at least three factors are taken into consideration: 1) the type of knowledge assessment (i.e., detailed versus abstract level); 2) the time of post-test (i.e., one week versus two weeks); and, 3) the type of immediacy manipulation (i.e., combination of immediacy cues vs. single immediacy cue). Future research could investigate these issues and manipulate these variables.

In conclusion, the findings of the current follow-up Study 2 are important as they confirm our findings of Study 1 that artificial agents with strong nonverbal behaviors, such as vocal expressiveness, can enhance learners' perceptions of psychological closeness with the artificial teacher (immediacy). However, in this Study 2, cognitive learning effects could not be established after the delay period of 12 days. Nonetheless, these findings are essential, as they highlight not only the necessity for continued research on cognitive performance over time associated with immediacy, but also the path that such future research could take, by identifying factors to be taken into consideration (i.e., type of assessment, time of post-test, method of immediacy manipulation). Overall, Study 2 complements our Study 1 by completing the picture on how an artificial agent's vocal expressiveness impacts immediate and delayed recall.

4.4 General Discussion

Innovative educational technology tools have a great promise for improving learning, yet they are often not used to their full potential. Attempts to utilize more of a technology's capabilities have led researchers to investigate instructional software tools such as artificial pedagogical agents. Such agents are designed to facilitate learning by providing instructional support and stimulating motivation in multimedia learning environments (Clark & Choi, 2005). Nonetheless, findings regarding their effectiveness for learning are mixed. One of the reasons for the non-significant effects of artificial agents on learning might be the fact that while research has investigated artificial teachers' visual nonverbal cues (i.e., Baylor & Kim, 2009), their vocal nonverbal cues (i.e., vocal expressiveness) have received little attention (Valetsianos, 2009).

The current work investigated the effects of artificial agent's vocal expressiveness on cognitive and affective learning. In addition, the underlying processes of such effects were also examined. Our findings of Study 1 showed that an artificial agent with strong vocal expressiveness increased affective learning. What is more, our findings of Study 1 revealed that vocal expressiveness is related to affective learning because it promotes nonverbal immediacy. Further, results of Study 1 provided evidence of motivation as a mediator of the path from immediacy to affective learning. Such findings provide support for motivation theory that argues on the important role of motivation as an underlying process that (primarily) influences affective learning.

Consistent with earlier literature, in our current research, cognitive learning was measured as perceived cognitive learning and actual cognitive leaning (recall at two points in time). Our findings showed that an artificial agent with strong vocal expressiveness increased perceived cognitive learning, and that this effect is only mediated by nonverbal immediacy (but not by attention). Concerning actual cognitive learning findings were mixed. Specifically, we found support for the effect of vocal expressiveness on immediate recall (on one of the two types of assessments), but we found no evidence for an effect on delayed recall. In addition, neither immediacy nor motivation and attention have been found to explain the effect of vocal expressiveness on immediate recall. The effect of vocal expressiveness on learners' immediate recall is unclear. This effect could be explained by the affect model theory (see Rodriguez et al., 1996), which identifies affective learning, as the mediating variable between teacher immediacy and students' cognitive learning. Future research could investigate whether the affect model could explain the effect of nonverbal immediacy cues on immediate recall.

Thereby, our findings provide evidence that nonverbal immediacy cues of an artificial agent, such as vocal expressiveness, might enhance immediate recall but not delayed recall after 12 days. Given the lack of studies on the effect of immediacy cues on delayed recall, and given the difference in measurements between the few existing studies, we argue that more research is needed that takes into account the three factors identified in Study 2: 1) the type of knowledge assessment (i.e., detailed versus abstract level); 2) the time delay of post-test (i.e., one week versus two weeks delay) and 3) the type of immediacy manipulation (i.e., combination of immediacy cues vs. single immediacy cue). In addition, the present work did not find support for the proposed underlying mechanism of arousal-attention. However, before we discard this theory, we argue that more research is required to include advanced measurements of both arousal (i.e., using skin conductance) and attention (i.e., using eye-tracking technology).

Furthermore, our current work also has practical implications. That is, our studies provide practical knowledge on how to optimally design nonverbal vocal expressiveness of artificial agents in order to facilitate learning. Our studies provide practical recommendations on how to combine vocal parameters of pitch (tone and variation) and speech rate, so as to create strong vocal expressiveness that artificial agents can use to enhance perceptions of closeness of their students and further increase leaning outcomes.

Overall, results of these two studies show that an artificial agent with strong vocal expressiveness has a positive influence on individuals' affective learning, perceived cognitive learning, and, to a smaller extent, actual cognitive learning (immediate recall). What is more, our findings revealed that vocal expressiveness is related to affective learning and perceived cognitive learning because it reduces psychological distance, thus, promotes immediacy. Finally, our findings support motivation theory which emphasizes the important role of motivation as an underlying process that influences affective learning. These findings are important as they increase our understanding of artificial agents as teachers and the way in which they should make use of their vocal expressiveness in order to affect learning. These studies can help create such agents that might one day be important aides in education and other related domains crucial to our society.

Chapter 5

General discussion

Pedagogical artificial agents have been studied for more than two decades. Yet, as described in the General Introduction, the effectiveness of using an artificial agent in a learning environment remains unclear. Given the contradicting findings of earlier research, the overall goal of the current thesis was to examine under which conditions and in which ways an artificial agent could facilitate learning. In order to attain our overall goal, I broke it down into three sub-goals, which were empirically examined and reported in this thesis.

In more detail, a crucial condition that been neglected by earlier research is the instructional method an artificial agent applies in the multimedia learning environment (Heidig & Clarebout, 2011; Schroeder & Gotch, 2015). Thus, the first sub-goal of the current thesis was to answer the fundamental question of whether an artificial agent that uses a particular instructional method (i.e., modeling) is effective for learning (Chapter 2). Next, another point of confusion is the debate about artificial agents' visibility in multimedia leaning settings. As shown in the General Introduction, two competing perspectives exist in the literature on whether the visibility of an artificial agent in multimedia settings hinders or augments learning (agents-as-complements versus agents-as-distractors) (Frechette & Moreno, 2010). Therefore, the second sub-goal of the current thesis was to examine the conditions under which the visual presence of the artificial agent (as a model) is beneficial for learning. Specifically, I hypothesized that the type of learning task (psychomotor versus cognitive) is a decisive condition for the inclusion of an artificial agent's visual presence (Chapter 3). Lastly, the third sub-goal of the current thesis was to examine conditions that increase the effectiveness of a visible artificial agent that acts as model for learning. Specifically, I argued that the effectiveness of an artificial agent as a model depends on the nonverbal behavior that it appears to exhibit. That is, the third sub-goal was to examine the effects of artificial model's vocal expressiveness (i.e., pitch tone, pitch variation and speech rate) as a powerful form of nonverbal behavior that strengthens students' learning (Chapter 4).

In the current chapter, I first discuss the findings and contributions of our empirical work in relation to our three sub-goals highlighted above (sections 5.1, 5.2. and 5.3). Then, I discuss the findings of our studies in relation to the overarching goal of this thesis (section 5.4). Next, I discuss directions for future research (section 5.5). Finally, after I specify practical contributions (i.e., design) and societal contributions that stem from our findings (section 5.6), I present the general conclusion of the thesis (section 5.7).

5.1 Sub-goal 1: Is modeling by an artificial agent effective?

Earlier research has found that behavioral modeling employed by human models is an effective instructional method. Specifically, past work in the domain of technological innovation (the focal domain of Chapter 2) revealed that behavioral modeling by a human model yields higher scores of computer self-efficacy and better task performance compared to other commonly used non-modeling instructional methods (i.e., Compeau & Higgins, 1995a; Compeau & Higgins, 1995b). In Chapter 2, I conducted two studies to examine behavioral modeling as a facilitating instructional role that an artificial agent can embody in a multimedia learning environment. In more detail, in Study 1 of Chapter 2, I investigated the effect of an artificial agent as a behavioral model as compared to two common, non-modeling instructional methods (agent-delivered instructional narration and no-agent, text-only instruction) on learners' beliefs of their computer-self efficacy and the system's perceived ease of use. In line with our hypothesis, the results of Study 1 showed that learners in the agent-delivered modeling condition reported higher computer self-efficacy, as compared to learners in the two non-modeling conditions. Additionally, Study 1 showed that learners in the agent-delivered modeling condition had improved perceptions of ease of use of the system, compared to learners in the agent-delivered instructional narration condition, but not when compared with the no-agent, text-only condition. Therefore, our second hypothesis was partially supported, as results were mixed. I attributed this finding to the different modalities used between the conditions. That is, the no-agent, text-only instruction condition delivered only visual information to participants, while the agent-delivered modeling condition delivered information through both visual and auditory modalities. I postulated that as a result of this, participants in the text-only condition were provided with less concrete system experience from the instructional video as compared to participants in the agent-delivered modeling condition and they might have judged the system's ease of use based on their earlier extensive experience with technologies in general.

In Study 1, I examined the effect of the artificial agent functioning as a behavioral model on learners' beliefs of their computer self-efficacy and system's ease of use. The next step was to conduct a follow-up study to extend the insights about the effects of agent delivered modeling, as compared to other non-modeling methods, on learners' declarative knowledge and task performance. An additional purpose of Study 2 was to further examine the mixed results regarding system's perceived ease of use we found in Study 1. Given the limitations identified in Study 1, in Study 2 we substituted the no-agent, text-only condition with the no-agent, voice-only narration condition to control for any modality effect (written vs oral) on participants' learning. In line with our hypothesis, results of Study 2 showed that participants in the agent-delivered modeling condition showed better task performance when using the system, compared to participants in the two non-modeling conditions. What is more, as hypothesized, participants in the agent-delivered modeling condition scored significantly higher on a declarative knowledge assessment compared to participants in the agent-delivered instructional narration. However, the results showed that this advantage was not present when agent-delivered modeling was compared to voice-only instructional narration. This unexpected finding on declarative knowledge assessment could indicate that learners who receive less visual information in a task (i.e., voice only instructional narration condition) rely more, and thus, pay more attention to verbal explanation in order to comprehend a given task. Such an effort to understand the task could have resulted in better verbal information processing and acquisition, as compared to the agent-delivered instructional narration (this is further discussed in Section 5.5.). Finally, in line with the findings of Study 1, results of Study 2 showed that participants in the agent-delivered modeling condition showed higher computer self-efficacy, as compared to participants in the two non-modeling conditions. However, although we overcame the modality effect, Study 2 did not provide evidence for a significant difference in system's perceived ease of use between the agent-delivered modeling and the two non-modeling conditions. The lack of success in reproducing the effect of agent-delivered modeling on participants' system's perceived ease of use, as compared to the agent-delivered instructional narration, as also to find an effect as compared to voice-only instructional narration, could be attributed to the smaller sample size as compared to that of Study 1. Due to the fact that the primary goal of Study 2 was to extend findings of Study 1 examining task performance, we ensured that the statistical power was adequate for realizing this goal. Thus, for practical reasons (i.e., related to task performance measurements), we were not able to further increase the power that was required to detect an effect on system's perceived ease of use. Besides other contributions, Study 2 provides evidence that it is the instructional approach of an artificial agent (i.e., modeling) that can positively influence learners' behavior (i.e., task performance) rather than the agent's mere presence (i.e., agent-delivered instructional narration condition).

Overall, findings of Study 1 and Study 2 showed that, similar to a human model, an artificial model can positively influence learners' motivational (computer self-efficacy and perceptions of ease of use of the system), cognitive (declarative knowledge) and, most importantly, skill-based (i.e., task performance) learning outcomes, as compared to other non-modeling instructional methods.

Finally, as discussed in the general introduction, past research reported contradicting findings on the effect of using an artificial agent on learning outcomes. However, the majority of past studies focussed either on artificial agent's presence or appearance (i.e., visible qualities). Studies on artificial agents' behavior are far less common. Our findings of Chapter 2 suggest that it is the artificial agent's behavior (i.e., behavioral modeling) rather than its mere presence that has a positive impact on learning. Therefore, as we have claimed in the General Introduction, examining the instructional method employed by an artificial agent, may help clarifying why many studies found no effects while, to the contrary, several other studies present positive effects.

5.2 Sub-goal 2: What is the value of an artificial agent's visual presence on learning?

As we discussed in the General Introduction, another point of confusion is the debate about artificial agents' visibility in multimedia learning settings. This debate mainly relates to whether it would be more effective if the instructional design were presented by simpler means of communication, rather than by an embodied character (Choi & Clark, 2006). The existing literature contains contradictory theories concerning the overall impact of an artificial agent's visual presence on learning. Specifically, theories such as social presence theory and social agency theory argue that the artificial agent's physical presence leads to well-formed mental models of concepts taught and better learning due to an increased motivation (i.e., Hoyt et al., 2003; Moreno et al., 2001). Nonetheless, findings of recent studies are inconclusive in terms of the motivational effect of the artificial agent's visual presence (i.e., Chen & Chou, 2015; Dinçer & Doğanay, 2017; Lin et al., 2020; Park, 2015). On the other hand, theories, such as cognitive load theory (Sweller, 1988; Sweller 2004), hold that that such on-screen presence can impose cognitive and affective distractions and, thus, hamper learning. However, studies on the effects of an artificial agent on cognitive load reported opposing results (i.e., Dinçer & Doğanay,2017; Frechette & Moreno, 2010; Moreno et al., 2001). We proposed that for solving this second inconsistency we need to have a closer look at the conditions under which an artificial agent's visual presence is relevant and, therefore, essential for goal achievement, or irrelevant and, thus, an unnecessary addition.

In Chapter 3, we conducted two studies to investigate whether the learning effect of a visually present artificial agent as a model is dependent on the type of learning task. According to Bandura (1986), a model is only effective when it is relevant to the modeled behavior. Given that the modeled behavior is determined by the learning task, we argue that the type of learning task determines the relevance of an artificial agent's visual presence. Specifically, in Study 1 of Chapter 3, we examined 1) the effects of the interaction between the on-screen visibility of an artificial model (presence versus absence) and type of task (psychomotor versus cognitive) on learning outcomes, and 2) whether an artificial agent's visual presence becomes more relevant for task performance of a psychomotor task as the level of complexity increases.

Overall, in line with our first hypothesis, our results revealed that learners provided with a psychomotor task (behavioral modeling) from a visible artificial agent had better task performance, and they further reported higher self-efficacy and affective beliefs, as compared to learners who received the same instructions but without the artificial model being visible to them. To the contrary, and as expected, under the condition of a cognitive task (cognitive modeling), the visual presence of the artificial model was not found to influence learners' task performance, self-efficacy and affective beliefs when compared to learners in a cognitive task who were confronted with a visual artificial agent. However, there was one surprising finding. Contrary to our hypothesis, the findings suggested that regardless of the type of task, learners showed better recall when the artificial model was not visually present. We explained this unexpected result in light of cognitive load theory and its concept of the redundancy effect (Sweller et al., 2011). That is, we assumed that under the condition of behavioral modeling, the artificial model's visual presence was not necessary for learners to recall task instructions and, rather, the model's visual presence might have caused an unnecessary increase of their extraneous cognitive load. To the contrary, this was not the case for task performance. We argue that in this case, optimum task performance was based on the successful integration of the two types of information provided: the visual demonstration and the auditory narration.

Overall, the findings of this study support our argument on the additional value of the visual presence of the artificial model being dependant on the learning task to be modeled. Furthermore, in line with our expectations, our findings revealed that, for the psychomotor task, the effect of the artificial model's visual presence on task performance was larger for the difficult level than for the easy level. Thus, the findings suggest that as the level of complexity of a psychomotor task increases, the visual information provided by the artificial model becomes more important for learners' construction of a more accurate mental model of the task, and, consequently, for better task performance.

In Study 1 (of Chapter 3), it was implicitly assumed that the visual presence of the artificial agent as a model has a different effect on learners' cognitive processes depending on the type of task it models. More specifically, we expected that the visual presence of an artificial model would reduce extraneous cognitive load when the learning task to be modeled was psychomotor (i.e., behavioral modeling). We argued that this would not be the case for modeling a cognitive task (i.e., cognitive modeling). However, this argument was not tested explicitly. This omission led to Study 2 that aimed not only to replicate but also to extent findings of Study 1. That is, Study 2 examined effects of the interaction between the on-screen visibility of an artificial model (presence vs. absence) and type of task (psychomotor vs. cognitive) on learners' performance-related cognitive load. In addition, given the unexpected finding in Study 1 regarding learners' recall for the psychomotor task, in Study 2 we aimed to investigate further the effect of the visibility of an artificial model on learners' perceived cognitive load of the recall test for the psychomotor task. In line with our hypothesis, results revealed that when it comes to the demonstration of a psychomotor task (behavioral modeling), the visual presence of the artificial model decreased learners' perceived cognitive load for the task performance, as compared to those who received the same psychomotor task instructions, but without the artificial model being visually present. Furthermore, under the condition of a cognitive task demonstration (cognitive modeling), the visual presence of the artificial model was not found to influence individuals' perceived cognitive load with respect to the task performance, which was in line with our expectations. Nonetheless, contrary to our second hypothesis, results did not provide evidence that the visual presence of the artificial model increases learners' cognitive load related to recall. Overall, findings of the two studies provide strong evidence that the visual presence of the artificial model enhanced learners' self-efficacy, affective beliefs and task performance, as it also minimized cognitive load associated with their task performance, for psychomotor tasks (behavioral modeling), but less so for cognitive tasks (cognitive modeling). Thus, results in Chapter 3 support the argument that the question of whether the visibility of artificial agents facilitate learning can only be answered by taking into consideration the specific conditions of their use. Findings of the current work confirm that the type of learning task that an artificial agent models is an important condition.

5.3 Sub-goal 3: What are the conditions that increase the effectiveness of a (visible) artificial agent for learning?

The third sub-goal of the current thesis is to examine the conditions that increase the effectiveness of a visible artificial agent that acts as a model for learning. For this, a closer inspection of the behavior of human teachers was taken into consideration. As sketched in the General Introduction, in traditional classroom settings with human teachers, various nonverbal forms of teacher behavior have been found to increase nonverbal immediacy (i.e., the teacher's psychological closeness created through nonverbal communication) and, subsequently, learning (Ellis, Carmon & Pike, 2016; Witt, Wheeless & Allen, 2004).

In Chapter 4, we conducted two studies to investigate effects of an artificial agent's vocal expressiveness (strong vs. weak vocal expressiveness) on affective and cognitive learning. Specifically, the aim of Study 1 was twofold. Firstly, we examined the effect of vocal expressiveness (strong vs. weak vocal expressiveness) of an artificial agent on affective and cognitive learning (immediate recall and perceived cognitive learning). Results supported our hypotheses showing that an artificial agent with strong vocal expressiveness increased affective and perceived cognitive learning. Partial support was found for actual cognitive learning (i.e., immediate recall). The study's second aim was to examine the underlying processes of the effect of vocal expressiveness on learning. Our findings revealed that vocal expressiveness is related to affective and perceived cognitive learning because it promotes nonverbal immediacy. Secondly, we tested whether motivation or attention explain the effect of immediacy on affective and cognitive outcomes. Results provided evidence of motivation as a mediator of the path from immediacy to affective learning, thus supporting motivational theory (Christopher, 1990). However, not supporting our expectations, results also showed that neither immediacy nor attention mediated the effect of vocal expressive-

ness on cognitive learning. Thus, arousal-attention theory was not supported (Kelly & Gorham, 1988; implications are discussed in Section 5.6).
In Study 1 (of Chapter 4), we found mixed effects regarding the effects of vocal expressiveness on actual cognitive learning (measured as immediate recall). Similarly, we did not find evidence supporting attention-arousal theory (Kelly & Gorham, 1988). This theory posits that immediacy stimulates arousal, which, thereby, affects attention and memory leading to greater cognitive learning (Kelly & Gorham, 1988). However, other psychological studies demonstrated the important role of arousal in altering both attention and consolidation of memories (Christianson & Loftus, 1991; Eysenck, 1976; Heuer & Reisberg, 1992; Revelle & Loftus, 1992). Such studies suggest that if arousal acts specifically on memory consolidation, its influence magnifies following a delay, as consolidation is a process that occurs over time. In line with these studies, we argued that a plausible reason for the mixed effects of vocal expressiveness on recall, as well as for the lack of evidence of immediacy as a mediator of vocal expressiveness on recall we found in Study 1, might be the fact that we examined immediate recall (i.e., measured immediately after treatment) rather than delayed recall (i.e., several days after treatment). Thus, in Study 2, we extended findings of Study 1 by examining the effect of an artificial agent's vocal expressiveness on delayed recall. In line with findings of Study 1, we found that strong vocal expressiveness increased immediacy, as compared to weak vocal expressiveness However, contrary to our hypothesis, findings showed that vocal expressiveness of an artificial model did not affect delayed recall (we provide an explanation of potential reasons of this finding in Section 5.5).

Overall, findings of Chapter 4 suggest that an artificial agent with strong vocal expressiveness increases affective learning, perceived cognitive learning and, to a smaller extent, actual cognitive learning (immediate recall). What is more, our findings revealed that vocal expressiveness is related to affective learning and perceived cognitive learning because it reduces psychological distance, thus, promotes immediacy. Finally, our findings support motivation theory which emphasizes the important role of motivation as an underlying process that influences affective learning.

5.4 Overall goal: Under which conditions can an artificial agent facilitate learning?

As discussed above, results of earlier research with regard to the overall effect of an artificial agent on learning are contradicting. In this thesis, we argued that in light of the great variety of artificial agents performing different tasks and roles in different contexts, as used in past studies, this issue is too broad to receive a simple answer. Therefore, we proposed that a more fruitful approach would be to ask under which conditions artificial agents can facilitate learning.

In order to identify such conditions, the current research shifted its attention to an artificial agent's pedagogical behavior, and, specifically, to the instructional method of modeling employed by an artificial agent. Our focus on the pedagogical behavior of an artificial agent is opposed to what the majority of earlier studies have focused on (that is, either presence or appearance). Our focus on a specific form of behavior of artificial agents reflects our overall position: it is mainly the artificial agent's pedagogical behaviors that make a difference in learning, and therefore, more consideration of the conditions under which such behavior facilitates learning is required.

Findings of the two studies in Chapter 2 provided evidence in favour of our position. That is, these findings suggest that it is the artificial agent's behavior (i.e., modeling) rather than its mere presence (artificial agent as an information source) that has a positive impact on learning. Therefore, as we have claimed in the General Introduction, examining the instructional method employed by an artificial agent, might help clarifying why some studies found no effects (Heidig & Clarebout, 2011; Martha, & Santoso, 2019). while, to the contrary, other studies present positive effects (Castro-Alonso, et al., 2021; Schroeder, Adesope, & Gilbert, 2013).

The main reason that studies on pedagogical behavior are less common, may be that they do not always technically require an embodied agent (Choi & Clark, 2006; VanLehn, 2011). In fact, one could argue that modeling that takes place in multimedia settings can occur without the actual visual presence of a model. Findings of our two studies in Chapter 3 confirmed our argument that the visual presence of an artificial agent can be beneficial for learning, but under certain conditions. In case of modeling by an artificial agent, this condition is the type of learning task to be modelled (cognitive versus psychomotor). Findings of Chapter 3 add to the findings in Chapter 2, by highlining conditions (i.e., type of task) under which pedagogical behavior (i.e., instructional method employed) necessitates the visual presence of an artificial agent for increased learning. In line with our results, we argue that the visual presence of an artificial agent is beneficial when it helps learners to construct a more accurate mental model of the learning task (such as in the case of behavioral modeling as opposed to cognitive modeling).

What is more, in the current thesis we aimed to go one step further by examining how the pedagogical behavior of a visible artificial agent can be strengthened by means of nonverbal behavior. The two studies of Chapter 4 suggest that an artificial agent with strong nonverbal, vocal expressiveness has a positive influence on individuals' perceptions of immediacy and, in turn, affective learning and perceived cognitive learning. What is more, our findings support motivation theory (Christopher, 1990) which emphasizes the important role of motivation as an underlying mechanism that relates vocal expressiveness to affective learning.

Overall, our studies combined show that an artificial agent that applies behavioral moeling can positively influence learners' beliefs (i.e., self-efficacy), affective and skill-based learning (i.e., task performance). Under the condition of a psychomotor task (implying behavioral modeling), the visibility of the agent has beneficial effects on learners' beliefs (i.e., self-efficacy), affective and skill-based learning (i.e., task performance) and minimizes task-performance cognitive load. What is more, under certain conditions (i.e., increased difficulty of a psychomotor task), the visual presence of an artificial agent becomes even more vital for enhancing learners' psychomotor skills. Lastly, the benefits of behavioral modeling for affective learning are further strengthened when the agent employs enhanced nonverbal behavior (i.e., strong vocal expressiveness).

However, there are also ambivalent findings. The current work provides limited evidence when it comes to the beneficial effect of an artificial agent as a model on cognitive learning. That is, findings in Chapter 2 showed that there was no significant difference in recall between participants in the agent-delivered modeling condition and no-agent, voice-only instructional narration. What is more, findings of Chapter 3 suggest that the sole visual presence of an artificial agent as a behavioral model may even have aversive effects on cognitive learning (i.e., recall of the instructions) independent of the type of leaning task being modelled (although we failed to replicate this finding). Studies in Chapter 4 provided only partial support with regard to the effect of an artificial agent's vocal expressiveness on immediate recall and no effect on delayed recall, both of which are examples of cognitive learning outcomes.

These findings suggest that the inclusion of an artificial agent in multimedia learning environments might not be necessary for enhancing participants' recall. Nonetheless, given that recall is only one example of a cognitive leaning, the current thesis cannot draw conclusions about the effect of an artificial agent on cognitive learning overall (see below for a discussion of cognitive learning and future research). However, the important contribution of our work is that it paves the way for future research to investigate conditions under which an artificial agent might have aversive effects on cognitive learning (see Section 5.5. below).

5.5 Directions of future research

Throughout all empirical chapters we identified specific limitations as also potential areas for future research that derive from these findings and limitations. In this section, we recognize a number of more general directions for future work that stem from the studies of the current thesis.

Firstly, the current studies assessed, amongst others, the effect of modeling by an artificial agent on recall, as an example of cognitive learning outcome. The results of our studies do not seem to suggest a strong effect on recall. Nonetheless, according to Bloom's taxonomy (Bloom, 1994) recalling important information is the first level out of the six levels that comprise the cognitive learning domain. The others are understanding, application, analysis, synthesis, and evaluation. We postulate that the inclusion of an artificial agent functioning as a model might be effective for higher levels of cognitive learning outcomes. Hence, we suggest future work to examine the effect of an artificial model on other levels of the cognitive learning domain. We argue that this is important because, due to the internet and continuous online connectivity, recalling information is becoming less vital (Dong, G., & Potenza, 2015; Firth et al., 2019). To the contrary, educational technologies necessitate the development of more advanced cognitive skills such as synthesis and evaluation of information (Michael & Godfrey, 2014).

What is more, all experiments presented in this thesis provide learners with a short duration of multimedia learning (~10 min). Thus, although the current research provided evidence for the effective of modeling by an artificial agent, future research is required to examine the effects of repeated and prolonged exposure to an artificial model on learning.

In addition, the current thesis stresses the necessity to re-evaluate the different theories that discuss the effectiveness of artificial agents' visual presence for learning. That is, our results in Chapter 3 do not provide evidence that the mere presence of an artificial agent has a positive impact on learning as claimed by earlier theories (i.e., social presence theory, Hoyt, Blascovich & Swinth, 2003; and social agency theory, Moreno et al., 2001). As we have argued, such theories are incomplete, and, therefore, future research could identify and examine other conditions under which the inclusion of an artificial agent's sole visual presence can facilitate learning. We hypothesize that a framework for such conditions could be related to learners' feelings of loneliness and isolation as a consequence of reduced social presence and psychological immediacy in online learning environments, compared to in-person instruction (Jeste, Lee & Cacioppo, 2020). In fact, earlier research has shown that social exclusion increases both attentiveness to social cues (Pickett, Gardner & Knowles, 2004) and attributions of human-likeness to artificial agents (Epley, Waytz, Akalis & Cacioppo, 2008). Further, it has been found that socially excluded people are more easily persuaded by an artificial agent to change their behavior (Ruijten, Midden & Ham, 2015).

Similarly, our findings do not provide support for the set of theories that claim that artificial agents are distractors of learning, such as cognitive load theory (Sweller 2004; Sweller, Ayres, & Kalyuga, 2011). That is, our work in Chapter 3, provides evidence against cognitive load theory, when the type of task being modeled is considered. More specifically, our findings reveal the effectiveness of modeling by an artificial in increasing learners' self-efficacy, affective beliefs, and task performance, while minimizing task performance-related cognitive load for psychomotor tasks as opposed to purely cognitive tasks. However, some support was found for cognitive learning. That is, the current research provided some evidence that an artificial model can negatively affect learners' recall regardless of the type of learning task it demonstrates. However, we could not to replicate this finding, as recall in Study 2 (of Chapter 3) was measured in a different way from Study 1 (of Chapter 3) (i.e., in Study 2 there were less multiple-choice questions and additional use of gap-filling questions). Therefore, we argue that future research could further examine the impact of the visibility of an artificial agent as a model on cognitive learning, by having as a reference how recall was measured in Study 1 (of Chapter 3). Furthermore, we propose that theories opposing the instructional benefit of an artificial agent, like cognitive load theory could benefit by specifying conditions under which the coexistence of visualizations and verbal explanations (as in the case of modeling by an artificial agent) can facilitate or hamper learning. Our findings suggest that such a condition might be the type of learning outcome (cognitive, psychomotor).

Finally, our findings in Chapter 4 provide some evidence on the effect of artificial agent's vocal expressiveness on immediate recall, but no evidence on delayed recall (measured after 12 days). Given the lack of studies on the effect of immediacy cues on delayed recall, and given the difference in measurements between the few existing studies (Comstock et al., 1995; Titsworth, 2001), we argue that more research is needed that takes into account the three factors identified in Study 2 (of Chapter 4): 1) the type of knowledge assessment (i.e., detailed versus abstract level); 2) the time of post-test (i.e., one week versus two weeks delay) and 3) the type of immediacy manipulation (i.e., combination of immediacy cues vs. single immediacy cue). In addition, the present work provide support for motivational theory, but not for the proposed underlying mechanism of arousal-attention. However, before we discard this theory, we argue that more research is required to include more direct measurements of arousal (e.g., using skin conductance) and attention (e.g., using eye-tracking technology). We deem that such more direct measurements are necessary in order to exclude the possibility that our findings are due to measurement limitations as it is difficult for learners to estimate their own level of arousal and attention subjectively.

5.6 Contributions to design and society

The current work contributes to the improvement of human-agent interaction design that takes place in multimedia learning environments. Below we describe several recommendations for interaction designers who wish to include artificial agent as educational tools. Firstly, the current research demonstrates the benefits of using artificial agents that act as models in order to improve learning. Secondly, our work clarifies the conditions under which including a visible artificial model in the learning environment is worthwhile in terms of time, money and energy invested in the design process. That is, an artificial agent as a model could be a valuable tool when the task to be modeled is psychomotor (behavioral modeling). This is because, as our results suggest, an artificial agent as a behavioral model can enhance learner's self-efficacy beliefs, affect and task performance, as also to minimize task performance-related cognitive load. Thirdly, our research suggests that interaction designers should also take into consideration the complexity of the psychomotor task to be modeled when deciding upon the inclusion of an artificial model. This is because as the complexity of a given psychomotor task increases, the visual presence of an artificial model becomes more vital for learning. Fourthly, our research provides practical knowledge on how to design nonverbal vocal expressiveness of artificial agents in order to further strengthen learning. That is, we provide a practical example on the combination of the vocal parameters of pitch (tone and variation) and speech rate, so as to create strong vocal expressiveness. Finally, the remaining two recommendations that stem from our work pertain to the cases in which designers should restrain from using an artificial agent in a learning environment. One such case is when a task to be modeled by an artificial agent is purely cognitive, or when the learning objective is exclusively of a cognitive nature (i.e., recall). Lastly, when modeling by an artificial agent cannot be implemented, designers should avoid using an artificial agent that acts as an information source, as it does not appear to have any beneficial effect on learning.

Besides practical contributions to the human-agent interaction design, the work presented in this thesis make particular contributions that are relevant to the educational community. In the current work we used insights from research on human teachers' instructional methods (i.e., modeling) and communication (i.e., non-verbal immediacy) and transfer them to research on pedagogical artificial agents. However, the transfer of insights the other way around might also be useful. We argue that the current insights might also be valuable for human teachers. Specifically, our findings suggest that the provision of opportunities for students to experience concepts through teacher's modeling for increased learning is important. In addition, our work recommends that teachers' nonverbal behavior (such as vocal expressiveness) are educational assets that can increase perceived cognitive learning, strengthen immediacy which in turn increases student motivation and, thereby, affective learning.

Finally, the strong evidence that the current thesis provides with regard to the beneficial effect of artificial agents on learning evokes ethical, moral, and philosophical questions regarding the actual employment of pedagogical artificial agents in multimedia learning environments. More specifically, future work could explore the properties that pedagogical artificial agents must have in order to be considered moral agents and whether their deployment and functioning is morally justified (i.e., Himma, 2009).

5.7 General conclusion

The primary aim of the current thesis was to shed light on the conditions under which artificial agents can facilitate learning. More specifically, this research focused on artificial agents' pedagogical behavior, and, specifically, on the instructional method they apply in a learning environment.

In sum, the findings of the current research reveal that modeling by an artificial agent is more effective as compared to other non-modeling methods in enhancing learners' beliefs of self-efficacy, affective learning and task performance (Chapter 2). What is more, our work also reveals the conditions under which the visual presence of an artificial model is beneficial for learning (Chapter 3). That is, the learning task being modeled is a decisive factor of whether the visual presence of an artificial model enhances learning. More specifically, the visual presence of an artificial model is beneficial for psychomotor tasks (i.e., behavioral modeling), as it increases self-efficacy beliefs, and affective and skill-based learning. However, this is not the case for cognitive tasks (i.e., cognitive modeling). Furthermore, the level of task complexity of a psychomotor task determines the importance of the on-screen visibility of an artificial agent. That is, as the level of the task complexity of a psychomotor task increases, the visual presence of an artificial agent becomes more essential for the improvement of learners' skill-based learning. Further, the current results show that the benefits of behavioral modeling by an artificial agent for affective learning are further strengthened when the agent employs enhanced nonverbal behavior (i.e., vocal expressiveness) (Chapter 4). Finally, our work provides ambivalent findings on the effects of modeling employed by an artificial agent on recall as an example of cognitive learning. However, such findings provide interesting directions for fruitful future research.

Overall, going back to the hypothetical scenario described in in the General Introduction, the current research attempted to answer the question of whether an artificial agent can be effective in teaching how to play the piano. The current thesis has shown that an artificial agent that acts as a model could effectively do it. Nonetheless, the process of learning piano includes both cognitive learning tasks (i.e., reading/writing music theory) and psychomotor learning tasks (i.e., playing the piano). This research suggests that the visual presence of an artificial model is not necessary for mastering such cognitive tasks but is essential for enhancing actual piano performance. Finally, an artificial agent modeling while making use of enhanced nonverbal behavior can further assist one to become a decent pianist, and, who knows, maybe the next Mozart.

References

Abu-Zaid, A., & Khan, T. A. (2013). Assessing declarative and procedural knowledge using multiple-choice questions. *Medical Education Online*, p.18.

Agarwal, R., Sambamurthy, V. & Stair, R., (2000). Research Report: The Evolving Relationship Between General and Specific Computer Self-Efficacy—An Empirical Assessment. *Information Systems Research*, *11*(4), pp.418-430.

Andersen, J. F. (1979). Teacher immediacy as a predictor of teaching effectiveness. *Annals of the International Communication Association*, *3*(1), pp.543-559.

Anderson, J. R. (1985). *Cognitive psychology and its implications*. New York: Freeman.

Anderson, L. W. & Burns, R. B. (1989). Research in Classrooms. *The study of teachers, teaching, and instruction.* Elmsford, NY: Pergamon Press.

Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology, 94*(2), p.416.

Atkinson, R. K., Mayer, R. E., & Merrill, M. M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology, 30*(1), pp.117-139.

Ayres, P., & Sweller, J. (2005). The split-attention principle. In R. E. Mayer (Ed.). *Cambridge handbook of multimedia learning* (pp. 135–146). New York: Cambridge University Press.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), p.191.

Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.

Bartlett Ellis, R., Carmon, A. & Pike, C., (2016). A review of immediacy and implications for provider-patient relationships to support medication management. *Patient Preference and Adherence*, 10, p.9.

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2008). Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International Journal of Social Robotics, 1*(1), pp.71-81.

Baylor, A. L., & Kim, S. (2009). Designing nonverbal communication for pedagogical agents: When less is more. *Computers in Human Behavior, 25*(2), pp.450-457.

Baylor, A. L., & Kim, Y. (2005). Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education*, *15*(1), p.95.

Baylor, A. L., & Plant, E. A. (2005). Pedagogical agents as social models for engineering: The influence of appearance on female choice. In C. K. Looi, G. McCalla, B. Bredeweg, & J. Breuker (Vol. Eds.), *Artificial intelligence in education: Supporting learning through intelligent and socially informed technology*: Vol. 125, (pp. 65–72). Amsterdam: IOS Press.

Baylor, A. L., & Ryu, J. (2003). The effects of image and animation in enhancing pedagogical agent persona. *Journal of Educational Computing Research, 28*(4), pp.373–394.

Beas, M. I., & Salanova, M. (2006). Self-efficacy beliefs, computer training and psycho-logical well-being among information and communication technology workers. *Computers in Human Behavior, 22*(6), pp.1043–1058.

Bente, G., Krämer, N., & Eschenburg, F. (2008). Is there anybody out there? Analyzing the effects of embodiment and nonverbal behavior in avatar mediated communication. In E. Konijn (Author), *Mediated interpersonal communication* (pp. 131-157). New York: Routledge.

Biswas, G., Leelawong, K., Schwartz, D., Vye, N., & The Teachable Agents Group at Vanderbilt. (2005). Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence, 19*(3-4), pp.363-392.

Bloom, B. (1994). Reflections on the Development and Use of the Taxonomy. In Anderson, L. Sosniak, L (Eds.) *Bloom's Taxonomy: A Forty-Year Retrospective*. Chicago: The National Society for the Study of Education, pp.1-8.

Bolkan, S. (2015). Students' Affective Learning as Affective Experience: Significance, Reconceptualization, and Future Directions. *Communication Education, 64*(4), pp.502-505.

Bouffard-Bouchard, T. (1990). Influence of self-efficacy on performance in a cognitive task. *The Journal of Social Psychology, 130*(3), pp.353–363.

Breitenstein, C., Lancker, D. V., & Daum, I. (2001). The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample. *Cognition and Emotion, 15*(1), pp.57-79.

Carlotto, T., & Jaques, P. A. (2016). The effects of animated pedagogical agents in an English-as-a-foreign-language learning environment. *International Journal of Human-Computer Studies, 95*, pp.15-26.

Castro-Alonso, J., Wong, R., Adesope, O. & Paas, F., (2021). Effectiveness of Multimedia

Pedagogical Agents Predicted by Diverse Theories: a Meta-Analysis. *Educational Psychology Review*.

Chandler, P., & Sweller, J. (1996). Cognitive load while learning to use a computer program. *Applied cognitive psychology*, *10*(2), pp.151-170.

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: A narrative overview. Procedia Computer Science, 136, pp.16-24.

Chen, C., & Chou, M. (2015). Enhancing middle school students' scientific learning and motivation through agent-based learning. *Journal of Computer Assisted Learning, 31*(5), pp.481-492.

Chesebro, J. L., & Mccroskey, J. C. (2000). The relationship between students' reports of learning and their actual recall of lecture material: A validity test. *Communication Education, 49*(3), pp.297-301.

Choi, S., & Clark, R. E. (2006). Cognitive and Affective Benefits of an Animated Pedagogical Agent for Learning English as a Second Language. *Journal of Educational Computing Research, 34*(4), pp.441-466.

Chou, C. Y., Chan, T. W., & Lin, C. J. (2003). Redefining the learning companion: the past, present, and future of educational agents. *Computers & Education*, *40*(3), pp.255-269.

Christianson, S., & Loftus, E. F. (1991). Remembering emotional events: The fate of detailed information. *Cognition & Emotion, 5*(2), pp.81-108.

Christophel, D. M. (1990). The relationships among teacher immediacy behaviors, student motivation, and learning. *Communication Education, 39*(4), pp.323-340.

Clarebout, G., Elen, J., Johnson, W. L., & Shaw, E. (2002). Animated pedagogical agents: An opportunity to be grasped?. *Journal of Educational multimedia and hypermedia*, *11*(3), pp.267-286.

Clark, R. E., & Choi, S. (2005). Five design principles for experiments on the effects of animated pedagogical agents. *Journal of Educational Computing Research, 32*(3), pp.209-225.

Clark, R. E., & Choi, S. (2007). The questionable benefits of pedagogical agents: Response to Veletsianos. *Journal of Educational Computing Research*, *36*(4), pp.379-381.

Cohen, D., 2011. *Teaching and its predicaments*. Cambridge, Mass.: Harvard University Press.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453–494). Hillsdale, NJ: Lawrence Erlbaum.

Compeau, D. R., & Higgins, C. A. (1995a). Application of social cognitive theory to training for computer skills. *Information Systems Research, 6*(2), pp.118–143.

Compeau, D. R., & Higgins, C. A. (1995b). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, pp.189–211

Comstock, J., Rowell, E., & Bowers, J. W. (1995). Food for thought: Teacher nonverbal immediacy, student learning, and curvilinearity. *Communication Education, 44*(3), 251-266.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, pp.319–340.

Davis, F. D. (1993). User acceptance of information technology: System characteristics, user perceptions and behavioral impacts. *International Journal of Man-Machine Studies, 38*(3), pp.475–487.

Davis, R. O. (2018). The impact of pedagogical agent gesturing in multimedia learning environments: A meta-analysis. *Educational Research Review*, 24, pp.193–209.

Davis, R., & Antonenko, P. (2017). Effects of pedagogical agent gestures on social acceptance and learning: Virtual real relationships in an elementary foreign language classroom. *Journal of Interactive Learning Research, 28*(4), pp.459-480

De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional science, 38*(2), pp.105-134.

Dehn, D. M., & Van Mulken, S. (2000). The impact of animated artificial agents: A review of empirical research. International *Journal of Human-Computer Studies, 52*(1), pp.1–22.

Dinçer, S. & Doğanay, A., (2017). The effects of multiple-pedagogical agents on learners' academic success, motivation, and cognitive load. *Computers & Education*, 111, pp.74-100.

Domagk, S. (2010). Do pedagogical agents facilitate learner motivation and learning outcomes? *Journal of Media Psychology, 22*(2), pp.84–97.

Dong, G., & Potenza, M. N. (2015). Behavioural and brain responses related to Internet search and memory. *European Journal of Neuroscience, 42*(8), pp.2546-2554.

Ellis, R. B., Carmon, A., & Pike, C. (2016). A review of immediacy and implications for provider–patient relationships to support medication management. *Patient Preference and Adherence,* 9.

Epley, N., Waytz, A., Akalis, S., & Cacioppo, J. T. (2008). When we need a human: Motivational determinants of anthropomorphism. *Social cognition, 26*(2), pp.143-155.

Eysenck, M. W. (1976). Arousal, learning, and memory. *Psychological bulletin*, *83*(3), p.389.

Firth, J., Torous, J., Stubbs, B., Firth, J. A., Steiner, G. Z., Smith, L., ... & Sarris, J. (2019). The "online brain": how the Internet may be changing our cognition. *World Psychiatry, 18*(2), pp.119-129.

Fogg, B. J. (2003). *Persuasive technology: Using computers to change what we think and do.* Amsterdam: Morgan Kaufmann.

Förster, J., Higgins, E. T., & Bianco, A. T. (2003). Speed/accuracy decisions in task performance: Built-in trade-off or separate strategic concerns? *Organizational Behavior and Human Decision Processes, 90*(1), pp.148–164.

Fountoukidou, S., Ham, J., Matzat, U., & Midden, C. (2019). Effects of an artificial agent as a behavioral model on motivational and learning outcomes. *Computers in Human Behavior*, *97*, pp.84-93.

Frechette, C., & Moreno, R. (2010). The Roles of Animated Pedagogical Agents' Presence and Nonverbal Communication in Multimedia Learning Environments. *Journal of Media Psychology, 22*(2), pp.61-72.

Frechette, C., Moreno, R. (2010). The roles of animated pedagogical agents' presence and nonverbal communication in multimedia learning environments. *Journal of Media Psychology, 22*(2), pp. 61–72.

Gist, M. E. (1989). The influence of training method on self-efficacy and idea generation among managers. *Personnel psychology, 42*(4), pp.787-805.

Gist, M. E., Schwoerer, C., & Rosen, B. (1989). Effects of alternative training methods on self-efficacy and performance in computer software training. *Journal of Applied Psychology, 74*(6), p.884.

Gorham, J. (1988). The relationship between verbal teacher immediacy behaviors and student learning. *Communication Education, 37*(1), pp.40-53.

Gorrell, J., & Capron, E. (1990). Cognitive modeling and self-efficacy: Effects on preservice teachers' learning of teaching strategies. *Journal of teacher education*, *41*(5), pp. 15-22.

Grabbe, J. W. (2015). Implications of experimental versus quasi-experimental designs. In K. D. Strang (Author), The Palgrave handbook of research design in business and management (pp. 141-152). New York City, NY: Palgrave Macmillan.

Gudykunst, W. B., Ting-Toomey, S., & Chua, E. (1988). *Culture and interpersonal communication*. Newbury Park: Sage Publications.

Ham, J., & Midden, C. (2014). A persuasive robot to stimulate energy conservation: The influence of positive and negative social feedback and task similarity on energy-consumption behavior. *International Journal of Social Robotics, 6*(2), pp.163–171.

Hamari, J., Koivisto, J., & Pakkanen, T. (2014). May). Do persuasive technologies persuade? -a review of empirical studies. *International conference on persuasive technology* (pp. 118–136). Cham: Springer.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139-183). North-Holland.

Hayes, A. F. (2018). Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation. *Communication monographs, 85*(1), pp.4-40.

Heidig, S. & Clarebout, G., (2011). Do pedagogical agents make a difference to student motivation and learning?. *Educational Research Review*, *6*(1), pp.27-54.

Hess, J. A., & Smythe, M. J. (2001). Is teacher immediacy actually related to student cognitive learning? *Communication Studies, 52*(3), pp.197-219.

Heuer F. Reisberg D. 1992. Emotion, arousal, and memory for detail. In *The Handbook of Emotion and Memory*, ed. S. Christianson, pp. 151–64. Hillsdale, NJ: Erlbaum.

Himma, K. E. (2009). Artificial agency, consciousness, and the criteria for moral agency: What properties must an artificial agent have to be a moral agent?. *Ethics and Information Technology, 11*(1), pp.19-29.

Hollands, F. & Tirthali, D., (2014). Resource requirements and costs of developing and delivering MOOCs. *The International Review of Research in Open and Distributed Learning*, *15*(5).

Hollien, H., & Shipp, T. (1972). Speaking Fundamental Frequency and Chronologic Age in Males. *Journal of Speech and Hearing Research, 15*(1), pp.155-159.

Hoyt, C. L., Blascovich, J., & Swinth, K. R. (2003). Social inhibition in immersive virtual environments. *Presence: Teleoperators & Virtual Environments*, *12*(2), pp.183-195.

Hsiao, T., Solomon, N. P., Luschei, E. S., & Titze, I. R. (1994). Modulation of fundamental frequency by laryngeal muscles during vibrato. *Journal of Voice, 8*(3), pp.224-229.

Jan, H. (2017). Teacher of 21st century: Characteristics and development. *Research on Humanities and Social sciences*, *7*(9), pp.50-54.

Jeste, D. V., Lee, E. E., & Cacioppo, S. (2020). Battling the modern behavioral epidemic of loneliness: suggestions for research and interventions. *JAMA psychiatry, 77*(6), 553-554.

Johnson, A. M., Ozogul, G., & Reisslein, M. (2015). Supporting multimedia learning with visual signalling and animated pedagogical agent: Moderating effects of prior knowledge. *Journal of Computer Assisted Learning, 31*(2), pp.97–115.

Johnson, A. M., Ozogul, G., Moreno, R., & Reisslein, M. (2013). Pedagogical agent signaling of multiple visual engineering representations: The case of the young female agent. *Journal of Engineering Education, 102*(2), pp.319–337.

Johnson, R. D., & Marakas, G. M. (2000). The role of behavioral modeling in computer skills acquisition: Toward refinement of the model. Information *Systems Research, 11*(4), pp.402–417.

Johnson, W. L., Rickel, J. W., & Lester, J. C. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education, 11*(1), pp.47–78.

Jones, S. (2017). Mediated immediacy: constructing authentic testimony in audio-visual media. *Rethinking History, 21*(2), pp.135-153.

Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need?. *Educational Psychology Review, 23*(1), pp.1-19.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 13*(4), pp.351-371.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/ aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74*(4), p.657.

Kansanen, P., (1999). Teaching as Teaching-Studying-Learning Interaction. *Scandinavian Journal of Educational Research*, *43*(1), pp.81-89.

Kasilingam G, & Chinnavan E. (2014). Assessment of learning domains to improve student's learning in higher education. *Journal of Young Pharmacists, 6* (1), pp.27-33.

Keiler, L., (2018). Teachers' roles and identities in student-centered classrooms. *International Journal of STEM Education*, *5*(1), pp.1-20.

Kelley, D. H., & Gorham, J. (1988). Effects of immediacy on recall of information. *Communication Education, 37*(3), pp.198-207.

Kennedy, J., Baxter, P., & Belpaeme, T. (2016). Nonverbal Immediacy as a Characterisation of Social Behaviour for Human–Robot Interaction. *International Journal of Social Robotics, 9*(1), pp.109-128.

Kim, C., & Baylor, A. L. (2008). A virtual change agent: Motivating pre-service teachers to integrate technology in their future classrooms. *Educational Technology & Society*, *11*(2), pp.309-321.

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), pp.220-234.

Kolb, D., (2014). *Experiential learning: Experience as the source of learning and development*. London: FT Press.

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology, 78*(2), p.311.

Krämer, N. C., & Bente, G. (2010). Personalizing e-learning. The social effects of pedagogical agents. *Educational Psychology Review, 22*(1), pp.71–87.

Lane, H. C. (2016). Pedagogical agents and affect: Molding positive learning interactions. In *Emotions, technology, design, and learning* (pp. 47-62). Academic Press.

Leppink, J., Paas, F., Van der Vleuten, C. P., Van Gog, T., & Van Merriënboer, J. J. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior research methods, 45*(4), pp.1058-1072.

Li, J., Kizilcec, R., Bailenson, J. & Ju, W., (2016). Social robots and virtual agents as lecturers for video instruction. *Computers in Human Behavior*, 55, pp.1222-1230.

Liew, T. W., Tan, S., Tan, T. M., & Kew, S. N. (2020). Does speaker's voice enthusiasm affect social cue, cognitive load and transfer in multimedia learning? *Information and Learning Sciences, 121*(3/4), pp.117-135.

Lin, L., Ginns, P., Wang, T., & Zhang, P. (2020). Using a pedagogical agent to deliver conversational style instruction: What benefits can you obtain? *Computers & Education, 143*, 103658.

Linek, S. B., Gerjets, P., & Scheiter, K. (2010). The speaker/gender effect: Does the speaker's gender matter when presenting auditory text in multimedia messages? *Instructional Science, 38*(5), pp.503-521.

Manoli, R., Chartaux-Danjou, L., Mecheri, H., Delecroix, H., Noulhiane, M., & Moroni, C. (2018). Extension to one week of verbal memory consolidation assessment. *Neuropsychology, 32*(3), p.366.

Marakas, G. M., Yi, M. Y., & Johnson, R. D. (1998). The multilevel and multifaceted character of computer self-efficacy: Toward clarification of the construct and an integrative framework for research. Information *Systems Research, 9*(2), pp.126–163.

Martha, A. S., & Santoso, H. (2019). The Design and Impact of the Pedagogical Agent: A Systematic Literature Review. *The Journal of Educators Online, 16*(1).

Martocchio, J. J., & Judge, T. A. (1997). Relationship between conscientiousness and learning in employee training: Mediating influences of self-deception and self-efficacy. *Journal of Applied Psychology, 82*(5), p.764.

Mayer, R. E. (2001). *Multimedia learning.* New York: Cambridge University Press.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of educational psychology, 90*(2), p.312.

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology, 95*(2), pp.419-425.

Mccroskey, J. C., Richmond, V. P., & Bennett, V. E. (2006). The Relationships of Student End-of-Class Motivation with Teacher Communication Behaviors and Instructional Outcomes. *Communication Education, 55*(4), pp.403-414.

Medawela, R. S., Ratnayake, D. R., Wijeyapala Abesinghe Mudiyanselage Udari Lakshika Abeyasinghe, Jayasinghe, R. D., & Marambe, K. N. (2018). Effectiveness of "fill in the blanks" over multiple choice questions in assessing final year dental undergraduates. *Educación Médica, 19*(2), pp.72-76.

Mehrabian, A. (1981). *Silent messages: Implicit communication of emotions and attitudes*. Belmont, CA: Wadsworth Pub.

Meij, H. V., Meij, J. V., & 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), pp.381-403.

Menges, R., Kumar, C., Müller, D., & Sengupta, K. (2017). April). Gazetheweb: A gaze-controlled web browser. *Proceedings of the 14th web for all conference on the future of accessible work* (pp. 25). ACM.

Michael, W & Godfrey, M. (2014). Development of Higher Order Cognitive Skills in Technology Mediated Learning: Activity Theory and Distance Education at Makerere University in Uganda. *Journal of Communication and Computer* 11, pp.229-238.

Mizuno, O., & Nakajima, S. Y. (1998). A new synthetic speech/sound control language. *In Fifth International Conference on Spoken Language Processing.*

Moos, D. & Azevedo, R., (2009). Learning With Computer-Based Learning Environments: A Literature Review of Computer Self-Efficacy. *Review of Educational Research*, *79*(2), pp.576-600.

Moreno, R. (2005). Multimedia learning with animated pedagogical agents. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 507–523). Cambridge, England: Cambridge University Press.

Moreno, R., & Mayer, R. E. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology, 97,* pp.117–125.

Moreno, R., Mayer, R., Spires, H. & Lester, J., (2001). The Case for Social Agency in Computer-Based Teaching: Do Students Learn More Deeply When They Interact With Animated

Pedagogical Agents?. *Cognition and Instruction*, *19*(2), pp.177-213.

Moreno, R., Reislein, M., & Ozogul, G. (2010). Using virtual peers to guide visual attention during learning: A test of the persona hypothesis. *Journal of Media Psychology: Theories, Methods, and Applications, 22*(2), p.52.

Mousavinasab, E., Zarifsanaiey, N., R. Niakan Kalhori, S., Rakhshan, M., Keikha, L. & Ghazi Saeedi, M., (2018). Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments*, *29*(1), pp.142-163.

Moustafa, A., Ben-Zvi-Assaraf, O. & Eshach, H., (2012). Do Junior High School Students Perceive Their Learning Environment as Constructivist?. *Journal of Science Education and Technology*, *22*(4), pp.418-431.

Paas, F. G., & Van Merriënboer, J. J. (1993a). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human factors, 35*(4), pp.737-743.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational psychologist, 38*(1), pp.63-71.

Park, S. (2015). The Effects of Social Cue Principles on Cognitive Load, Situational Interest, Motivation, and Achievement in Pedagogical Agent Multimedia Learning. *Journal of Educational Technology & Society*, *18*(4).Payr, S. (2003). The virtual university's faculty: An overview of educational agents. *Applied artificial intelligence*, *17*(1), pp.1-19.

Pickett, C. L., Gardner, W. L., & Knowles, M. (2004). Getting a cue: The need to belong and enhanced sensitivity to social cues. *Personality and social psychology bulletin, 30*(9), pp.1095-1107.

Plant, E. A., Baylor, A. L., Doerr, C. E., & Rosenberg-Kima, R. B. (2009). Changing middle-school students' attitudes and performance regarding engineering with computer-based social models. *Computers & Education, 53*(2), pp.209–215.

Prozesky, D. R. (2000). Communication and effective teaching. *Community Eye Health, 13*(35), p.44.

Raja, F. & Khan, N., (2018). Comparing Traditional Teaching Method and Experiential Teaching Method using Experimental Research. *Journal of Education and Educational Development*, *5*(2), p.276.Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television and new media like real people and places*. New York: Cambridge University Press.

Revelle, W., & Loftus, D. A. (1992). The implications of arousal effects for the study of affect and memory. In S.-Å. Christianson (Ed.), *The handbook of emotion and memory: Research and theory* (p. 113–149). Lawrence Erlbaum Associates, Inc.

Rey, G. D., & Steib, N. (2013). The personalization effect in multimedia learning: The influence of dialect. *Computers in Human Behavior, 29*(5), pp.2022-2028.

Richmond, V. P., Gorham, J. S., & Mccroskey, J. C. (1987). The Relationship Between Selected Immediacy Behaviors and Cognitive Learning. *Annals of the International Communication*

Association, 10(1), pp.574-590.
Richmond, V. P., Mccroskey, J. C., & Johnson, A. D. (2003). Development of the nonverbal immediacy scale (NIS): Measures of self-and other-perceived nonverbal immediacy. *Communication Quarterly, 51*(4), pp.504-517.
Rodríguez, J. I., Plax, T. G., & Kearney, P. (1996). Clarifying the relationship between teacher nonverbal immediacy and student cognitive learning: Affective learning as the central causal mediator. *Communication education, 45*(4), pp.293-305.
Rosenberg-Kima, R. B., Baylor, A. L., Plant, E. A., & Doerr, C. E. (2008). Interface agents as social models for female students: The effects of agent visual presence and appearance on female students' attitudes and beliefs. *Computers in Human Behavior, 24*(6), pp.2741–2756.
Ruijten, P. A., Midden, C. J., & Ham, J. (2015). Lonely and susceptible: the influence of social exclusion and gender on persuasion by an artificial agent. *International Journal of Human-Computer Interaction, 31*(11), pp.832-842.
Schneider, G., & Winters, J. P. (2001). *Applying use cases: A practical guide*. Pearson Education.
Schneider, S., Nebel, S., Pradel, S., & Rey, G. D. (2015). Introducing the familiarity mechanism: A unified explanatory approach for the personalization effect and the examination of youth slang in multimedia learning. Computers in Human Behavior, 43, 129-138.
Schroeder, N. L., & Adesope, O. O. (2014). A systematic review of pedagogical agents' persona, motivation, and cognitive load implications for learners. *Journal of Research on Technology in Education, 46*(3), pp.229-251.
Schroeder, N. L., & Gotch, C. M. (2015). Persisting issues in pedagogical agent research. *Journal of Educational Computing Research*, *53*(2), pp.183-204.
Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013). How Effective are Pedagogical Agents for Learning? A Meta-Analytic Review. *Journal of Educational Computing Research, 49*(1), pp.1-39.
Scott, M. D., Yates, M. P., & Wheeless, L. R. (1975, April). An exploratory investigation of the effects of communication apprehension in alternative systems of instruction. *In annual convention of the International Communication Association,* Chicago.
Sequeira, A., (2012). Introduction to Concepts of Teaching and Learning.
Servilha, E. A., & Costa, A. T. (2015). Conhecimento vocal e a importância da voz como recurso pedagógico na perspectiva de professores universitários. *Revista CEFAC, 17*(1), pp.13-26.
Simonds, B. K., Meyer, K. R., Quinlan, M. M., & Hunt, S. K. (2006). Effects of Instructor Speech Rate on Student Affective Learning, Recall, and Perceptions of Nonverbal Immediacy, Credibility, and Clarity. *Communication Research Reports, 23*(3), pp.187-197.
Swanson, R. & Falkman, S., (1997). Training delivery problems and solutions: Identification of novice trainer problems and expert trainer solutions. *Human Resource Development Quarterly*, *8*(4), pp.305-314.
Sweeney, M., Maguire, M., & Shackel, B. (1993). Evaluating user-computer interaction: A

framework. *International Journal of Man-Machine Studies, 38*(4), pp.689–711.
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science, 12*(2), pp.257-285.
Sweller, J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instructional science*, *32*(1), pp.9-31.
Sweller, J. (2011). Cognitive load theory. In J. Mestre & B. Ross (Eds.), *The psychology of learning and motivation: Cognition in education* (Vol. 55, pp. 37–76). Oxford: Academic.
Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and instruction*, *12*(3), pp.185-233.
Sweller, J., Ayres, P., & Kalyuga, S. (2011). Intrinsic and extraneous cognitive load. In *Cognitive load theory* (pp. 57-69). Springer, New York, NY.
Szymanski, D. M. (1988). Determinants of selling effectiveness: The importance of de-clarative knowledge to the personal selling concept. *Journal of Marketing*, p.64–77.
Titsworth, B. S. (2001). The effects of teacher immediacy, use of organizational lecture cues, and students' notetaking on cognitive learning. *Communication education, 50*(4), pp.283-297.
Utari, A. Y. S. (2013). the usage of multiple choice and gap filling in measuring grade-schoolers' understanding of grammar. *Journal of English Language Teaching in Indonesia, 1*(1).
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), pp.381-403.
VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46*(4), pp.197-221.
Veletsianos, G. & Miller, C., (2008). Conversing with pedagogical agents: A phenomenological exploration of interacting with digital entities. *British Journal of Educational Technology*, *39*(6), pp.969-986.
Veletsianos, G. (2009). The impact and implications of virtual character expressiveness on learning and agent-learner interactions. *Journal of Computer Assisted Learning, 25*(4), pp.345-357.
Veletsianos, G., Miller, C., & Doering, A. (2009). EnALI: A research and design framework for virtual characters and pedagogical agents. *Journal of Educational Computing Research, 41*(2), pp.171–194.
Venkatesh, V. (2000). Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Information Systems Research, 11*(4), pp.342–365.
Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences, 39*(2), pp.273–315.
Venkatesh, V., & Davis, F. D. (1996). A model of the antecedents of perceived ease of use: *Development and test. Decision Sciences, 27*(3), pp.451–481.
Wang, F., Li, W., Mayer, R. E., & Liu, H. (2018). Animated pedagogical agents as aids in

multimedia learning: Effects on eye-fixations during learning and learning outcomes. *Journal of Educational Psychology, 110*(2), p.250.

Westlund, J. M., Jeong, S., Park, H. W., Ronfard, S., Adhikari, A., Harris, P. L., . . . Breazeal, C. L. (2017). Flat vs. Expressive Storytelling: Young Children's Learning and Retention of a Social Robot's Narrative. *Frontiers in Human Neuroscience*, 11.

Witt, P. & Wheeless, L., (2001). An experimental study of teachers' verbal and nonverbal immediacy and students' affective and cognitive learning. *Communication Education*, *50*(4), pp.327-342.

Witt, P., Wheeless, L. & Allen, M., (2004). A meta-analytical review of the relationship between teacher immediacy and student learning. *Communication Monographs*, *71*(2), pp.184-207.

Woo, H. (2009). Designing multimedia learning environments using animated pedagogical agents: Factors and issues. *Journal of Computer Assisted Learning, 25*(3), pp.203-218.

Woolf, B. P. (2009). *Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning.* Burlington, MA: Morgan Kaufmann.

Wouters, P., Paas, F., & van Merriënboer, J.J.G. (2008). How to optimize learning from animated models: A review of guidelines based on cognitive load. *Review of Educational Research*, 78(3), pp.645-675.

Yi, M. Y., & Davis, F. D. (2003). Developing and validating an observational learning model of computer software training and skill acquisition. *Information Systems Research, 14*(2), pp.146–169.

List of Figures

List of Tables

Appendices

The following appendices are included on the next pages:

A. Questionnaires of Chapter 2

This appendix contains the questionnaires as used in Study 1 (specific self-efficacy, system's perceived ease of use and general self-efficacy) and Study 2 (declarative knowledge test and user booklet) of Chapter 2.

B. Questionnaires of Chapter 3

This appendix contains the questionnaires as used in Study 1 (recall test for both psychomotor and cognitive tasks and affective learning) and Study 2 (performance-related cognitive load for both psychomotor and cognitive tasks, recall test both psychomotor and cognitive tasks and recall-related cognitive load for the psychomotor task) of Chapter 3.

C. Questionnaires of Chapter 4

This appendix contains the questionnaires as used in Study 1 (nonverbal vocal immediacy, affective learning, motivation, attention, recall test) of Chapter 4.

Appendix A Questionnaires of Chapter 2

A.1 Questionnaire of Study 1

Specific computer self-efficacy

- I believe I have the ability to scroll up and down on a webpage using the GazeTheWeb browser.
- I believe I have the ability to begin a web search using the GazeTheWeb browser.
- I believe I have the ability to correct errors made when typing a search term using the GazeTheWeb keyboard.
- I believe I have the ability to open hyperlinks using the GazeTheWeb browser.
- I believe I have the ability to type a search term using the GazeTheWeb keyboard.

System's perceived ease of use

- My interaction with GazeTheWeb will be clear and understandable.
- Interacting with GazeTheWeb will not require a lot of my mental effort.
- I will find it easy to get GazeTheWeb to do what I want it to do.
- Overall, I will find GazeTheWeb easy to use.

General self-efficacy

I could complete a job using a software package…

- …If there was no one around to tell me what to do as I go.
- …If I had only the software manuals for reference.
- …If I had seen someone else using it before trying it myself.
- …if I could call someone for help when I got stuck.
- …If someone else had helped me get started.

- …If I had a lot of time to complete the job for which the software was provided.
- …If I had just the built-in help facility for assistance.
- …If somebody else showed me how to do it first.

A.2 Questionnaire of Study 2

Multiple-choice questions (for declarative knowledge assessment)

1. Which of the following is NOT an icon of GazeTheWeb?

A. A “T” icon.
B. A finger-point button.
C. A diamond icon.
D. A “Y” icon.

2. Which of the follow propositions is true?

A. While scrolling, the scroll buttons change color form orange to brown indicating the scroll progress.
B. While scrolling, the scroll buttons change color form orange to brown indicating the number of most relevant page results.
C. While scrolling, the scroll buttons change color form orange to blue indicating the scroll progress.
D. While scrolling, the scroll buttons change color form orange to blue indicating the number of most relevant page results.

3. What happens when the “hyperlink navigation” function is activated?

A. The web page starts zooming in.
B. The eye icon changes to a finger-point button.
C. A new tab is open for every link present on-screen.
D. You can copy the URL of the desired link.

4. The color of the cursor on GazeTheWeb is white, so as…

A. ...to be easier to find it on screen.
B. ...to prevent eye strain.
C. ...to be aesthetically pleasing.
D. ...to increase the productivity.

5. The right-hand side panel of GazeTheWeb contains…

A. …buttons such as the tab overview, the going back and the forward button.
B. …buttons to interact with a web page (e.g. link selection, automatic scrolling button).
C. …buttons to apply a query to search engine or directly start the search.

D. …the GazeTheWeb logo and the input field.

6. The backspace icon is located…

A. ...at the left-hand side of the virtual keyboard.
B. ...at the right-hand side of the virtual keyboard.
C. ...in the middle of the virtual keyboard.
D. ...at the bottom of the virtual keyboard.

7. To deactivate the “automatic scrolling” function you must…

A. ...select another button and wait for automatic swap among functions.
B. ...focus your gaze on the icon related to the function you want to deactivate.
C. ...move back to the homepage.
D. ...activate the keyboard and wait for automatic deactivation of any previously activated function.

8. The left-hand side panel of GazeTheWeb contains…

A. ...buttons such as the tab overview, the going back and the forward button.
B. ...buttons to interact with a web page (e.g. link selection, scrolling button).
C. ...buttons to apply a query to search engine or directly start the search.
D. ...the GazeTheWeb logo and the input field.

9**. What is the difference between icons with the arrows on the right-hand side and the left-hand side of the keyboard?**

A. The arrows on the right-hand side move the cursor to the previous/next letter while the arrows on the left-hand side move the cursor to the previous/next word.
B. The arrows on the right-hand side move the cursor to the previous/next word while the arrows on the left-hand side move the cursor to the previous/next letter.
C. the arrows on the right-hand side move the cursor only to the previous letter while the arrows on the left-hand side move the cursor only to the next letter.
D. the arrows on the right-hand side move the cursor only to the previous word while the arrows on the left-hand side move the cursor only to the next word.

10. Which icon activates the “automatic scrolling” function?

A. A left-oriented arrow.
B. A square containing a number.
C. A diamond.
D. None of the previous: automatic scrolling is activated by default.

11. The finger-point button…

A. ...activates the keyboard.
B. ...activates the copy-paste function.
C. ...allows to pass from a tab page to another.
D. ...allows hyperlink navigation.

12. Besides automatic scrolling, scroll buttons are also located…

A. ...at the top and the bottom of a page.
B. ...at the left and right side of a page.
C. ...in the middle of a page.
D. ...there are no such scroll buttons.

13. Similar to other icons, how do you deactivate the zoom icon once it is selected?

A. it is deactivated automatically once it is used.
B. By focusing the eyes on the same icon.
C. By moving back to the homepage.
D. By focusing forward to the next page.

14. Which of the GazeTheWeb icons below would you use to start the search directly after you have typed your search term on the keyboard?

A.  B. 

C.  D. 

15. Which of the GazeTheWeb icons below would you use to apply the search term on the search engine after you have typed it on the keyboard?

A.  B. 

C.  D. 

User booklet (for task performance assessment)

Instructions: Conduct a web search using GazeTheWeb

Your task involves performing a web search using GazeTheWeb, as explained in the instructional video you watched before.

Specifically:
Your mission is to type *hello world* and then, to open the “hello word “program-Wikipedia page. Once it opens, find the last paragraph (located almost at the bottom of the page, just before the references section).
Find the last two words of this last paragraph and write them down in the sentence below:

“*It is significantly more useful for developers, however, as it provides an example of how to create a deb package, either traditionally or using debhelper, and the version of hello used, GNU Hello, serves as an example of how to write a* _______ ____________”.

Tips before you start:

- Please perform the task as fast and as accurately as possible. Both errors and speed are equally important for a performance to be considered successful.
- You can ask experimenters’ assistance when you think that their intervention is necessary. However, please ask for assistance only after you have tried yourself independently.
- The above does not apply in case you encounter any problems with the eye-tracker. In case you face any difficulty in operating the system with your eyes call the experimenter immediately.

Appendix B Questionnaires of Chapter 3

B.1 Questionnaire of Study 1

Recall test for psychomotor task

1. What is the hand gesture for moving a block to the right?

a) Using the right hand, release fingers wide and then make a fist.
b) Using the left hand, make a fist with the left hand, and then release fingers wide.
c) Using the right hand, make a fist, and then release fingers wide.
d) Using the left hand, release fingers wide and then make a fist.
e) I do not know.

2. What is the hand gesture for moving a block to the left?

a) Using the left hand, release fingers wide and then make a fist.
b) Using the right hand, make a fist, and then release fingers wide.
c) Using the right hand, release fingers wide and then make a fist.
b) Using the left hand, make a fist with the left hand, and then release finger wide.
e) I do not know.

3. What is the hand gesture for rotating a block clockwise?

a) Using the right hand turn your palm of up towards the ceiling, and then down towards the floor.
b) Using the right hand, turn your palm down towards the floor and then up towards the ceiling.

c) Using the left hand, turn your palm of up towards the ceiling, and then down towards the floor.
d) Using the left hand, turn your palm down towards the floor and then up towards the ceiling.
e) I do not know.

4. What is the hand gesture for rotating a block counterclosckwise?
a) Using the left hand, turn your palm of up towards the ceiling, and then down towards the floor.
b) Using the right hand turn your palm of up towards the ceiling, and then down towards the floor.
c) Using the left hand, turn your palm down towards the floor and then up towards the ceiling.
d) Using the right hand, turn your palm down towards the floor and then up towards the ceiling.
e) I do not know

Recall test for cognitive task

1. What is the math operation for moving a block to the right?
c) L+X
b) M-L
c) L+M
d) L-M
e) I do not know

2. What is the math operation for moving a block to the left?
a) L+M
b) L+X
c) M-L
d) L-M
e) I do not know

3. What is the math operation for rotating a block to the clockwise?
a) X + L = (+) R * 90
b) X + L = -R * 90
c) X + R = -L * 90
d) X+R= (+) L*90
e) I do not know

4. What is the math operation for rotating a block counterclosckwise?
   a) X + L = (+) R * 90
   b) X+R= (+) L*90
   c) X + L = -R * 90
   d) X + R = -L * 90
   e) I do not know

Affective learning (for both psychomotor and cognitive tasks)

I feel the content of the video, pertained to playing Tetris using hand movements/using math calculations was:

- "Bad – Good"
- "Worthless – Valuable"
- "Negative – positive"

I feel that the instructor of the video on how to play Tetris using hand movements/using math calculations was:

- "Bad – Good"
- "Worthless – Valuable"
- "Negative – positive"

B.2 Questionnaire of Study 2

Performance-related cognitive load for psychomotor task

- How much mental effort did you invest in completing the performance test?
- How easy or difficult was it to complete the performance test?
- In terms of learning how to play the Tetris game by performing the gestures...
  - o ... the demonstrations and instructions in this video were very ineffective.
  - o ... the demonstrations and instructions in this video were very unclear.
  - o ... the demonstrations and instructions in this video used unclear movements and language.

Recall test for psychomotor task (gap filling questions)

"Now I will explain how you can rotate a block. To do this: hold one ___ (1) next to you with your ___ (2) close to your ___ (3) and your ___ (4) 90 degrees to your ___ (5) arm."

Recall-related cognitive load for psychomotor task

- How much mental effort did you invest in completing the recall test?
- How easy or difficult was it to complete the recall test?
- In terms of recalling the gestures...

o ... the demonstrations and instructions in this video were very ineffective.
o ... the demonstrations and instructions in this video were very unclear.
o ... the demonstrations and instructions in this video used unclear movements and language.

Performance-related cognitive load for cognitive task

- How much mental effort did you invest in completing the performance test?
- How easy or difficult was it to complete the performance test?
- In terms of learning how to play the Tetris game by solving math calculations...
  o ... the demonstrations and instructions in this video were very ineffective.
  o ... the demonstrations and instructions in this video were very unclear.
  o ... the demonstrations and instructions in this video used unclear movements and language.

Recall test for cognitive task (gap filling questions)

"L is the line score, and, ___ (1) is the number of times that you want to rotate a block. This number gets a ___ (2) sign for a ___ (3) rotation and a ___ (4) sign for ___ (5) rotation".

Appendix C Questionnaires of Chapter 4

C.1 Questionnaire of Study 1

Nonverbal vocal immediacy

The artificial teacher...

- ...has an unpleasant / annoying voice - Has a pleasant voice.
- ...uses an inexpressive / dull voice - Uses an expressive / energetic voice.
- ...has a boring / unanimated voice - Has an enthusiastic / animated voice.
- ...has an unappealing/ unengaging voice - Has an appealing / engaging voice.
- ...has an unfriendly voice - Has a friendly voice.

Affective learning

- I feel the content of the instructional video is:
  - o "Bad – Good"
  - o "Worthless – Valuable"
  - o "Negative – positive"

- I feel that the teacher I had during the instructional video is:
  - o Bad – Good"
  - o "Worthless – Valuable"
  - o "Negative – positive"

- My likelihood of taking future video tutorials (assuming they were available) with this specific teacher is:
  - o "Bad – Good"
  - o "Worthless – Valuable"
  - o "Negative – positive"

Motivation

During the instructional video I felt...

- "Motivated – Unmotivated"
- "Interested – Uninterested"
- "Involved – Uninvolved"
- "Not stimulated – Stimulated"
- "Inspired – Uninspired"
- "Unenthused – Enthused"
- "Excited - Not excited"
- "Aroused - Not aroused"
- "Not fascinated – Fascinated"

Attention

During the instructional video:

- I paid close attention to the instructional video.
- I was able to concentrate on the video.
- The video held my attention.
- I was absorbed by the presented software activity.

Recall test

A) Gap-filling questions

- Hello, this is Eric, and, in this tutorial, we will cover the basics of 1) ___. This is a new web browser that you can control using 2) ___.
- To move around in the text faster and to correct possible typing errors you can use the icons with the 3) ___.
- GazeTheWeb can capture your eye movements through a 4) ___ 5) ___.
- Page Navigation is possible by looking at the scroll buttons, located on the 6) ___ and 7) ___of the page.
- While scrolling, these buttons change color from 8) ___ to 9) ___ indicating my scroll progress.

B) Multiple choice questions

1. The right-hand side panel of GazeTheWeb contains…

A. Buttons such as the tab overview, the going back and the forward button.
B. Buttons to interact with a web page (e.g. link selection, automatic scrolling button).
C. Buttons to apply a query to search engine or directly start the search.
D. The GazeTheWeb logo and the input field.
E. I don't know.

2.The left-hand side of GazeTheWeb contains …

A. ...buttons such as the tab overview, the going back and the forward button.
B. ...buttons to interact with a web page (e.g. link selection, automatic scrolling button).
C. ...buttons to apply a query to search engine or directly start the search.
D. ...the GazeTheWeb logo and the input field.
E. I don't know.

3. Which of the following is NOT an icon of GazeTheWeb?

A. A "T" icon.
B. A finger-point button.
C. A diamond icon.
D. A "Y" icon.
E. I don't know.

4. To begin a web search on GazeTheWeb, you should focus your eyes on the T button until it changes color…

A. ...from red to brown.
b. …from orange to blue.
C.…from orange to brown.
D.…from grey to blue.
E. I don't know.

5. The backspace icon is located…

A. ...at the left-hand side of the virtual keyboard.
B. ...At the right-hand side of the virtual keyboard.
C. ...in the middle of the virtual keyboard.
D. ...at the bottom of the virtual keyboard.
E. ...I don't know.

6. What is the difference between icons with the arrows on the right-hand side and the left-hand side of the keyboard?

A. The arrows on the right-hand side move the cursor to the previous/next letter

while the arrows on the left-hand side move the cursor to the previous/next word.
B. The arrows on the right-hand side move the cursor to the previous/next word while the arrows on the left-hand side move the cursor to the previous/next letter.
C. The arrows on the right-hand side move the cursor only to the previous letter while the arrows on the left-hand side move the cursor only to the next letter.
D. The arrows on the right-hand side move the cursor only to the previous word while the arrows on the left-hand side move the cursor only to the next word.
E. I don't know.

7. Which icon activates the "automatic scrolling" function?

A. A left-oriented arrow.
B. A square containing a number.
C. A diamond.
D. None of the previous: automatic scrolling is activated by default.
E. I don't know.

8. What happens when the "hyperlink navigation" function is activated?

A. The web page starts zooming in.
B. The eye icon changes to a finger-point button.
C.A new tab is open for every link present on-screen.
D. You can copy the URL of the desired link.
E. I don't know.

9. how do you deactivate the zoom icon once it is selected?

A. it is deactivated automatically once it is used.
B. By focusing the eyes on the same icon.
C. By moving back to the homepage.
D. By focusing forward to the next page.
E. I don't know.

10. The finger-point button…

A. ...activates the keyboard.
B. ...activates the copy-paste function.
C. ...allows to pass from a tab page to another.
D. ...allows hyperlink navigation.
E. I don't know.

11. The text selection icon contains...

A. ...pencil.
B. ...the capital letter A.
C. ...the letters ABC.
D. ...the numbers 123.
E. I don't know.

12. Immediately after the text selection button has been clicked, a message appears on screen asking....

A. ...to look at the end point of the text selection.
B. ...to copy text to clipboard.
C. ...to look at the starting point of the text selection.
D. ...to look both at the starting and end point of the text selection.
E. I don't know.

13. The clock button...

A. ...is used to bookmark a page.
B. ...shows the tabs that are currently open.
C. ...is used to cancel the action and go back to the navigation panel.
D. ...shows the history of all the actions performed within the GazeTheWeb environment.
E. I don't know.

14. How would you access bookmarks that have been already saved?

A. By focusing on the pencil button and then on the star-shaped button.
B. By focusing on the pencil button and then on the agenda button.
C. By focusing on the star-shaped button and then on the pencil button.
D. By focusing on the agenda button and then on the pencil button.
E. I don't know.

15. Which of the GazeTheWeb icons below would you use to start the search directly after you have typed your search term on the keyboard?

A. 

B. 

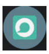
C. 

D. 

16. To create a bookmark, which of the gaze the web icons below would you use?

A. 

B. 

C. 

D. 

17. To reload a tab, which of the gaze the web icons below would you use?

A. A chatbox.
B. A pencil.
C. A star.
D. A clock.
E. I don't know.

18. Which of the GazeTheWeb icons below would you use to apply the search term on the search engine after you have typed it on the keyboard?

A. 

B. 

C. 

D. 

Artificial agent's likeability

- "Dislike – Like"
- "Unfriendly – Friendly"
- "Unkind – Kind"
- "Unpleasant – Pleasant"
- "Awful – Nice"

Summary

Understanding artificial agents as facilitators of learning

The role of the teacher has transformed over time from traditionally being disseminator of information to facilitator of learning. This transformation, coupled with the increased availability and sophistication of technology in recent decades, motivated the question of whether technology can become the teacher itself. This question led to a new line of research examining the use of embedded artificial agents (so called pedagogical agents)-anthropomorph virtual characters that serve various instructional functions in multimedia learning environments.

Despite artificial agents' vast potential as educational tools, findings regarding their overall effectiveness for learning are mixed. In the current thesis, we argued that in light of the great variety of artificial agents used in past studies, performing different tasks and roles in different contexts, studying their effectiveness in general is too broad to receive a simple answer. Instead, we proposed that a more fruitful approach would be to explore under which conditions artificial agents can facilitate learning. Thus, the overall goal of the current thesis was to examine under which conditions and in which ways an artificial agent can facilitate learning.

Despite the fact that human teachers' instructional methods have been found to have a tremendous impact on students' learning, the instructional methods an artificial agent applies in the multimedia learning environment has received little attention in earlier research. The current work, in line with voices echoing that artificial agents may be effective due to their pedagogy rather than merely their appearance, argues that more research is warranted

on artificial agents' instructional method. Thus, the research question, examined in Chapter 2, was whether the instructional method of modeling (learning by observing another's behavior) employed by an artificial agent is effective for learning. To answer this question, we conducted two experimental studies, in which modeling by an artificial agent was compared to other commonly used non-modeling instructional methods: a) agent-delivered instructional narration (=agent as a source of information), b) no agent, text-only instruction, and c) no agent, voice-only instructional narration. Overall, according to our hypotheses, findings of Chapter 2 show that an artificial model can positively influence learners' motivational (i.e., computer self-efficacy and perceptions of ease of use of the system), cognitive (i.e., declarative knowledge) and, most importantly, skill-based (i.e., task performance) learning outcomes, as compared to other popular non-modeling instructional methods.

The negligence of systematic examination of artificial agents' instructional method has been attributed to the argument that it does not technically require an embodied agent. In fact, two competing perspectives exist in the literature on whether the visibility of an artificial agent in multimedia settings hinders or augments learning ("agents-as-complements" versus "agents-as-distractors"). In this thesis, we argued that it is crucial to investigate this issue further, because in contrast to modeling taking place in classrooms, modeling taking place in multimedia settings can occur without the actual visual presence of a model. Therefore, the research question examined in Chapter 3 was whether the positive effects of modeling by an artificial agent on learning depend on the visual presence of the artificial agent. Specifically, we examined whether the type of learning task (psychomotor versus cognitive) is a decisive condition for the inclusion of an artificial agent's visual presence. Specifically, the first experimental study of Chapter 3 aimed to examine effects of the interaction between the on-screen visibility of an artificial model (presence vs. absence) and type of task to be modelled (psychomotor vs. cognitive) on learning outcomes (recall, affective beliefs, and task performance). Thus, in Study 1, it was implicitly assumed that the visual presence of the artificial agent has a different effect on learners' cognitive processes depending on the type of task it models. Therefore, in the second experimental study of Chapter 3, we aimed to extend findings, by examining whether learners' perceived cognitive load changes depending on the match between the visibility of the artificial model and the type of task. Overall, confirming our hypotheses, findings of Chapter 3 show that the visual presence of the artificial model enhances learners' self-efficacy, affective beliefs and task performance, and that visual presence also minimizes cognitive load associated with task performance for psychomotor tasks (behavioral modeling), but less so for cognitive tasks (cognitive modeling).

The final goal of the current thesis was to examine the conditions that increase the effectiveness of a visible artificial agent for learning. Specifically, the research question examined in Chapter 4 was whether and how an artificial models' nonverbal behavior can increase learning outcomes (affective and cognitive learning). To answer this question, we conducted two experimental studies, in which an artificial model showing strong vocal expressiveness (i.e., higher pitch tone, more pitch variation, higher speech rate) was compared to an artificial model showing weak vocal expressiveness (i.e., lower pitch tone, less pitch variation and lower speech rate). Overall, confirming our hypotheses, findings of Chapter 4 showed that an artificial agent with strong vocal expressiveness increased affective learning, perceived cognitive learning and, to a smaller extent, actual cognitive learning (immediate recall). What is more, our findings revealed that vocal expressiveness is related to affective learning and perceived cognitive learning because it reduces psychological distance, and thereby promotes immediacy. Finally, our findings motivation theory which argues that strong nonverbal cues, such as vocal expressiveness, increase perceptions of immediacy, which has a positive effect on motivation leading to enhanced affective learning.

The work presented in the current thesis revealed that modeling by an artificial agent is more effective than other non-modeling methods in enhancing learners' beliefs of self-efficacy, affective learning, and task performance. What is more, our work also revealed the conditions under which the visual presence of an artificial model is beneficial for learning. That is, the learning task being modeled is a decisive factor of whether the visual presence of an artificial model enhances learning. More specifically, the visual presence of an artificial model is beneficial for psychomotor tasks as it increases self-efficacy beliefs, and affective and skill-based learning. However, this is not the case for cognitive tasks. Further, the current results showed that the benefits of behavioral modeling by a visible artificial agent for affective learning are strengthened when the agent employs enhanced nonverbal behavior (i.e., vocal expressiveness).

Overall, the findings of the current thesis are important as they increase our understanding of the conditions under which artificial agents as teachers can facilitate learning. This work can help create such agents that might one day be important aides in education and other related domains crucial to our society.

Acknowledgements

This dissertation is the end result of a long path in which I was lucky enough to have been surrounded by exceptional people. Without all their support, this project would have not be the same. For this reason I dedicate the following lines to the many people who accompanied me within this journey.

First and foremost I would like to express my deepest gratitude to my promotor Cees Midden and daily supervisors Jaap Ham and Uwe Matzat for the continuous support of my research endeavours, for their patience, motivation and immense knowledge. Your guidance, helped throughout my research and writing of the thesis. I could not imagine having better advisors and mentors during my PhD years.

Cees, from all the teachers I met during my student years, you are by far the most important. During our collaboration, I admired your ability to actively listen to my research ideas and explanations of findings, and your incredible skill to ask the right questions that would direct my thought process to interesting and unexpected paths. My mind never reached so far as it did during our discussions. Further, I truly appreciate that you empowered me to take charge of my PhD. The sense of full control that you gave me over my work provided me with the right recipe for success: freedom of choice and responsibility for the consequences. For all the above reasons and for many more I feel honoured and grateful to have met and worked with you for such a long time.

Jaap, I still remember the first time I met you during my initial interview for the PhD position. I remember how at ease I felt to introduce myself to you and to present my earlier work and my research ideas. Years later, I am still very grateful to you for providing me with this opportunity. I had the chance to spend more time and work with you closer, as besides my PhD studies, we have also worked together on the MAMEM project. We have travelled together for project meetings and conferences and it was always such a pleasure to spend time with you and get to know you better. I truly admired your sense of humour, your positive and optimistic attitude, your faith in me despite the difficulties I faced during my early PhD years, your support and your ability to find solutions to various practical problems I had faced.

Uwe, I feel a deep sense of gratitude because half of the studies presented in this thesis were realized because of your effective supervision. You were the one that paid attention to some of my initial research conceptions, introduced my topic to students and helped me co-supervise student projects together with you to explore my research ideas. During this time, I admired your ability to keep a sharp eye for details, your precision, punctuality, encouragement and the psychological support you gave me during the challenging moments I faced during the early years of my PhD. You were always available to meet and listen to me, revise my work and provide me with invaluable feedback.

Special thanks goes to the members of my committee defense – Eric Postma, Ellen Christiansen and Panos Markopoulos, who spent their precious time reading my thesis in full, for their positive words on my thesis and their interesting questions and discussions. Many thanks also to Professor van Houtum for accepting to be the chair of my defense.

I would like to thank colleagues from the European project MAMEM. It was an amazing experience being part of such an interesting project and to have been surrounded by such smart people. My Special thanks goes to Spiros for his effective co-ordination and management of the project, to Chandan for the great laughs we had and to Dario for the amazing person that he is.

Much of my gratitude goes to colleagues and friends from the HTI group for all the nice moments in and out the University. In particular, I would like to thank Shengnan, Alain and Sima. My dear Shengnan, thank you so much for all your support, particularly during my first year as a PhD student. I truly admired your kind heart, lively personality and sharp mind and I am very grateful to you for staying by my side during all these happy and challenging times. Alain, I consider myself very lucky to have shared the same office with you for such a long time. Thank you for introducing me to the Dutch culture in your own unique way, for translating Dutch news to English for me and for your incredible sense of humour. Sima, thank you for the amazing collaboration we had for some parts of the MAMEM project, your hard work and for your patience when I was coming up with crazy ideas, for our trip to Berlin, and for your effort to keep in touch with me despite the distance we had during the last 2 years. I am also thankful to Peter for always having the time to listen to me and providing me with solutions to practical issues I experienced and to Hanne and Chao for our unforgettable trip to a far, cold place called Kitchener. Many thanks to our secretaries, Anita and Renata and technical staff for being so helpful, patient and supportive. I would also like to thank the students I supervised, who chose to work with me and helped me collect part of the data used in some of the studies of this thesis. Finally, I am grateful to Veronica for designing the cover and the lay-out of this thesis and for treating me like a sister.

To all my friends I met outside the University, thank you for your support and the time we spent together. Specifically, most of my gratitude goes to Vangelis and Danai for their friendship and support all these years. I am very thankful for our Wednesday movie nights, for our trip to Paris, our Greek dinners and gatherings and for becoming "housemates" the last months of my PhD. You are both incredible people and I wish you will be part of my life for many years to come. I would also like to thank Daniela, Giannis and Makis for being such good friends and for making my time in Eindhoven fun and enjoyable.

I am thankful to my parents in Greece for always supporting my choices, for encouraging me to be independent and to explore my physical and mental space with excitement rather than fear. Μπαμπά, μαμά, ευχαριστώ που είστε πάντα δίπλα μου, για την αγάπη και τη φροντίδα σας. I would also like to thank my sister and brother for helping me in every possible way and motivating me to reach as far as I can. You are my best friends and my dearest people! Further, I would like to thank my grandmother for the endless support and assistance since the day I was born. Γιαγιάκα, σ' ευχαριστώ για ότι έχεις κάνει για μένα από την μέρα που γεννήθηκα. Θα σου είμαι ευγνώμων για πάντα. Special thanks goes to my daughter, my little Ariadni, who motivates me to become a better version of myself every day.

Giorgos, I am eternally grateful to you because whatever I am today it is because of you. You have this incredible way of making my heart beat fast and my mind run wild. Life is exciting with you. Μίου.

Curriculum Vitae

Sofia Fountoukidou was born on 02-06-1986 in Thessaloniki, Greece. After finishing high school in 2004 in Thessaloniki, she started the Psychology, Education & Philosophy program at the University of Ioannina in Ioannina, Greece. In 2009 she received her Bachelor of Science degree, and in 2011 she also obtained a diploma in Psychology from Glasgow Caledonian University in Glasgow, Scotland. In 2015 she obtained the Master of Science degree in Psychology under two specializations, Clinical Psychology and Social & Organizational Psychology, from Leiden University in Leiden, the Netherlands. In October 2015, she started a PhD project at Eindhoven University of Technology in Eindhoven, the Netherlands, of which the results are presented in this dissertation. Her PhD project was part of an EU-funded research project, titled Multimedia Authoring & Management using your Eyes & Mind (MAMEM). Besides her PhD work, she also contributed to this research project by, amongst other, helping to develop a gamified training app, where various persuasive strategies were incorporated to enhance user engagement and technology acceptance. During her time as a PhD student she presented her work at a number of national and international conferences. In 2018, she received the Best Paper Award at the 13th International Conference on Persuasive Technology, in Waterloo, Ontario, Canada. She also published several papers in conferences and in journals. Since 2021, she is employed at Cambridge University Press, in Cambridge, England, as a Senior User Experience Researcher.