META-ANALYSIS

Effectiveness of Multimedia Pedagogical Agents Predicted by Diverse Theories: a Meta-Analysis

Juan C. Castro-Alonso [1] · Rachel M. Wong [2] · Olusola O. Adesope [2] · Fred Paas [3,4]

Abstract

Multimedia pedagogical agents are on-screen characters that allow users to navigate or learn in multimedia environments. Several agents' characteristics may moderate their instructional effectiveness, including appearance, gender, nonverbal communication, motion, and voice. Here, we conducted a meta-analysis to test hypotheses from diverse theories predicting the effects of these agents' characteristics. We tested predictions of cognitive load theory, cognitive theory of multimedia learning, computers are social actors, social agency theory, uncanny valley, and the action observation network. Our meta-analysis of 32 effect sizes ($N = 2104$) revealed a small overall effect ($g+ = 0.20$), showing that learning with multimedia pedagogical agents was more effective than learning without these agents. As predicted by the redundancy effect of cognitive load theory and the coherence principle of cognitive theory of multimedia learning, 2D agents ($g+ = 0.38$) tended to be more effective than 3D agents ($g+ = 0.11$). As predicted by the computers are social actors hypothesis, most of the agents' characteristics, including nonverbal communication, motion, and voice, appeared not to moderate their effectiveness. We conclude that multimedia pedagogical agents help learning through multimedia, and that students may be able to learn similarly from different types of agents.

Keywords Multimedia animated pedagogical agent · Cognitive load theory and cognitive theory of multimedia learning · Computers are social actors and social agency theory · Uncanny valley and action observation network · Meta-analysis

✉ Juan C. Castro-Alonso
jccastro@ciae.uchile.cl

Rachel M. Wong
rachel.wong@wsu.edu

Olusola O. Adesope
olusola.adesope@wsu.edu

Fred Paas
paas@essb.eur.nl

Extended author information available on the last page of the article

Learning is a complex process that involves not only individual cognitive processing by students, but also social exchanges between students and instructors (e.g., Gehlbach 2010). As described by Burgoon et al. (2000), when learners and instructors share a physical space, they exchange rich social information (e.g., anthropomorphism, nonverbal communication, and voice) that can aid the learning process. However, as learners and instructors are moving away from physically coincident dialogs, into interactions through different media (e.g., smartphones, computers, and virtual reality), the rich exchange of information may diminish, resulting in hindered communication and learning (Burgoon et al. 2000; see also Sbarra et al. 2019; Syrjämäki et al. 2020).

In the present meta-analysis, we investigated if this diminishing effect could be observed when the media instructor was a multimedia pedagogical agent. A pedagogical agent, as defined by Heidig and Clarebout (2011), is a lifelike on-screen character who allows users to navigate or learn in multimedia environments. A multimedia environment, according to Mayer (2014b), presents images and texts. In consequence, a multimedia pedagogical agent has an image (appearance) and conveys textual information, for example, as narrations.

Multimedia pedagogical agents can range in complexity from simple static characters that provide information to complex and dynamic animated three-dimensional agents that narrate information while gesturing (see Schroeder et al. 2013). Although these agents can also range in behavior, they are independent, not avatar or replica agents controlled by a human. Despite the negative predictions of media learning by Burgoon et al. (2000), there is mounting evidence for the positive effects of multimedia pedagogical agents on learning with multimedia modules (e.g., Schroeder et al. 2013; Wang et al. 2018; Wang et al. 2020; Yılmaz and Kılıç-Çakmak 2012).

However, there is also somewhat inconclusive evidence for the instructional success of these agents (e.g., Lin et al. 2020; Moreno et al. 2001; Wang and Antonenko 2017; Yi et al. 2019), which could be partially explained by the diminished social information (Burgoon et al. 2000) provided by the agent. Investigations about specific agents' characteristics, such as how they convey social information, could provide suggestions on how to improve their instructional effectiveness (e.g., Lawson et al. 2021). This need for finer-grained analyses about multimedia pedagogical agents has been suggested before (e.g., Heidig and Clarebout 2011; Schroeder and Adesope 2014).

The primary aim of the present study was to fill this gap by conducting a contemporary meta-analysis of the moderating effects of different agents' characteristics on their instructional effectiveness, predicted by diverse theories. We focused on the agent's characteristics of appearance, nonverbal communication, motion, and voice, which have not been investigated in a meta-analysis with updated data. A previous meta-analysis by Schroeder et al. (2013) also considered appearance, motion, and voice, but it was conducted with studies until 2011, whereas the present meta-analysis included studies in the period 2012–2019.

A secondary aim was to investigate the effects of other moderating variables, external to the pedagogical agents, that have shown effects in previous meta-analyses about multimedia environments (Alpizar et al. 2020; Castro-Alonso et al. 2019b; Rey et al. 2019) and pedagogical agents (Davis 2018; Schroeder et al. 2013).

Agents' Characteristics

In the present study, we considered the following agents' characteristics or social cues: appearance, nonverbal communication, motion, and voice. As describe next, we considered different degrees in which multimedia pedagogical agents could show these characteristics.

Appearance

Concerning the appearance of the agent, in the present meta-analysis, we compared the effects of two-dimensional (2D) agents versus three-dimensional (3D) agents, like previous studies (e.g., Davis 2018; Dinçer and Doğanay 2017; Yılmaz and Kılıç-Çakmak 2012). We tested two hypotheses to predict how these two levels of appearance affected the effectiveness of multimedia agents. One hypothesis was based on cognitive load theory (see Sweller et al. 2011, 2019) and cognitive theory of multimedia learning (see Mayer 2014a). Another hypothesis was based on computers are social actors (see Nass and Moon 2000; Nass and Steuer 1993).

Both the redundancy effect of cognitive load theory (Castro-Alonso et al. 2019a; Kalyuga and Sweller 2014) and the coherence principle of cognitive theory of multimedia learning (Mayer and Fiorella 2014; Mayer et al. 2008) promote a design that avoids nonessential visual information, because having to coordinate redundant with essential information may surpass the limits of working memory (see Oberauer et al. 2018) and interfere with learning. For example, a photograph or highly realistic illustration may show more information than what is relevant for the learning task (e.g., Brucker et al. 2014; Menendez et al. 2020). Following this rationale, an agent with a 2D appearance would be more effective than a more visually complex 3D or 3D-like agent.

The prediction of computers are social actors is different, and it is based on the findings that humans behave socially with computers as they would behave with other humans (e.g., Nass et al. 1997; Nass and Steuer 1993). This means that humans respond to similarly looking entities (other humans) as they respond to less alike entities (computer screens). As multimedia agents look more similar to humans than computer screens, humans' mindless response observed for computers (cf. Langer et al. 1978) can be expected for all multimedia agents. In other words, the appearance of the agent would not influence its effectiveness significantly, and an agent with a 2D appearance would be as effective as one with a 3D appearance.

Previous literature about pedagogical agents tends to be more supportive of the computers are social actors hypothesis. For example, the meta-analysis of multimedia pedagogical agents by Schroeder et al. (2013) revealed no significant differences in the effectiveness of the agents associated with their form or appearance. Similarly, the meta-analysis of gesturing pedagogical agents by Davis (2018) showed that the effect of appearance (humanoid vs. cartoonish) did not significantly moderate the learning measures of retention and transfer.

Moreover, Yılmaz and Kılıç-Çakmak (2012) provided evidence opposing the predictions of the redundancy effect of cognitive load theory and the coherence principle of cognitive theory of multimedia learning. In the study, 70 eighth grade students (54% females) learned about living organisms through pedagogical agents with different appearance. The retention test revealed that the human pedagogical agent was more effective than the simpler cartoon character.

In the present meta-analysis, we also explored the effect of the agents' gender, acknowledging positive findings with specific gender of the agent (e.g., Kim et al. 2007) or when the gender of agents and participants was considered (e.g., Dinçer and Doğanay 2017; Krämer et al. 2016; Makransky et al. 2019; Shiban et al. 2015).

Nonverbal Communication

We considered gesturing, eye gaze, and facial expression as three nonverbal communication cues that agents may use. Regarding gesturing and eye gaze, we compared agents without any of these capabilities versus agents that could either gesture or eye gaze. Concerning facial

expression, we compared three levels, namely, static facial expressions, simple dynamic expressions (e.g., only lip sync or smiling), and complex dynamic expressions (e.g., both lip sync and smiling).

As with appearance, for nonverbal communication, we tested the hypothesis based on the redundancy effect and the coherence principle, and the hypothesis of computers are social actors. We also tested a third hypothesis, based on social agency theory. A fourth hypothesis was aligned with the uncanny valley perspective.

The redundancy effect of cognitive load theory and the coherence principle of cognitive theory of multimedia learning predict that showing simultaneously too much distracting information—which could include gesturing, eye gaze, and facial expression that do not convey meaning—may be counterproductive (Baylor and Kim 2009; J. Moon and Ryu 2020; see also Stull et al. 2018b).

Based on the computers are social actors hypothesis, we predicted that pedagogical agents would be similarly effective, irrespective of executing or not gesturing and eye gaze, and irrespective of their levels of facial expression. Findings aligned with this prediction were reported by Craig et al. (2002), who observed that gesturing agents were as effective as nongesturing agents for 135 psychology undergraduates learning about lightning formation.

Regarding social agency theory (see Moreno et al. 2001; see also Mayer 2014c; Mayer et al. 2003), it describes that social cues made by instructors can trigger in learners a disposition of being in social exchange with the teachers, in which both parties expect a productive communication. This expectancy produces in students more willingness to learn. In other words, social agency theory predicts that instructors exhibiting human social cues will be more effective than instructors showing fewer of these signals. Hence, a multimedia pedagogical agent communicating with nonverbal social cues would trigger larger social effects beneficial for learning than an instructor not using these social signals (see Mayer et al. 2020; Sinatra et al. 2021).

For the present meta-analysis, testing the social agency theory we expected that multimedia pedagogical agents showing gesturing, eye gaze, and more facial expression would be more effective than agents not showing or showing less of these social cues (e.g., Li et al. 2019; Mayer and DaPra 2012; Wang et al. 2018). Under the social agency theory framework, Mayer and DaPra (2012) described this phenomenon as the embodiment effect, which predicts that pedagogical agents would be more effective if they show more dynamic social cues, such as gesturing, eye gaze, and facial expression (see also Wang et al. 2018).

Considering these three dynamic cues separately, gesturing has considerable evidence showing its positive effects on comprehension, memory, and learning (see Dargue et al. 2019; Hostetter 2011). For multimedia agents, the positive effects of gesturing (e.g., Cook et al. 2017) can be partially explained by their signaling function, as signaling devices, such as the limbs of pedagogical agents (Alpizar et al. 2020), are usually effective for multimedia learning (e.g., Li et al. 2019). Regarding eye gaze, studies with videos of human instructors and pedagogical agents have reported that showing their eye gaze is an effective method to promote learning (Beege et al. 2017b; Mayer et al. 2020; Stull et al. 2018a; Wang et al. 2018). Concerning the separate effects of facial expression, Schneider et al. (2018) reported two experiments in which texts were supplemented with pictures of robot agents either not showing or showing facial characteristics (e.g., smile and eyes). Results revealed that including the facial features for the images of the robots was beneficial for learning.

In the present meta-analysis, we also tested a fourth hypothesis, specifically for facial expression: the uncanny valley. This phenomenon was coined in robotics half a century ago

(see Mori et al. 1970/2012). The phenomenon is defined as a valley because it is the lowest degree of affinity that a human can experience with another agent. When the agent is in the middle–top zone between totally robotic (e.g., industrial robot) or full human, a sense of eeriness or uncanniness is supposed to be evoked. In contrast, the extremes of totally robotic or totally human produce higher affinity and lower eeriness. Although this perspective has not been employed much in educational research, we used it to test a hypothesis for facial expression. We predicted that an agent exhibiting simple dynamic facial expressions (e.g., unnatural lip sync) would evoke more eeriness and less learning outcomes than total static facial expressions or totally human and complex dynamic facial expressions. In this line, previous research with multimedia agents has shown these detrimental outcomes when the agents approach but fail to imitate the social naturality of humans (e.g., Ikeda et al. 2017; Tinwell et al. 2011; Veletsianos 2012).

Motion

In this meta-analysis, we compared static to animated agents, in order to investigate whether the unnatural versus natural or biological motion of the agent would moderate its effects on learning (see Schroeder et al. 2013; see also Cracco et al. 2018; Shimada and Oki 2012; Williams et al. 2019). We tested two hypotheses, one based on the action observation network and one based on the computers are social actors.

The action observation network describes natural or biological movement as more straightforward to process by humans than unnatural jerky kinematics (see Press 2011). For example, Shimada and Oki (2012) measured the sensorimotor brain activity of 14 participants (7% females) who watched videos of a computer character grasping and moving an object. Results showed that when participants watched videos that included two pauses in the action, brain activity was attenuated. In other words, the unnatural jerky motion deactivated sensorimotor processing. In two experiments with a total of 99 adult participants (61% females), Williams et al. (2019) observed that videos of human actors moving fluently (biologically) led to greater attentional engagement than videos of the actors moving rigidly (robot-like). Testing the action observation network hypothesis, we predicted that pedagogical agents showing natural or animated motion would be more effective than agents with no motion. This result was observed in the meta-analysis about pedagogical agents by Schroeder et al. (2013).

The computers are social actors hypothesis predicts that either static nonfluent, or fluent natural motion in agents will be equally effective. Although this prediction has not been supported by research about multimedia pedagogical agents, the meta-analysis by Cracco et al. (2018) showed that modeling a movement in either static images or a video did not influence the strength of automatic imitation. Additionally, the study by Shimada (2010) reported that the type of movement of the entities was not as influential as a mismatch between the movement and the performer (e.g., nonfluent movements by a human or fluent movements by a robot).

Voice

Regarding the agents' characteristic of voice, we compared the narration of the agents as either synthesized by a machine or recorded by a human (see Sinatra et al. 2021). We tested the predictions of the voice principle of social agency theory and of computers are social actors.

The voice principle predicts that spoken narrations are more effective for learning if delivered in a human voice rather than a machine synthesized voice (see Mayer 2014c). The

seminal study of the voice principle was conducted by Mayer et al. (2003) with 40 psychology undergraduates studying lightning formation. For both measures of retention and transfer, students in the human voice group outperformed students in the machine voice group.

Additional supporting evidence for the voice principle is provided in a study by Atkinson et al. (2005), who reported two experiments in which students were randomly assigned to learn math topics through animated pedagogical agents supplemented with either human voice or machine voice. Experiment 1 tested 50 undergraduates (18% females), and experiment 2 investigated 40 high school students (50% females). Results showed higher scores of retention, near transfer, and far transfer when the agent spoke in human voice rather than in machine voice (see also Mayer and DaPra 2012; Veletsianos 2012). Analogous findings of better transfer performance with gesturing agents speaking human rather than machine voices were revealed in the meta-analysis by Davis (2018).

The effects on the voice of the agent can also be predicted by the computers are social actors hypothesis. This perspective would predict no differences between machine and human voices. Recently, Chiou et al. (2020) investigated 98 adult participants (47% females) studying multimedia modules about the formation of lightning. Results showed that a pedagogical agent delivering the narrations in a high-quality machine voice was as effective as an agent narrating with a human voice (see also Craig and Schroeder 2017). Also, the meta-analysis of pedagogical agents by Schroeder et al. (2013) and the retention measures on the meta-analysis of gesturing agents by Davis (2018) showed no significant differences between machine and human voices.

Note that the supporting evidence for the voice principle was arguably produced because the machine synthesized voices were perceived as less natural than human voices. With current technology, a high-quality machine synthesized voice that can be heard as natural as a human voice may not produce these differences due to the source of the voice. Thus, current technology may balance the predictions from the voice principle and the hypothesis of computers are social actors.

Nonagents' Characteristics

Besides agents' characteristics, in the present meta-analysis, we also considered other variables that could impact the effectiveness of the agents. As such, we investigated the two broad learning domains of science, technology, engineering, and mathematics (STEM) versus non-STEM, and also the effects of several disciplines (e.g., the STEM disciplines of biology, physics, and computing; and the non-STEM disciplines of English, education, and law). Learning domain and discipline were analyzed, because previous meta-analyses (Alpizar et al. 2020; Castro-Alonso et al. 2019b; Schroeder et al. 2013; Sundararajan and Adesope 2020) have shown that they are critical moderating variables for multimedia learning.

We also investigated the effects of the developmental stage of the participants, as age could moderate the effects of pedagogical agents (e.g., Beege et al. 2017a; Davis 2018; Schroeder et al. 2013). Also, we investigated different proportions of females participating in the studies, acknowledging that students' gender can sometimes influence learning through multimedia modules (e.g., Castro-Alonso et al. 2019b; Wong et al. 2015).

To investigate the pacing of the multimedia, as in previous meta-analyses (Adesope and Nesbit 2012; Alpizar et al. 2020; Schroeder et al. 2013; Sundararajan and Adesope 2020), we compared system-paced versus learner-paced multimedia, based on the understanding that

system-paced multimedia cannot be controlled by the learner, so it generally includes more transient information that is difficult to understand (e.g., Castro-Alonso et al. 2018; Rey et al. 2019). We also analyzed the different languages spoken by the pedagogical agents, and the countries in which the studies were conducted, as in previous meta-analyses about multimedia learning (e.g., Alpizar et al. 2020; Sundararajan and Adesope 2020).

Lastly, as in recent meta-analyses about multimedia learning (e.g., Alpizar et al. 2020; Sundararajan and Adesope 2020), we investigated two methodological features. First, regarding type of control, we compared unmatched controls, in which not all the variables were tightly matched (see descriptions of these biases in Castro-Alonso et al. 2016), to matched controls, where the only difference between the treatment and the control groups was the inclusion or exclusion of the pedagogical agent. Second, concerning randomization, we compared quasi-experiments (randomly assigning classes or groups of students to conditions) and experiments (randomly assigning individual students to conditions).

Method

Throughout the present meta-analysis, we followed the PRISMA guidelines for conducting meta-analyses (e.g., Adesope et al. 2017; Lipsey and Wilson 2001; Moher et al. 2009). Considering the previous meta-analysis of pedagogical agents by Schroeder et al. (2013), which only included studies until 2011, our present study included studies published since 2012. In addition, our analysis was broader than that by Davis (2018), who only included gesturing pedagogical agents.

Selection Criteria

For the present meta-analysis, a study was deemed eligible for inclusion if it:

(1) was published between 2012 and 2019;
(2) was written in English;
(3) compared, in a between-subjects design, the learning effects of a condition including a multimedia pedagogical agent with a control condition without a pedagogical agent. We excluded studies in which the agents were coinciding physically with the students, rather than in exchanges through multimedia. We excluded studies with avatar or replica agents (agents controlled by a human);
(4) investigated school students, university students, or adults. We excluded studies with samples less representative of the whole population, such as infants, preschool students, elderly adults, clinical samples, and participants with learning disabilities;
(5) depicted either a STEM or a non-STEM task. We excluded studies in which the agent modeled a manipulative or a procedural task not included in formal STEM or non-STEM curricula;
(6) included only agents communicating through narrations or no speaking. We excluded studies in which the agents only provided on-screen text;
(7) reported measurable outcomes of instructional performance, such as retention and transfer tests; and
(8) included enough data to calculate effect sizes.

Literature Search and Selection of Studies

We used three combined queries as keywords to conduct a comprehensive and systematic search on electronic databases. The first query was *pedagogical agent* OR *animated agent* OR *multimedia agent* OR *virtual agent* OR *conversational agent*. The second query was (*human-like* OR *humanlike* OR *anthropomorphism* OR *anthropomorphic* OR *humanness*) AND (*humanoid* OR *cartoon-like* OR *computer* OR *machine*). The third query was *uncanny valley*. The selected databases were as follows: (a) Web of Science (Social Sciences Citation Index, SSCI; and Arts & Humanities Citation Index, A&HCI); (b) ProQuest (ERIC); (c) PsycARTICLES; (d) PsycINFO (APA); and (e) ProQuest (Dissertations and Theses). The search was conducted in August 2019, and it produced a total of 2979 articles. Following the removal of 618 duplicates, we reviewed 2361 studies.

The eight inclusion criteria were applied across three filtering phases to determine eligibility for further examination. The first filtering phase was conducted by one author of the present study, who screened the titles and abstracts of the studies to discard those blatantly contravening the criteria, including nonempirical works, studies not measuring learning outcomes, and interventions without pedagogical agents. This phase retained 217 studies. The second filtering phase was conducted by two authors of the present study, who screened in parallel all the abstracts of the remaining articles, to discard further studies not meeting the criteria. Disagreements between both authors were discussed until consensus was reached. Interrater agreement was high ($k = 0.94$). This phase retained 56 studies.

In the third filtering phase, we applied the selection criteria to the full-text copies. When data was not available, we asked the corresponding authors for the information. In this last phase, we retained 20 studies. We also searched the reference sections of two meta-analyses about pedagogical agents (Davis 2018; Guo and Goh 2015). These meta-analyses allowed us to incorporate one eligible study (Schroeder and Adesope 2013) meeting all inclusion criteria. Next, the selected articles were carefully read to extract relevant data for the meta-analysis. All authors agreed on the relevant information and the coding. In total, $n = 21$ articles and $k = 32$ independent comparisons were included in the meta-analysis. A summary of this selection of studies is provided in Fig. 1. Descriptive information of the 32 independent comparisons is shown in Table 1.

Extraction and Calculation of Effect Sizes

We extracted Cohen's d for each independent comparison. Cohen's d effect size is the standardized estimate of the difference in a comparison between participants who learned with pedagogical agents compared with participants who learned without these agents, divided by the pooled standard deviations of the two groups. However, as Cohen's d may be biased by unequal sample sizes across studies, Hedges' $g+$ was used to provide an unbiased estimate of effect sizes (Hedges and Olkin 1985). A positive $g+$ effect size indicates here an advantage of learning with pedagogical agents over no agents. Conversely, a negative $g+$ indicates that it is more advantageous to learn without a pedagogical agent than with an agent.

Data were analyzed using Comprehensive Meta-Analysis (CMA) version 3 (Borenstein et al. 2013) and IBM™ SPSS™ version 26 for Windows. As there were no outliers, all 32 independent effect sizes from the 21 studies were included for subsequent analyses.
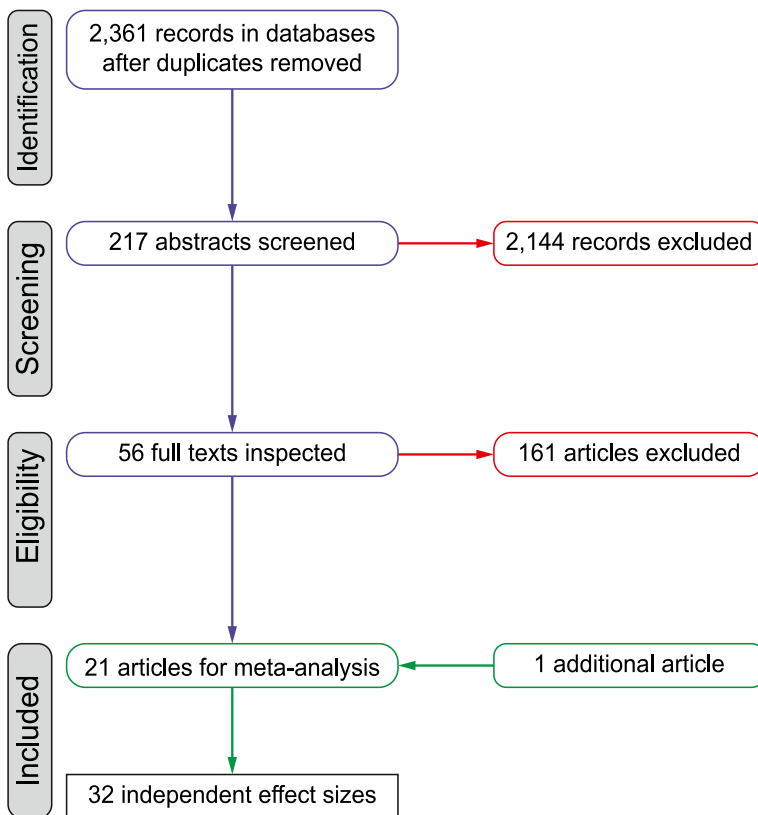
Fig. 1 Flow diagram of the selection of studies

Moderating Variables

Agents' Characteristics

The moderating variables regarding agent's characteristics are presented in Table 2. Concerning appearance, the agents were coded as *2D* or *3D agents*. The gender of the agent was also coded as *female*, *male*, or *other* (e.g., no gender, either). Regarding the agents' nonverbal communication: (a) gesturing was coded as *no* (absent) or *yes* (present); (b) eye gaze was coded as *no* or *yes*, as in gesturing; and (c) facial expression was coded as *static*, *simple dynamic* (e.g., only lip sync or smiling), and *complex dynamic* (e.g., both lip sync and smiling). We coded motion as either *static* or *animated*. Concerning voice, we coded the narration of the agents as either *no*, *machine*, or *human*.

Nonagents' Characteristics

Concerning the nonagent characteristic of domain, we coded *STEM* and *non-STEM* studies (see Table 3). These categories included several disciplines (e.g., the STEM disciplines of physics and computing, and the non-STEM disciplines of education and law). The developmental stage of the participants in the samples was categorized into the following categories based on the reported education level: *elementary school* (USA Grades k–5), *middle school*

Table 1 Descriptive information and effect sizes for the coded studies (separated by domain and school vs. postsecondary education)

Study	Agents' characteristics							Nonagents' characteristics							g^+	n
	Appe[a]	Gendr	Gestur	Gaze	FcExp	Motion	Voice	Discipline	%Fem	Pacing	Language	Cntry	Contrl	Rndm		
STEM, school																
Dinçer and Doğanay 2017	3D	Male	Yes	Yes	Cmplx	Animat	Human	Computing	Low	Learner	Other	Other	Match	Indivl	0.408	174
Johnson et al. 2013	3D	Fem	Yes	Yes	Cmplx	Animat	Human	Physics	Modt	Learner	English	USA	Match	Indivl	−0.002	200
Johnson et al. 2015 (1)	3D	Male	Yes	Yes	Cmplx	Animat	Human	Physics	Modt	Learner	English	USA	Match	Indivl	0.495	124
Johnson et al. 2015 (2)	3D	Male	Yes	Yes	Cmplx	Animat	Human	Physics	Modt	Learner	English	USA	Match	Indivl	−0.297	126
Osman and Lee 2014	2D	Other	Yes	Yes	Cmplx	Animat	Human	Chemistry	Modt	Learner	English	Other	Unmat	Group	0.403	127
van der Meij et al. 2015 (1)	3D	Fem	No	No	Cmplx	Animat	Machn	Physics	Modt	Learner	Other	Other	Match	Indivl	0.203	30
van der Meij et al. 2015 (2)	3D	Fem	No	No	Cmplx	Animat	Machn	Physics	Modt	Learner	Other	Other	Match	Indivl	−0.184	31
Yılmaz and Kılıç-Çakmak 2012 (1)	2D	Other	No	No	Static	Static	Machn	Biology	Modt	Learner	Other	Other	Unmat	Group	0.831	26
Yılmaz and Kılıç-Çakmak 2012 (2)	3D	Fem	No	No	Static	Static	Human	Biology	Modt	Learner	Other	Other	Match	Group	1.446	26
STEM, postsecondary																
Fountoukidou et al. 2019 (1)	3D	Male	No	No	Static	Animat	Human	Computing	Low	System	English	Other	Match	Indivl	−0.160	49
Fountoukidou et al. 2019 (2)	3D	Male	Yes	Yes	Cmplx	Animat	Human	Computing	Low	System	English	Other	Match	Indivl	0.208	50
Huang and Mayer 2016	2D	Fem	No	Yes	Simple	Animat	Human	Math	Modt	Learner	English	USA	Unmat	Indivl	0.241	54
Kim and Bennekin 2016	3D	Fem	Yes	No	Cmplx	Animat	Human	Math	High	Learner	English	USA	Match	Group	0.625	38
Lin et al. 2013 (1)	3D	Fem	No	Yes	Cmplx	Animat	Human	Physics	Modt	Learner	English	USA	Match	Indivl	−0.034	67
Lin et al. 2013 (2)	3D	Fem	No	Yes	Cmplx	Animat	Human	Physics	Modt	Learner	English	USA	Match	Indivl	0.046	68
Mayer and DaPra 2012 (1)	3D	Male	No	No	Simple	Static	Human	Chemistry	High	System	English	USA	Match	Indivl	−0.204	43
Mayer and DaPra 2012 (2)	3D	Male	Yes	Yes	Cmplx	Animat	Human	Chemistry	High	System	English	USA	Match	Indivl	0.138	45
Tegos and Demetriadis 2017	2D	Male	No	Yes	Simple	Static	Machn	Computing	Low	Learner	Other	Other	Match	Indivl	0.642	72
Wang et al. 2018 (1)	2D	Fem	Yes	Yes	Static	Animat	Human	Biology	High	System	Other	Other	Match	Indivl	0.809	51
Wang et al. 2018 (2)	2D	Fem	No	Yes	Static	Static	Human	Biology	High	System	Other	Other	Match	Indivl	0.114	56
Wang et al. 2018 (3)	2D	Fem	Yes	Yes	Static	Animat	Human	Biology	High	System	Other	Other	Match	Indivl	0.695	53
Non-STEM, school																
Chen and Chen 2014	2D	Other	No	No	Simple	Animat	No	Language	Modt	Learner	Other	Other	Match	Group	0.059	43
Nielen et al. 2018	2D	Male	Yes	No	Static	Animat	Human	Language	Modt	Learner	Other	Other	Match	Indivl	0.029	97
Non-STEM, postsecondary																
Ali et al. 2015	3D	Male	No	No	Cmplx	Animat	Human	English	N/R	Learner	English	Other	Match	Indivl	0.902	40
Carlotto and Jaques 2016 (1)	2D	Fem	No	No	Static	Animat	Human	English	N/R	Learner	English	Other	Match	Indivl	0.309	55

Table 1 (continued)

Study	Agents' characteristics							Nonagents' characteristics							g^+	n
	Appe[a]	Gendr	Gestur	Gaze	FcExp	Motion	Voice	Discipline	%Fem	Pacing	Language	Cntry	Contrl	Rndm		
Carlotto and Jaques 2016 (2)	2D	Fem	Yes	Yes	Cmplx	Animat	Human	English	N/R	Learner	English	Other	Match	Indivl	0.323	52
Park 2015 (1)	3D	Male	No	No	Simple	Animat	Human	Law	Low	Learner	English	USA	Match	Indivl	0.534	47
Park 2015 (2)	3D	Male	No	No	Simple	Animat	Human	Law	Low	Learner	English	USA	Match	Indivl	− 0.030	46
Romero-Hall et al. 2014 (1)	3D	Male	No	Yes	Simple	Animat	Human	History	High	Learner	English	USA	Unmat	Indivl	− 0.457	33
Romero-Hall et al. 2014 (2)	3D	Male	No	Yes	Cmplx	Animat	Human	History	High	Learner	English	USA	Unmat	Indivl	− 1.175	33
Schroeder 2017	3D	Male	No	No	Simple	Static	Human	Education	High	Learner	English	USA	Match	Indivl	− 0.006	69
Schroeder and Adesope 2013	3D	Male	Yes	No	Cmplx	Animat	Machn	Education	High	Learner	English	USA	Match	Indivl	− 0.076	79

The studies presenting numbers under parentheses indicate different groups compared per study

[a] Abbreviations in the column headings are (in column order): Appe = appearance, Gendr = gender, Gestur = gesturing, FcExp = facial expression, %Fem = percentage of female participants, Cntry = country, Contrl = control, Rndm = randomization. Abbreviations in the columns are (in alphabetical order): Animat = animated, Cmplx = complex dynamic, Fem = female, Indivl = individual, Machn = machine, Match = matched, Modt = moderate, N/R = not reported, Simple = simple dynamic, Unmat = unmatched

Table 2 Overall Effect and Weighted Mean Effect Sizes for Agents' Characteristics

Moderator	N	k	Effect size		95% CI		Test of heterogeneity		
			g+	SE	LCI	UCI	Q_B	df	p
Overall effect	2104	32	0.20**	0.07	0.07	0.34	69.76	31	<.001
Appearance									
2D	686	11	0.38**	0.11	0.16	0.60			
3D	1418	21	0.11	0.08	-0.06	0.27			
Total between levels							3.88	1	.049
Gender									
Female	781	13	0.30**	0.11	0.08	0.52			
Male	1127	16	0.10	0.10	-0.10	0.29			
Other	196	3	0.39	0.24	-0.07	0.85			
Total between levels							2.54	2	.28
Gesturing									
No	888	19	0.15	0.10	-0.04	0.34			
Yes	1216	13	0.26*	0.10	0.06	0.46			
Total between levels							0.62	1	.43
Eye gaze									
No	719	15	0.23*	0.11	0.01	0.44			
Yes	1385	17	0.19*	0.09	0.00	0.37			
Total between levels							0.09	1	.76
Facial expression									
Static	413	8	0.42**	0.15	0.14	0.71			
Simple dynamic	407	8	0.13	0.14	-0.16	0.41			
Complex dynamic	1284	16	0.14	0.10	-0.04	0.33			
Total between levels							2.91	2	.23
Motion									
Static	292	6	0.37*	0.17	0.04	0.71			
Animated	1812	26	0.17*	0.08	0.02	0.32			
Total between levels							1.15	1	.28
Voice									
No	43	1	0.06	0.42	-0.77	0.88			
Machine	238	5	0.26	0.19	-0.11	0.64			
Human	1823	26	0.20**	0.08	0.05	0.35			
Total between levels							0.22	2	.89

*p < .05; **p < .01

(USA grades 6–8), *high school* (USA grades 9–12), or *postsecondary* (usually participants enrolled in university, aged 18 and above).

Regarding percentage of females, we coded the proportion of females in the studies as *not reported*, *low*, *moderate*, and *high*. Specifically, if the percentage of female participants was less than 50% of the overall sample, we coded these studies as *low*. If the percentage of female participants was between 50 and 75% of the overall sample, we coded these studies as *moderate*, and finally, we coded studies as *high* if the percentage of female participants was above 75%.

Concerning the pacing of the multimedia, we coded the studies as *system-paced* or *learner-paced* multimedia. We also coded whether the language (spoken by the agent and shown as texts on the multimedia) was *English* or *Other* (e.g., Dutch, Chinese, Turkish). Regarding countries, the effects were reported in studies conducted in the USA (k = 15), the Netherlands (k = 5), China or Turkey (each k = 3), Brazil or Malaysia (each k = 2), and Greece or Taiwan (each k = 1). To obtain a more evenly distributed comparison between sample sizes, we coded the countries in which the studies were conducted as: *USA* or *Other*.

Table 3 Weighted mean effect sizes for nonagents' characteristics

Moderator	N	k	Effect size		95% CI		Test of heterogeneity		
			g+	SE	LCI	UCI	Q_B	df	p
Domain									
STEM	1510	21	0.27**	0.08	0.11	0.44			
Non-STEM	594	11	0.06	0.12	-0.17	0.30			
Total between levels							1.94	1	.16
Discipline									
Biology	212	5	0.67**	0.16	0.35	1.01			
Computing	345	4	0.33*	0.15	0.04	0.62			
Math	92	2	0.41	0.24	-0.07	0.88			
Chemistry	215	3	0.19	0.18	-0.16	0.54			
Physics	646	7	0.04	0.11	-0.18	0.25			
English	147	3	0.49*	0.20	0.09	0.88			
Law	93	2	0.25	0.25	-0.24	0.75			
Language	140	2	0.04	0.21	-0.38	0.46			
Education	148	2	-0.04	0.21	-0.45	0.36			
History	66	2	-0.80**	0.29	-1.37	-0.22			
Total between levels							27.50	9	.001
Participants' development									
Elementary school	43	1	0.06	0.43	-0.79	0.91			
Middle school	773	7	0.28	0.15	-0.01	0.57			
High school	188	3	0.20	0.25	-0.29	0.68			
Postsecondary	1100	21	0.18*	0.08	0.00	0.36			
Total between levels							0.44	3	.93
Proportion of females									
Not reported	147	3	0.49*	0.24	0.03	0.96			
Low	438	6	0.29	0.16	-0.02	0.61			
Moderate	1019	13	0.18	0.11	-0.03	0.39			
High	500	10	0.09	0.13	-0.17	0.34			
Total between levels							2.63	3	.45
Pacing of the multimedia									
System-paced	347	7	0.24	0.16	-0.06	0.55			
Learner-paced	1757	25	0.19*	0.08	0.04	0.35			
Total between levels							0.07	1	.79
Language									
English	1445	21	0.10	0.08	-0.06	0.26			
Other	659	11	0.41**	0.12	0.18	0.64			
Total between levels							4.83	1	.03
Country									
USA	1072	15	0.02	0.09	-0.16	0.19			
Other	1032	17	0.38**	0.09	0.21	0.55			
Total between levels							8.40	1	.004
Type of control									
Unmatched	273	5	0.04	0.19	-0.33	0.41			
Matched	1831	27	0.23**	0.08	0.08	0.38			
Total between levels							0.85	1	.36
Randomization									
Groups	260	5	0.56**	0.18	0.20	0.93			
Individuals	1844	27	0.15*	0.07	0.01	0.29			
Total between levels							4.42	1	.04

$*p < .05$; $**p < .01$

For the two methodological features investigated, we coded the control groups as either *unmatched* (not all the variables were tightly matched) or *matched* (the variables were

matched). Concerning randomization, we coded as *groups* the quasi-experimental studies that randomized classes or groups of students, and as *individuals* the experiments in which each student was randomly assigned to a condition.

Results

Both tables of effects (see Tables 2 and 3) present the same structure. They include the following data: the number of participants (N) in each category; the number of effect sizes (k); the weighted mean effect size ($g+$) and its standard error (SE); the 95% confidence intervals (95% CI), including the lower (LCI) and upper (UCI) limits; and the results of a test of heterogeneity (Q_B) with its associated degrees of freedom (df) and probability (p).

Overall Effects

As shown in the first row of Table 2, the analysis of the 32 independent effect sizes from 21 studies showed an overall effect of $g+ = 0.20$ ($p = .003$; $SE = 0.07$; LCI = 0.07, UCI = 0.34) across a diverse sample of participants ($N = 2104$). In other words, we observed that between-subject comparisons showed a statistically significant positive effect of learning with multimedia pedagogical agents compared to learning without these agents. The effect can be regarded as small (Cohen 1988). The distribution of the 32 effect sizes is shown in Fig. 2, and their forest plot is provided in Fig. 3.

Further analysis revealed that significant heterogeneity existed, $Q_B (31) = 69.76$, $p < .001$, and there was moderate variability within the sample, $I^2 = 55.56\%$. This implies that the total variability that can be attributed to true heterogeneity or between-studies variability was approximately 56%. Due to this heterogeneous distribution, moderator analyses were conducted.

Moderator Analyses for Agents' Characteristics

Agents' Appearance and Gender

Concerning the appearance of the agent, our analysis revealed a significant difference between both groups, $Q_B (1) = 3.88$, $p = .049$ (see Table 2). We observed a small to medium effect size ($g+ = 0.38$) favoring 2D agents over nonagents. However, there were no significant effects of 3D agents versus no agents. This higher outcome associated with 2D than 3D agents tends to support the hypothesis based on the redundancy effect of cognitive load theory and the coherence principle of cognitive theory of multimedia learning. The computers are social actors hypothesis is not as supported by these findings.

We also examined the effects of the agents' gender on learning performance. As observed in Table 2, results showed no significant differences between the groups, $Q_B (2) = 2.54$, $p = .28$. Although female agents were associated with a small–medium positive effect size ($g+ = 0.30$), this was not statistically different from male agents or other categories of gender (e.g., either or no gender).

Fig. 2 Distribution of 32
independent effect sizes obtained
from 21 articles (M = 0.21, SD =
0.48)
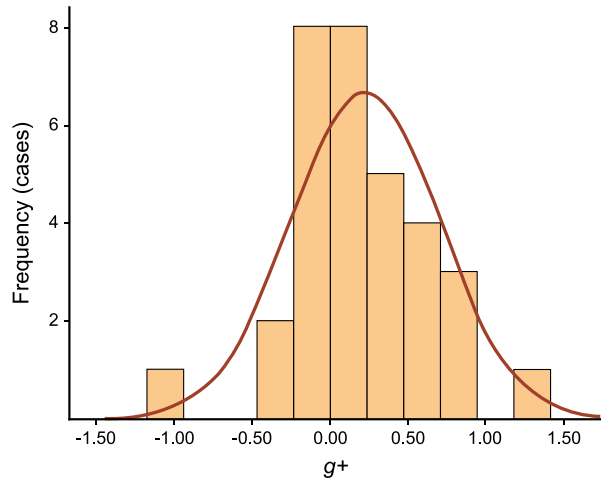
Agents' Nonverbal Communication

We included three separate categories for nonverbal communication: gesturing, eye gaze, and
facial expression. Within gestures, as the test of heterogeneity (Q_B) in Table 2 shows, there
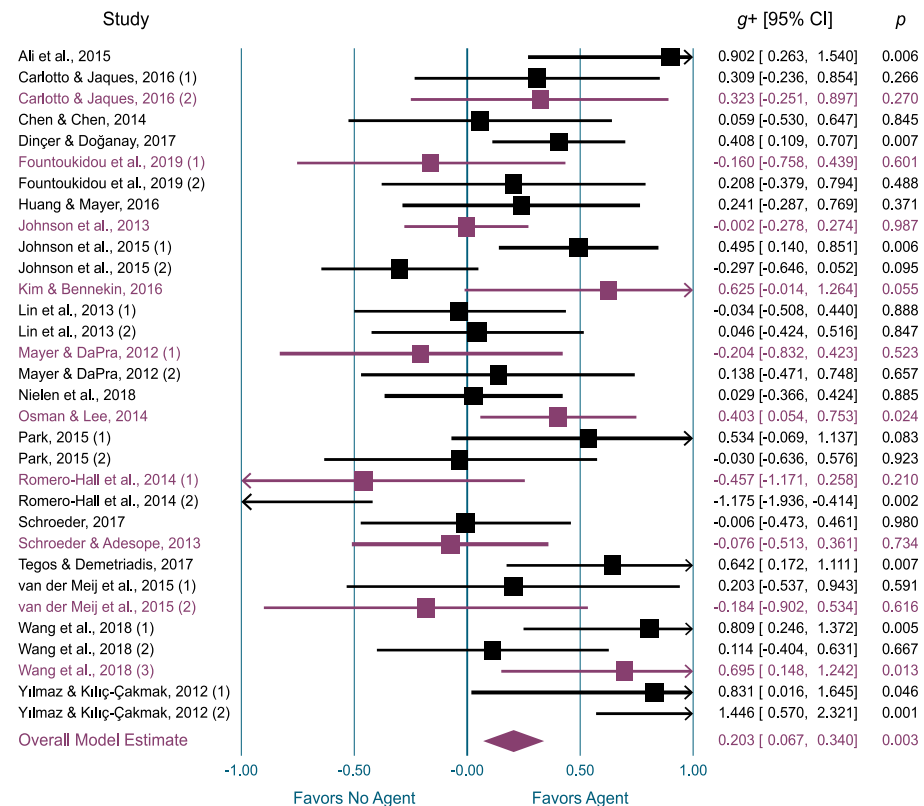
Fig. 3 Forest plot of the 32 effect sizes from the 21 articles. Color coding was used to help with reading the plot

were no significant differences between no gesturing and gesturing conditions ($p = .43$). Although the use of gesturing was associated with a small significant positive effect size ($g+ = 0.26$), it was not significantly different from the no gesturing group.

Regarding eye gaze, we coded whether agents were incapable or capable of gazing. Both groups showed that the agent condition outperformed the no agent condition with small effect sizes. As shown in Table 2, there were no significant differences between both groups ($p = .76$). In other words, both no gaze ($g+ = 0.23$) and gaze agents ($g+ = 0.19$) were associated with higher learning than the conditions without agents.

Within facial expression, we coded for static expression, simple dynamic expressions (e.g., only lip sync), and complex dynamic expressions (e.g., lip sync and smiling). As shown in Table 2 (see Q_B), there were no significant differences between the three groups ($p = .23$). It was observed that static facial expressions in agents were associated with a moderately significant effect size ($g+ = 0.42$), but this effect was not significantly different from the other two conditions of facial expressions.

In all, the analyses of the agents' nonverbal communication showed the same results for the three variables. For all the variables—gesturing, eye gaze, and facial expression— there were nonsignificant differences between the groups. These null differences tend to support the computers are social actors hypothesis to a larger extent than the other three hypotheses: (a) the redundancy effect and the coherence principle, (b) social agency theory, and (c) uncanny valley.

Agents' Motion and Voice

Concerning the motion of the agents, we investigated the benefits of learning with static or animated agents. Results showed positive and statistically significant effects of learning with both static ($g+ = 0.37$) and animated ($g+ = 0.17$) agents, being a small–medium size for static and small for animated. Findings also indicated that there were no significant differences between both types, $p = .28$ (see Table 2). This indicates that pedagogical agents are effective for learning regardless of whether they are static or animated. This null difference supports the computers are social actors hypothesis more than the predictions of the action observation network.

About the voice of the agents, we compared studies with agents without voice, machine-synthetized narrations, and human-recorded voices. As revealed by the analysis (see bottom of Table 2), there were no significant differences between the three groups ($p = .89$). Although human voice was associated with a small significant effect size ($g+ = 0.20$) favoring agents with human voice versus no agent conditions, the effect was not significantly different from the machine voice and the no voice groups. These findings provide support for the computers are social actors hypothesis.

Moderator Analyses for Nonagents' Characteristics

Domain and Discipline

We examined the effect of learning with pedagogical agents depending on the domain of the educational intervention. The learning materials used in the studies were categorized as either STEM or non-STEM relevant. As shown in Table 3, although studies that utilized STEM learning materials produced a statistically significant weighted mean effect ($g+ = 0.27$), our analysis revealed no significant difference between the two groups ($p = .16$).

In contrast, the results for the type of discipline indicated significant differences between the disciplines, $p = .001$ (see Table 3). For the STEM disciplines, we observed that biology ($g+ = 0.67$) and computing ($g+ = 0.33$) were associated with positive effect sizes, being medium–large for biology and small–medium for computing. For non-STEM tasks, the only significant positive effect was observed in a medium effect size for English ($g+ = 0.49$). Unexpectedly, studies in history were associated with a *negative* and significantly large effect size ($g+ = -0.80$).

Participants' Development

We also investigated the benefits of learning with pedagogical agents depending on participants' developmental stage, coded by their educational level. Analyzing studies with elementary school, middle school, high school, and postsecondary students (see Table 3), we observed that only studies conducted with postsecondary students were associated with a small positive effect size ($g+ = 0.18$). However, there were no significant differences between the four groups ($p = .93$).

Proportion of Females

We examined the effects of learning with pedagogical agents across samples with not reported, low, moderate, and high proportion of female participants. The moderator analysis revealed no significant differences between groups ($p = .45$). In other words, the similar confidence intervals between the low, moderate, and high categories indicated that pedagogical agents could be used effectively in educational settings with varying proportions of male and female participants.

Pacing of the Multimedia

The pacing of the multimedia presentation was coded as either system-paced or learner-paced. Although learner-paced studies were associated with a small significant positive effect size ($g+ = 0.19$), our analysis revealed no significant differences between both groups ($p = .79$, see Table 3).

Language and Country

Concerning the language of the multimedia and the agent (see Table 3), we compared English versus Other. It was observed that the effects were significantly different between both groups ($p = .03$). The English language showed no significant effects, but Other (e.g., Dutch, Chinese, Turkish) showed a small–medium significant effect size ($g+ = 0.41$) favoring pedagogical agents on multimedia modules.

Regarding country, we also grouped the studies into two groups, namely, USA and Other. There was a significant difference between the groups ($p = .004$). Studies coded as Other (e.g., the Netherlands, China, Turkey, Brazil) showed a small to medium significant effect size ($g+ = 0.38$) for using agents. However, studies where participants were drawn from the USA were associated with a nonsignificant effect size.

Methodological Features

We were also interested in the methodology to design the control group without the pedagogical agent that was compared to the treatment condition with the agent. We compared unmatched and matched controls. Even though the studies using matched controls were associated with a small significant positive effect size ($g+ = 0.23$), as shown in Table 3, our analysis showed no significant differences between both groups ($p = .36$).

We also considered the type of randomization as another methodological variable, analyzing the level of randomization in either groups or individuals. As shown in the bottom of Table 3, the comparison showed significant differences, $p = .04$. Randomization at the level of groups showed a medium effect size ($g+ = 0.56$) favoring the inclusion of agents in these quasi-experiments. In contrast, for the experimental studies, which randomized at the level of individuals, the effect was small ($g+ = 0.15$).

Publication Bias

A primary concern in all meta-analyses, including ours, is the potential for publication bias, as studies with statistically significant results are more likely to be published than studies with no significant findings (e.g., Franco et al. 2014; Rosenthal 1979). To examine the influence of publication bias on our results and to ascertain the validity of the present meta-analysis, four publication bias tests were conducted using the CMA software. First, we examined the funnel plot distribution (see Fig. 4). The fairly symmetrical distribution indicates that publication bias is unlikely in our sample. Second, to corroborate our findings, we also examined the results of Egger's linear regression test (Egger et al. 1997). The results showed that publication bias would not influence the interpretation of our results based on the present sample of studies, intercept = 0.41, $t(30) = 0.45$, $p = .66$.

Third, we also conducted the classic fail-safe N test to identify the additional studies with a null effect necessary to raise the p value greater than .05. The analysis revealed that 129 studies would be needed. The result of this test is greater than the $5k + 10$ limit recommended by Rosenthal (1979). Fourth, we examined Duval and Tweedie's trim-and-fill approach (Duval and Tweedie 2000) and results indicate that zero studies were needed to be trimmed or adjusted, indicating an absence of publication bias. Overall, consistent results from the four tests suggest that the validity of our findings is not threatened by publication bias.

Fig. 4 Funnel plot distribution of the 32 effect sizes
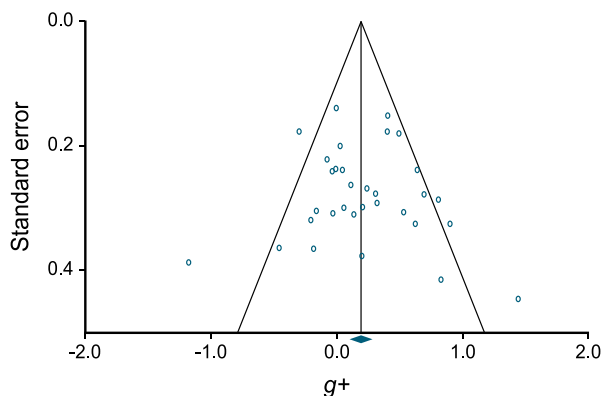
Discussion

We conducted a contemporary meta-analysis to investigate how different characteristics in multimedia pedagogical agents could influence their instructional effectiveness, as predicted by diverse theories. We also investigated other moderating variables that were not related to the multimedia agents. Overall, the meta-analysis of 32 independent effect sizes from 21 articles ($N = 2104$) revealed a small effect ($g+ = 0.20$), indicating a positive result of learning with multimedia pedagogical agents compared to learning without these agents. This finding, which seems not to be threatened by publication bias, mirrors similar positive sizes in previous meta-analyses (Davis 2018; Schroeder et al. 2013). However, the size of this overall effect was moderated by several variables, discussed next.

Agents' Characteristics

Investigating the moderating role of the agents' appearance, we observed that 2D agents ($g+ = 0.38$) tended to be more effective than 3D agents ($g+ = 0.11$). Note that this result is from comparisons across different studies. As in all moderator analyses, this outcome does not necessarily suggest that we included single studies comparing 2D and 3D agents (see also "Limitations and Future Directions").

A higher effectiveness of 2D agents could be explained by the redundancy effect of cognitive load theory (Castro-Alonso et al. 2019a; Kalyuga and Sweller 2014) and the coherence principle of cognitive theory of multimedia learning (Mayer and Fiorella 2014; Mayer et al. 2008). Both theories have supported educational material that do not overload working memory with unnecessary detail or nonessential information for learning. Following this rationale, an agent with a 2D, cartoonish, or simpler appearance would be more effective than a more detailed and visually complex 3D agent.

This result somewhat contrasts with previous meta-analyses of multimedia pedagogical agents (Davis 2018; Schroeder et al. 2013) showing no significant effects attributed to the appearance of the agent. This null difference, which is more supportive of the computers are social actors hypothesis (Y. Moon and Nass 1996; Nass et al. 1997; Nass and Steuer 1993), was observed in our meta-analysis regarding gender, as we observed that the gender of the agents did not moderate their effects. This is also contrasting with previous findings (e.g., Y. Kim et al. 2007) that have shown an influence of the specific gender of the pedagogical agent.

Concerning the agents' nonverbal communication, the three variables investigated (gesturing, eye gaze, and facial expression) showed that their presence or absence (e.g., doing gesturing or eye gaze versus not showing these cues) or their degree (facial expression: static vs. simple dynamic vs. complex dynamic) did not influence the effectiveness of the agents. These nonsignificant moderator effects can be explained by the computers are social actors hypothesis, as agents displaying or not social nonverbal communication were equally effective. This aligns with the study by Craig et al. (2002), in which gesturing pedagogical agents were as instructionally effective as nongesturing agents.

Also, the agents' motion did not show significant differences between their levels. As such, the motion displayed by static and animated agents did not moderate their effects, supporting previous findings outside multimedia pedagogical agents literature (Cracco et al. 2018; Shimada 2010), but not replicating the previous meta-analysis by Schroeder et al. (2013) that showed more effectiveness in animated agents.

Concerning the agents' voice, machine and human voices were equally effective, echoing recent finding with high-quality machine voices (Chiou et al. 2020; Craig and Schroeder 2017) and a previous meta-analysis with machine voices of less quality (Schroeder et al. 2013). These voice results must be interpreted cautiously, as most of the agents used human voices ($k = 26$) rather than machine voice ($k = 5$) or no voice ($k = 1$). However, both motion and voice findings tend to support the computers are social actors hypothesis. In summary, most of the agents' characteristics, including nonverbal communication, motion, and voice, supported the predictions of the computers are social actors hypothesis.

Nonagents' Characteristics

When assessing nonagents' characteristics, we observed that the domain of the multimedia topic (STEM versus non-STEM) did not moderate the results. In contrast, the meta-analysis of pedagogical agents by Schroeder et al. (2013) showed that the agents were more effecting teaching science and math, compared to humanities.

Our meta-analysis showed that the disciplines moderated the results. Specifically, it was more effective to learn with multimedia agents in the STEM disciplines of biology ($g+ = 0.67$) and computing ($g+ = 0.33$), and the non-STEM discipline of English ($g+ = 0.49$). Unpredictably, the non-STEM discipline of history ($g+ = -0.80$) revealed that it was *less* effective to learn with these agents. However, caution needs to be exercised in interpreting these findings of domain and discipline, considering the wide confidence intervals across most of the disciplines.

Three nonagents' characteristics did not show moderating effects, namely, participants' development, percentage of females, and pacing of the multimedia. In other words, the agents were similarly effective with (a) elementary school, middle school, high school, or postsecondary students; (b) the different proportion of female participants; and (c) either system-paced or learner-paced multimedia modules. These findings suggest that the pedagogical agents are effective in different conditions, independently of other multimedia effects reported for age (Beege et al. 2017a; Davis 2018; Schroeder et al. 2013), female percentage (Castro-Alonso et al. 2019b), or pacing (Rey et al. 2019).

Concerning the language of the pedagogical agent and the multimedia lesson, it was observed that agents in English materials ($g+ = 0.10$) were less effective than agents in multimedia modules in other languages ($g+ = 0.41$), such as Dutch, Chinese, and Turkish, echoing a recent meta-analysis that also showed language effects on multimedia learning (Sundararajan and Adesope 2020). Regarding the country of the published article, studies conducted in the USA ($g+ = 0.02$) showed a smaller influence of the pedagogical agent than studies conducted in other countries ($g+ = 0.38$), such as the Netherlands, China, Turkey, and Brazil. Future research could determine the reasons for these differences regarding language and country.

Concerning methodological features of the included articles, the type of control condition, either unmatched or matched, did not moderate the effects of the multimedia agent. In contrast, there was an effect in the type of randomization, as quasi-experiments that randomly assigned conditions at the levels of groups ($g+ = 0.56$) tended to show larger effects than experiments randomizing per individual ($g+ = 0.15$). Explanations of these effects also warrant further investigation.

Implications

A first implication of these findings is to endorse the inclusion of pedagogical agents in instructional multimedia materials. Specifically, the most recommended multimedia

pedagogical agents would be those with a simpler 2D appearance. In contrast, agents whose appearance is of 3D models or more visually complex would be less recommended. Other characteristics of the agents, including gender, nonverbal communication (gesturing, eye gaze, and facial expression), motion, and voice, seem to be less relevant than appearance, as predicted by the computers are social actors hypothesis (Moon and Nass 1996; Nass and Steuer 1993).

A second implication of these results is theoretical and merits further research. We observed that the degree of several social cues in the pedagogical agents tended not to influence their instructional effectiveness. This might suggest that participants from recent studies (2012–2019), who are accustomed to the multimedia environments with pedagogical agents, learn equally well with multimedia agents that display different degrees of social cues. This could imply a revision or extension to the theories used here to predict the effectiveness of these cues.

Limitations and Future Directions

A first limitation of this meta-analysis was that some of the coded variables produced very small number of studies within groups (e.g., *other* gender $k = 3$, *no* voice $k = 1$, *elementary school* $k = 1$) whose results could not be confidently used to make inferences for the population. A future direction of research is to conduct an analysis considering only these variables. For example, the gender characteristic could be investigated, as there is evidence of the effects of gender on multimedia learning (e.g., Castro-Alonso et al. 2019b; Heo and Toomey 2020; Wong et al. 2018) and visuospatial processing (e.g., Castro-Alonso and Jansen 2019; Lauer et al. 2019).

A second limitation we have noticed concerns the separate analyses of the moderating variables, without considering that the effects of different agents' characteristics could interact. Future research could investigate if agents showing simultaneously different social capabilities (e.g., gesturing, eye gaze, and animated motion) produce better (Mayer and DaPra 2012; Wang et al. 2018) or worse (Baylor and Kim 2009; Moon and Ryu 2020) learning outcomes than those showing fewer social cues. Also, future factorial studies could investigate the interactions between different social cues in the agents (e.g., appearance and motion, cf. Shimada 2010).

Thirdly, although it is not a limitation of the current study, it is a feature of moderator analyses of meta-analyses. It must be noted that these moderator analyses are conducted overall, so they include comparisons across different studies. For example, if a moderator analysis shows a significant difference between 2D and 3D agents, this does not necessarily mean that single studies showed this difference by comparing participants studying with either 2D or 3D agents. Future investigations could compare in a single study some of the moderating variables presented here (e.g., 2D vs. 3D agents, static vs. complex facial expression, STEM vs. non-STEM domain).

Conclusion

As humans are progressively disconnecting from nature (e.g., Kesebir and Kesebir 2017) and from learning via direct person to person communication (e.g., Burgoon et al. 2000; Sbarra et al. 2019), contemporary human learners may be equally prepared to learn from either face-to-face or mediated instructors (e.g., multimedia pedagogical agents). Analogously, learners may be

equipped to learn similarly from any type of multimedia pedagogical agent, as predicted by the computers are social actors hypothesis. We found support for this prediction, because our meta-analysis revealed that most of the agents' social behavior (e.g., gesturing, facial expression, motion, and voice) tended to be irrelevant for their instructional effectiveness.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

References marked with an asterisk indicate studies included in the meta-analysis

Adesope, O. O., & Nesbit, J. C. (2012). Verbal redundancy in multimedia learning environments: A meta-analysis. *Journal of Educational Psychology, 104*(1), 250–263. https://doi.org/10.1037/a0026147.

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research, 87*(3), 659–701. https://doi.org/10.3102/0034654316689306.

*Ali, A. Z. M., Segaran, K., & Hoe, T. W. (2015). Effects of verbal components in 3D talking-head on pronunciation learning among non-native speakers. *Educational Technology & Society, 18*(2), 313–322.

Alpizar, D., Adesope, O. O., & Wong, R. M. (2020). A meta-analysis of signaling principle in multimedia learning environments. *Educational Technology Research and Development, 68*(5), 2095–2119. https://doi.org/10.1007/s11423-020-09748-7.

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), 117–139. https://doi.org/10.1016/j.cedpsych.2004.07.001.

Baylor, A. L., & Kim, S. (2009). Designing nonverbal communication for pedagogical agents: When less is more. *Computers in Human Behavior, 25*(2), 450–457. https://doi.org/10.1016/j.chb.2008.10.008.

Beege, M., Schneider, S., Nebel, S., Mittangk, J., & Rey, G. D. (2017a). Ageism – Age coherence within learning material fosters learning. *Computers in Human Behavior, 75*, 510–519. https://doi.org/10.1016/j.chb.2017.05.042.

Beege, M., Schneider, S., Nebel, S., & Rey, G. D. (2017b). Look into my eyes! Exploring the effect of addressing in educational videos. *Learning and Instruction, 49*, 113–120. https://doi.org/10.1016/j.learninstruc.2017.01.004.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2013). *Comprehensive meta-analysis (version 3.0) [computer software]*. Englewood: Biostat.

Brucker, B., Scheiter, K., & Gerjets, P. (2014). Learning with dynamic and static visualizations: Realistic details only benefit learners with high visuospatial abilities. *Computers in Human Behavior, 36*, 330–339. https://doi.org/10.1016/j.chb.2014.03.077.

Burgoon, J. K., Bonito, J. A., Bengtsson, B., Cederberg, C., Lundeberg, M., & Allspach, L. (2000). Interactivity in human–computer interaction: A study of credibility, understanding, and influence. *Computers in Human Behavior, 16*(6), 553–574. https://doi.org/10.1016/S0747-5632(00)00029-7.

*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*, 15–26. https://doi.org/10.1016/j.ijhcs.2016.06.001.

Castro-Alonso, J. C., & Jansen, P. (2019). Sex differences in visuospatial processing. In J. C. Castro-Alonso (Ed.), *Visuospatial processing for education in health and natural sciences* (pp. 81–110). Cham: Springer. https://doi.org/10.1007/978-3-030-20969-8_4.

Castro-Alonso, J. C., Ayres, P., & Paas, F. (2016). Comparing apples and oranges? A critical look at research on learning from statics versus animations. *Computers & Education, 102*, 234–243. https://doi.org/10.1016/j.compedu.2016.09.004.

Castro-Alonso, J. C., Ayres, P., Wong, M., & Paas, F. (2018). Learning symbols from permanent and transient visual presentations: Don't overplay the hand. *Computers & Education, 116*, 1–13. https://doi.org/10.1016/j.compedu.2017.08.011.

Castro-Alonso, J. C., Ayres, P., & Sweller, J. (2019a). Instructional visualizations, cognitive load theory, and visuospatial processing. In J. C. Castro-Alonso (Ed.), *Visuospatial processing for education in health and natural sciences* (pp. 111–143). Cham: Springer. https://doi.org/10.1007/978-3-030-20969-8_5.

Castro-Alonso, J. C., Wong, M., Adesope, O. O., Ayres, P., & Paas, F. (2019b). Gender imbalance in instructional dynamic versus static visualizations: A meta-analysis. *Educational Psychology Review, 31*(2), 361–387. https://doi.org/10.1007/s10648-019-09469-1.

*Chen, Z.-H., & Chen, S. Y. (2014). When educational agents meet surrogate competition: Impacts of competitive educational agents on students' motivation and performance. *Computers & Education, 75*, 274–281. https://doi.org/10.1016/j.compedu.2014.02.014.

Chiou, E. K., Schroeder, N. L., & Craig, S. D. (2020). How we trust, perceive, and learn from virtual humans: The influence of voice quality. *Computers & Education, 146*, 103756. https://doi.org/10.1016/j.compedu.2019.103756.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Erlbaum.

Cook, S. W., Friedman, H. S., Duggan, K. A., Cui, J., & Popescu, V. (2017). Hand gesture and mathematics learning: Lessons from an avatar. *Cognitive Science, 41*(2), 518–535. https://doi.org/10.1111/cogs.12344.

Cracco, E., Bardi, L., Desmet, C., Genschow, O., Rigoni, D., De Coster, L., et al. (2018). Automatic imitation: A meta-analysis. *Psychological Bulletin, 144*(5), 453–500. https://doi.org/10.1037/bul0000143.

Craig, S. D., & Schroeder, N. L. (2017). Reconsidering the voice effect when learning from a virtual human. *Computers & Education, 114*, 193–205. https://doi.org/10.1016/j.compedu.2017.07.003.

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology, 94*(2), 428–434. https://doi.org/10.1037//0022-0663.94.2.428.

Dargue, N., Sweller, N., & Jones, M. P. (2019). When our hands help us understand: A meta-analysis into the effects of gesture on comprehension. *Psychological Bulletin, 145*(8), 765–784. https://doi.org/10.1037/bul0000202.

Davis, R. O. (2018). The impact of pedagogical agent gesturing in multimedia learning environments: A meta-analysis. *Educational Research Review, 24*, 193–209. https://doi.org/10.1016/j.edurev.2018.05.002.

*Dinçer, S., & Doğanay, A. (2017). The effects of multiple-pedagogical agents on learners' academic success, motivation, and cognitive load. *Computers & Education, 111*, 74–100. https://doi.org/10.1016/j.compedu.2017.04.005.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629.

*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*, 84–93. https://doi.org/10.1016/j.chb.2019.03.013.

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science, 345*(6203), 1502–1505. https://doi.org/10.1126/science.1255484.

Gehlbach, H. (2010). The social side of school: Why teachers need social psychology. *Educational Psychology Review, 22*(3), 349–362. https://doi.org/10.1007/s10648-010-9138-3.

Guo, Y. R., & Goh, D. H.-L. (2015). Affect in embodied pedagogical agents: Meta-analytic review. *Journal of Educational Computing Research, 53*(1), 124–149. https://doi.org/10.1177/0735633115588774.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review, 6*(1), 27–54. https://doi.org/10.1016/j.edurev.2010.07.004.

Heo, M., & Toomey, N. (2020). Learning with multimedia: The effects of gender, type of multimedia learning resources, and spatial ability. *Computers & Education, 146*, 103747. https://doi.org/10.1016/j.compedu.2019.103747.

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin, 137*(2), 297–315. https://doi.org/10.1037/a0022128.

*Huang, X., & Mayer, R. E. (2016). Benefits of adding anxiety-reducing features to a computer-based multimedia lesson on statistics. *Computers in Human Behavior, 63*, 293–303. https://doi.org/10.1016/j.chb.2016.05.034.

Ikeda, T., Hirata, M., Kasaki, M., Alimardani, M., Matsushita, K., Yamamoto, T., Nishio, S., & Ishiguro, H. (2017). Subthalamic nucleus detects unnatural android movement. *Scientific Reports, 7*(1), 17851. https://doi.org/10.1038/s41598-017-17849-2.

*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), 319–337. https://doi.org/10.1002/jee.20009.

*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), 97–115. https://doi.org/10.1111/jcal.12078.

Kalyuga, S., & Sweller, J. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 247–262). New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.013.

Kesebir, S., & Kesebir, P. (2017). A growing disconnection from nature is evident in cultural products. *Perspectives on Psychological Science, 12*(2), 258–269. https://doi.org/10.1177/1745691616662473.

*Kim, C., & Bennekin, K. N. (2016). The effectiveness of volition support (VoS) in promoting students' effort regulation and performance in an online mathematics course. *Instructional Science, 44*(4), 359–377. https://doi.org/10.1007/s11251-015-9366-5.

Kim, Y., Baylor, A. L., & Shen, E. (2007). Pedagogical agents as learning companions: The impact of agent emotion and gender. *Journal of Computer Assisted Learning, 23*(3), 220–234. https://doi.org/10.1111/j.1365-2729.2006.00210.x.

Krämer, N. C., Karacora, B., Lucas, G., Dehghani, M., Rüther, G., & Gratch, J. (2016). Closing the gender gap in STEM with friendly male instructors? On the effects of rapport behavior and gender of a virtual agent in an instructional interaction. *Computers & Education, 99*, 1–13. https://doi.org/10.1016/j.compedu.2016.04.002.

Langer, E. J., Blank, A., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action: The role of "placebic" information in interpersonal interaction. *Journal of Personality and Social Psychology, 36*(6), 635–642. https://doi.org/10.1037/0022-3514.36.6.635.

Lauer, J. E., Yhang, E., & Lourenco, S. F. (2019). The development of gender differences in spatial reasoning: A meta-analytic review. *Psychological Bulletin, 145*(6), 537–565. https://doi.org/10.1037/bul0000191.

Lawson, A. P., Mayer, R. E., Adamo-Villani, N., Benes, B., Lei, X., & Cheng, J. (2021). Recognizing the emotional state of human and virtual instructors. *Computers in Human Behavior, 114*, 106554. https://doi.org/10.1016/j.chb.2020.106554.

Li, W., Wang, F., Mayer, R. E., & Liu, H. (2019). Getting the point: Which kinds of gestures by pedagogical agents improve multimedia learning? *Journal of Educational Psychology, 111*(8), 1382–1395. https://doi.org/10.1037/edu0000352.

*Lin, L., Atkinson, R. K., Christopherson, R. M., Joseph, S. S., & Harrison, C. J. (2013). Animated agents and learning: Does the type of verbal feedback they provide matter? *Computers & Education, 67*, 239–249. https://doi.org/10.1016/j.compedu.2013.04.017.

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. https://doi.org/10.1016/j.compedu.2019.103658.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks: SAGE.

Makransky, G., Wismer, P., & Mayer, R. E. (2019). A gender matching effect in learning with pedagogical agents in an immersive virtual reality science simulation. *Journal of Computer Assisted Learning, 35*(3), 349–358. https://doi.org/10.1111/jcal.12335.

Mayer, R. E. (2014a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 43–71). New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005.

Mayer, R. E. (2014b). Introduction to multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 1–24). New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.002.

Mayer, R. E. (2014c). Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 345–368). New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.017.

*Mayer, R. E., & DaPra, C. S. (2012). An embodiment effect in computer-based learning with animated pedagogical agents. *Journal of Experimental Psychology: Applied, 18*(3), 239–252. https://doi.org/10.1037/a0028616.

Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 279–315). New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.015.

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), 419–425. https://doi.org/10.1037/0022-0663.95.2.419.

Mayer, R. E., Griffith, E., Jurkowitz, I. T. N., & Rothman, D. (2008). Increased interestingness of extraneous details in a multimedia science presentation leads to decreased learning. *Journal of Experimental Psychology: Applied, 14*(4), 329–339. https://doi.org/10.1037/a0013835.

Mayer, R. E., Fiorella, L., & Stull, A. (2020). Five ways to increase the effectiveness of instructional video. *Educational Technology Research and Development, 68*(3), 837–852. https://doi.org/10.1007/s11423-020-09749-6.

Menendez, D., Rosengren, K. S., & Alibali, M. W. (2020). Do details bug you? Effects of perceptual richness in learning about biological change. *Applied Cognitive Psychology, 34*(5), 1101–1117. https://doi.org/10.1002/acp.3698.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Journal of Clinical Epidemiology, 62*(10), 1006–1012. https://doi.org/10.1016/j.jclinepi.2009.06.005.

Moon, Y., & Nass, C. (1996). How "real" are computer personalities?: Psychological responses to personality types in human-computer interaction. *Communication Research, 23*(6), 651–674. https://doi.org/10.1177/009365096023006002.

Moon, J., & Ryu, J. (2020). The effects of social and cognitive cues on learning comprehension, eye-gaze pattern, and cognitive load in video instruction. *Journal of Computing in Higher Education.* https://doi.org/10.1007/s12528-020-09255-x.

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (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), 177–213. https://doi.org/10.1207/S1532690XCI1902_02.

Mori, M., MacDorman, K. F., & Kageki, N. (1970/2012). The uncanny valley. *IEEE Robotics and Automation Magazine, 19*(2), 98–100. https://doi.org/10.1109/MRA.2012.2192811.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153.

Nass, C., & Steuer, J. (1993). Voices, boxes, and sources of messages: Computers and social actors. *Human Communication Research, 19*(4), 504–527. https://doi.org/10.1111/j.1468-2958.1993.tb00311.x.

Nass, C., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. *Journal of Applied Social Psychology, 27*(10), 864–876. https://doi.org/10.1111/j.1559-1816.1997.tb00275.x.

*Nielen, T. M. J., Smith, G. G., Sikkema-de Jong, M. T., Drobisz, J., van Horne, B., & Bus, A. G. (2018). Digital guidance for susceptible readers: Effects on fifth graders' reading motivation and incidental vocabulary learning. *Journal of Educational Computing Research, 56*(1), 48–73. https://doi.org/10.1177/0735633117708283.

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., Donkin, C., Farrell, S., Hitch, G. J., Hurlstone, M. J., Ma, W. J., Morey, C. C., Nee, D. E., Schweppe, J., Vergauwe, E., & Ward, G. (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin, 144*(9), 885–958. https://doi.org/10.1037/bul0000153.

*Osman, K., & Lee, T. T. (2014). Impact of interactive multimedia module with pedagogical agents on students' understanding and motivation in the learning of electrochemistry. International *Journal of Science and Mathematics Education, 12*(2), 395–421. https://doi.org/10.1007/s10763-013-9407-y.

*Park, S. (2015). The effects of social cue principles on cognitive load, situational interest, motivation, and achievement in pedagogical agent multimedia learning. *Educational Technology & Society, 18*(4), 211–229.

Press, C. (2011). Action observation and robotic agents: Learning and anthropomorphism. *Neuroscience & Biobehavioral Reviews, 35*(6), 1410–1418. https://doi.org/10.1016/j.neubiorev.2011.03.004.

Rey, G. D., Beege, M., Nebel, S., Wirzberger, M., Schmitt, T. H., & Schneider, S. (2019). A meta-analysis of the segmenting effect. *Educational Psychology Review, 31*(2), 389–419. https://doi.org/10.1007/s10648-018-9456-4.

*Romero-Hall, E., Watson, G., & Papelis, Y. (2014). Using physiological measures to assess the effects of animated pedagogical agents on multimedia instruction. *Journal of Educational Multimedia and Hypermedia, 23*(4), 359–384.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638.

Sbarra, D. A., Briskin, J. L., & Slatcher, R. B. (2019). Smartphones and close relationships: The case for an evolutionary mismatch. *Perspectives on Psychological Science, 14*(4), 596–618. https://doi.org/10.1177/1745691619826535.

Schneider, S., Nebel, S., Beege, M., & Rey, G. D. (2018). Anthropomorphism in decorative pictures: Benefit or harm for learning? *Journal of Educational Psychology, 110*(2), 218–232. https://doi.org/10.1037/edu0000207.

*Schroeder, N. L. (2017). The influence of a pedagogical agent on learners' cognitive load. *Educational Technology & Society, 20*(4), 138–147.

*Schroeder, N. L., & Adesope, O. O. (2013). How does a contextually-relevant peer pedagogical agent in a learner-attenuated system-paced learning environment affect cognitive and affective outcomes? *Journal of Teaching and Learning with Technology, 2*(2), 114–133.

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), 229–251. https://doi.org/10.1080/15391523.2014.888265.

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), 1–39. https://doi.org/10.2190/EC.49.1.a.

Shiban, Y., Schelhorn, I., Jobst, V., Hörnlein, A., Puppe, F., Pauli, P., & Mühlberger, A. (2015). The appearance effect: Influences of virtual agent features on performance and motivation. *Computers in Human Behavior, 49*, 5–11. https://doi.org/10.1016/j.chb.2015.01.077.

Shimada, S. (2010). Deactivation in the sensorimotor area during observation of a human agent performing robotic actions. *Brain and Cognition, 72*(3), 394–399. https://doi.org/10.1016/j.bandc.2009.11.005.

Shimada, S., & Oki, K. (2012). Modulation of motor area activity during observation of unnatural body movements. *Brain and Cognition, 80*(1), 1–6. https://doi.org/10.1016/j.bandc.2012.04.006.

Sinatra, A. M., Pollard, K. A., Files, B. T., Oiknine, A. H., Ericson, M., & Khooshabeh, P. (2021). Social fidelity in virtual agents: Impacts on presence and learning. *Computers in Human Behavior, 114*, 106562. https://doi.org/10.1016/j.chb.2020.106562.

Stull, A. T., Fiorella, L., Gainer, M. J., & Mayer, R. E. (2018a). Using transparent whiteboards to boost learning from online STEM lectures. *Computers & Education, 120*, 146–159. https://doi.org/10.1016/j.compedu.2018.02.005.

Stull, A. T., Fiorella, L., & Mayer, R. E. (2018b). An eye-tracking analysis of instructor presence in video lectures. *Computers in Human Behavior, 88*, 263–272. https://doi.org/10.1016/j.chb.2018.07.019.

Sundararajan, N., & Adesope, O. (2020). Keep it coherent: A meta-analysis of the seductive details effect. *Educational Psychology Review, 32*(3), 707–734. https://doi.org/10.1007/s10648-020-09522-4.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5.

Syrjämäki, A. H., Isokoski, P., Surakka, V., Pasanen, T. P., & Hietanen, J. K. (2020). Eye contact in virtual reality – A psychophysiological study. *Computers in Human Behavior, 112*, 106454. https://doi.org/10.1016/j.chb.2020.106454.

*Tegos, S., & Demetriadis, S. (2017). Conversational agents improve peer learning through building on prior knowledge. *Educational Technology & Society, 20*(1), 99–111.

Tinwell, A., Grimshaw, M., Nabi, D. A., & Williams, A. (2011). Facial expression of emotion and perception of the Uncanny Valley in virtual characters. *Computers in Human Behavior, 27*(2), 741–749. https://doi.org/10.1016/j.chb.2010.10.018.

*van der Meij, H., van der Meij, J., & Harmsen, R. (2015). Animated pedagogical agents effects on enhancing student motivation and learning in a science inquiry learning environment. *Educational Technology Research and Development, 63*(3), 381–403. https://doi.org/10.1007/s11423-015-9378-5.

Veletsianos, G. (2012). How do learners respond to pedagogical agents that deliver social-oriented non-task messages? Impact on student learning, perceptions, and experiences. *Computers in Human Behavior, 28*(1), 275–283. https://doi.org/10.1016/j.chb.2011.09.010.

Wang, J., & Antonenko, P. D. (2017). Instructor presence in instructional video: Effects on visual attention, recall, and perceived learning. *Computers in Human Behavior, 71*, 79–89. https://doi.org/10.1016/j.chb.2017.01.049.

*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), 250–268. https://doi.org/10.1037/edu0000221.

Wang, J., Antonenko, P., & Dawson, K. (2020). Does visual attention to the instructor in online video affect learning and learner perceptions? An eye-tracking analysis. *Computers & Education, 146*, 103779. https://doi.org/10.1016/j.compedu.2019.103779.

Williams, E. H., Cristino, F., & Cross, E. S. (2019). Human body motion captures visual attention and elicits pupillary dilation. *Cognition, 193*, 104029. https://doi.org/10.1016/j.cognition.2019.104029.

Wong, M., Castro-Alonso, J. C., Ayres, P., & Paas, F. (2015). Gender effects when learning manipulative tasks from instructional animations and static presentations. *Educational Technology & Society, 18*(4), 37–52.

Wong, M., Castro-Alonso, J. C., Ayres, P., & Paas, F. (2018). Investigating gender and spatial measurements in instructional animation research. *Computers in Human Behavior, 89*, 446–456. https://doi.org/10.1016/j.chb.2018.02.017.

Yi, T., Yang, X., Pi, Z., Huang, L., & Yang, J. (2019). Teachers' continuous vs. intermittent presence in procedural knowledge instructional videos. *Innovations in Education and Teaching International, 56*(4), 481–492. https://doi.org/10.1080/14703297.2018.1470020.

*Yılmaz, R., & Kılıç-Çakmak, E. (2012). Educational interface agents as social models to influence learner achievement, attitude and retention of learning. *Computers & Education, 59*(2), 828–838. https://doi.org/10.1016/j.compedu.2012.03.020.

Affiliations

Juan C. Castro-Alonso[1] · **Rachel M. Wong**[2] · **Olusola O. Adesope**[2] · **Fred Paas**[3,4]

[1]   Center for Advanced Research in Education, Institute of Education, Universidad de Chile, Santiago, Chile

[2]   College of Education, Washington State University, Pullman, WA, USA

[3]   Department of Psychology, Education, and Child Studies, Erasmus University Rotterdam, Rotterdam, The Netherlands

[4]   School of Education/Early Start, University of Wollongong, Wollongong, Australia