

# Role of Race and Gender of Pedagogical Agents in Multimedia Learning

Journal of Educational Computing Research  
2025, Vol. 0(0) 1–40  
© The Author(s) 2025



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/07356331251322469

[journals.sagepub.com/home/jec](https://journals.sagepub.com/home/jec)



Fangzheng Zhao<sup>1</sup> , Richard E. Mayer<sup>1</sup>,  
Nicoletta Adamo-Villani<sup>2</sup>,  
Christos Mousas<sup>2</sup> , Minsoo Choi<sup>2</sup>, and  
Klay Hauser<sup>2</sup>

## Abstract

Do students learn from video lessons presented by pedagogical agents of different racial and gender types equivalently to those delivered by a real human instructor? How do the race and gender of these agents impact students' learning experiences and outcomes? In this between-subject design study, college students were randomly assigned to view a six 9-minute video lesson on chemical bonds, presented by pedagogical agents varying in gender (male, female) and race (Asian, Black, White), or to view the original lesson with a real human instructor. In comparing learning with a human instructor versus with a pedagogical agent of various races and genres, ANOVAs revealed no significant differences in learning outcomes (retention and transfer scores) or learner emotions, but students reported a stronger social connection with the human instructor over pedagogical agents. Students reported stronger positive emotions and social connections with female agents over male agents. Additionally, there was limited evidence of a race-matching effect, with White students showing greater positive emotion while learning with pedagogical agents of the same race. These findings highlight the limitations of pedagogical agents compared to human instructors in video lessons, while partially reflecting gender stereotypes and intergroup bias in instructor evaluations.

---

<sup>1</sup>University of California Santa Barbara, Goleta, CA, USA

<sup>2</sup>Purdue University, West Lafayette, IN, USA

## Corresponding Author:

Fangzheng Zhao, Department of Psychological and Brain Sciences University of California, Santa Barbara, Santa Barbara, CA 93106, USA.

Email: [f\\_zhao@ucsb.edu](mailto:f_zhao@ucsb.edu)

**Keywords**

pedagogical agents, race, gender, multimedia learning, learning experiences and outcomes

**Introduction***Objective and Rationale*

Consider a learning scenario in which a student sits at a computer screen that delivers a video lesson on a science topic. The instructor presenting the lesson on the screen is either a real human instructor or a pedagogical agent of a particular gender (male, female) and race (Asian, Black, White). Do the students learn better with the real human instructor than the pedagogical agents? Can the gender and race of the pedagogical agents affect students' learning experiences (i.e., affective and social processes) and learning outcomes (i.e., cognitive processes)? This present study addresses these issues.

The motivation for this study arises from the potential influence of the instructor's race and gender on learners' emotional, social, and cognitive processing during learning from video lectures. By using pedagogical agents, we can replicate the original lesson delivered by a human instructor, altering only the instructor's race and gender. This approach significantly enhances diversity in online education. Unlike human instructors, whose identities are fixed, pedagogical agents offer flexibility, allowing for customizable appearances and voices that provide a more inclusive learning experience for students from diverse backgrounds. Therefore, this study aimed to investigate whether pedagogical agents are as effective as, or more effective than, human instructors for students, and whether students react differently to pedagogical agents of certain genders or races.

In terms of theoretical implications, this research also seeks to examine whether theories such as the Media Equation Theory (Reeves & Nass, 1996), the Alliance Hypothesis (Taylor et al., 1978), and the Matching Hypothesis (Kalick & Hamilton, 1996) can be extended into learning with pedagogical virtual agents. The Media Equation Hypothesis suggests that learners interact with pedagogical agents in the same way they do with real humans, indicating that pedagogical agents in video lessons could be just as effective as human instructors, while also subject to similar racial and gender biases. Theories based on intergroup bias, such as the Alliance Hypothesis and the Matching Hypothesis, suggest that students tend to favor instructors who share their racial or gender characteristics, raising the important question of whether such preferences extend to pedagogical agents. In terms of practical implications, the findings of this investigation are intended to provide insights into the gender and racial design of animated pedagogical agents, underscoring their potential to promote inclusivity and diversity in multimedia learning environments.

## Theoretical Background

*Media Equation Theory.* Media Equation Theory, introduced by Reeves and Nass (1996; 2005), argues that people tend to instinctively respond to computers and media as if they were real human beings. This response is rooted in the human brain's natural tendency to apply social and cognitive behaviors to non-human entities (Rosenthal-Von Der Putten et al., 2014). As a result, under certain conditions, people could interact with digital agents in virtual settings as if they were interacting with actual people. A range of studies (Lawson et al., 2021a, 2021b, 2021c; Lawson & Mayer, 2021; Nass et al., 1997; Nass & Steuer, 1993; Zhao & Mayer, 2023a, 2023b) has provided support for this theory. These investigations have examined whether behaviors typically seen in human-to-human or human-environment interactions—such as social connections, emotional reactions, and gender-based responses—are similarly reflected in interactions between humans and technology.

For example, one study by Nass and Steuer (1993) explored social connections and emotional reactions in polite interactions with computers. Participants were asked to evaluate a computer's performance in teaching a lesson, and were randomly assigned to provide feedback either on the same computer they had used for the lesson or on a different computer in another room. Interestingly, those who evaluated on the same computer gave significantly more favorable responses than those who evaluated on a different computer, suggesting a form of politeness toward the machine. This phenomenon was later replicated in both text- and voice-based interactions, indicating that people tend to act politely toward computers, as if avoiding hurting the computer's "feelings."

In a more recent study by Zhao & Mayer (2023b), two experiments were conducted to explore the effect of both the lesson design (cartoon vs. original slides) and voice types (machine vs. human) on learning experiences and outcomes. Participants viewed either cartoon-like slides or original line-drawn slides, narrated by either a machine-synthesized voice or a human voice. The findings revealed no significant interaction between lesson design and voice type, suggesting that both types of voices led to similar effects on students' emotions, social connections with instructors, and learning outcomes.

Another study investigated gender stereotypes by assigning computer voices to topics typically associated with either men or women (Nass et al., 1997). The results showed that participants easily detected gender differences and applied human-like stereotypes to the computer agents, based solely on the voice. This demonstrates that people engage in social interactions with technology, whether through text or voice, and that these behaviors are triggered by subtle cues in the interaction.

Additionally, a more recent study by Zhao et al. (2024) asked participants to observe virtual agents of different races and genders and then identify their characteristics. Although this study did not investigate the effect of those various virtual agents on students' learning, the findings indicated that individuals could accurately recognize the race and gender of virtual agents, much like they do with real people.

In summary, several research studies supporting the Media Equation Hypothesis show that individuals can form social connections with technological entities, such as computers, machine voices, and virtual agents. However, the specific influence of race and gender of virtual agents on students' learning remains an area that requires further exploration, particularly in terms of their effectiveness compared to human instructors.

Grounded in the Media Equation Theory, which suggests that people respond to technological entities as they would to real humans, this study explores how pedagogical virtual agents with diverse racial and gender identities influence students' learning experiences (i.e., affective and social dimensions) and academic performance (i.e., cognitive processes). Specifically, it seeks to expand the theoretical framework by investigating whether learners engage with virtual agents in video-based learning environments in ways that parallel human-to-human interactions. Given the theory's premise that non-human entities can elicit human-like responses, this study utilizes sophisticated 3D modeling techniques to create pedagogical agents based on an original female human instructor. These agents represent six distinct identities: Asian Female, Asian Male, Black Female, Black Male, White Female, and White Male.

Guided by the Media Equation Theory, this research seeks to determine whether students taught by these different types of pedagogical agents achieve equivalent or superior learning outcomes compared to those taught by a real human instructor. Furthermore, this study examines whether the racial and gender identities of the agents influence students' learning experiences, including their affective feelings and social connections with instructors, as suggested by the previous findings of human-like social and emotional interactions with media (Lawson et al., 2021a, 2021b, 2021c; Lawson & Mayer, 2021; Zhao & Mayer, 2023a, 2023b). By addressing these questions, the study not only extends the Media Equation Theory into the domain of multimedia education but also provides practical insights into the design and application of diverse pedagogical agents. The findings can potentially promote a more inclusive and equitable learning environment, benefiting students from various cultural and social backgrounds.

*Racial and Gender Stereotypes in Instructor Evaluations.* The influence of gender and racial stereotypes on instructor evaluations has been widely studied, with evidence suggesting that students' perceptions of their instructors are deeply shaped by these biases (Eagly & Wood, 2012; Kierstead et al., 1988).

Specifically, Social Role Theory (Eagly & Wood, 2012) posits that gender roles are shaped mainly by the division of labor within society, with women typically perceived as nurturing and emotionally expressive, while men are viewed as assertive, dominant, and competent. These stereotypical perceptions are often extended to the classroom, influencing how students perceive and evaluate their instructors based on gender. A study by Kierstead et al. (1988) explored the impact of gender stereotypes on the evaluations of instructors. Their findings revealed that female instructors who demonstrated warmth—such as by smiling or engaging in informal social interactions with students—received more favorable evaluations. These behaviors reinforced the

stereotype of women as nurturing and approachable, suggesting that female instructors are often held to different standards.

In contrast, male instructors' evaluations were largely unaffected by similar behaviors, highlighting a gender bias in how professional competence is perceived. Female instructors, therefore, may be expected to display warmth to receive positive evaluations, while male instructors are judged more on authority and competence. For example, [Renström and colleagues \(2021\)](#) examined how gender stereotypes influence student evaluations of teaching (SET), showing that female lecturers are evaluated based on adherence to traditional gender roles. Feminine traits (e.g., nurturing) are linked to higher likability, while masculine traits (e.g., assertiveness) are associated with competence. However, research by [Khokhlova et al. \(2023\)](#) challenges earlier findings by showing that male virtual instructors in higher education received significantly higher ratings than female virtual instructors, particularly on traits such as enthusiasm and expressiveness—traits typically associated with female instructors.

In addition to gender, racial stereotypes also heavily influence instructor evaluations by students. For example, a study conducted by [Campbell \(2023\)](#) focused on racial disparities in teacher evaluations within North Carolina's Teacher Evaluation System (TES), specifically among female teachers. The findings highlight that Black female teachers consistently received lower classroom observation ratings compared to White female teachers, despite exhibiting similar levels of teaching effectiveness. In addition, the intersection of race and gender in student evaluations is further evidenced by [Reid \(2010\)](#), who analyzed ratings from [RateMyProfessors.com](#). Minority faculty, particularly Black and Asian professors, received lower ratings in overall quality, helpfulness, and clarity compared to their White counterparts. These evaluations were often skewed by racial stereotypes, despite the actual effectiveness of the instructor. The results also showed that gender played a less pronounced role overall, but Black male professors were evaluated more harshly than other groups, which highlights the compounded impact of both race and gender in instructor evaluations.

Furthermore, [Anderson \(2010\)](#) examined the double-layered impact of race and gender on student perceptions of professors. The study found that female professors, particularly Latina professors, were rated higher in warmth when teaching traditionally feminine courses such as composition. However, Latina professors faced polarized evaluations depending on their teaching style, a phenomenon known as response amplification. Students either excessively praised or penalized them based on whether their teaching conformed to stereotypical expectations, revealing how both race and gender intersect to influence student perceptions.

As another example, [Basow et al. \(2013\)](#) examined how race and gender stereotypes influence both student evaluations and academic performance. The study found that Black and female professors consistently received lower evaluations compared to their White and male counterparts, despite delivering similar teaching performance. Additionally, female professors were expected to display nurturing behaviors in order to receive positive evaluations, while male professors were rewarded for demonstrating authority and competence. These stereotypes not only skewed evaluations but also

affected students' academic outcomes. Specifically, students performed better when their professors aligned with societal expectations. For example, male or White professors were perceived as more competent and, therefore, more effective in enhancing student performance. In contrast, minority professors, especially Black women, faced biases that undermined their credibility and effectiveness, ultimately negatively impacting student learning outcomes.

In conclusion, while the literature demonstrates that racial and gender stereotypes significantly influence student evaluations of instructors, there is no clear consensus on how these biases affect students' learning experiences (i.e., affective and social processes) and learning outcomes (i.e., cognitive processes). To address this gap, the present study aims to explore which gender and racial types of pedagogical agents are most preferred by students and result in the best learning experiences and outcomes, contributing to a deeper understanding of the impact of these stereotypes in educational settings.

Expanding on prior research, this study examines the design and implementation of pedagogical agents that vary in both race and gender. Using state-of-the-art 3D modeling, a female human instructor in an instructional video was digitally transformed into six virtual agents, encompassing three racial backgrounds (Asian, Black, and White) and two gender identities (Female and Male). The resulting six agents—Asian Female, Asian Male, Black Female, Black Male, White Female, and White Male—serve as the foundation for analyzing how race and gender differences in virtual instructors may shape students' learning experiences, including affective and social interactions, as well as cognitive learning outcomes. By incorporating these diverse pedagogical agents, this study builds upon existing work on instructor stereotypes, providing further insight into their influence on students' perceptions and engagement in digital learning environments.

This study is informed by insights from Social Role Theory and prior research (Eagly & Wood, 2012; Kierstead et al., 1988), which highlight how gender and racial stereotypes influence the perception and evaluation of instructors. By incorporating diverse virtual agents, this study provides a nuanced exploration of whether students' interactions with these agents also reflect the biases commonly observed in evaluations of human instructors. Furthermore, this approach addresses a critical gap in the existing literature by investigating how virtual agents of varying racial and gender identities may differentially influence the effects of stereotypes in multimedia learning environments, helping us to identify the most preferred and effective pedagogical agent type for enhancing student learning. The broader impact of this research could guide the development of more inclusive and effective multimedia learning environments, tailoring pedagogical agents to meet the diverse needs of students and potentially reducing the negative effects of stereotypes on learning.

*Alliance Hypothesis and Intergroup Bias Theory.* According to the Alliance Hypothesis, humans are naturally inclined to form alliances with individuals who share similar physical characteristics, such as race and gender, promoting cooperation and collective

success (Taylor et al., 1978). This theory argues that racial and gender categorization plays a role of visual cue that influences the form of alliance, subsequently shaping social behavior and group dynamics (Kurzman et al., 2001; Pietraszewski, 2021; Taylor et al., 1978). Evolutionary theories such as kin selection and reciprocal altruism (Eberhard, 1975; Michod, 1982) provide explanations for the Alliance Hypothesis, which proposes that humans have an innate tendency to align with genetically similar individuals, thereby increasing the likelihood of mutual support and collaboration.

Numerous studies provide support for the Alliance Hypothesis, demonstrating that individuals naturally form groups with clear boundaries based on superficial and context-dependent cues like race, gender, clothing, or speech. For instance, Taylor et al. (1978) showed that participants categorized others based on race and gender while observing a group discussion, making more errors within racial or gender categories than between them, suggesting a tendency to minimize differences within groups and exaggerate differences between them. Similarly, additional research (Cosmides et al., 2003; Pietraszewski, 2009, 2016, 2021) revealed that people can quickly and automatically identify racial and gender categories, often associating these characteristics with alliances.

Pietraszewski's (2021) study, for example, demonstrated that when participants had no clear team memberships, they defaulted to race as a basis for forming alliances. However, when team affiliations were clear, race categorization was temporarily suppressed, though gender continued to strongly influence alliance formation. Together, these findings emphasize that people tend to minimize within-group differences and exaggerate between-group differences, using race and gender as key alliance markers.

Building on this, Intergroup Bias Theory suggests that these alliance formations can extend into intergroup biases, emphasizing that individuals tend to favor their own group (the in-group) over other groups (the out-group) based on characteristics such as race and gender, particularly when social hierarchies or group statuses are perceived as unstable (Hewstone et al., 2002). For example, Social Dominance Theory (SDO) explains that men often exhibit higher levels of intergroup bias due to their greater social dominance orientation. Men, compared to women, are more likely to promote hierarchical distinctions between groups, reflecting stronger gender-based intergroup biases (Sidanius et al., 2000). Similarly, racial biases also emerge, with individuals showing a tendency to favor their racial in-group. Social dominance theorists have found that individuals with high SDO across racial groups tend to support systems that maintain societal hierarchies, which corresponds with heightened racial bias (Hewstone et al., 2002). These findings illustrate how both race and gender influence intergroup bias, shaping social behavior and preferences, especially when group status is perceived as unstable or at risk.

However, previous studies on the Alliance Hypothesis and Intergroup Bias have primarily focused on real human interactions rather than virtual agents, and they have not extensively explored educational contexts. This leaves a gap in research regarding the application of these theories to pedagogical agents, particularly pedagogical agents

in educational settings. Specifically, there is limited understanding of how these theories influence students' learning experiences (i.e., affective and social processes) and learning outcomes (i.e., cognitive processes) in multimedia environments taught by pedagogical agents of different racial and gender types. In this study, we intended to extend the scope of the Alliance Hypothesis and Intergroup Bias by investigating whether students exhibit similar tendencies and biases when evaluating their instructors—specifically, pedagogical agents with varying racial and gender characteristics. We aimed to explore whether students show more positive emotions and stronger social connections with agents who share their race and/or gender, and if so, whether these tendencies and biases influence their learning outcomes.

*Matching Hypothesis.* In line with the Alliance Hypothesis and Intergroup Bias Theory, the Matching Hypothesis posits that individuals are inclined to form relationships with others who share similar characteristics, including demographic factors such as gender and race (Berscheid & Reis, 1998; Kalick & Hamilton, 1996; Murstein, 1980). Numerous studies have supported this hypothesis, emphasizing its impact on educational settings. Specifically, when extended to educational contexts, the Matching Hypothesis posits that students learn better with instructors of the same race and gender as them.

*Gender-Matching Effect.* Research on the gender-matching effect in students' learning (Makransky et al., 2019; Zhao & Mayer, 2023a) has shown that students tend to perform better and experience stronger social connections when their instructor, including pedagogical agents, matches their gender. For example, a study by Zhao and Mayer (2023a) investigated how the emotional tone (happy vs. sad) and gender (male vs. female) of machine voices influenced learners' emotions, social connection with the instructor, and learning outcomes in multimedia lessons. The results provided partial evidence for the gender matching effect, demonstrating that female learners responded more positively to and built stronger social connections with female machine voices than male voices. While the gender-matching design of machine voices did not significantly affect learning outcomes, it enhanced learners' emotional experiences and social connections with instructors of the same gender.

Similarly, in another study by Makransky et al. (2019), middle school students were tasked with learning about laboratory safety in an immersive virtual reality environment. The participants were divided into two groups: one group interacted with a female virtual agent (Marie), and the other with a drone designed to serve as a male role model. The findings revealed that girls performed significantly better with the female agent, achieving higher retention and transfer scores, while boys also performed better when interacting with the male-representing drone. These results suggest that the gender-specific design of pedagogical agents can improve learning outcomes, especially when the agent's gender aligns with the learner's gender.

Solanki and Xu (2018) also revealed that female instructors positively influence female students' motivation, serving as role models who enhance engagement and foster identity congruence. Specifically, while female students generally underperform



compared to male students in these courses, the presence of female instructors slightly reduces this performance gap, which helps narrow gender disparities in interest and persistence in STEM fields.

Overall, these studies underscore the potential important role of gender-matching effect in educational settings, where gender congruence between students and instructors can enhance both learning experiences (i.e., affective and social processes) and academic performance (i.e., cognitive processes).

**Race-Matching Effect.** Previous research on the race-matching effect has shown that students, particularly those from minority groups, tend to benefit when their instructors share their racial background (de Albuquerque Rocha et al., 2024; Egalite et al., 2015; Gershenson et al., 2021; Harbatkin, 2021; Joshi & James, 2022). In their book, *Teacher Diversity and Student Success: Why Racial Representation Matters in the Classroom*, Gershenson and colleagues (2021) emphasize the critical role of teacher racial diversity in fostering equitable educational outcomes, suggesting that racial representation among teachers is essential for closing the racial and ethnic achievement gap and enhancing student success, particularly for students of color. They highlight the benefits of same-race teacher-student matching, such as improved test scores, graduation rates, attendance, and relationships, and reduced behavioral infractions.

In addition, prior research findings have also highlighted the positive impact of race-matching between students and instructors on learning performances. For instance, Egalite et al. (2015) utilized a large dataset from the Florida Department of Education to examine the academic performance of students in grades 3–10. Their findings revealed that students, especially Black and Asian/Pacific Islander students, performed better in both math and reading when their teachers matched their racial or ethnic background. This race congruence was particularly beneficial for minority students in lower-performing schools. Additionally, Harbatkin (2021) investigated the impact of race matching on course grades, demonstrating that Black students, in particular, achieved better academic outcomes when they were taught by teachers of the same race. The effects were especially notable for lower-performing students, suggesting that race matching may help mitigate achievement gaps.

Overall, these studies highlight the positive impact of race-matching between students and instructors, suggesting that a more racially diverse and inclusive teaching workforce could enhance academic outcomes and better meet the needs of students from diverse backgrounds, particularly those from underrepresented groups.

**Research Gaps.** While previous studies supporting the Matching Hypothesis in educational settings have primarily focused on its impact on cognitive processes measured by students' learning outcomes, it is equally important to explore how race and gender-matching designs influence students' learning experiences, indicated by students' emotional experiences and social connections during learning. To address this gap, the present study includes a supplemental exploratory analysis that investigates not

only the effect of gender and race-matching pedagogical agents on students' learning outcomes but also on their overall learning experiences.

### *Theory and Predictions*

Pedagogical agents, with their flexibility to adapt race, gender, and emotional expressions through facial cues and voice tones, provide a unique opportunity to create more diverse, inclusive, and customizable learning experiences. This flexibility may lead to different learning experiences (i.e., affective and social processes) and learning outcomes (i.e., cognitive processes when compared to human instructors). Since this study is in its exploratory phase and prior research has shown inconsistent findings regarding the impact of race and gender stereotypes on instructor evaluations and learning outcomes, we have chosen to frame our investigation through research questions rather than specific hypotheses.

These research questions are grounded in the framework of the Cognitive-Affective Model of Learning with Media (Lawson & Mayer, 2021; Mayer, 2022; Moreno & Mayer, 2007; Zhao & Mayer, 2023a, 2023b, 2024). According to this model, meaningful learning with media follows a cascading process that incorporates affective, social, and cognitive processing, ultimately resulting in improved learning outcomes. Specifically, in this study, **learning experiences** include the affective processes and social processes students engage in during learning (i.e., ratings of their felt emotions and perceived connections with the instructor), while **learning outcomes** indicate the cognitive processes measured by their posttest performance (i.e., retention and transfer test scores). Therefore, this research is guided by six main research questions, with two questions addressing each type of process—cognitive processing, affective processing, and social processing—as outlined in the Cognitive-Affective Model of Learning with Media. This alignment ensures a consistent approach to investigating how various agents or the original human instructor impact students' overall learning experiences (including affective and social processes) and outcomes (cognitive processes) in video-based learning.

#### *Learning Outcomes: Cognitive Processes*

*Research Question 1.* Do students achieve different learning outcomes from a video lesson delivered by pedagogical agents compared to a real human instructor? Specifically, we aimed to assess differences in retention and transfer posttest scores across the seven video lesson conditions (i.e., Asian female agent, Asian male agent, Black female agent, Black male agent, White female agent, White male agent, and the original human instructor). Our goal was to compare the effectiveness of the original human instructor with the various types of pedagogical agents in terms of their impact on students' learning outcomes. According to the Media Equation Hypothesis, we expected that pedagogical agents could be as effective, or perhaps even more effective, than human instructors in enhancing students' learning experiences (i.e., affective and social processes) and learning outcomes (i.e., cognitive processes).

*Research Question 2.* Are the learners' learning outcomes affected by the pedagogical agents' gender and/or race? Specifically, we are particularly interested in exploring the race and gender of the pedagogical agents to determine which combination of race and gender is most effective in facilitating student learning.

#### *Learning Experiences: Affective Processes*

*Research Question 3.* Do pedagogical agents lead to different felt emotions of the learners compared with a real human instructor? Specifically, we aimed to assess differences in students' ratings of their own positive and negative emotions during learning across the seven video lesson conditions (i.e., Asian female agent, Asian male agent, Black female agent, Black male agent, White female agent, White male agent, and the original human instructor). Our goal was to compare the impact of the original human instructor with the various pedagogical agents on students' emotional experiences during learning.

*Research Question 4.* Are the learners' felt emotions affected by pedagogical agents' gender and race? Specifically, we aimed to investigate the race and gender of the pedagogical agents to determine which combination is most effective in enhancing students' emotional experiences during learning.

#### *Learning Experiences: Social Processes*

*Research Question 5.* Are the learners' social connections with the instructor different for pedagogical agents compared with the real human instructor? Our specific goal was to examine differences in students' ratings of perceived social connections with the instructors across the seven video lesson conditions (i.e., Asian female agent, Asian male agent, Black female agent, Black male agent, White female agent, White male agent, and the original human instructor). We aimed to compare the influence of the original human instructor versus the various pedagogical agents on how socially connected students felt to their instructors. Based on research into racial and gender stereotypes in instructor evaluations, we anticipated that female pedagogical agents may be generally preferred, as female instructors are often perceived as more supportive. Similarly, White pedagogical agents might be favored due to their perceived credibility, which is often rated higher than that of minority instructors.

*Research Question 6.* Are the learners' social connections with the instructor affected by the pedagogical agents' gender and race? Specifically, we aimed to investigate the race and gender of the pedagogical agents to determine which combination best facilitates the development of a strong social connection between students and the agent.

For the exploratory analysis, we further examined the potential impact of gender and race matching on students' learning experiences (i.e., affective and social processes) and learning outcomes (i.e., cognitive processes) when interacting with pedagogical agents of varying gender and race. Building on the Alliance Hypothesis, Intergroup Bias Theory, and the Matching Hypothesis, we expected students to exhibit a

preference for pedagogical agents that share their own race or gender, potentially demonstrating a race and gender matching effect on both learning experiences and outcomes. Additionally, we compared the overall effectiveness of pedagogical agents against a real human instructor by consolidating data from all pedagogical agent groups and comparing it to the data from the original human instructor group.

## Method

### *Participants and Design*

The participants were 229 undergraduate students from the psychology subject pool at a university in California, in which they fulfilled a course requirement by participating. Concerning gender, 153 identified as female, 72 identified as male, 3 identified as other gender, and 1 preferred not to say. Concerning race and ethnicity, 60 identified as Asian, 5 identified as Black, 72 identified as Hispanic/Latino, 4 identified as Indian, 49 as White, 39 as other ethnicity, and 2 preferred not to say. The mean age was 19.25 years ( $SD = 1.49$ ). The mean prior knowledge score was 6.80 ( $SD = 2.87$ ), indicating low prior knowledge of the lesson topic (maximum possible score = 13).

To ensure the sample size was sufficient for detecting meaningful effects, an a priori power analysis was conducted using G\*Power 3.1.9.4. The analysis was based on a one-way ANOVA with seven groups, an alpha level of 0.05, power of 0.80, and a medium effect size ( $f = 0.25$ ). The results indicated that a minimum total sample size of 231 participants was required to achieve the desired power level. The actual sample size in this study ( $N = 229$ ) closely approximated this requirement, resulting in an actual power of approximately 80.97%, confirming that the sample size was sufficient for the planned analyses.

In a between-subjects design based on the characteristics of the instructor in a 9-min video lecture on chemical bonding participants were randomly assigned to one of seven groups: 33 participants received the original lesson with a real human instructor (original group), 32 viewed the lesson with an Asian female agent, 33 viewed the lesson with an Asian male agent, 33 viewed the lesson with a Black female agent, 32 viewed the lesson with a Black male agent group, 34 viewed the lesson with a White female agent, and 32 viewed the lesson with a White male agent group. The dependent measures in this study included posttest scores (i.e., retention and transfer scores), ratings of learners' felt emotions, and ratings of learner's partnership connections with the instructor.

### *Materials*

The materials included a perceived prior knowledge questionnaire, seven versions of a 9-minute video lesson about chemical bonds (i.e., original lesson, Asian female agent lesson, Asian male agent lesson, Black female agent lesson, Black male agent lesson, White female lesson, and White male lesson), Positive and Negative Affect Schedule

(i.e., PANAS; to measure the students' felt emotion), Agent Persona Instrument (i.e., to assess the social partnership connection of learners with the instructor), and demographics survey. All research materials were published as a Qualtrics survey and presented on Dell or iMac desktop computers in individual cubicles in a research lab.

*Perceived Prior Knowledge Questionnaire.* The perceived prior knowledge questionnaire assessed participants' knowledge of chemistry before taking the video lesson. The first question asked participants to rate how much knowledge they think they have about chemistry (i.e., "Please rate your knowledge of chemistry:"), on a scale of 1 (very little) to 5 (very much). The second question asked participants to check any of nine statements about their experience or understanding of chemistry knowledge that applied to them (i.e., "I have taken a chemistry class before." or "I know what endothermic means."). The perceived prior knowledge score was calculated by summing the rating of the first question and the number of items chosen in the second question together. The Cronbach's alpha showed an acceptable reliability level,  $\alpha = 0.78$ . We used a questionnaire in order to prevent the possibility of a testing effect, in which the act of taking a test is a learning event that can affect the process and outcome of subsequent learning (Brown et al., 2014; Roediger & Karpicke, 2006).

*Video Lessons.* There were seven versions of a 9-minute video lesson that provided an introduction to three types of chemical bonds and their formation processes. Each of the seven versions of the video involved an instructor standing at a podium next to a series of slides as they lectured. All versions maintained the same script and slides arranged in the same way, ensuring content consistency across the groups.

The original version was excerpted from a UCI Open course lecture on general chemistry available on YouTube (<https://youtu.be/6GjYGd-k32U?t=92>; Brindley, 2013), spanning from 1:32 to 10:46, with the UCI logo and the professor's name obscured for this study. The instructor was a White female.

Six modified versions of the original video were created by replacing the human instructor with a 3D animated pedagogical agent that was either Asian female, Asian male, Black female, Black male, White female, or White male. Versions with a female agent used the same voice as in the original version whereas versions with a male agent were given a corresponding synthesized male voice, while the script remained unaltered. Example screenshots of the original video lesson and modified video lesson are shown in Figure 1. Access to the seven versions of the video lectures is available at the following link: [https://osf.io/zy2tm/?view\\_only=9e6e67ec00434d15aa052be8d66ccbc8](https://osf.io/zy2tm/?view_only=9e6e67ec00434d15aa052be8d66ccbc8).

The six animated agents were selected based on Zhao and colleagues' previous research (2024), which explored the degree to which people could relate to and recognize the race/ethnicity and gender of animated pedagogical agents intended to represent different races/ethnicities and genders. It was found that participants could more accurately identify the agents' gender (female vs. male) and race (Asian, Black, and White) rather than their ethnicity (Indian and Hispanic). Therefore, in this study, we



**Figure 1.** Example screenshots of the original video lesson and a modified video lesson with a pedagogical agent.

*Note.* The figure on the top panel is the screenshot of the original video lesson presented by a human instructor, and the figure on the bottom panel is the example screenshot of the example screenshot of a modified video lesson presented by a pedagogical agent (White male agent).

used agents from Zhao and colleagues' earlier work (2024) that were found to be superior in representing the various racial types (Asian, Black, White) and genders (female and male). Specifically, the six agents selected—Asian female, Asian male, Black female, Black male, White female, and White male—were chosen based on two main criteria: first, the agent supported high accuracy in participants' being able to identify their racial and gender types; and second, the agent supported high ratings on scales for human-likeness and likability. In short, within each of the six categories of agents, we selected the one that Zhao and colleagues (2024) found displayed the most recognizable racial and gender characteristics and was rated highest in likability and human-likeness. Example screenshots of the six agents are shown in Figure 2.

To integrate the pedagogical agents into the video lecture, we processed the original video. For the motion of the pedagogical agents, we extracted the instructor's movements from the original video and applied them to the pedagogical agents. Specifically, using a deep learning-based video-to-motion extraction method, we obtained the instructor's motion data. Then, we retargeted it to the pedagogical agents.



**Figure 2.** Example screenshots of the six agents.

*Note.* The figures displayed on the top panel (from left to right) represent the Asian female agent, the White female agent, and the Black female agent. The figures on the bottom panel (from left to right) are the Asian male agent, the White male agent, and the Black male agent.

Additionally, to enhance the realism of the pedagogical agents, we applied eye movements, such as blinking and gaze shifts, and lip-sync animation based on the original lecture's script. Furthermore, we added motions like pressing a button when the slide image changed to make the pedagogical agents appear more natural. Regarding the voice of the pedagogical agents, we aimed to preserve the instructor's voice from the original video. To achieve this, we extracted the audio file from the original lecture and used it for female pedagogical agents. For male pedagogical agents, we employed Praat (Boersma, 2011) to convert the audio file to a male voice. We normalized it to the recommended 23 loudness units relative to full scale (LUF) for consistent perceived loudness (EBU-Recommendation, 2011). Moreover, we designed a 3D virtual classroom as the environment for the video lesson. Specifically, we placed a podium in front of the pedagogical agent and positioned a slide screen on the right side of the pedagogical agent.

*Positive and Negative Affect Schedule (PANAS).* To measure the positive and negative emotions experienced by students, the study utilized the Positive and Negative Affect Schedule (PANAS), a well-established instrument in psychology for measuring individuals' emotional states and moods (Plass et al., 2020; Watson et al., 1988). The PANAS comprises 20 emotion-related words that express either positive (e.g.,

“excited”;  $\alpha = 0.91$ ) or negative emotions (e.g., “nervous”;  $\alpha = 0.83$ ), and instructs participants to rate how strongly they experienced each listed emotion on a 5-point Likert scale from 1 (“very slightly or not at all”) to 5 (“extremely”). To make this survey more relevant to video learning, we slightly modified the survey instructions by prompting participants to reflect on their emotions in relation to their experience during the video lesson. Scores for positive and negative felt emotions were calculated based on the sum of the ratings for the 10 positive and 10 negative words, respectively.

*Agent Persona Inventory (API).* The Agent Persona Instrument (API) evaluated participants’ perceptions of their social connection with the instructor, based on a previous study by [Ryu and Baylor \(2005\)](#). This survey aimed to explore whether learners could build a partnership connection (i.e., social connection) with the instructor (i.e., the human instructor vs. the various 3D agents) and which types of instructors were most effective in fostering a stronger partnership connection. The API survey comprised 23 items, categorized into four sub-scales. The first sub-scale encompassed ten items that evaluated the instructor’s ability to facilitate learning (e.g., “The instructor helps me concentrate on the presentation.”;  $\alpha = 0.94$ ). The second sub-scale consisted of four items that assessed the instructor’s credibility (e.g., “The instructor is knowledgeable.”;  $\alpha = 0.91$ ). The third sub-scale included five items measuring the instructor’s human-like character (e.g., “The instructor’s emotion feels natural.”;  $\alpha = 0.85$ ). Lastly, the fourth sub-scale encompassed five items that measured the instructor’s level of engagement (e.g., “The instructor is friendly.”;  $\alpha = 0.89$ ).

*Posttest.* The posttest included both a retention test and a transfer test, with all questions and grading rubrics accessible in the OSF repository: [https://osf.io/zy2tm/?view\\_only=9e6e67ec00434d15aa052be8d66ccbc8](https://osf.io/zy2tm/?view_only=9e6e67ec00434d15aa052be8d66ccbc8). The retention test consisted of 10 multiple-choice questions designed to evaluate students’ ability to recall the conceptual knowledge presented in the lesson, aligning with the cognitive theory of multimedia learning’s selection process ([Mayer, 2022](#)). In the retention test, participants selected the correct answer from four options per question, each reflecting the fundamental concepts of chemical bonding discussed in the video lesson. An example question from the assessment is: “What type of bonding is ionic bonding?” or “What is ionization energy?” The question is followed by four multiple-choice options: (A) Metal + Metal, (B) Non-Metal + Non-Metal, (C) Metal + Non-Metal, and (D) None of the above. The correct answer, in this case, is (C), Metal + Non-Metal, which accurately describes the nature of ionic bonding. No time limit was required for the retention test responses. Participants earned one point for each correct answer, yielding a maximum possible score of 10. Cronbach’s alpha was 0.67.

The transfer test consisted of five short-answer questions aimed at assessing students’ ability to apply their acquired knowledge in a new situation. This process necessitates the selection, organization, and integration of information to form a mental model that connects with their prior knowledge, in line with the cognitive theory of multimedia learning ([Mayer, 2022](#)). The transfer questions are shown in [Table 1](#).



**Table 1.** Transfer Test Items.

Transfer test:	<ol style="list-style-type: none"> <li>1. What can be done to either the metal or non-metal atoms to disrupt the formation of an ionic bond?</li> <li>2. To disrupt the formation of an ionic bond, what can be done to the three energies of the bond formation process, namely ionization energy, electron affinity, and plus energy?</li> <li>3. If you discovered a new type of atom that possesses extra electrons, what types of bonding could be generated with this atom? Why?</li> <li>4. If you observed a parallel universe with different rules governing the energy of bonds, where more energy in a bond result in greater stability, and discovered an ionic bond between one metal and one non-metal atom, what would happen to the sum of ionization energy, electron affinity, and plus energy? Please explain your answer.</li> <li>5. Suppose you observed a parallel universe with different chemical structures of the atoms, and discovered that non-metal atoms in this universe possess extra electrons, while metal atoms do not have enough electrons: In light of these observations, how might ionic bonding and covalent bonding work in this parallel universe (if the rules of both bonding are the same)? Please explain your answer.</li> </ol>
----------------	---

Participants had 3 minutes to finish each short answer question, and could not go back to previous items. The grading rubric for the transfer test included a set of acceptable answers for each question, with the total score derived from the total number of acceptable responses provided (with a maximum possible score of 24). Students earned one point for each acceptable answer, indicating a successful application of knowledge to a new context (i.e., indicative of deep learning). Students were not expected to generate all possible correct answers; hence, even a score of 1 on any question can be considered as an indication of successful knowledge transfer. To maintain inter-rater reliability, two independent evaluators reviewed the answers, and any scoring discrepancies were resolved through discussion to reach a consensus. The inter-rater reliability was high,  $r = 0.95$ . Cronbach's alpha was 0.61.

*Demographics Questionnaire.* Finally, the demographic questionnaire requested participants to indicate their age, gender, and race/ethnicity (i.e., White/Caucasian, Black/African American, Asian, Hispanic/Latino, Indian, and Other).

### **Procedure**

Participants were solicited through SONA, a computer-based automated participant scheduling system, with up to four participants scheduled for each session. Each participant was randomly assigned to either the original video lesson or one of the six modified versions. The study took place in a psychology laboratory equipped with five Dell or iMac desktop computers, each stationed within a visually isolated cubicle. The

experiment was conducted using Qualtrics to present all material on the computer screen.

Upon their arrival, participants were directed to a computer displaying the initial Qualtrics survey page, where they signed a consent form outlining the study's general objectives. Following this, they completed the perceived prior knowledge questionnaire. Then, participants were instructed to attentively watch one of the seven versions of the 9-minute video lesson about chemical bonds corresponding to their treatment group, with instructions that it would be followed by a test. Depending on their assigned group, they viewed either the original video lesson or one of six versions of modified video lessons. These video lessons played without the option to stop or rewind. After watching the video, participants proceeded to fill out the PANAS and API surveys, followed by the posttest. This posttest consisted of 10 multiple-choice questions assessing retention, with no time limit, and five short-answer questions evaluating transfer knowledge, with a three-minute time limit per question. Finally, participants completed a short demographic questionnaire, read a debriefing form describing the study, and were thanked for their contribution. The procedure of this study is illustrated in [Figure 3](#).

We received IRB approval and followed guidelines for research with human subjects.

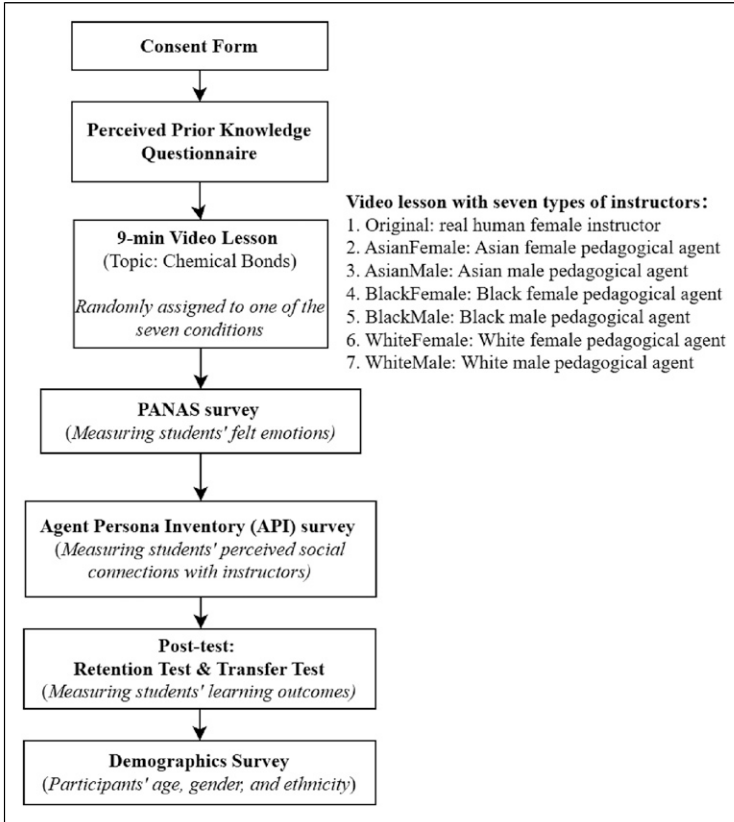
## Results

### *Did the Groups Differ on Basic Characteristics?*

As the initial analysis of participants' basic characteristics, we found no significant differences among the seven groups in terms of age,  $F(6, 222) = 1.62, p = .14$ ; prior knowledge level,  $F(6, 222) = 0.59, p = .74$ ; proportion of participants in each gender category,  $\chi^2(24) = 22.57, p = .55$ ; and proportion of participants in each race category,  $\chi^2(36) = 40.82, p = .27$ . Therefore, we can conclude that the seven groups did not differ in basic characteristics.

### *Research Question 1: Did Students Achieve Different Learning Outcomes from a Video Lesson Delivered by Pedagogical Agents Compared to a Real Human Instructor?*

According to the Media Comparison Hypothesis, people can treat computer-generated characters in the same way they treat real humans, suggesting that learning outcomes should be equivalent to lessons with human and virtual instructors. One-way ANOVAs were performed to assess differences in retention and transfer posttest scores across the seven video lesson conditions (i.e., Asian female agent, Asian male agent, Black female agent, Black male agent, White female agent, White male agent, and the original human instructor). Subsequently, a post-hoc Dunnett test was applied, as needed, to contrast each modified video lesson group (i.e., video lessons featuring pedagogical agents) against the original lesson group. [Table 2](#) presents the mean scores and standard



**Figure 3.** Experiment procedure.

deviations on the retention and transfer tests across the groups. Table 2 also includes Cohen's  $d$ s for all the comparisons between each experimental condition (i.e., video lesson presented by pedagogical agents) and the original condition (video lesson with a human instructor).

The analysis revealed no significant differences in retention test scores among the groups,  $F(6, 222) = 0.86, p = .53, \eta^2 = 0.02$ . Despite the Asian female agent group achieving the highest average retention score ( $M = 7.81, SD = 1.98$ ) and the original lesson group recording a relatively lower score ( $M = 7.06, SD = 2.15, d = 0.36$ ), the post-hoc Dunnett test indicated no significant differences between the modified lesson groups with animated agents and the original lesson group. Similarly, no significant differences were found in the transfer test scores across the groups,  $F(6, 222) = 0.38, p = .89, \eta^2 = 0.01$ . Even though the Black female agent group ( $M = 4.03, SD = 2.85$ ) obtained the highest average transfer score and the original lesson group the lowest ( $M = 3.24, SD = 2.45$ ), these differences were not statistically significant in the post-hoc Dunnett test.

**Table 2.** The Mean and Standard Deviation of the Retention and Transfer Scores by Groups.

	Retention			Transfer		
	<i>M</i>	<i>SD</i>	<i>d</i>	<i>M</i>	<i>SD</i>	<i>d</i>
AsianFemale	7.81	1.98	0.36	3.69	2.42	0.18
AsianMale	6.82	2.13	-0.11	3.61	2.59	0.15
BlackFemale	7.30	2.08	0.11	4.03	2.85	0.30
BlackMale	7.28	1.80	0.11	3.50	2.27	0.11
WhiteFemale	6.85	2.31	-0.09	3.71	2.34	0.20
WhiteMale	7.06	2.14	0.00	3.97	2.60	0.29
Original	7.06	2.15	-	3.24	2.45	-
Total	7.17	2.09	-	3.68	2.49	-

Note. The *d* indicates the effect size between each of the first six conditions and the Original condition.

The findings indicate that video lessons delivered by pedagogical agents did not produce significantly different learning outcomes in terms of retention and transfer compared to lessons delivered by a real human instructor. These results align with the Media Equation Hypothesis, which suggests that learners treat computer-generated characters as they would human instructors, resulting in equivalent educational impacts. While certain groups, such as the Asian female agent group for retention and the Black female agent group for transfer, exhibited slightly higher average scores than the original lesson group, these differences were not statistically significant. This consistency across conditions supports that pedagogical agents can serve as effective alternatives to human instructors, at least for learning outcomes assessed in this context. Moreover, the negligible effect sizes underscore the robustness of this equivalence, suggesting that pedagogical agents could provide scalable instructional solutions without compromising learning efficacy.

### **Research Question 2: Are the Learners' Learning Outcomes Affected by the Pedagogical Agents' Gender and/or Race?**

To investigate which pedagogical agent was the most effective in facilitating learning for students, we concentrated on data from the six groups exposed to the modified video lessons. We applied 2 (gender: female vs. male) x 3 (race: Asian, Black, and White) ANOVAs to explore the main effects and interactions of the agents' gender and racial types on students' outcomes in both retention and transfer tests.

The analysis revealed no significant main effects for the gender,  $F(1, 190) = 0.82$ ,  $p = .37$ ,  $\eta^2 = 0.004$ , or race of the pedagogical agents,  $F(2, 190) = 0.61$ ,  $p = .55$ ,  $\eta^2 = 0.01$ , as well as no significant interaction effects between these two factors on retention test scores,  $F(2, 190) = 1.54$ ,  $p = .22$ . Similarly, for transfer test scores, there were no significant main effects for gender,  $F(1, 190) = 0.11$ ,  $p = .75$ ,  $\eta^2 = 0.001$ , or race of the pedagogical agents,  $F(2, 190) = 0.10$ ,  $p = .91$ ,  $\eta^2 = 0.001$ , and

no significant interaction effects were observed between these factors,  $F(2, 190) = 0.41$ ,  $p = .67$ ,  $\eta^2 = 0.004$ .

In conclusion, these findings demonstrated that the learning performance of students, in terms of retention and transfer, was not influenced by agents of different genders or races, nor by the interaction between these factors. This suggests that the effectiveness of pedagogical agents in facilitating learning is consistent across diverse gender and racial representations. These results underscore the potential for using diverse pedagogical agents in educational settings without concerns about adverse impacts on learning outcomes, thereby supporting inclusive representation in instructional design.

### *Research Question 3: Did Pedagogical Agents Lead to Different Felt Emotions of the Learners Compared with a Real Human Instructor?*

Using one-way ANOVAs, we analyzed participants' ratings of felt positive and negative emotions experienced during learning across the seven video lesson conditions (i.e., Asian female agent, Asian male agent, Black female agent, Black male agent, White female agent, White male agent, and the original human instructor). To further examine differences, we employed post-hoc Dunnett tests to compare each modified video lesson featuring pedagogical agents against the original lesson group. Table 3 displays the mean and standard deviations for the ratings of positive and negative emotions across these groups. Table 3 also consists of Cohen's  $d$ s for all the comparisons between each experimental condition (i.e., video lesson presented by pedagogical agents) and the original condition (video lesson with a human instructor).

The analysis indicated no significant differences in the ratings of positive emotions across the groups,  $F(6, 222) = 1.87$ ,  $p = .09$ ,  $\eta^2 = 0.05$ , and similarly, no significant

**Table 3.** The Mean and Standard Deviation of the Ratings of Felt Positive and Negative Emotion by Groups.

	Positive Emotion			Negative Emotion		
	<i>M</i>	<i>SD</i>	<i>d</i>	<i>M</i>	<i>SD</i>	<i>d</i>
AsianFemale	2.54	0.86	0.18	1.29	0.41	0.10
AsianMale	2.04	0.70	-0.48	1.33	0.52	0.18
BlackFemale	2.32	0.92	-0.08	1.26	0.33	0.03
BlackMale	2.18	0.72	-0.28	1.13	0.19	-0.41
WhiteFemale	2.22	0.78	-0.22	1.31	0.48	0.14
WhiteMale	2.03	0.65	-0.50	1.29	0.44	0.10
Original	2.39	0.77	-	1.25	0.37	-
Total	2.25	0.78	-	1.27	0.40	-

Note. The  $d$  indicates the effect size between each of the first six conditions and the Original condition.

differences were found in the ratings of negative emotions,  $F(6, 222) = 0.87, p = .52, \eta^2 = 0.02$ , across groups. The post-hoc Dunnett tests also failed to reveal any significant differences in the ratings of both positive and negative emotions between the original lesson group and each of the six modified video lesson groups with pedagogical agents.

In summary, the findings demonstrate that the presence of pedagogical agents did not lead to a significantly different effect on the learners' emotional experiences compared to a real human instructor. This outcome aligns with the Media Equation Hypothesis, suggesting that participants respond similarly to video lessons irrespective of whether they are taught by pedagogical agents or human instructors. Despite slight variations in mean ratings across conditions, these differences were not statistically significant, suggesting that the emotional impact of the instructional medium remains comparable. These findings provide further evidence that pedagogical agents can effectively replicate the emotional feelings during learning typically associated with human instructors, supporting their viability as scalable, emotionally neutral alternatives in educational settings.

#### ***Research Question 4: Are the Learners' Felt Emotions Affected by Pedagogical Agents' Gender and Race?***

We employed 2 (Gender: Female vs. Male) x 3 (Race: Asian, Black, White) ANOVAs to explore the impact of the gender and race of pedagogical agents on learners' felt emotions. The analysis for positive emotion ratings revealed a significant main effect for gender,  $F(1, 190) = 6.28, p = .01, \eta^2 = 0.03$ , suggesting that learners experienced more positive emotions with female pedagogical agents ( $M = 2.36, SD = 0.86$ ) compared to male pedagogical agents ( $M = 2.08, SD = 0.68, d = 0.36$ ). However, no significant main effect was found for race,  $F(2, 190) = 0.83, p = .44, \eta^2 = 0.01$ , nor was there an interaction effect between gender and race,  $F(2, 190) = 0.97, p = .38, \eta^2 = 0.01$ . Moreover, regarding negative felt emotion ratings, there were no significant main effects for gender,  $F(1, 190) = 0.45, p = .51, \eta^2 = 0.002$ , or race,  $F(2, 190) = 1.56, p = .22, \eta^2 = 0.02$ , and no significant interaction effect between the two factors,  $F(2, 190) = 0.72, p = .45, \eta^2 = 0.01$ .

In summary, the findings indicate that while the variations in gender and race of the pedagogical agents did not significantly influence the students' negative emotions, female pedagogical agents, regardless of their race, led to more positive felt emotions for learners, compared with male pedagogical agents. This pattern may reflect the influence of gender stereotypes, as Social Role Theory suggests that women are often perceived as more nurturing and emotionally expressive, traits that could foster more positive emotional responses from learners (Eagly & Wood, 2012; Kierstead et al., 1988). However, while gender stereotypes appear to influence affective experiences, the results suggest that racial biases may be less pronounced in this particular context, which contrasts with findings from prior research (Basow et al., 2013; Reid, 2010). The absence of significant effects for race and the interaction between gender and race

underscores the intricate interplay between race and gender in shaping learners' emotional responses.

### **Research Question 5: Are the Learners' Social Connections with the Instructor Different for Pedagogical Agents Compared with the Real Human Instructor?**

We conducted one-way ANOVAs to determine whether there were differences in the learners' perceived social connections with the instructor when comparing modified video lesson groups featuring various pedagogical agents to the original video lesson group with a real human instructor. These analyses evaluated learners' ratings on the four subscales of the Agent Persona Instrument (API)—facilitating learning, credibility, human-likeness, and engagement—across the seven groups. When necessary, post-hoc LSD tests were performed to compare each pair of the groups. Table 4 displays the mean and standard deviations for the ratings of the learners' perceived social connection across these groups. Table 4 also reports Cohen's  $d$  for each comparison between each experimental condition (i.e., video lesson presented by pedagogical agents) and the original condition (video lesson with a human instructor).

Although our analysis did not reveal any significant differences in learners' perceptions of how well the video lessons facilitated learning across the various groups,  $F(6, 222) = 1.53, p = .17, \eta^2 = 0.04$ , significant differences were observed in the credibility ratings among these groups,  $F(6, 222) = 3.12, p = .01, \eta^2 = 0.08$ . Specifically, learners in the original lesson group ( $M = 5.79, SD = 0.88$ ) rated their instructor as

**Table 4.** The Mean and Standard Deviation of the Ratings of Four Subscales of API Survey (Perceived Social Connection) by Groups.

	Facilitative			Credible			Human-Like			Engaging		
	<i>M</i>	<i>SD</i>	<i>d</i>	<i>M</i>	<i>SD</i>	<i>d</i>	<i>M</i>	<i>SD</i>	<i>d</i>	<i>M</i>	<i>SD</i>	<i>d</i>
AF	4.70	1.33	0.17	5.71	0.85	-0.09	4.63*	1.25	-0.78	4.67	1.40	-0.17
AM	4.05	1.13	-0.33	5.15*	1.07	-0.65	4.46*	1.39	-0.87	3.97*	1.40	-0.68
BF	4.31	1.58	-0.11	5.22*	1.47	-0.47	4.26*	1.31	-1.07	3.95*	1.51	-0.67
BM	4.17	1.07	-0.24	5.07*	0.91	-0.80	4.23*	1.30	-1.10	4.08*	1.02	-0.69
WF	4.25	1.19	-0.17	5.26*	1.20	-0.50	4.41*	1.37	-0.91	4.15*	1.35	-0.56
WM	3.84	1.22	-0.48	4.91*	0.81	-1.04	3.74*	1.35	-1.49	3.74*	1.26	-0.89
O	4.47	1.41	-	<b>5.79</b>	0.88	-	<b>5.52</b>	1.03	-	<b>4.91</b>	1.36	-
Total	4.26	1.29	-	5.30	1.08	-	4.47	1.37	-	4.21	1.38	-

Note. The  $d$  indicates the effect size between each of the first six conditions and the Original condition. Asterisk(\*) represents a significant difference from the **bolded** condition. AF indicates the Asian Female group, AM indicates the Asian Male group, BF indicates the Black Female group, BM indicates the Black Male group, WF indicates the White Female group, WM indicates the White Male group, and O indicates the original group.

significantly more credible compared to the Asian male agent ( $M = 5.15$ ,  $SD = 1.07$ ,  $p = .02$ ,  $d = -0.65$ ), Black female agent ( $M = 5.22$ ,  $SD = 1.47$ ,  $p = .03$ ,  $d = -0.47$ ), Black male agent ( $M = 5.07$ ,  $SD = 0.91$ ,  $p = .01$ ,  $d = -0.80$ ), White female agent ( $M = 5.26$ ,  $SD = 1.20$ ,  $p = .04$ ,  $d = -0.50$ ), and White male agent ( $M = 3.91$ ,  $SD = 0.81$ ,  $p < .001$ ,  $d = -1.04$ ) groups. However, there was no significant difference in the credibility ratings between the original lesson group and the Asian female agent group ( $M = 5.71$ ,  $SD = 0.85$ ,  $p = .77$ ,  $d = -0.09$ ). Additionally, the Asian female agent was rated as significantly more credible than the Asian male ( $p = .03$ ), Black male ( $p = .02$ ), and White male agents ( $p = .003$ ). While the Asian female agent group also gave higher credibility ratings compared to the Black female and White female agent groups, these differences were not statistically significant.

Significant differences were found in the ratings of human-likeness among the seven groups,  $F(6, 222) = 5.76$ ,  $p < .001$ ,  $\eta^2 = 0.14$ . The original lesson, featuring a real human instructor, received the highest ratings for human-likeness ( $M = 5.52$ ,  $SD = 1.03$ ), outperforming each of the six modified video lessons with pedagogical agents. Specifically, the Asian female agent ( $M = 4.63$ ,  $SD = 1.25$ ,  $p = .01$ ,  $d = -0.78$ ), the Asian male agent ( $M = 4.46$ ,  $SD = 1.39$ ,  $p = .001$ ,  $d = -0.87$ ), the Black female agent ( $M = 4.26$ ,  $SD = 1.31$ ,  $p < .001$ ,  $d = -1.07$ ), the Black male agent ( $M = 4.23$ ,  $SD = 1.30$ ,  $p < .001$ ,  $d = -1.10$ ), the White female agent ( $M = 4.41$ ,  $SD = 1.37$ ,  $p < .001$ ,  $d = -0.91$ ), and the White male agent ( $M = 3.74$ ,  $SD = 1.35$ ,  $p < .001$ ,  $d = -1.49$ ), all received lower ratings in comparison to the original group with a human instructor. Moreover, we also found that the Asian female agent was perceived as more human-like compared to both the Asian male agent ( $p = .04$ ) and the White male agent ( $p = .01$ ). No significant differences in human-like ratings were observed among the other pairs.

For the ratings of how engaging the instructor was, there was a significant difference among the groups,  $F(6, 222) = 3.22$ ,  $p = .01$ ,  $\eta^2 = 0.08$ . The original lesson, featuring a real human instructor, achieved significantly higher ratings for being engaging ( $M = 4.91$ ,  $SD = 1.36$ ) compared to the Asian male agent ( $M = 3.97$ ,  $SD = 1.40$ ,  $p = .02$ ,  $d = -0.68$ ), Black female agent ( $M = 3.95$ ,  $SD = 1.51$ ,  $p = .03$ ,  $d = -0.67$ ), Black male agent ( $M = 4.08$ ,  $SD = 1.02$ ,  $p = .01$ ,  $d = -0.69$ ), White female agent ( $M = 4.15$ ,  $SD = 1.35$ ,  $p = .04$ ,  $d = -0.56$ ), and White male agent ( $M = 3.74$ ,  $SD = 1.26$ ,  $p < .001$ ,  $d = -0.89$ ). However, while the ratings of being engaging for the Asian female agent group ( $M = 4.67$ ,  $SD = 1.40$ ,  $p = .77$ ,  $d = -0.17$ ) were lower than the original lesson group, the difference was not significant. Additionally, the Asian female agent was rated as significantly more engaging compared to the Asian male agent ( $p = .03$ ), Black male agent ( $p = .02$ ), and White male agent ( $p = .003$ ). Among the other pairs, no significant differences in engagement ratings were observed.

In conclusion, while students perceived the pedagogical agents in the modified lessons as equally capable of facilitating learning as the real human instructor in the original lesson—consistent with the Media Equation Hypothesis—they consistently rated the human instructor higher in credibility, human-likeness, and engagement.



These findings suggest that learners still tend to build a stronger social connection with a human instructor than with a virtual instructor. Notably, the Asian female agent received the highest ratings for credibility, human-likeness, and engagement among all the pedagogical agent groups. This finding suggests that the perceived social connection with the Asian female agent was less negatively impacted compared to other agents, which may be attributed to the interplay of cultural and gender-based stereotypes (Eagly & Wood, 2012; Kierstead et al., 1988).

### ***Research Question 6: Are the Learners' Social Connections with the Instructor Affected by the Pedagogical Agents' Gender and Race?***

To explore the effect of pedagogical agents' gender and race on learners' perceptions of social connection with their instructor, we restricted our analysis to data from the six modified video lesson groups. We applied 2 (Gender: Female vs. Male) x 3 (Race: Asian, Black, White) ANOVAs to determine the interactive effect of different gender and racial types of pedagogical agents on the learners' perceived social connections. These social connection ratings were assessed by four subscales of the Agent Persona Instrument (API), namely facilitating learning, credibility, human-likeness, and being engaging.

The analysis revealed significant main effects for the influence of the pedagogical agents' gender on ratings of facilitating learning,  $F(1, 190) = 4.93, p = .03, \eta^2 = 0.03$ , and ratings of credibility,  $F(1, 190) = 5.17, p = .02, \eta^2 = 0.03$ . Specifically, female agents were rated as significantly more facilitative of learning ( $M = 4.42, SD = 1.38$ ) than male agents ( $M = 4.02, SD = 1.13, d = 0.32$ ). Additionally, female agents were considered significantly more credible ( $M = 5.39, SD = 1.21$ ) compared to their male counterparts ( $M = 5.05, SD = 0.93, d = 0.31$ ). However, no significant main effect was found for the race on facilitating learning ratings,  $F(2, 190) = 1.15, p = .32, \eta^2 = 0.01$ , or credibility ratings,  $F(2, 190) = 1.91, p = .15, \eta^2 = 0.02$ . In addition, there were no significant interaction effects between the agents' gender and race on facilitating learning ratings,  $F(2, 190) = 0.67, p = .52, \eta^2 = 0.01$ , nor on credibility ratings,  $F(2, 190) = 0.59, p = .56, \eta^2 = 0.01$ .

Moreover, for the ratings of being engaging, although there is a marginally significant main effect for gender,  $F(1, 190) = 3.00, p = .09, \eta^2 = 0.02$ , showing that female agents ( $M = 4.25, SD = 1.44$ ) were rated to be more engaging than the male agents ( $M = 3.93, SD = 1.23, d = 0.25$ ), there was no significant main effect for race,  $F(2, 190) = 1.44, p = .24, \eta^2 = 0.02$ , and no significant interaction between the two factors,  $F(2, 190) = 1.60, p = .20, \eta^2 = 0.02$ . Similarly, the analysis did not reveal any significant main effects for gender,  $F(1, 190) = 2.32, p = .13, \eta^2 = 0.01$ , or race,  $F(2, 190) = 2.14, p = .12, \eta^2 = 0.02$ , and no significant interaction effect on ratings of human-likeness,  $F(2, 190) = 1.04, p = .35, \eta^2 = 0.01$ .

In summary, the findings suggest that learners' social connections with their instructors were affected by varying the gender of the instructor, with female instructors rated as more facilitative of learning, more credible, and more engaging

than their male counterparts. This aligns with Social Role Theory (Eagly & Wood, 2012), which suggests that societal expectations cast women as nurturing and approachable, traits that enhance evaluations of teaching effectiveness (Kierstead et al., 1988). In contrast, learners' social connections with their instructors were not affected by the varying racial identities of pedagogical agents, in contrast to previous findings that highlighted racial biases in instructor evaluations (Basow et al., 2013; Reid, 2010).

### *Exploratory Analysis 1: Was There a Gender-Matching Effect?*

In this exploratory analysis, we aimed to investigate whether there was a gender-matching effect on various dependent variables. To this end, we excluded three participants who identified their gender as "Other" or preferred not to disclose their gender and then conducted a 2 (student gender: female vs. male) x 2 (pedagogical agent gender: female vs. male) ANOVA to explore the interaction effects on students' felt emotion ratings, social connection ratings, and learning outcome scores. We viewed an interaction as indicating the potential of a gender-matching effect.

For ratings of felt emotion, we did not find any significant interaction effects between instructors' gender and students' gender on students' felt positive emotion,  $F(1, 189) = 0.37, p = .54, \eta^2 = 0.002$ , or negative felt emotion,  $F(1, 189) = 0.25, p = .62, \eta^2 = 0.001$ . For the ratings of social connection, there was also no significant interaction effect on facilitating learning,  $F(1, 189) = 1.83, p = .18, \eta^2 = 0.01$ , credibility,  $F(1, 189) = 1.22, p = .27, \eta^2 = 0.01$ , human-like,  $F(1, 189) = 2.13, p = .15, \eta^2 = 0.01$ , and engaging,  $F(1, 189) = 0.16, p = .69, \eta^2 = 0.001$ . Additionally, we also found no significant interaction effect on retention scores,  $F(1, 189) = 0.31, p = .58, \eta^2 = 0.002$ , or on transfer scores,  $F(1, 189) = 0.07, p = .79, \eta^2 = 0.000$ .

In conclusion, our findings did not indicate evidence for a gender-matching effect on students' learning outcomes, emotional responses, and perceived social connections with instructors. Specifically, students did not respond differently when they received a lesson presented by a pedagogical agent of the same versus different gender as themselves. The lack of significant interaction effects suggests that gender congruence between students and pedagogical agents does not play a meaningful role in shaping learning experiences or outcomes in this context. This finding underscores the robustness of pedagogical agents in supporting diverse learner populations without requiring gender alignment to achieve similar educational outcomes.

### *Exploratory Analysis 2: Was There a Race-Matching Effect?*

In this exploratory analysis, we aimed to explore whether there was a race-matching effect on various dependent variables. Due to the insufficient recruitment of Black participants ( $N = 5$ ), we categorized participants into two groups based on their racial identities: Asian students ( $N = 56$ ) and Non-Asian students ( $N = 140$ ), as well as White students ( $N = 42$ ) and Non-White students ( $N = 154$ ). Consequently, we conducted two

separate 2 (student race: Asian vs. Non-Asian, White vs. Non-White) x 3 (pedagogical agent race: Asian, Black, White) ANOVAs. These analyses aimed to investigate whether students of different racial types respond differently to pedagogical agents of varying racial types, and whether there is a preference for pedagogical agents of their own race.

**Asian Versus Non-Asian Students.** For comparing the race-matching effect between Asian and Non-Asian students, we did not find a significant interaction effect between students' race and pedagogical agents' race on students' felt positive emotion,  $F(2, 190) = 0.19, p = .83, \eta^2 = 0.002$ , felt negative emotion,  $F(2, 190) = 0.12, p = .88, \eta^2 = 0.001$ , facilitating learning,  $F(2, 190) = 0.26, p = .77, \eta^2 = 0.003$ , human-likeness,  $F(2, 190) = 1.22, p = .30, \eta^2 = 0.01$ , being engagement,  $F(2, 190) = 1.30, p = .28, \eta^2 = 0.01$ , retention score,  $F(2, 190) = 0.45, p = .64, \eta^2 = 0.005$ , or transfer score,  $F(2, 190) = 0.41, p = .67, \eta^2 = 0.004$ .

However, we did find one marginally significant interaction between students' race and pedagogical agents' race on the ratings of being credible,  $F(2, 190) = 3.00, p = .05, \eta^2 = 0.03$ . To be more specific, the Non-Asian students did not reveal any significant differences across the race of the virtual instructors (Asian, Black, and White) ( $F(2, 137) = 0.52, p = .59, \eta^2 = 0.01$ ). This suggests that the race of the pedagogical agents did not significantly influence the ratings of Non-Asian participants. Conversely, a one-way ANOVA for Asian participants suggested a significant effect,  $F(2, 53) = 4.66, p = .01, \eta^2 = 0.15$ . Given the significant result, post-hoc LSD pairwise comparisons were conducted. These revealed that Asian pedagogical agents ( $M = 5.57, SD = 0.91$ ) were rated significantly higher than Black agents ( $M = 4.53, SD = 1.37, p = .004, d = 0.89$ ) and that the Asian agents were rated higher, although not significantly, than White agents ( $M = 4.93, SD = 0.87, p = .06, d = 0.72$ ). No significant differences were found between Black ( $M = 4.53, SD = 1.37$ ) and White pedagogical agents ( $M = 4.93, SD = 0.87, p = .26, d = 0.35$ ).

In summary, although we did not find a race-matching effect between Asian and Non-Asian students on felt emotion, facilitative learning, human-like, engaging, and learning outcomes, the results partially revealed the presence of a racial-matching effect among Asian participants on credibility ratings, reflecting a preference for pedagogical agents that share their racial type.

**White Versus Non-White Students.** For comparing the race-matching effect between White and Non-White students, there was no significant interaction effect between students' race and pedagogical agents' race on students' felt negative emotion,  $F(2, 190) = 0.27, p = .76, \eta^2 = 0.003$ , facilitating learning,  $F(2, 190) = 2.23, p = .11, \eta^2 = 0.02$ , credibility,  $F(2, 190) = 1.54, p = .22, \eta^2 = 0.02$ , human-likeness,  $F(2, 190) = 1.07, p = .35, \eta^2 = 0.01$ , retention score,  $F(2, 190) = 0.01, p = .99, \eta^2 = 0.000$ , and transfer score,  $F(2, 190) = 0.09, p = .91, \eta^2 = 0.001$ .

However, we identified a significant interaction between the students' racial backgrounds and the racial types of the pedagogical agents affecting the ratings of felt

positive emotion,  $F(2, 190) = 3.00, p = .05, \eta^2 = 0.03$ . Although the one-way ANOVA tests did not reveal significant effects across the racial types (Asian, Black, and White) of the pedagogical agents for both Non-White students,  $F(2, 151) = 2.32, p = .10, \eta^2 = 0.03$ , and White students,  $F(2, 39) = 0.82, p = .45, \eta^2 = 0.04$ , there was a noticeable trend. White students rated their positive emotions highest when engaging with a White virtual agent instructor ( $M = 2.64, SD = 0.80$ ), compared to Black ( $M = 2.54, SD = 1.06, d = 0.11$ ) and Asian ( $M = 2.17, SD = 0.82, d = 0.58$ ) pedagogical agents. Conversely, Non-White students reported the lowest positive emotion ratings in response to White pedagogical agents ( $M = 2.01, SD = 0.65$ ), relative to Black ( $M = 2.13, SD = 0.68, d = 0.18$ ) and Asian ( $M = 2.31, SD = 0.82, d = 0.41$ ) pedagogical agents.

Moreover, there was also a significant interaction between the students' racial backgrounds and the racial types of the pedagogical agents affecting the ratings of engagement,  $F(2, 190) = 4.09, p = .02, \eta^2 = 0.04$ . To be more specific, the further one-way ANOVA conducted for Non-White participants revealed a statistically significant difference in ratings of engagement based on the pedagogical agents' racial types,  $F(2, 151) = 3.99, p = .02, \eta^2 = 0.03$ . The subsequent pairwise comparisons showed that the Non-White students rated the Asian agent instructors ( $M = 4.42, SD = 1.33$ ) to be more engaging than Black agent instructors ( $M = 3.72, SD = 1.24, p = .01, d = 0.54$ ) and White agent instructors ( $M = 3.89, SD = 1.33, p = .04, d = 0.40$ ), while there was no significant difference on ratings of engaging between Black and White agents ( $p = .50$ ). Conversely, for White students, the one-way ANOVA did not indicate a statistically significant overall effect across different racial types of pedagogical agents,  $F(2, 39) = 1.60, p = .22, \eta^2 = 0.04$ . However, the pairwise comparisons identified a trend showing that the White students gave higher ratings of engaging for the Black agent instructors ( $M = 4.74, SD = 1.11$ ) compared with White agent instructors ( $M = 4.23, SD = 1.27, d = 0.43$ ) and Asian agent instructors ( $M = 3.82, SD = 1.87, d = 0.60$ ).

In summary, these findings indicate that the race of pedagogical agents differentially impacted the positive emotions experienced by Non-White and White students. Although White students rated Black pedagogical agents as more engaging, contradicting the race-matching hypothesis, they generally reported feeling most positive when taking a lesson with White pedagogical agents.

Overall, the findings for both White versus Non-White and Asian versus Non-Asian students offer some evidence that the race of pedagogical agents differentially impacts the positive emotions experienced by students from different racial backgrounds. Notably, there is evidence suggesting that students tend to prefer pedagogical agents who share their racial identity, which partially supports the Matching Hypothesis and Intergru Bias Theory.

### ***Exploratory Analysis 3: Pedagogical Agents Versus Human Instructor***

Finally, we consolidated the data from all six pedagogical agent groups into a single category labeled *Pedagogical Agent* group. We then conducted a *t* test to compare students' felt emotions, perceived social connections with instructors, and learning

outcomes between this *Pedagogical Agent* group and the *Original Lesson* group. This analysis aimed to further explore the overall differences between video lessons featuring pedagogical agents and those with human instructors, thereby providing additional testing of the media equation hypothesis.

The results indicated no significant differences between the pedagogical agents' groups and the original lesson group in terms of students' felt positive emotion,  $t(227) = -1.18, p = .24, d = -0.22$ ; or felt negative emotion,  $t(227) = 0.34, p = .74, d = -0.06$ ). Similarly, no significant differences were observed in students' learning outcomes between the two groups on retention score,  $t(227) = 0.31, p = .76, d = 0.06$ , or transfer score:  $t(227) = 1.09, p = .28, d = 0.20$ .

However, in terms of perceived social connections with instructors, although there was no significant difference in the ratings of facilitative learning between the two groups,  $t(227) = -1.02, p = .31, d = -0.19$ , there were significant differences in ratings of credibility,  $t(227) = -2.83, p = .01$ , human-likeness,  $t(227) = -5.01, p < .001$ , and engagement,  $t(227) = -3.21, p = .002$ . To be more specific, the human instructor in the original lesson was rated to be more credible ( $M = 5.79, SD = 0.88, d = 0.53$ ), more human-like ( $M = 5.52, SD = 1.03, d = 0.94$ ), and more engaging ( $M = 4.91, SD = 1.36, d = 0.60$ ), than the pedagogical agents in the modified video lessons ( $M = 5.22, 4.29, 4.09, SD = 1.09, 1.34, 1.35$ , respectively).

Overall, the results indicate that pedagogical agents were generally perceived as less credible, less human-like, and less engaging than the real human instructor. This suggests that inconsistent with the Media Equation Hypothesis, students tend to form stronger social connections with real human instructors than with pedagogical agents, while such preference did not appear to impact their felt emotions or learning outcomes.

## Discussion

### *Empirical Implications*

The findings of this study offer insights into the impact of pedagogical agents with varying race and gender types on students' learning experiences (i.e., affective and social processes) and learning outcomes (cognitive processes), in comparison to a real human instructor. These insights are framed within the Cognitive-Affective Model of E-learning, addressing cognitive, affective, and social processing.

*Learning Outcomes: Cognitive Processing.* First, regarding the cognitive processing and learning outcomes, the results demonstrated no significant differences in retention and transfer posttest scores between students who learned from pedagogical agents and those who learned from a real human instructor. This finding is consistent with the Media Equation Hypothesis (Reeves & Nass, 1996), suggesting that pedagogical agents can facilitate learning as effectively as real instructors. However, the negligible differences in learning outcomes align with previous research (e.g., Zhao & Mayer, 2023a) suggesting that computer-generated instructors (i.e., machine voice narrators)

may not inherently outperform human instructors in promoting cognitive learning. This underscores the need for further exploration of pedagogical agent design to achieve meaningful improvements in cognitive processing.

Furthermore, contrary to the expectations derived from Social Role Theory (Eagly & Wood, 2012), there was no significant impact of the gender or race of the pedagogical agents on students' retention and transfer learning outcomes, indicating that the agents' race and gender did not influence students' cognitive processes. These findings do not align with the previous studies that suggest gender and race stereotypes in educational settings (Basow et al., 2013; Campbell, 2023; Kierstead et al., 1988; Reid, 2010; Renström et al., 2021). For example, Basow et al. (2013) found that Black and female instructors consistently received lower evaluations compared to their White and male counterparts, which in turn impacted students' academic performance, as students performed better when their instructors aligned with societal expectations. Such discrepancies may stem from the controlled, brief nature of the lessons in this study, suggesting that real-world settings or longer exposure to pedagogical agents might yield different results.

*Learning Experiences: Affective Processes.* Second, in terms of affective processing, the study found no significant differences in positive or negative emotions between students who learned from various types of pedagogical agents and those taught by a real human instructor. This finding is in line with the Media Equation Hypothesis (Reeves & Nass, 1996). Furthermore, female pedagogical agents elicited more positive emotions compared to male agents, regardless of race. This finding aligns with Social Role Theory (Eagly & Wood, 2012) as well as previous research on gender stereotypes (e.g., Anderson, 2010; Basow et al., 2013; Renström and colleagues, 2021) suggesting that female instructors are often perceived as more nurturing, leading to more positive emotional experiences for students. However, race did not significantly influence emotional responses, nor did the interaction between gender and race, indicating that the affective processing during multimedia learning was largely unaffected by these factors. The absence of significant racial effects on students' felt emotions might highlight the diminished salience of race in digital education environments compared to face-to-face education settings. Notably, there was also no significant effect on negative emotions, likely due to the emotional neutrality of the video lessons.

*Learning Experiences: Social Processes.* Third, in terms of social processing, while both pedagogical agents and human instructors were rated similarly for facilitating learning, the real human instructor was perceived as more credible, human-like, and engaging. This finding challenges the Media Equation Hypothesis by showing that real human instructors may still have an advantage in fostering stronger social connections with students, which indicates a gap in the design of pedagogical agents to match the social rapport of human instructors. Among the different race and gender types of pedagogical agents, female agents received higher ratings for facilitative learning, credibility, and

engagement than male agents, indicating that students tended to build stronger social connections with female agents. This finding echoes findings on gender stereotypes favoring women in teaching roles (Renström et al., 2021), suggesting that gender stereotypes may influence social dynamics in learning, with female instructors perceived as more approachable and engaging. However, race stereotypes were not supported, as there was no significant main effect of agent race or interaction between agents' gender and race. Overall, these findings suggest that gender, more than race, shapes social perceptions in virtual learning environments.

*Exploratory Insights on Matching Effects.* Additionally, exploratory analyses examined the potential effects of gender and race-matching, alongside a comparison of the overall effectiveness of pedagogical agents versus a human instructor. For the exploratory analysis investigating gender-matching effects, the results revealed no significant interactions between students' gender and the gender of the pedagogical agents across emotional responses, social connections, or learning outcomes. This suggests that gender congruence did not significantly impact students' experiences or performance, contrasting with prior research highlighting the gender-matching effect in educational settings (e.g., Makransky et al., 2019; Zhao & Mayer, 2023a). A possible explanation for this discrepancy is that the female pedagogical agents in this study may have been designed with features that made them appear more authentic or approachable compared to the male agents. These design characteristics could have led students to prefer the female agents, irrespective of gender congruence, thereby diminishing any potential gender-matching effect.

The exploratory analysis of the race-matching effect provided limited evidence. The results indicate that the race of pedagogical agents had a differential impact on the positive emotions experienced by Non-White and White students. Notably, White students generally reported feeling most positive when taught by White pedagogical agents, which aligns with research on intergroup bias (Taylor et al., 1978). However, they found Black agents to be more engaging, which contradicts the Matching Hypothesis (Kalick & Hamilton, 1996), while Asian students rated Black agents as less credible, which aligns with the previous findings of the racial stereotypes of minority groups (Basow et al., 2013). Overall, the findings did not consistently support the Matching Hypothesis, underscoring the complexity of the race-matching effect. Specifically, while some students may naturally gravitate toward instructors of similar racial backgrounds, others appear to be more influenced by stereotypes or contextual cues. A potential reason for the limited evidence supporting the race-matching effect could be the straightforward nature of the survey questions used to rate instructors, which may have prompted participants to provide biased responses shaped by social desirability or cultural norms rather than genuine evaluations of the pedagogical agents. These findings highlight the need for further research to untangle the interplay of race, gender, and intergroup biases in shaping learner's learning experiences and outcomes with pedagogical agents.

## *Theoretical Implications*

From a theoretical perspective, this study makes contributions by extending the application of key frameworks such as the Media Equation Theory (Reeves & Nass, 1996), the Social Role Theory (Eagly & Wood, 2012), the Alliance Hypothesis (Taylor et al., 1978), and the Matching Hypothesis (Kalick & Hamilton, 1996) to pedagogical agents.

The Media Equation Theory posits that learners interact with virtual agents in ways similar to how they interact with real humans. While our findings partly support this theory—showing that pedagogical agents were equally effective as human instructors in terms of cognitive processes (i.e., learning outcomes) and emotional processes (i.e., felt emotions)—the real human instructor was perceived as more credible, and engaging, suggesting that certain aspects of social connection may still be better facilitated by human instructors. This challenges the assumption that virtual agents can fully replicate the social rapport built by real human instructors, emphasizing the need for further research into how pedagogical agents can be optimized to enhance social connections.

In addition, Social Role Theory (Eagly & Wood, 2012) and previous studies of race stereotypes posit that students perceive and evaluate their instructors based on gender and race (Basow et al., 2013; Eagly & Wood, 2012; Kierstead et al., 1988; Reid, 2010). Specifically, female instructors are often perceived as more nurturing and emotionally supportive than male instructors, while male instructors are associated with competence and authority (Eagly & Wood, 2012; Kierstead et al., 1988; Renström et al., 2021). Similarly, racial stereotypes can influence perceptions of credibility and competence, with minority instructors, such as Black instructors, often evaluated less favorably than their White counterparts (Basow et al., 2013; Campbell, 2023; Reid, 2010). These biases can extend to pedagogical agents, as seen in this study, where female agents were rated more favorably than male agents in terms of positive emotions and perceived social connections, reflecting gendered expectations. However, the lack of consistent race effects suggests that racial stereotypes may manifest differently in virtual settings compared to face-to-face interactions, potentially moderated by the digital nature of pedagogical agents. This underscores the importance of designing agents that challenge rather than reinforce societal biases, creating more inclusive and equitable learning environments.

Furthermore, the Alliance Hypothesis (Taylor et al., 1978) and Matching Hypothesis (Kalick & Hamilton, 1996) suggest that learners are likely to favor instructors who share their racial or gender characteristics, raising critical questions about whether these preferences are equally applicable to virtual instructors. Our findings challenge these assumptions, showing limited evidence for the race and gender-matching effects, suggesting that pedagogical agents do not consistently benefit from shared characteristics in the way human instructors might. Notably, some findings (e.g., Asian students rating Black agents as less credible) align with the influence of racial stereotypes, suggesting that implicit biases may also shape learner preferences. This



underscores the need for a nuanced interpretation of intergroup dynamics, suggesting that contextual factors, such as the nature of learning tasks and the design of survey instruments, might impact these effects.

Overall, this study not only sheds light on the complexities of applying these theories to pedagogical agents with different racial and gender characteristics but also emphasizes the importance of designing more effective and inclusive educational technologies. The findings related to the Media Equation Theory provide a foundation for understanding the equivalent effect of affective processes (i.e., felt emotions) and cognitive processes (i.e., learning outcomes) between human and virtual instructors. However, the observed influence of race and gender stereotypes on students' evaluations of various pedagogical agents suggests that learners' social and emotional responses to agents might be influenced by implicit biases on gender. Meanwhile, the limited support for the Alliance Hypothesis and Matching Hypothesis points to the need for a more nuanced understanding of how intergroup dynamics operate in virtual learning environments. These findings underscore the importance of designing pedagogical agents that account for societal biases and stereotypes to foster a more inclusive and equitable learning environment, thereby better supporting students' affective, social, and cognitive processing during learning. Future research should focus on enhancing the human-like qualities of pedagogical agents and exploring how to mitigate the influence of stereotypes, ultimately bridging the gap between theoretical predictions and practical applications.

### *Practical Implications*

The findings suggest that while pedagogical agents can deliver similar learning outcomes as real human instructors, there are still key areas where human instructors outperform virtual agents, especially in fostering social connections. This highlights the need for further refinement in the design of pedagogical agents to enhance their credibility, human-likeness, and engagement, thereby bridging the gap between virtual agents and human instructors.

Additionally, the study reveals that students respond more positively to female pedagogical agents, which indicates the importance of considering gender dynamics when designing virtual instructors in multimedia lessons. Based on the results of the study, it would be recommended to provide a more supportive, warmth, and emotionally expressive female pedagogical agent in a lesson for a better learning experience (i.e., enhanced affective and social processing) and learning outcomes (i.e., enhanced cognitive processing) for students.

Furthermore, while race-matching effects were not consistently found, the findings still suggest that pedagogical agents should be designed with cultural sensitivity. This approach can prevent the reinforcement of stereotypes and support the development of a more inclusive learning environment. The design process should account for diverse cultural representations while avoiding oversimplified or stereotypical portrayals, ensuring that agents resonate positively with students from varied backgrounds.

Overall, the findings of this study offer valuable guidance for educators and developers, highlighting the importance of optimizing the design of pedagogical agents to foster positive emotions and stronger social connections with students. Additionally, the study emphasizes the need to address potential biases and stereotypes related to instructors' race and gender, ensuring that these agents foster more inclusive learning environments that meet the needs of students from diverse backgrounds. By doing so, we can create more effective educational tools that better serve a broad range of learners.

### *Limitations and Future Directions*

Due to the limited number of Black students recruited in this study, the analysis of the Matching Hypothesis was restricted to comparisons between Asian and White students, not including Black students. This limitation affects the generalizability of the findings. Future studies should aim to recruit an equal number of samples from each race type to better understand how race-matching effects manifest across various racial categories, including Black students.

Moreover, although the power analyses have suggested a relatively sufficient sample size for this study, the smaller sample size within each experimental group might still limit the statistical power to detect subtle differences between groups. This could potentially reduce the credibility of the findings. Additionally, the possibility of social desirability bias and individual differences in aesthetic preferences may have influenced participants' questionnaire responses, potentially impacting the accuracy and reliability of the data. To mitigate these limitations, future studies could expand the sample size and incorporate well-designed qualitative methods, such as participant interviews before and after the experiment, to provide deeper insights and strengthen the interpretation of the results.

In addition, the use of the PANAS scale to measure students' felt emotions during learning may not have been entirely suitable. PANAS is designed to be more adept at assessing general daily emotions rather than assessing those emotions during the learning process. Therefore, the lack of significant effects for positive and negative emotions in this study might be attributed to the inadequacy of the scale for this context. Future research should consider utilizing other standardized emotion measurement tools that were specifically designed for educational settings to more accurately gauge emotional responses during learning.

Moreover, this study employed only one style of virtual pedagogical agent, raising concerns about the generalizability of the findings to other styles of agents. The results may be specific to the specific agent design used in this study and may not apply broadly to all pedagogical agents. To enhance the generalizability of future research, it would be valuable to explore a wider variety of agent designs, such as less realistic or more playful, cartoonish avatars. This would help determine whether different styles of agents produce varying effects on students' learning experiences and outcomes—in other words, their affective, social, and cognitive processing.

Another limitation is that the posttest in this study was administered immediately following the lesson without evaluating long-term learning performance. Future research should incorporate delayed posttests, conducted days or weeks after the initial lesson, to gain deeper insights into the lasting effects of pedagogical agents on students' knowledge retention and transfer.

Finally, the lesson in this study was relatively short (i.e., 9 minutes) and focused on a fundamental chemistry topic (i.e., the formation of chemical bonds). As the design and effectiveness of a pedagogical agent can be influenced by the type of knowledge content, and the brief exposure to the learning materials may not be sufficient to produce a significant impact on students, the brevity of the video and the narrow focus raise concerns about the generalizability of the findings to more formal lectures that cover a broader range of subjects. Future research should explore longer, more complex lessons that better reflect real classroom experiences. Additionally, broadening the scope of the learning materials to include other STEM topics, such as mathematics or biology, as well as non-STEM topics, such as history or social sciences, could offer a more comprehensive understanding of whether and how different types of pedagogical agents uniquely influence learning across a variety of content areas.

## Conclusion

In conclusion, this study investigated whether pedagogical agents are as effective as, or more effective than, human instructors for students and whether students show a preference for pedagogical agents of certain genders or races. The findings of this study reveal that while pedagogical agents can deliver comparable learning outcomes to real human instructors, human instructors still foster stronger social connections, challenging the Media Equation Theory. Female agents were more positively received, highlighting gender stereotypes in student-instructor interactions. However, limited evidence was found for the race-matching hypothesis, suggesting the complexity of intergroup biases in learning. These findings emphasize the need for inclusive, well-designed pedagogical agents and call for further research to explore their role in diverse learning environments.

## Author Contributions

Fangzheng Zhao and Richard E. Mayer both contributed to developing the design of the study, interpreting the results, and writing the manuscript. Fangzheng Zhao took responsibility for creating the materials, running the participants, and tabulating the data. Nicoletta Adamo-Villani, Christos Mousas, Minsoo Choi, and Klay Hauser contributed to developing the 3D virtual agents, created by Reallusion's Character Creator, and transform the original female human instructor into six distinct types of pedagogical agents.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was supported by Grant 2201019 from the National Science Foundation.

## Ethical Statement

### *Ethical Approval*

The research was approved by the Institutional Review Board (IRB) at the University of California, Santa Barbara, and followed the standards for research involving human subjects of the American Psychological Association.

### *Informed Consent*

Participants provided informed consent.

## ORCID iDs

Fangzheng Zhao  <https://orcid.org/0000-0002-3575-9039>

Christos Mousas  <https://orcid.org/0000-0003-0955-7959>

## References

- Anderson, K. J. (2010). Students' stereotypes of professors: An exploration of the double violations of ethnicity and gender. *Social Psychology of Education, 13*(4), 459–472. <https://doi.org/10.1007/s11218-010-9121-3>
- Basow, S., Codos, S., & Martin, J. (2013). The effects of professors' race and gender on student evaluations and performance. *College Student Journal, 47*(2), 352–363.
- Berscheid, E., & Reis, H. T. (1998). Attraction and close relationships. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 193–281). McGraw-Hill.
- Boersma, P. (2011). Praat: Doing phonetics by computer. [Computer program]. <https://www.praat.org/>
- Brindley, A. (2013). Chem 1A general chemistry: Lecture 8 - chemical bonds [Video]. *UCI Open*. <https://youtu.be/6GjYGd-k32U?t=92>
- Brown, P. C., Roediger, H. L., III, & McDaniel, M. A. (2014). Make it stick. The science of successful learning. *The Journal of Educational Research, 108*(4), 346. <https://doi.org/10.1080/00220671.2015.1053373>
- Campbell, S. L. (2023). Ratings in black and white: A quantcrit examination of race and gender in teacher evaluation reform. *Race, Ethnicity and Education, 26*(7), 815–833. <https://doi.org/10.1080/13613324.2020.1842345>
- Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences, 7*(4), 173–179. [https://doi.org/10.1016/S1364-6613\(03\)00057-3](https://doi.org/10.1016/S1364-6613(03)00057-3)
- de Albuquerque Rocha, S., de Arruda Raposo, I. P., Fonseca Pereira Oliveira Gomes, S. M., & Firmino Costa da Silva, D. (2024). Representation in Brazilian classrooms: Racial matching

- between students and teachers and school performance. *Education Economics*, 1–19. <https://doi.org/10.1080/09645292.2024.2404226>
- Eagly, A. H., & Wood, W. (2012). Social role theory. *Handbook of Theories of Social Psychology*, 2, 458–476. <https://doi.org/10.4135/9781446249222.n49>
- Eberhard, M. J. W. (1975). The evolution of social behavior by kin selection. *The Quarterly Review of Biology*, 50(1), 1–33. <https://doi.org/10.1086/408298>
- Egalite, A. J., Kisida, B., & Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, 45, 44–52. <https://doi.org/10.1016/j.econedurev.2015.01.007>
- Gershenson, S., Hansen, M. J., & Lindsay, C. A. (2021). *Teacher diversity and student success: Why racial representation matters in the classroom* (Vol. 8). Harvard Education Press.
- Harbatkin, E. (2021). Does student-teacher race match affect course grades? *Economics of Education Review*, 81(5), Article 102081. <https://doi.org/10.1016/j.econedurev.2021.102081>
- Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, 53(1), 575–604. <https://doi.org/10.1146/annurev.psych.53.100901.135109>
- Joshi, P. R., & James, M. C. (2022). An ethnic advantage: Teacher-student ethnicity matching and academic performance in Nepal. *Education Inquiry*, 15(2), 164–187. <https://doi.org/10.1080/20004508.2022.2073055>
- Kalick, S. M., & Hamilton, T. E. (1996). The matching hypothesis re-examined. *Journal of Personality and Social Psychology*, 51(4), 673–682. <https://doi.org/10.1037/0022-3514.51.4.673>
- Khokhlova, O., Lamba, N., & Kishore, S. (2023). Evaluating student evaluations: Evidence of gender bias against women in higher education based on perceived learning and instructor personality. *Frontiers in Education*, 8, Article 1158132. <https://doi.org/10.3389/feduc.2023.1158132>
- Kierstead, D., D'Agostino, P., & Dill, H. (1988). Sex role stereotyping of college professors: Bias in students' ratings of instructors. *Journal of Educational Psychology*, 80(3), 342–344. <https://doi.org/10.1037/0022-0663.80.3.342>
- Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences*, 98(26), 15387–15392. <https://doi.org/10.1073/pnas.251541498>
- Lawson, A., Mayer, R. E., Adamo-Villani, N., Benes, B., Lei, X., & Cheng, J. (2021a). The positivity principle: Do positive instructors improve learning from instruction video lectures? *Educational Technology Research & Development*, 69(6), 3101–3129. <https://doi.org/10.1007/s11423-021-10057-w>
- Lawson, A. P., & Mayer, R. E. (2021). The power of voice to convey emotion in multimedia instructional messages. *International Journal of Artificial Intelligence in Education*, 32(4), 971–990. <https://doi.org/10.1007/s40593-021-00282-y>
- Lawson, A. P., Mayer, R. E., Adamo-Villani, N., Benes, B., Lei, X., & Cheng, J. (2021b). Do learners recognize and relate to the emotions displayed by virtual instructors? *International Journal of Artificial Intelligence in Education*, 31(1), 134–153. <https://doi.org/10.1007/s40593-021-00238-2>

- Lawson, A. P., Mayer, R. E., Adamo-Villani, N., Benes, B., Lei, X., & Cheng, J. (2021c). Recognizing the emotional state of human and virtual instructors. *Computers in Human Behavior*, *114*(6), Article 106554. <https://doi.org/10.1016/j.chb.2020.106554>
- 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. (2022). *Multimedia learning* (3rd ed.). Cambridge University Press.
- Michod, R. E. (1982). The theory of kin selection. *Annual Review of Ecology and Systematics*, *13*(1), 23–55. <https://doi.org/10.1146/annurev.es.13.110182.000323>
- Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments: Special issue on interactive learning environments: Contemporary issues and trends. *Educational Psychology Review*, *19*(3), 309–326. <https://doi.org/10.1007/s10648-007-9047-2>
- Murstein, B. I. (1980). Mate selection in the 1970s. *Journal of Marriage and Family*, *42*(4), 777–792. <https://doi.org/10.2307/351824>
- 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>
- 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>
- Pietraszewski, D. (2009). *Erasing race with cooperation: Evidence that race is a consequence of coalitional inferences*. University of California Santa Barbara.
- Pietraszewski, D. (2016). Priming race: Does the mind inhibit categorization by race at encoding or recall? *Social Psychological and Personality Science*, *7*(1), 85–91. <https://doi.org/10.1177/1948550615602934>
- Pietraszewski, D. (2021). The correct way to test the hypothesis that racial categorization is a byproduct of an evolved alliance-tracking capacity. *Scientific Reports*, *11*(1), 3404. <https://doi.org/10.1038/s41598-021-82975-x>
- Plass, J. L., Homer, B. D., MacNamara, A., Ober, T., Rose, M. C., Pawar, S., Hovey, C. M., & Olsen, A. (2020). Emotional design for digital games for learning: The effect of expression, color, shape, and dimensionality on the affective quality of game characters. *Learning and Instruction*, *70*(4), Article 101194. <https://doi.org/10.1016/j.learninstruc.2019.01.005>
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people*. Cambridge University Press.
- Reid, L. D. (2010). The role of perceived race and gender in the evaluation of college teaching on RateMyProfessors.Com. *Journal of Diversity in Higher Education*, *3*(3), 137–152. <https://doi.org/10.1037/a0019865>
- Renström, E. A., Gustafsson Sendén, M., & Lindqvist, A. (2021). Gender stereotypes in student evaluations of teaching. *Frontiers in Education*, *5*, Article 571287. <https://doi.org/10.3389/educ.2020.571287>

- Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological science*, *17*(3), 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>
- Rosenthal-Von Der Pütten, A. M., Schulte, F. P., Eimler, S. C., Sobieraj, S., Hoffmann, L., Maderwald, S., Brand, M., & Krämer, N. C. (2014). Investigations on empathy towards humans and robots using fMRI. *Computers in Human Behavior*, *33*, 201–212. <https://doi.org/10.1016/j.chb.2014.01.004>
- Ryu, J. E. E. H. E. O. N., & Baylor, A. L. (2005). The psychometric structure of pedagogical agent persona. *Technology, Instruction, Cognition and Learning*, *2*(4), 291–315.
- Sidanius, J., Levin, S., Liu, J., & Pratto, F. (2000). Social dominance orientation, anti-egalitarianism and the political psychology of gender: An extension and cross-cultural replication. *European Journal of Social Psychology*, *30*(1), 41–67. [https://doi.org/10.1002/\(sici\)1099-0992\(200001/02\)30:1<41::aid-ejsp976>3.0.co;2-o](https://doi.org/10.1002/(sici)1099-0992(200001/02)30:1<41::aid-ejsp976>3.0.co;2-o)
- Solanki, S. M., & Xu, D. (2018). Looking beyond academic performance: The influence of instructor gender on student motivation in STEM fields. *American Educational Research Journal*, *55*(4), 801–835. <https://doi.org/10.3102/0002831218759034>
- Taylor, S. E., Fiske, S. T., Etoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*, *36*(7), 778–793. <https://psycnet.apa.org/doi/10.1037/0022-3514.36.7.778>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. <https://doi.org/10.1037//0022-3514.54.6.1063>
- Zhao, F., & Mayer, R. E. (2023a). Role of emotional tone and gender of computer-generated voices in multimedia lessons. *Educational Technology Research & Development*, *71*(4), 1449–1469. <https://doi.org/10.1007/s11423-023-10228-x>
- Zhao, F., & Mayer, R. E. (2023b). Benefits of turning the illustrations in a narrated slideshow into cartoons: An extension of the positivity principle. *Learning and Instruction*, *86*, Article 101779. <https://doi.org/10.1016/j.learninstruc.2023.101779>
- Zhao, F., & Mayer, R. E. (2024). Limitations of disembodied computer-generated voice to convey emotion in multimedia lessons. *International Journal of Human-Computer Interaction*, Advance online publication. <https://doi.org/10.1080/10447318.2024.2371681>
- Zhao, F., Mayer, R. E., Adamo-Villani, N., Mousas, C., Choi, M., Lam, L., Mukanova, M., & Hauser, K. (2024). Recognizing and relating to the race/ethnicity and gender of animated pedagogical agents. *Journal of Educational Computing Research*, *62*(3), 675–701. <https://doi.org/10.1177/07356331231213932>

## Author Biographies

**Fangzheng Zhao** is a fifth year PhD Candidate in Dr. Richard Mayer's Lab at University of California, Santa Barbara. Her research focuses on multimedia learning,

to be more specific, exploring various strategies to improve the effectiveness of video learning or game-based learning.

**Richard E. Mayer** is Distinguished Professor of Psychology at the University of California, Santa Barbara. His research interests are in applying the science of learning to education, with current projects on multimedia learning, computer-supported learning, learning with pedagogical agents, and learning in virtual reality.

**Nicoletta Adamo-Villani** is a Professor of Computer Graphics Technology and Purdue University Faculty Scholar. She is an award-winning animator and graphic designer and creator of several 2D and 3D animations that aired on national television. Her area of expertise is in character animation and character design and her research interests focus on the application of 3D animation technology to education, HCI (Human Computer Interaction), and visualization.

**Christos Mousas** is an Associate Professor in the Department of Computer Graphics Technology and director of the Virtual Reality Lab at Polytechnic Institute, Purdue University. His research revolves around virtual reality, virtual humans, computer graphics & animation, and human-computer interaction. Specifically, he uses computer graphics and computational tools to design virtual reality experiences, and aspects of cognitive and experimental psychology to understand the way humans interact in virtual environments.

**Minsoo Choi** is a Ph.D. student in the Department of Computer Graphics Technology at Purdue University. His research interests include Virtual Reality, Human-Computer Interaction, and Computer Animation. Specifically, in his research, Minsoo focuses on how people interact with virtual agents in virtual environments.

**Klay Hauser** is from West Lafayette Indiana. He studied at Purdue University for his Bachelor of Science in Animation. He is currently studying at Purdue University for his Masters in Computer Graphics Technology.