

August – 2024

# Video Lectures With AI-Generated Instructors: Low Video Engagement, Same Performance as Human Instructors

Selay Arkün-Kocadere<sup>1</sup> and Şeyma Çağlar-Özhan<sup>2,\*</sup>

<sup>1</sup>Department of Computer Education and Instructional Technology, Hacettepe University; <sup>2</sup>Department of Computer Technology and Information Systems, Bartın University; \*Corresponding author

## Abstract

Via AI video generators, it is possible to create educational videos with humanistic instructors by simply providing a script. The characteristics of video types and features of instructors in videos impact video engagement and, consequently, performance. This study aimed to compare the impact of human instructors and AI-generated instructors in video lectures on video engagement and academic performance. Additionally, the study aimed to examine students' opinions on both types of videos. Convergent-parallel approach mixed method was used in this study. A total of 108 undergraduate students participated: 48 in the experimental group, 52 in the control group, and eight in the focus group. While the experimental group (AI-generated instructor) and control group (human instructor) watched 10 minutes of two videos each in two weeks, the students in the focus group watched both types of videos with human and AI-generated instructors. Data were collected through the Video Engagement Scale (VES) after the experimental process, and the Academic Performance Test as a pretest and posttest was administered in both groups. The findings of the experimental part revealed that learners' video engagement was higher in the course with the human instructor compared to the course with the AI-generated instructor. However, the instructor type did not have a significant effect on academic performance. The results based on the qualitative part showed that students thought the AI-generated instructor caused distraction, discomfort, and disconnectedness. However, when the video lesson topic was interesting or when students focused on the video with the intention of learning, these feelings could be ignored. In conclusion, even in today's conditions, there is no difference in performance between human and AI-generated instructors. As AI technology continues to develop, the difference in engagement is expected to disappear, and AI-generated instructors could be used effectively in video lectures.

*Keywords:* generative AI, human instructor, AI-generated instructor, video lecture, video engagement

## Introduction

AI developments, whose place in the education system has been discussed and subject to research for some time, have vivified with generative AI. In the *EDUCAUSE Horizon Report* (Pelletier et al., 2023), generative AI was considered one of the top technologies that will shape the learning and teaching process. While AI can be defined as the simulation of human intelligence processes by computer systems, generative AI is defined as “the production of previously unseen synthetic content, in any form and to support any task, through generative modeling” by García-Peñalvo and Vázquez-Ingelmo (2023) as a result of their systematic mapping study. The form of this content can be visual, text, audio, or video. While there are concerns about the application of generative AI in education (Grassini, 2023; Lambert & Stevens, 2023), it has a huge potential waiting to be explored. When considered in the context of education, it is possible to produce instructional content with generative AI, give automated feedback, individualize the learning process, or support courses with conversational educational agents (Bozkurt, 2023; Lo, 2023; Pelletier et al., 2023). The focus of this study is on video lectures produced with generative AI.

Instructional videos, which could be used as part of larger lessons and either in online or face-to-face classrooms, include both visual and verbal material; the verbal part could be in the form of voice and/or text (Fiorella & Mayer, 2018). A real instructor or pedagogical agent often delivers the verbal part in instructional videos. For a while, massive open online courses (MOOCs) have been popular, and they mainly consist of video lectures; on the other hand, with the COVID-19 pandemic and recent learning movements such as hybrid or flipped classrooms, even face-to-face courses include video lectures. In this respect, it can be said that video lectures are frequently used for education, especially for online learning.

There are many challenges in the video lecture producing process with instructors (Crook & Schofield, 2017). AI video generators have started to be used as a solution. While AI video generators can be used to automate video editing, enhance video features such as resolution, perform automatic video summarization, generate subtitles for spoken content, and perform translation, they can also generate video from text/image and manipulate facial expressions, swap faces, and create deepfake videos. Through AI-generated video platforms, individuals with no video editing skills can create highly convincing synthetic video lectures, with AI-generated humanistic avatars and voices, simply by providing a script (Pellas, 2023). As expected, this is less time-consuming and costly than producing a video lecture through traditional means (Pellas, 2023). There are different types of video lecture productions with AI. Having an agent presenting slides with the real instructor’s face and voice is an interesting example of these productions (Dao et al., 2021).

It is known that the effect of video lecture types on video engagement differs (Chen & Thomas, 2020). Especially, video lectures in which the instructor is present have been the subject of research from several perspectives. Even the instructor’s presence in the video lecture is controversial (Alemdag, 2022). In a large-scale study in which 4,466 learners from 10 highly rated MOOCs participated, Hew (2018) found that having the instructor’s face in a video is preferred, but it is not necessary that it be there throughout the video. Recent studies have shown that instructor presence is critical in different video lecture types. For example, in a course that mainly includes declarative knowledge, performance was higher with instructor voice-over handbook videos as compared to videos showing the instructor and a whiteboard or slides, while in a procedural knowledge intensive course, performance was higher with videos that showed the instructor

in front of a whiteboard (Urhan & Kocadere, 2024). Horovitz and Mayer (2021) studied how the virtual and human instructor being happy or bored in the video affects the learning process and outcomes. Although students perceived the emotional state of both instructors, the learning outcome did not differ. Learners responded to the emotional state of instructors in a similar way, but learners were found to be better at recognizing the human instructor's emotions.

Whether virtual or human, factors such as how much of the instructor's body is seen, the tone of their voice, displaying enthusiasm, showing empathy, and using non-verbal cues have been shown to have different effects on video engagement and learning (Dai et al., 2022; Verma et al., 2023). Even the context in which video is used has an impact on its effect; students in online courses engage with videos more than students in blended courses (Seo et al., 2021). With the perspective that teachers' physical characteristics are effective on students' learning and engagement, Daniels and Lee (2022) stated that AI can be used to create computer-generated teachers customized by race, gender, age, voice, language, and ethnicity when developing online courses.

In the research, AI-generated pedagogical agents in human form are described using terms such as AI-generated avatar, animated agent, virtual human, digital human, virtual tutor, and synthetic virtual instructor. However, not all of these terms are used in the same sense. For example, a virtual tutor can even be an animal cartoon character. Studies about AI-generated instructors that look and sound like humans are not very common in the literature. One example in this context is the study of Leiker et al. (2023). They examined the impact of using generative AI to create video lectures with synthetic virtual instructors. In the comparison, no difference was found between improvement and learning experience perception. As a result of the study, it was stated that AI-generated learning videos have the potential to replace those produced in the traditional method.

Vallis et al. (2023) used AI-generated avatars for presenting videos and online activities in a course. Through qualitative research, they examined students' perceptions of learning with AI-generated avatars. Although students drew attention to the avatar's less social and personal nature, they found the AI-generated avatar appropriate for lecture delivery.

Daniels and Lee (2022) examined the impact of avatar teachers on student learning and engagement via survey and interview. The research had mixed results; while some students found the avatar teacher engaging, others found it distracting. The importance of using appropriate teacher voice in the videos was expressed. However, some students stated that the teacher's physical characteristics had no effect as long as content was delivered well. In addition, some stated that the presence of a human teacher fosters a feeling of connection, establishing confidence in the information presented during the video lecture.

While the literature has not yet reached saturation even for human avatars (Beege et al., 2023), and it is not clear which agents with which characteristics should be included in videos for which conditions (Henderson & Schroeder, 2021), AI-based agents have come to the surface and are being used. This study aimed to analyse the difference between video lectures which include human and AI-generated instructors from the video engagement and academic performance perspective. It is possible to say that the higher the video engagement, the higher the success (Ozan & Ozarslan, 2016; Soffer & Cohen, 2019). For this reason, it is

essential to produce videos with high engagement. Whether this is possible with AI-generated instructors is an issue that needs to be examined. From this point of view, we posed these research questions :

1. Does the type of instructor (AI-generated or human) affect the academic performance of participants?
2. Does the type of instructor (AI-generated or human) affect participants' engagement in video lectures?
3. What are students' views about AI-generated and human instructors in video lectures?

## Method

### Research Model

A convergent parallel approach was used in the study based on mixed methods. In this approach, quantitative and qualitative data are collected independently and combined in the interpretation phase to support the quantitative data and understand the subject more deeply (Creswell & Clark, 2011). In the quantitative dimension of this study, an experimental design was used to examine the effect of instructor types on the video engagement and academic performances of participants. In the qualitative dimension of the research, a focus group was formed to examine the factors that caused the results. The collected quantitative and qualitative data were analyzed and interpreted.

### Study Group

The participants were undergraduate students in the 2<sup>nd</sup> and 3<sup>rd</sup> grade at the Department of Computer Technologies and Information Systems at a state university in Bartın, Turkey. Initially, there were 112 participants, 32 female and 80 male, all between the ages of 19 and 22. Participants had similar backgrounds and previously had not taken any courses on gamification, the subject of the video lecture. On the voluntary participation form signed by participants, it was verified whether they had taken a gamification course before. We then assigned 52 participants to a control group and 52 others to an experimental group. Four people in the experimental group were eventually excluded from the study because they could not complete the data collection process. Eight students participated in the focus group.

### Procedure

Participants were randomly assigned to control and experimental groups according to whether their student numbers were odd or even. The control group watched video lectures that featured a human instructor while the experimental group had video lectures delivered by an AI-generated instructor. The methodological process of the research is summarized in Table 1.

**Table 1**

*Summary of Research Method Comparing Student Responses to Human and AI-Generated Instructors in Video Lectures*

Group	Participants	Pretest	Video lecture type	Posttest
Experimental	48 students randomly assigned	Academic performance	With AI-generated instructor	VES and academic performance
Control	52 students randomly assigned	Academic performance	With human instructor	VES and academic performance
Focus	8 volunteer students		With AI-generated instructor and with human instructor	Interview

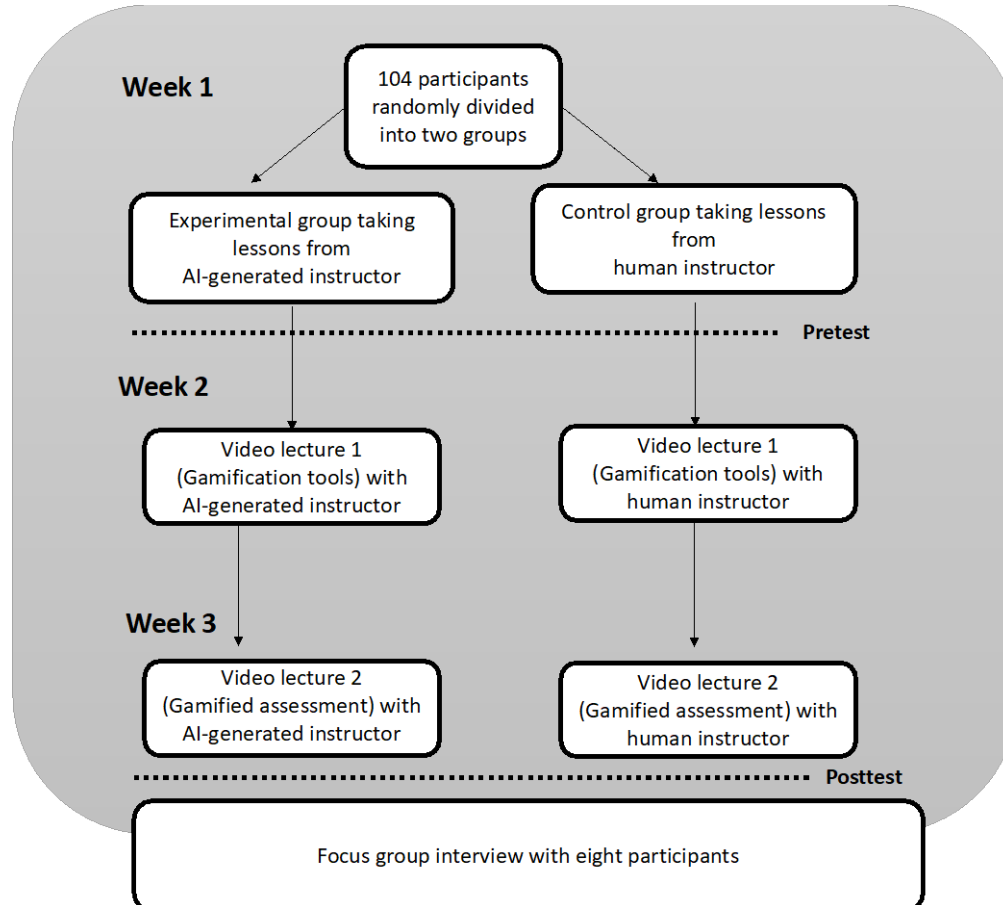
*Note.* VES = Video Engagement Scale. There was no pretest given to the focus group.

After different video types were developed, implementation began. The research included two different processes: the experimental process and the focus group interviews. The experimental process covered 3 weeks. In the first week, students were assigned to experimental and control groups. An academic performance test was applied to both groups. In the second and third weeks, videos featuring an AI-generated instructor were shown to participants in the experimental group, and videos featuring a human instructor were shown to the control group. Both groups watched the videos in the computer laboratory while wearing headphones. After completing the videos, the academic performance test and video engagement scale were administered to participants.

Then, a focus group consisting of eight people was formed. This focus group was shown both types of videos in the computer laboratory. At the end of the application, data regarding participants' opinions about the two different video types were collected. The implementation process is summarized in Figure 1.

**Figure 1**

*Implementation Process of the Study Comparing Student Responses to Human and AI-Generated Instructors in Video Lectures*

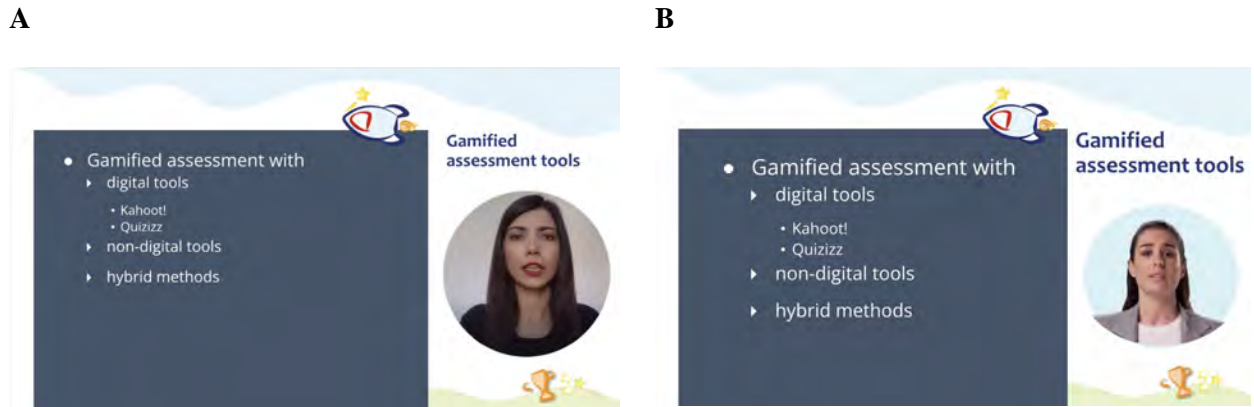


## Videos Lectures

This study used two different videos explaining gamified assessment and gamification tools as the lecture material. Within the scope of the study, two types of videos were developed: one with a human instructor and the other with an AI-generated instructor. During the video development process for both types of video lectures, the text was first scripted and then the visuals and texts were combined in a presentation form. Then, an AI-based instructor video was added to the prepared presentation for the experimental group, and a human instructor video for the control group. The videos are of the talking-head type. Lessons were prepared in Turkish. The screenshots, shown in Figure 2, have been translated into English in this article for the purpose of reader comprehension.

## Figure 2

*Screenshots from the Control and Experimental Groups Video Lecture on Gamified Assessment Tools*



*Note.* Panel A: Screenshot from the control group video with a human instructor. Panel B: Screenshot from the experimental group video with an AI-generated instructor.

The video with the AI-based instructor prepared for the experimental group was created with Vidext. Vidext is a generative AI tool that can produce videos in 27 languages from text by selecting different avatars (Ferre, 2023). The avatar chosen in video lectures created with Vidext exhibits various body movements during the lecture, just like real instructors. Their facial expressions differ, and they can display different postures (Ferre, 2023). Vidext is a commercial software, and only the trial version is free. It contains many design templates, and it is stated that it produces videos 10 times faster than other video generators (Marín, 2024).

In the control group, the prepared learning material was presented by a human instructor. The same script used in the experimental group was voiced, and the human instructor recorded the video. The human instructor is a lecturer from a different university who, therefore, is not known to the students who took part in the research. In summary, the same presentation and the same script were used in both video types. The instructors added to the presentations were similar in terms of their location on the screen and design appearance. Taking into consideration the studies in the literature (e.g., Hew, 2018; Manasrah et al., 2021; Ozan & Ozarslan, 2016), short videos (approximately 10 minutes) were prepared to engage the learners more easily.

## Instruments

In this study, the Video Engagement Scale (VES) was administered to the participants after the experimental process, and data were collected by applying the Academic Performance Test before and after the experimental process. After this, a focus group interview was held with students to explore their opinions about the two different types of videos. Information about these data collection tools is explained under the subheadings that follow.

## **VES**

Within the scope of this study, the Video Engagement Scale developed by Visser et al. (2016) and adapted into Turkish by Deryakulu et al. (2019) was used to determine the level of engagement of learners while watching the educational videos described in the video lecture section. This 7-point Likert-type scale consists of 15 questions. The lowest score that can be obtained from the scale is 15 and the highest is 105.

The scale has a 5-factor structure: (a) attention, (b) going into a narrative world, (c) identity, (d) empathy, and (e) emotion. In the original study group, the internal consistency coefficient (Cronbach's alpha) for the factors was 0.57, 0.73, 0.87, 0.78, and 0.69, respectively. The scale's overall consistency was 0.90. The attention factor measures how much individuals focus on the video by disconnecting from the outside world. The going into a narrative world factor measures the extent to which one is immersed in the narrative world of the video. The identity factor measures the extent to which the identity of the character in the video, that is, the teacher, is adopted by the learners. The empathy factor measures the extent to which learners experience similar feelings to the teacher in the video. The emotion factor measures what kind of emotions the video lecture evokes in the individual. In this study, the Cronbach's alpha of the factors was 0.64, 0.68, 0.88, 0.89, and 0.81, respectively.

## ***Academic Performance Test***

As explained previously, the video lectures cover gamification tools and gamified assessment topics. In order to determine the performance of students in these subjects, we prepared an academic performance test that included 10 questions, five related to each topic in the video. The items were multiple-choice, and each question had four possible answers. The test was structured so that each correct answer equaled 1 point. The lowest score that could have been obtained from the test was 0, and the highest score was 10.

For the academic performance test, the researchers prepared a multiple-choice question pool containing 14 questions based on the video lectures' content. A subject matter expert was asked to evaluate the scope and structure of the draft academic performance test. Following expert opinion, four questions were removed, and expressions that could lead to misunderstandings were edited.

## ***Semi-Structured Interview Form***

As Roulston (2010) suggested, the main interview questions were limited in number, and the opinions of learners were detailed with follow-up questions. The form included three basic questions for each type of video:

1. What do you think about the video you watched?
2. How did the video you watched make you feel?
3. How did the video you watched affect your learning?

These were followed by another question to compare the types:

4. When you compare video lectures, what would you say?



## Data Analysis

In this study, the academic performance pretest was first compared between the groups to evaluate whether the experimental and control groups had similar knowledge levels about gamification. According to the results of the independent samples *t*-test, there was no significant difference between the pretest academic performances of the groups ( $t_{(98)} = 0.18, p > .05$ ). Then, posttests for the variables were compared. Parametric tests were preferred since the data for the groups were normally distributed and homogeneous.

In this study, independent samples *t*-test and analysis of covariance (ANCOVA) were used to compare the video engagement and academic performance of the experimental and control groups. ANCOVA was applied to the data of the academic performance test used for the pretest and posttest. In this way, the effect of the pretest on the posttest was intended to be controlled (Büyüköztürk, 1998). The data collected with the VES used only for the posttest was analyzed with an independent samples *t*-test, and the differences between the groups were reported.

In the qualitative dimension of the research, participants were asked for their opinions about two different types of videos. Qualitative data were examined with content analysis. Content analysis focuses on making repeatable and valuable inferences from a text (Krippendorff, 2018). The analysis is carried out in four stages: (a) data coding, (b) finding themes, (c) organizing the codes and themes, and (d) interpreting the findings (Yıldırım & Şimşek, 2006). The data were analyzed and, independently, an external researcher also reviewed the figures. Reliability between coders was calculated with Cohen's kappa coefficient, and it was found that a significant level of agreement was reached (Cohen's  $\kappa = .61, p < .05$ ). The results were then confirmed by asking additional questions to the focus group participants. To increase trustworthiness, data from interviews were reported with quotations and participant IDs.

## Findings

### Comparison of Findings on Video Engagement

The first research question was "Does the type of instructor (AI-generated/human) affect participants' engagement in the video lectures?" To answer the question, the groups' responses to the VES were analyzed with an independent sample *t*-test.

In Table 2, the averages in the experimental and control groups for the subdimensions of video engagement, attention, going into a narrative world, identity, empathy, and emotion are compared. In this context, a significant difference was found between the average of the experimental group ( $X_E = 11.13$ ) and the average of the control group ( $X_C = 13.12$ ) in the attention dimension ( $t_{(98)} = 3.04, p < .05$ ). A significant difference was also found between the average of the experimental group ( $X_E = 10.81$ ) and the average of the control group ( $X_C = 12.27$ ) in the going into a narrative world dimension ( $t_{(98)} = 1.99, p < .05$ ). A significant difference was found as well between the average of the experimental group ( $X_E = 7.60$ ) and the average of the control group ( $X_C = 9.73$ ) in the identity dimension ( $t_{(98)} = 2.51, p < .05$ ). For the empathy dimension, a significant difference was found between the experimental group ( $X_E = 7.50$ ) and control group averages ( $X_C = 11.90$ ), where  $t_{(98)} = 5.22$  and  $p < .05$ . Finally, for the emotion dimension, the difference between the

experimental group ( $X_E = 9.73$ ) and control group averages ( $X_C = 12.42$ ) is statistically significant ( $t_{(98)} = 3.89, p < .05$ ). A significant difference was seen in all subdimensions, and it was found that the total video engagement score of the control group ( $X_C = 59.46$ ) was significantly higher than the experimental group ( $X_E = 46.77$ ), where  $t_{(98)} = 3.89$  and  $p < .05$ .

**Table 2**

*Comparing Video Engagement Between Groups*

VES subdimension	Group	$X$	$df$	$t$	$p$
Attention	Experimental	11.13	98	3.04	.03
	Control	13.12			
Going into a narrative world	Experimental	10.81	98	1.99	.049
	Control	12.27			
Identity	Experimental	7.60	98	2.51	.014
	Control	9.73			
Empathy	Experimental	7.50	98	5.22	.00
	Control	11.90			
Emotion	Experimental	9.73	98	3.30	.001
	Control	12.42			
Total (video engagement)	Experimental	46.77	98	3.89	.00
	Control	59.46			

*Note.* Experimental  $n = 48$ . Control  $n = 52$ . VES = video engagement scale.

**Comparison of Findings on Academic Performance**

The academic performance test scores of the groups were analyzed with ANCOVA to answer, “Does the type of instructor (AI-generated/human) affect the academic performance of the participants?” Table 3 shows the adjusted average academic performance scores. According to the analysis results displayed in Table 4, no significant difference was observed between the adjusted academic performance averages of the groups ( $F_{(\text{group-error})} = 0.638, p > .05$ ).

However, it was determined that the posttest average of the experimental group ( $X_E = 7.29$ ) was higher than the pretest average ( $X_E = 5.53$ ), and the posttest average of the control group ( $X_C = 7.35$ ) was higher than the pretest average ( $X_C = 5.47$ ). Students learned from both types of video lessons; there was no significant difference between them in terms of performance.

**Table 3**

*Adjusted Average Academic Performance Scores for Control and Experimental Groups*

Group	Pretest			Posttest		
	<i>N</i>	<i>X</i>	<i>SD</i>	<i>N</i>	<i>X</i>	<i>SD</i>
Experimental	48	5.53	1.92	48	7.29	1.82
Control	52	5.47	1.42	52	7.35	1.69

**Table 4**

*Comparing Academic Performance Between Groups*

Variance source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Pretest	75.053	1	75.053	31.892	.000
Group	1.501	1	1.501	0.638	.426
Error	225.922	96	2.353		
Total	5660	100			

### Students' Views on the Different Types of Video Lectures

When the qualitative data were analyzed, it was seen that the opinions could be gathered under themes of distraction, discomfort, disconnectedness, and learning. The themes were named from the perspective of AI-generated instructor lectures, and while the first three are negative, learning is a theme that can be interpreted as positive. These are shown in Table 5 and discussed below.

#### **Distraction**

Six students expressed that the AI-generated instructor's speech and animation caused distraction due to low human-likeness. They mentioned that the voice was digital, there was a lack of emphasis, and that there were pronunciation mistakes in the speech. There were also complaints about the AI-generated instructor's animation, which included movements, gestures, and expressions. Participant 1 said, "The voice in the AI version was very digital, it sounded like I was listening to the woman in the navigation system. I thought it was too synthetic. It was hard for me to focus."

Participant 2 also found it hard to focus:

It is more difficult to focus to the AI instructor, the voice always goes in the same rhythm. Besides, there were some pronunciation mistakes. That's why I couldn't pay full attention to the subject. Also, the teacher animation created with AI was distracting for me.

While two students mentioned their displeasure just about the speech, and four students mentioned similar issues with both the speech and animation, Participant 6 was not bothered by the speech but was unsatisfied with the animation of the AI-generated instructor.

I was also distracted by the AI, if there was no animation of the AI-generated instructor, just the synthetic voice, it might not have felt so different from the human version. Even though the voice sounded human-like in the animated videos, the movements, gestures, and facial expressions on the screen felt unnatural. I think it would be more effective without the animation.

### ***Discomfort***

Two students mentioned a discomfort similar to the disturbance caused by the *uncanny valley effect* in the literature. One even used this expression directly. The uncanny valley describes a situation in which a machine appears to be almost human, yet not quite, which can cause a sense of unease (Cambridge Dictionary, n.d.). The AI-generated instructor was very similar to a human, but there was also an unfamiliarity, making students uncomfortable. For example, Participant 7 said, “The AI-instructor made me uncomfortable at first, but then I got used to it. ... Our brains tend to fear things that are very human-like but not a human, that might be the reason that I felt uncomfortable.” Participant 8 added, “I had a bit of an uncanny valley effect, I felt uncomfortable, the AI instructor acts like a human, but she is not.”

### ***Disconnectedness***

Although a recorded human video was used in the study, and the instructor in the videos was not recognized by the students, the AI-generated instructor was considered to be insincere, not having human characteristics, and, therefore, not giving a sense of personal connection. Participant 6 discussed this phenomenon.

The real instructor gives a more intimate feeling, more personalized than the AI. It is not possible to get this personalization feeling in the AI version. When there is a real person in the video, even if it is a recording, I can empathize as if someone is talking to me considering my previous learning experiences.

Participant 8 agreed with this assessment.

Seeing the teacher’s face helped me to focus on the video. I followed her voice by watching her face. I know there is a real person behind her, it is easy to follow her. In AI, I got detached, there was no connection feeling.

### ***Learning***

In addition to the negative opinions about AI-generated videos, four students made positive comments about being able to learn from the AI instructor. Some students stated that they could learn when the subject was interesting enough or that they could learn if they focused on learning instead of AI-human instructor comparison. These views seem to be meaningful in explaining the significant improvement in the academic performance in both groups. Participant 7 said, “The subject of the video was very interesting for me. As such, my initial bad feeling about the AI-generated instructor disappeared, and I was able to watch the AI video and learn the topic.”

Participant 8 expressed a similar sentiment:

The AI instructor seemed synthetic, but the fact that I had this knowledge and that I had also watched the real human version may have been effective in making me feel this way. If I had only watched the AI version instead of watching both and making a comparison, I think I would have learned the subject easily because my focus would have been on learning.

**Table 5**

*Participant Opinions About the AI-Generated Instructor*

Participant	Negative theme			Positive theme
	Distraction		Discomfort	Disconnectedness
	Speech	Animation		Learning
1	x			
2	x	x		x
3	x			
4	x	x		x
5	x	x		
6		x		x
7	x	x	x	x
8			x	x
Total <i>n</i>	6	5	2	4
Total %	87.5		25	25

Students' opinions support the quantitative findings. The reason for the low video engagement in the AI-generated instructor could be the distraction, which is directly related to the attention dimension of the VES. In addition, it seems easier for students to engage in videos with the human instructor, since they feel more connected and the video is more personalized. In the disconnectedness theme, one student mentioned empathy, which might indicate the empathy dimension of the VES. Another theme that emerged was related to students' uncomfortable feelings about the AI-generated instructor. The reason for this was expressed as both having an avatar very similar to a human and knowing that it was not human. On the other hand, when they were interested in the content or watched the video intending to learn, they did not have difficulty in learning. This is also a consistent finding with the academic performance test results.

## Discussion

According to the experimental part of the study, the engagement of learners in the video course with the human instructor was higher than in the course with the AI-generated instructor. However, instructor type did not have a significant effect on academic performance. In the focus group interview, it was emphasized that the AI-generated virtual instructor made participants feel distracted, uncomfortable, and disconnected. In this context, like in the study of Vallis et al. (2023), the most frequently cited negative aspect of AI-generated instructors is that they cause distraction. It has been stated that the source of this situation is speech and animation. In fact, it is known that first impressions are more positive for a human

than for a virtual instructor (Miller et al., 2023). It is thought that this problem can be solved with technological developments, thanks to synthesis engines that are getting better at creating reliable faces that don't cause uncanny valleys (Nightingale & Farid, 2022). On the other hand, the discomfort and disconnectedness felt in the video lecture with the AI-generated instructor might be corrected by improving the emotional characteristics of the avatar. Indeed, studies have found that adding emotional expressions to the virtual instructor positively affected empathy and the uncanny valley effect (Higgins et al., 2023). However, it was determined that these negative situations can be ignored by learners who find the video lecture's subject matter interesting or simply focus on learning.

The internal validity of the attention dimension of the VES used in this study is not at the desired level, which can be considered a limitation. However, this study indicated that the engagement in video lectures with the AI-generated virtual instructor is lower than that of the human instructor, and the main source of this is the synthetic voice, speech, and animation. For learners' engagement in a video lecture, they are expected to be able to focus, adopt the identity of the teacher, identify the teacher's emotions and feel similar feelings, and experience various emotions while being immersed in the narrative world of the lecture (Visser et al., 2016). In this context, although the AI-generated instructor used human-like gestures, her unnatural voice, speech, and gestures may have caused participants to have difficulty experiencing emotions, empathizing, adopting the instructor's identity, and focusing on the video. The close relationship of emotional factors with engagement has been shown in different studies. Pan et al. (2023) found that affective scaffolding predicted engagement positively. In a study examining the effect of the instructor's presence and absence on engagement in a video lecture, it was concluded that the instructor's presence gives an intimate and personal feel, which leads to a higher level of engagement (Guo et al., 2014). Similarly, a teacher's closeness in communication with a student in the classroom, that is, immediacy behavior and sincere behavior, have a high level of impact on engagement (Hu & Wang, 2023; Wang & Kruk, 2024). The results of these studies, that a human instructor increases engagement by providing an intimate feeling, seem consistent with our findings. In addition to these situations, participants stated that the learners who found the subject interesting could focus on the video without being affected by the instructor's type. According to Shoufan's (2019) study, the factors that affect university-level students' liking of an educational video are technical presentation, content, efficiency, the speaker's voice, and the interestingness of the video. Therefore, the content itself and subject being interesting may cause the instructor's type (AI-generated or human) to be ignored by learners.

Additionally, this study found that instructor type affected video engagement but did not affect academic performance. Leiker et al. (2023) compared those who watched a video lecture featuring an AI-generated virtual instructor with those who watched the traditionally produced instructor video and found that the learning of both groups improved, while no difference was found between the two groups in terms of learning. Studies investigating the effect of the human versus synthesized voice on academic performance in AI-based lectures are frequently encountered (e.g., Chiou et al., 2020). A systematic review conducted between 2010–2021 examining the impact of pedagogical agent features on academic performance and motivation found that the voice type did not affect academic performance (Dai et al., 2022). Another study found that the impact of using human voice and synthesized voice on retention did not differ (Davis et al., 2019). It is known that academic performance does not depend only on instructor characteristics. Within the scope of this study, the learning content prepared for both video lectures, the visuals used, and the

additional resources recommended were all the same. Therefore, whether the channel transmitting the information was AI-generated or human, instructor type did not prevent performance improvement and did not cause a difference between the groups.

## Conclusion

In this study, the effect of AI-generated virtual and human instructors in video lectures on university students' video engagement and academic performance was investigated. One hundred and twelve university students participated in the study conducted with mixed methods. In summary, this study found that engagement differs in favor of the human instructor in video lectures where an AI-generated instructor and a human instructor are used. Despite this, it was found that both types of instructors improved academic performance, and it was concluded that academic performance did not differ between groups. This shows that AI-generated video lectures, produced quickly and at a low cost, can be used instead of videos with human instructors provided certain conditions and improvements are met and made.

## Implications and Recommendations

Many studies conducted on the multimedia principle to improve the conditions in question show that many features, such as instructors' presence in the video lecture, their position on the screen, gestures, tone of voice, and whether they are loved or familiar, are related to behavioral, cognitive, and emotional factors of learning. In this context, in AI-generated video lectures, which have an increasing interest in educational research, structural equation modeling studies can be carried out to see the effect of instructor characteristics on the students' learning from a holistic perspective. In particular, without forgetting the close relationship of engagement with learning, it can be investigated how affective factors such as empathy, emotion, and identity that the human instructor evokes in the learner can be achieved in video lectures, independent of the instructor's characteristics. In this context, AI-generated instructors can be added to video lectures based on emotional design and their effects can be discovered. Additionally, in this study, the importance of the instructor sharing a common denominator with the learner was ignored. Considering this, AI-generated instructors can be used in future studies and their effects on engagement can be examined. Finally, it should not be forgotten that content delivery is a very limited part of the learning process and interaction is the main element. Therefore, investigating the situations in which AI-generated instructors interact with students could be meaningful for future studies.

## References

- Alemdag, E. (2022). Effects of instructor-present videos on learning, cognitive load, motivation, and social presence: A meta-analysis. *Education and Information Technologies*, 27(9), 12713–12742. <https://doi.org/10.1007/s10639-022-11154-w>
- Beege, M., Schroeder, N. L., Heidig, S., Rey, G. D., & Schneider, S. (2023). The instructor presence effect and its moderators in instructional video: A series of meta-analyses. *Educational Research Review*, Article 100564. <https://doi.org/10.1016/j.edurev.2023.100564>
- Bozkurt, A. (2023). Unleashing the potential of generative AI, conversational agents and chatbots in educational praxis: A systematic review and bibliometric analysis of GenAI in education. *Open Praxis*, 15(4), 261–270. <https://doi.org/10.55982/openpraxis.15.4.609>
- Büyükoztürk, Ş. (1998). Analysis of covariance (A comparative study with analysis of variance). *Ankara University Journal of Faculty of Educational Sciences (JFES)*, 31(1), 93–100.
- Cambridge Dictionary. (n.d.). Uncanny valley. In Cambridge online dictionary. Retrieved May 3, 2024, from <https://dictionary.cambridge.org/dictionary/english/uncanny-valley>.
- Chen, H.-T. M., & Thomas, M. (2020). Effects of lecture video styles on engagement and learning. *Educational Technology Research and Development*, 68(5), 2147–2164. <https://eric.ed.gov/?id=EJ1269149>
- Chiou, E. K., Schroeder, N. L., & Craig, S. D. (2020). How we trust, perceive, and learn from virtual humans: The influence of voice quality. *Computers & Education*, 146, Article 103756. <https://doi.org/10.1016/j.compedu.2019.103756>
- Creswell, J. W., & Clark, V. P. (2011). *Mixed methods research*. SAGE Publications.
- Crook, C., & Schofield, L. (2017). The video lecture. *The Internet and Higher Education*, 34, 56–64. <https://doi.org/10.1016/j.iheduc.2017.05.003>
- Dai, L., Jung, M. M., Postma, M., & Louwerse, M. M. (2022). A systematic review of pedagogical agent research: Similarities, differences and unexplored aspects. *Computers & Education*, Article 104607. <https://doi.org/10.1016/j.compedu.2022.104607>
- Daniels, D., & Lee, J. S. (2022). The impact of avatar teachers on student learning and engagement in a virtual learning environment for online STEM courses. In P. Zaphiris & A. Ioannou (Eds.), *Learning and collaboration technologies. Novel technological environments. HCII 2022. Lecture Notes in Computer Science, Vol. 13329* (pp. 158–175). Springer International Publishing. [https://doi.org/10.1007/978-3-031-05675-8\\_13](https://doi.org/10.1007/978-3-031-05675-8_13)
- Dao, X.-Q., Le, N.-B., & Nguyen, T.-M.-T. (2021). AI-powered MOOCs: Video lecture generation. In X. Jiang (Chair), *IVSP '21: Proceedings of the 2021 3rd International Conference on Image, Video*



- and Signal Processing* (pp. 95–102). Association for Computing Machinery. <https://doi.org/10.1145/3459212.3459227>
- Davis, R. O., Vincent, J., & Park, T. (2019). Reconsidering the voice principle with non-native language speakers. *Computers & Education*, *140*, Article 103605. <https://doi.org/10.1016/j.compedu.2019.103605>
- Deryakulu, D., Sancar, R., & Ursavaş, Ö. F. (2019). Adaptation, validity and reliability study of the video engagement scale. *The Journal of Educational Technology Theory and Practice*, *9*(1), 154–168. <https://doi.org/10.17943/etku.439097>
- Ferre, A. G. (2023, November 27). Videos con inteligencia artificial: mejores herramientas para impactar. [Videos with artificial intelligence: better tools to impact]. *Vidext Technologies*. <https://blog.vidext.io/videos-con-inteligencia-artificial-mejores-herramientas-para-impactar>
- Fiorella, L., & Mayer, R. E. (2018). What works and doesn't work with instructional video. *Computers in Human Behavior*, *89*, 465–470. <https://doi.org/10.1016/j.chb.2018.07.015>
- García-Peñalvo, F. J., & Vázquez-Ingelmo, A. (2023). What do we mean by GenAI? A systematic mapping of the evolution, trends, and techniques involved in Generative AI. *International Journal of Interactive Multimedia and Artificial Intelligence*, *8*(4), 7–16. [DOI:10.9781/ijimai.2023.07.006](https://doi.org/10.1016/j.ijimai.2023.07.006)
- Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, *13*(7), Article 692. <https://doi.org/10.3390/educsci13070692>
- Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In M. Sahami (Chair), *L@S '14: Proceedings of the first ACM Conference on Learning@Scale* (pp. 41–50). Association for Computing Machinery. <https://doi.org/10.1145/2556325.2566239>
- Henderson, M. L., & Schroeder, N. L. (2021). A systematic review of instructor presence in instructional videos: Effects on learning and affect. *Computers and Education Open*, *2*, Article 100059. <https://doi.org/10.1016/j.caeo.2021.100059>
- Hew, K. F. (2018). Unpacking the strategies of ten highly rated MOOCs: Implications for engaging students in large online courses. *Teachers College Record*, *120*(1), 1–40. <https://eric.ed.gov/?id=EJ1162815>
- Higgins, D., Zhan, Y., Cowan, B. R., & McDonnell, R. (2023, August). Investigating the effect of visual realism on empathic responses to emotionally expressive virtual humans. In A. Chapiro, A. Robb, F. Durupinar, Q. Sun, & L. Buck (Eds.), *SAP '23: ACM Symposium on Applied Perception 2023* (pp. 1–7). Association for Computing Machinery. <https://doi.org/10.1145/3605495.3605799>

- Horovitz, T., & Mayer, R. E. (2021). Learning with human and virtual instructors who display happy or bored emotions in video lectures. *Computers in Human Behavior*, *119*, Article 106724. <https://doi.org/10.1016/j.chb.2021.106724>
- Hu, L., & Wang, Y. (2023). The predicting role of EFL teachers' immediacy behaviors in students' willingness to communicate and academic engagement. *BMC Psychology*, *11*, Article 318. <https://doi.org/10.1186/s40359-023-01378-x>
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage Publications.
- Lambert, J., & Stevens, M. (2023, September 9). ChatGPT and generative AI technology: A mixed bag of concerns and new opportunities. *Computers in the Schools*, 1–25. <https://doi.org/10.1080/07380569.2023.2256710>
- Leiker, D., Gyllen, A. R., Eldesouky, I., & Cukurova, M. (2023). Generative AI for learning: Investigating the potential of learning videos with synthetic virtual instructors. In N. Wang, G. Rebolledo-Mendez, N. Matsuda, O. C. Santos, & V. Dimitrova (Eds.), *Artificial Intelligence in Education: 24<sup>th</sup> International Conference proceedings. Posters and late breaking results, workshops and tutorials, industry and innovation tracks, practitioners, doctoral consortium and blue sky*. (pp. 523–529). Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-36336-8\\_81](https://doi.org/10.1007/978-3-031-36336-8_81)
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, *13*(4), Article 410. <https://doi.org/10.3390/educsci13040410>
- Manasrah, A., Masoud, M., & Jaradat, Y. (2021). Short videos, or long videos? A study on the ideal video length in online learning. In K. M. Jaber (Ed.), *2021 International Conference on Information Technology (ICIT)* (pp. 366–370). IEEE. <https://www.doi.org/10.1109/ICIT52682.2021.9491115>
- Marín Orozco, D. (2024). *Aplicación de inteligencias artificiales en los procesos de diseño y creación en la malla curricular del programa de Diseño Visual* [Application of artificial intelligence in design and creation processes in the curriculum of the visual design program] [Doctoral dissertation, University of Caldas]. Universidad de Caldas Repositorio Institucional. <https://repositorio.ucaldas.edu.co/handle/ucaldas/19777>
- Miller, E. J., Foo, Y. Z., Mewton, P., & Dawel, A. (2023). How do people respond to computer-generated versus human faces? A systematic review and meta-analyses. *Computers in Human Behavior Reports*, Article 100283. <https://doi.org/10.1016/j.chbr.2023.100283>
- Nightingale, S. J., & Farid, H. (2022). AI-synthesized faces are indistinguishable from real faces and more trustworthy. *Proceedings of the National Academy of Sciences*, *119*(8), Article e2120481119. <https://doi.org/10.1073/pnas.2120481119>
- Ozan, O., & Ozarslan, Y. (2016). Video lecture watching behaviors of learners in online courses. *Educational Media International*, *53*(1), 27–41. <https://doi.org/10.1080/09523987.2016.1189255>

- Pan, Z., Wang, Y., & Derakhshan, A. (2023). Unpacking Chinese EFL students' academic engagement and psychological well-being: The roles of language teachers' affective scaffolding. *Journal of Psycholinguistic Research*, 52(5), 1799–1819. <https://doi.org/10.1007/s10936-023-09974-z>
- Pellas, N. (2023). The influence of sociodemographic factors on students' attitudes toward AI-generated video content creation. *Smart Learning Environments*, 10(1), Article 57. <https://doi.org/10.1186/s40561-023-00276-4>
- Pelletier, K., Robert, J., Muscanell, N., McCormack, M., Reeves, J., Arbino, N., & Grajek, S. (with Birdwell, T., Liu, D., Mandernach, J., Moore, A., Porcaro, A., Rutledge, R., & Zimmern, J.). (2023). *EDUCAUSE horizon report, teaching and learning edition*. EDUCAUSE. <https://library.educause.edu/resources/2023/5/2023-educause-horizon-report-teaching-and-learning-edition>
- Roulston, K. (2010). *Reflective interviewing: A guide to theory and practice*. Sage Publications. <https://doi.org/10.4135/9781446288009>
- Seo, K., Dodson, S., Harandi, N. M., Roberson, N., Fels, S., & Roll, I. (2021). Active learning with online video: The impact of learning context on engagement. *Computers & Education*, 165, Article 104132. <https://doi.org/10.1016/j.compedu.2021.104132>
- Shoufan, A. (2019). What motivates university students to like or dislike an educational online video? A sentimental framework. *Computers & Education*, 134, 132–144. <https://doi.org/10.1016/j.compedu.2019.02.008>
- Soffer, T., & Cohen, A. (2019). Students' engagement characteristics predict success and completion of online courses. *Journal of Computer Assisted Learning*, 35(3), 378–389. <https://doi.org/10.1111/jcal.12340>
- Urhan, S., & Kocadere, S. A. (2024). The effect of video lecture types on the computational problem-solving performances of students. *Educational Technology & Society*, 27(1), 117–133. [https://doi.org/10.30191/ETS.202401\\_27\(1\).RP08](https://doi.org/10.30191/ETS.202401_27(1).RP08)
- Vallis, C., Wilson, S., Gozman, D., & Buchanan, J. (2023, June 6). Student perceptions of AI-generated avatars in teaching business ethics: We might not be impressed. *Postdigital Science and Education*, 1–19. <https://doi.org/10.1007/s42438-023-00407-7>
- Verma, N., Getenet, S., Dann, C., & Shaik, T. (2023). Characteristics of engaging teaching videos in higher education: A systematic literature review of teachers' behaviours and movements in video conferencing. *Research and Practice in Technology Enhanced Learning*, 18, Article 040. <https://doi.org/10.58459/rptel.2023.18040>
- Visser, L. N. C., Hillen, M. A., Verdam, M. G. E., Bol, N., de Haes, H. C. J. M., & Smets, E. M. A. (2016). Assessing engagement while viewing video vignettes; Validation of the Video Engagement Scale

(VES). *Patient Education and Counseling*, 99(2), 227–235.

<https://doi.org/10.1016/j.pec.2015.08.029>

Wang, Y., & Kruk, M. (2024). Modeling the interaction between teacher credibility, teacher confirmation, and English major students' academic engagement: A sequential mixed-methods approach.

*Studies in Second Language Learning and Teaching*, 1–31. <https://doi.org/10.14746/ssllt.38418>

Yıldırım, A., & Şimşek, H. (2006). *Qualitative research methods in social sciences*. Seçkin Press.

