Student Reactions to AI-Replicant Professor in an Econ101 Teaching Video

Rosa-García, Alfonso

Universidad de Murcia

11 February 2024

Student Reactions to AI-Replicant Professor in an Econ101 Teaching Video

Alfonso Rosa García

Universidad de Murcia

alfonso.rosa@um.es

Abstract

This study explores student responses to AI-generated educational content, specifically a teaching video delivered by an AI-replicant of their professor. Utilizing ChatGPT-4 for scripting and Heygen technology for avatar creation, the research investigates whether students' awareness of the AI's involvement influences their perception of the content's utility. With 97 participants from first-year economics and business programs, the findings reveal a significant difference in valuation between students informed of the AI origin and those who were not, with the former group valuing the content less. This indicates a bias against AI-generated materials based on their origin. The paper discusses the implications of these findings for the adoption of AI in educational settings, highlighting the necessity of addressing student biases and ethical considerations in the deployment of AI-generated educational materials. This research contributes to the ongoing debate on the integration of AI tools in education and their potential to enhance learning experiences.

Keywords: AI-Generated Content; Virtual Avatars; Student Perceptions; Technology Adoption

JEL Codes: I23; O33

1. Introduction

Generative artificial intelligence, a type of artificial intelligence (AI) capable of creating new content, has emerged as a multifaceted tool in various domains, with the launch of ChatGPT in November 2022 as a major milestone. Based on GPT-3.5, a sophisticated Large Language Model (LLM), ChatGPT showcased its proficiency in producing text akin to human writing. Other recent advancements in AI have broadened its applications, notably in image creation (with Dall-e, Midjourney, and Stable Diffusion being among the most successful) and other emerging fields such as music composition (Audiocraft by Meta or MusicLM, by Alphabet) and voice synthesis (as Voicebox, by Meta). A notable application of these technologies is the creation of highly realistic avatars (like those from Heygen), which accurately replicate a person's image, facial expressions, and voice. Generative AI, particularly LLMs like ChatGPT, has transitioned from being novel innovations to offer services ranging from creative ideation and feedback to more complex tasks like coding, data analysis, and mathematical problem-solving. In education, the adoption of these tools is a subject of intense debate, encompassing various viewpoints on their appropriateness and the best practices for their integration into educational settings. This paper contributes to this discourse by examining the interaction between students and AI-generated content, addressing key concerns about the adoption of these tools and the ethical considerations they entail.

The significance of AI tools is growing, with their reliability and efficiency driving rapid adoption across multiple sectors, including academia. Korinek (2023) argues that these tools could revolutionize economic research by automating routine tasks, thereby enhancing productivity. He considers that the integration of LLMs will free up researchers from mundane tasks, allowing them to focus on more complex and innovative work. Brynjolfsson et al. (2023) find evidence of the productivity increases that generative AI generate, studying their use by customer support agents, a benefit which is important for novel and non-skilled wokers. Similarly, Dell'Acqua et al. (2023) find a notable increase in productivity in an experiment performed in a consultancy firm upon adopting GPT-4, with the most substantial improvements observed among consultants performing below average. Charness et al. (2023) delve into the multifaceted impact of LLMs on scientific practices, emphasizing their role in enhancing experimental design, implementation, and data analysis, and advocating for a governance framework that balances their benefits and risks. Horton (2023) introduces 'homo silicus', proposing the potential of LLMs as simulated economic agents to generate novel insights in social sciences through simulation.

These studies underscore the growing relevance of AI in academic and professional settings and encourage a deeper exploration of how individuals will interact with these new tools. Various studies have examined how interacting with AI agents, typically algorithms, differs from human interaction, by analyzing standard economic experiments. Ishowo-Oloko et al. (2019) found that algorithms are able to enhance cooperation levels, but the effect disappears when participants are aware they are interacting with a computer, suggesting a reluctance to engage with AI entities. Plaks et al. (2022) observed that the likelihood of cooperation with a robot varies depending on the emotions it displays. Upadhyaya and Galizzi (2023) found that participants in a trust game displayed lower levels of reciprocity with bots compared to humans. Farjam and Kirchkamp (2019) noted that in experimental asset markets, bubbles were less frequent when participants interacted with algorithmic entities or believed they were doing so, suggesting that awareness of AI interaction may alter behavior, potentially in anticipation of more rational responses. This body of evidence points to a general skepticism towards AI, highlighting the need for careful consideration of AI's role in educational contexts.

The increasing relevance of generative artificial intelligence, particularly its application in educational settings, alongside the observed type of human interaction with AI entities, lead us to ask ourselves how students will response to AI entities. The technological advancements in this field facilitate the creation of content that mirrors human output with remarkable accuracy, potentially altering the nature of interactions. Such AI-generated content, encompassing everything from text to voice and non-verbal cues, closely emulates human communication, reminiscent of the interactions with "replicants" as depicted in Philip K. Dick’s novel "Do Androids Dream of Electric Sheep?" and popularized by the film adaptation "Blade Runner", where distinguishing them from humans was nearly impossible. This raises intriguing questions about the reception of AI-generated educational materials by students, especially in scenarios where the distinction between AI-generated and human-generated content blurs. It calls for an exploration of acceptance and skepticism towards AI in educational contexts.

In our study, we produced a teaching video using ChatGPT-4 from OpenAI for the script and Heygen's technology for the professor's avatar. This AI-generated content closely resembled the face, voice, and non-verbal language of the actual professor of the students, making it quite difficult to distinguish that the content was not a recording of the professor. This educational video is performed, instead, by a “replicant” which copies the professor. This set up provides an ideal opportunity to assess student acceptance of AI-generated educational materials.

The participants were divided into two groups, who watch the video: one group was aware of the AI-created content, while the other was not. The findings revealed a significant discrepancy in the perception of the content based on awareness of its AI origin. Students informed about the AI

involvement tended to value the content less compared to those who were not informed, indicating a potential bias against AI-generated content. This reluctance to embrace AI-crafted educational materials raises important questions about the integration of AI in educational settings. It emphasizes the necessity to understand the underlying reasons behind student acceptance or skepticism towards AI-generated content. This paper explores the implications of our findings for educators and the broader implementation of AI in education, including both its advantages and drawbacks.

In the next section, we describe the Methodology we followed to create the video content and then we present the results of our study. Next, we discuss how to escalate the production of this type of content, that can potentially generate a large set of educational contents, we discuss the implications of our results and the ethical issues that this type of content generation rises.

2. Methodology

This section outlines the methodology employed to assess student perceptions of AI-generated educational content. This work is part of a project of the professor, which aims to complement the teaching with additional, complement materials. The research aimed to investigate whether students' awareness of AI involvement influenced their evaluation of such materials.

2.1. Participants and Procedure

The study involved 97 participants from 1st-year Bachelor programs in Business and Economics, enrolled in Econ 101. At the start of a class session, students were asked to voluntarily evaluate new educational materials by completing forms. Initially, they received a document explaining the economic concept of elasticity of demand, generated using ChatGPT-4, without being informed about this. Subsequently, participants watched a video featuring an AI-replicated professor avatar explaining the same concept. The video, created using the Heygen service, can be viewed here: https://www.youtube.com/watch?v=CWBIOXeo-nE. In the form they completed, half of the participants were informed about the AI-generated nature of the video content, while the others were not, with two different forms distributed accordingly (see Appendix A). This division facilitated comparative analysis between informed and uninformed groups. Technical difficulties arose during the video presentation to Business program students, requiring the distribution of the video's URL for subsequent viewing on mobile devices via the virtual campus platform. Conversely, Economics program students faced no such obstacles, enabling direct projection of the video in the classroom setting.

Data collection occurred anonymously through the completion of forms by participants, capturing their assessments of both the text and video content. These forms included Likert scale ratings of the content's overall utility, along with opportunities for qualitative feedback on their consideration of advantages and disadvantages of delivering such supplementary materials.

2.2. Data analysis

Data obtained included the group of the student (Business or Economics), if was informed about the AI generation of the content and valuation of the text (TxtVal) and the video (VidVal) in a 1 to 7 Likert-scale. From this, we also calculate the differential value of each student of video with respect to text (VidVal-TxtVal). We also included an evaluation from 0 to 10 about how positive are advantages or disadvantages described by the student in the open question (TxtFeedback and VidFeedback, respectively). This last evaluation was performed using ChatGPT-4 (in Appendix

B we describe the prompt we use and provide the URL to the conversation where it was generated).

Following data collection, the focus of the analysis was to compare the valuation of text and video content between the two experimental groups: those informed about the AI nature of the video and those who were not. The significance of observed differences was assessed using t-tests. The analysis was conducted using GPT-4[1] and checked with Stata 17.

3. Results

3.1. Descriptive Statistics

The study collected data from 97 participants, consisting of students from first-year Bachelor in Economics (53 students) and Business (44 students) programs. Among the students of Economics (Business), 27 (21) were informed on the IA nature of the materials, and 26 (23) were not. The valuation of the text material (TxtVal) had an average of 5.71 and a standard deviation of 1.11, and the valuation of the video material (VidVal) had an average of 5.90 and a standard deviation of 1.24, both proceeding from a Likert scale from 1 to 7. The difference between the valuation of video with respect to text has an average of 0.19 with a standard deviation of 1.72. Among the 97 participants, 77 gave a response to the open question about advantages and disadvantages of the text material and 73 of the video material. The average valuation of these advantages and disadvantages, according to ChatGPT-4, was of 8.05 for text (TxtFeedback) and 7.42 for video (VidFeedback), with a standard deviation of 1.04 and 1.67, respectively, measured in a 0 to 10 scale.

3.2. Valuation of contents

As expected, the valuation of the text did not differ among the groups informed or non-informed of the AI nature of the video, as can be seen in Table 1. However, equal valuation of the video is rejected at 5%, with the video materials being valued more by students non informed of the AI generation of the video (6.14) than by the students informed about that (5.65). When data are disaggregated by groups, video valuation is higher in both groups when students are not aware of the AI nature of the material, although it is significant at 5% only for students of Economics.

We analyze next the differences between the valuation of the video and the valuation of the text for each student. We find that the difference in favor of the video is higher when students are not aware of the AI nature of the content (0.35) than when they are aware (0.02), although differences are not significant (p-value=0.354). In the Economics group, the difference between both valuations is significantly different between those who were and not aware of the nature of AI (p-value=0.026), but not in the case of Business, where the sign is even reverse (p-value=0.674).

Finally, in Table 3 we analyze the differences in how positive students are in the open question with respect to the advantages and disadvantages of this type of material, after such positiveness being assessed by ChatGPT-4. There is no deferences between the positiveness of the advantages and disadvantages of text material for the two treatments, but there is a significant difference for the video materials, being the average positiveness higher in the non-informed about AI

[1] One advantage of using ChatGPT 4 for data analysis is the transparency it provides, since the analysis can be shared: https://chat.openai.com/share/b5f598dc-629f-4bfc-bef3-a36e07208e80

treatrment than in the informed treatment (an average of 7.83 versus 7.05, with a p-value of 0.047). Differences in the Economics and in the Business group are relevant, of more than half point, but not significant (p-values>0.1).

	TxtVal		VidVal		TxtVal - Economics		VidVal - Economics		TxtVal - Business		VidVal - Business	
	No AI	AI	No AI	AI	No AI	AI	No AI	AI	No AI	AI	No AI	AI
Mean	5.80	5.63	6.14	5.65	5.46	5.67	6.35	5.70	6.17	5.57	5.91	5.57
Std Dev	1.04	1.20	1.14	1.30	1.07	1.00	0.75	1.17	0.89	1.43	1.44	1.47
N	49	48	49	48	26	27	26	27	23	21	23	21
t	0.751		**2.009**		-0.722		**2.373**		1.692		0.778	
p-value	0.454		**0.047**		0.473		**0.021**		0.098		0.441	

Table 1. Comparison of Student Valuations for Text and Video Content, considering awareness of AI.

	No AI	AI	Economics		Business	
			No AI	AI	No AI	AI
Mean	0.35	0.02	0.88	0.04	-0.26	0.00
Std Dev	1.56	1.87	1.11	1.53	1.79	2.28
N	49	48	26	27	23	21
t	0.932		**2.301**		-0.424	
p-value	0.354		**0.025**		0.674	

Table 2. Analysis of students' overvaluation of Video content relative to Text, considering awareness of AI.

	TxtFeedback		VidFeedback		TxtFeedback – Economics		VidFeedback - Economics		TxtFeedback - Business		VidFeedback - Business	
	No AI	AI	No AI	AI	No AI	AI	No AI	AI	No AI	AI	No AI	AI
Mean	8.10	8.00	7.83	7.05	7.86	7.95	7.90	7.21	8.37	8.07	7.73	6.79
Std Dev	1.08	1.00	1.36	1.86	1.24	1.09	1.17	1.56	0.83	0.88	1.62	2.33
N	40	37	35	38	21	22	20	24	19	15	15	14
t	0.420		**2.020**		-0.274		1.637		1.023		1.279	
p-value	0.675		**0.047**		0.785		0.109		0.314		0.212	

Table 3. Evaluation of students' positivity towards advantages and disadvantages of text and video materials, considering awareness of AI

4. Insights into utilizing AI replicants for teaching purposes.

In considering the integration of AI replicants for educational purposes, it is crucial to weigh their potential benefits against ethical concerns. AI replicants, capable of mimicking instructors' voices, faces, and non-verbal communication, offer a promising avenue for enhancing student engagement and accessibility to educational materials. Traditionally, creating such materials has been complex, but advancements in AI technology now streamline this process, making it more accessible and cost-effective. This accessibility suggests that the use of AI replicants in educational videos can become increasingly common

4.1. How to make teaching materials with a professor's AI replicant.

Customized avatars have reached a level of detail making them nearly indistinguishable from real humans. This advancement opens the possibility of automating the creation of video educational materials, presented by the same professor who teaches the course. This approach is able to significantly benefits students, fostering a sense of closeness to the material. Traditionally, creating such materials has been a complex task, but AI simplification considerably eases this process. Specifically, it is a relatively rapid and cost-effective method. An accessible approach, as suggested to students, is the creation of video FAQs. The process involves:

1. **Content Selection**: The professor chooses the video content to explain.
2. **Writing Scripts**: The professor can draft the scripts personally, accelerating the process with the aid of a Large Language Model like ChatGPT and then, check everything is correct.
3. **Customized Avatar Creation**: Platforms like Heygen facilitate this step. The professor needs to upload a short video, no more than two minutes long, where she is speaking. The avatar's quality varies with the video's length, but it is nearly indistinguishable from the professor.
4. **Explanatory Video Production**: Using the professor's AI replica, explanatory videos on selected topics can be generated automatically, resulting in a teaching video where several concepts can be explained by the AI replicant according to the script..

These developments indicate that the automatic creation of such educational materials is not only technologically feasible but also easily achievable and moderately priced. The widespread availability and affordability of this technology is imminent, suggesting that the use of such videos will become increasingly common.

4.2. Ethical concerns

The potential desensitization of the professor-student relationship, alongside the ease of production and scalability of AI-generated content, raises important ethical concerns. While AI-generated materials offer accessibility and customization benefits, there is a risk of diminishing the human element in education. Therefore, it is essential to balance the benefits of AI-generated materials with the ethical considerations they entail. This includes transparently labeling AI-generated content to ensure users are aware of its origins. However, our study suggests that overemphasizing AI involvement may negatively influence material valuation. Thus, there is a need for nuanced approaches to disclosing the AI nature of content creation. Educators and content creators must weigh the benefits of AI-generated materials against the necessity of transparency and the potential impact on student perception.

Despite these concerns, supervised AI-generated materials have the potential to provide tailored educational experiences that align with contemporary learning preferences, particularly among digitally native generations. Therefore, understanding student perceptions and acceptance of AI content is crucial. Our study sheds light on initial reactions but also underscores the need for further investigation into how students' attitudes evolve over time. As AI technology continues to advance, the automatic creation of educational materials is poised to become even more prevalent. Future research should explore innovative ways to leverage AI-generated content to enhance learning outcomes while addressing ethical concerns. Additionally, investigating how students' attitudes towards AI content evolve over time can provide valuable insights for educators and policymakers. By staying attuned to these developments, educators can ensure that the integration of AI into educational settings is both ethical and effective.

5. Conclusion

The increasing ability of generative artificial intelligence is going to potentially change many different aspects of our daily live. Education is one of the sectors that is going to be more affected, with several types of different impacts that need to be analyzed. This investigation also showcases the application of AI in the research methodology itself, leveraging ChatGPT-4 not only in crafting educational scripts but also in analyzing qualitative feedback and performing data analysis. This meta-use of AI illustrates the technology's expansive potential. In our study, we have analyzed for the first time how students react to an AI replicant of their professor, an avatar able to mimic the voice, face, and non-verbal communication of the professor of the students. We find that students value this type of materials, but that their valuation is dependent on their awareness of the nature of the content: students value significantly less the material when they are conscious of the AI nature of its creation. Our study has found that students value less the AI materials only based on their knowledge of its origin, even when the material offered for them is indistinguishable of a material created by a human.

This suggests that there is a tradeoff between informing or not about the nature of this content. Ethical considerations require to inform recipients about this fact, although we must be conscious about the potential problems that this information will create in how students value the material. With the improvements of generative artificial intelligence, it seems likely that AI generated materials will become more and more frequent, covering different types as texts, audio and video. In such a new world, where individuals will interact with entities able to replicate humans, this research is also a call for a further understanding on how students, but also the general public, are going to interact with these new entities.

References

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at work. *NBER Papers*, No. w31161.

Charness, G., Jabarian, B., & List, J. A. (2023). Generation next: Experimentation with ai. *NBER Papers,* No. w31679.

Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. *Harvard Business School Technology & Operations Management Unit Working Paper*, (24-013).

Farjam, M., & Kirchkamp, O. (2018). Bubbles in hybrid markets: How expectations about algorithmic trading affect human trading. *Journal of Economic Behavior & Organization*, *146*, 248-269.

Horton, J. J. (2023). Large language models as simulated economic agents: What can we learn from homo silicus? *NBER Papers*, No. w31122).

Ishowo-Oloko, F., Bonnefon, J. F., Soroye, Z., Crandall, J., Rahwan, I., & Rahwan, T. (2019). Behavioural evidence for a transparency–efficiency tradeoff in human–machine cooperation. *Nature Machine Intelligence*, 1(11), 517-521.

Korinek, A. (2023). Generative AI for economic research: Use cases and implications for economists. *Journal of Economic Literature*, *61*(4), 1281-1317.

Plaks, J. E., Rodriguez, L. B., & Ayad, R. (2022). Identifying psychological features of robots that encourage and discourage trust. *Computers in Human Behavior*, *134*, 107301.

Upadhyaya, N., & Galizzi, M. M. (2023). In bot we trust? Personality traits and reciprocity in human-bot trust games. *Frontiers in Behavioral Economics*, *2*, 1164259.

Appendix A

We include here the forms fill by the students. Originals were in Spanish. Form 1 corresponds to Text evaluation, Form 2 corresponds to Video evaluation, non-AI informed and Form 3 corresponds to Video evaluation, AI informed. The experimental treatment difference is in the second paragraph of Forms 2 and 3.

ASSESSMENT ON NEW MATERIALS

This survey aims to evaluate your perception of the usefulness of new teaching materials. For this purpose, please read the following paragraph about elasticity:

"Elasticity of demand measures how the quantity demanded of a good responds to a change in its price. It is calculated as the percentage change in the quantity demanded divided by the percentage change in price. If the absolute value is greater than 1, the demand is elastic; if it is less than 1, it is inelastic. Income elasticity measures how the quantity demanded changes when the consumer's income varies. A positive value indicates a normal good, and a negative value indicates an inferior good. Cross-price elasticity measures how demand responds to a change in the price of another related good. A positive value indicates substitute goods and a negative value indicates complementary goods."

Written material with information like this could complement the currently available material. It would consist of a set of written materials with brief descriptions of concepts and brief aspects of different topics.

Please answer the following questions:

1. Rate from 1 to 7, where 1 is "Not at all useful" and 7 is "Extremely useful", how useful you think such material would be for your learning.

1 - 2 - 3 - 4 - 5 - 6 - 7

2. What advantages and disadvantages do you think having such material would have? (Optional)

ASSESSMENT ON NEW MATERIALS

This survey aims to evaluate your perception of the usefulness of new teaching materials. For this purpose, please watch the following video about elasticity.

A document with videos like this could complement the currently available material. It would consist of a series of videos featuring the teacher explaining brief descriptions of concepts and brief aspects of different topics.

Please answer the following questions:

1. Rate from 1 to 7, where 1 is "Not at all useful" and 7 is "Extremely useful", how useful you think such material would be for your learning.

1 - 2 - 3 - 4 - 5 - 6 - 7

2. What advantages and disadvantages do you think having such material would have? (Optional)

ASSESSMENT ON NEW MATERIALS

This survey aims to evaluate your perception of the usefulness of new teaching materials. For this purpose, please watch the following video about elasticity.

A document with videos like this could complement the currently available material. It would consist of a series of videos featuring an artificial intelligence-generated model that mimics the image and voice of the teacher, explaining brief descriptions of concepts and brief aspects of different topics.

Please answer the following questions:

1. Rate from 1 to 7, where 1 is "Not at all useful" and 7 is "Extremely useful", how useful you think such material would be for your learning.

1 - 2 - 3 - 4 - 5 - 6 - 7

2. What advantages and disadvantages do you think having such material would have? (Optional)

Appendix B

Here we described how we assigned a value between 0 and 10 to the positiveness with respect to advantages and disadvantages of the material in the response to the open question by the students.

We used ChatGPT-4 on January 25th, 2024. We introduce the following prompt (translated from Spanish):

> "I'm going to provide you with a series of opinions about the advantages of educational material. I want you to give me a rating, from 0 to 10, for each opinion, on whether it appears positive (10) or negative (0), regarding the material."

Followed by the answers of the students in blocks of 10 firt and 20 then. The full chat where it was produced can be found in the following URLs:

- Valuation of text content: https://chat.openai.com/share/da3c50c3-1ae8-423a-a8b3-6f49e9cd314b
- Valuation of video content: https://chat.openai.com/share/895f8c11-4382-4c68-ad5d-9f72f1d80c14