future internet

MDPI

Article

Novel Video Understanding Approach for Embodied Learning of Robotics Technology

Hongbo Zhang [1,*,†], Benjamin Li [2,†] and Gregory Rushton [3,†]

[1] Department of Engineering Technology, Middle Tennessee State University, Murfreesboro, TN 37132, USA
[2] Department of Data Science, Middle Tennessee State University, Murfreesboro, TN 37132, USA
[3] Tennessee STEM Education Center, Middle Tennessee State University, Murfreesboro, TN 37132, USA; gregory.rushton@mtsu.edu
* Correspondence: hbzhang@mtsu.edu
† These authors contributed equally to this work.

Abstract

Embodied learning involves the use of the physical embodiment of hands-on experiences, including gestures, body language, and gaze, during the instructional process for facilitation of the learning outcomes of robotics technology. Understanding the embodiment process is however challenging. In this research, large language model-based video understanding was used for the study of the effectiveness of embodied learning of robotics technology. Embodied and conventional videos were randomly selected, and the user comments were correlated with the transcript and summary of the videos. Results showed that there were higher numbers of user comments correlated with video content for the embodied learning-centered robotics instructional videos than the conventional learning-centered approach in terms of user sentiment and logical reasoning. The sentiment analysis of the video comments showed that the use of embodied learning was effective in achieving engagement in learning robotics, yielding fewer numbers of negative comments in comparison to the conventional learning videos. The embodied learning-centered videos were also helpful to enhance the logical reasoning of students. This user study shows that embodied learning is effective in engaging students, granting more positive sentiments toward the videos. Similarly, the logical reasoning of the students was also enhanced through the use of embodied learning for learning robotics technology.

Keywords: large language models; sentiment analysis; logical reasoning; robotics in education

check for updates

Academic Editors: Jim Prentzas and Ioannis Hatzilygeroudis

Received: 3 January 2026
Revised: 3 February 2026
Accepted: 6 February 2026
Published: 19 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Embodied learning involving human embodiment components such as hands-on learning, gesture, gaze, body language, gaze, and physical interaction through the learning process is an important topic in STEM education. Embodied learning involves the use of tangible objects for teaching, which emphasizes the interaction between the tangible objects and learners to facilitate the learning process. Specifically, the use of gesture and body language is critical for experiential learning due to its unique functions related to anthropology, linguistics, psychology, and education [1–4]. Gesture is one of the most fundamental visual communication approaches for learning technology [5]. Studies have shown that with the use of effective gestures and body language, levels of learning engagement, confidence, happiness, and memory retention have increased [6,7]. Through the use of embodied learning, teachers' instructional activities can be automatically converted to linguistic hints to

facilitate learning and increase the effectiveness of instruction outcomes [8–10]. The use of embodied learning systems is important for learning both concepts and procedures, as they create a learning environment to foster intrinsic motivation among learners through the use of the embodiment components [11]. Embodied learning is recognized as an effective instructional method through the use of motor-based learning, thus benefiting various types of education, particularly special education [7].

Previous studies have shown that embodied learning can lead to better academic performance, increased engagement, and more positive attitudes towards learning. For example, a study found that the active embodied learning strategy significantly improved student performance in STEM fields compared to traditional lecture-based instruction. Other studies have demonstrated that physical engagement in learning activities resulted in better retention and understanding of scientific concepts. These findings underscore the importance of embodied learning to transform educational practices and improve learning outcomes [12,13].

Despite its benefits, the implementation of embodied learning faces several challenges, including the need for specialized equipment; teacher training; hands-on learning facilities, spaces, and resources; and relevant curriculum design. Emphasis on embodied learning may also decrease the use of conventional teaching tools such as Power-point slides, which leads to the reduced use of text or diagram-based information when teaching robotics; therefore, this could compromise learning outcomes. It is also unknown whether embodied learning has led to any significant benefits for the logical reasoning of students in learning robotics technology. Without addressing these research questions, it becomes uncertain as to whether embodied learning-centered teaching should be adopted in STEM education.

The logical reasoning capability of students is crucial for robotics STEM education. However, the training and improvement of the logical reasoning of students are known to be difficult [14,15]. The use of videos to facilitate the learning of logical reasoning is promising. This therefore motivates us to study the differentiation of conventional learning and embodied learning videos to determine how they impact the logical reasoning of learners. The use of embodied learning-centered videos might be able to facilitate improvement of the logical reasoning capability of students due to the involvement of tangible objects, gaze, and gestures, which are not present in conventional learning-centered videos. However, it is still unknown whether these elements are crucial for the improvement of the logical reasoning capability of students.

Addressing these challenges faced by embodied learning-centered teaching requires a collaborative effort among educators, researchers, and policymakers to develop effective strategies and resources for integrating embodied learning into robotics education. It is therefore desired for exploring the long-term impacts of embodied learning, identifying best practices for its implementation, and leveraging emerging technologies to enhance the effectiveness of STEM education [16,17]. The multi-factor nature of embodied learning has however made it difficult to implement embodied learning in reality, where the heuristic teaching strategy is mostly used in practice, thus causing difficulties in improving the quality of robotics education and robotics workforce training.

The previous studies investigating the effectiveness of embodied learning have been mostly associated with experiential learning of tangible objects [4,18]. There is a lack of studies investigating the effectiveness of embodied learning videos in teaching. Despite the lack of such studies, the use of videos for assisting teaching is known to be effective in engaging student's interests and cognitive attention [19,20]. The combination of videos and video-related questions for video-based learning has been shown to be effective in stimulating student's interests in underlying topics [21]. The use of videos in teaching has been known to be effective not only in engaging student's interests but also facilitating their

learning outcomes [22]. The visual elements of the learning content are fundamental to stimulating the learner's interests, enhancing the encoding of information into the long-term memory, and transforming learning experiences into long-term academic interests [23,24]. Since embodied learning videos are associated with strong visual elements, it is therefore expected that the use of embodied learning videos maybe effective in assisting learning of underlying subjects.

This research was motivated by the need to evaluate the effectiveness of embodied learning compared to traditional and non-embodied methods. Conventional learning videos often rely on passive instructional methods, including lectures and visual demonstration, which may not engage students as effectively as embodied learning activities. By comparing the two approaches, our research aimed to determine whether videos incorporating embodied learning principles can result in higher levels of student engagement and more positive educational outcomes. Through the process, both hands-on embodied learning videos and conventional videos were used in this research. A systematic approach in selecting embodied learning and conventional learning videos was used for choosing balanced video content, length, number of user comments, and user sentiments toward the videos to ensure the conclusion was unbiased. Specifically, this study sought to assess whether hands-on, embodied learning videos elicit more user comments, increased logical reasoning, and higher levels of user engagement compared to conventional learning videos. Additionally, the research aimed to document the benefits of embodied learning in robotics education by analyzing user comments and conducting sentiment and feedback content analysis using large language models (LLMs). This comparative analysis will provide valuable insights into the effectiveness of embodied learning in enhancing educational experiences and outcomes in robotics education [25,26]. With our research, a user study was also conducted to examine the differences between embodied learning videos and conventional, lecture-centered videos in terms of their impacts on students learning robotics technology.

Our research is the first study using LLMs for an investigation of the effectiveness of using embodied learning videos for robotics education. The project aims to achieve several key objectives. First, it seeks to assess whether the embodied learning videos elicit more user comments and higher levels of user engagement compared to the conventional learning videos. This will involve analyzing the quantity and quality of user comments on each type of video to gauge user interests, interactions, and participations in the discussion after watching the videos. Second, the project intends to systematically research and document the educational benefits of embodied learning in robotics education through the extraction of the user comments and applying sentiment analysis to understand the emotional and cognitive responses elicited by both embodied and conventional learning-centered videos. Third, the project aims to compare the sentiment distribution (positive, negative, neutral) of comments on embodied learning videos versus conventional learning videos. Fourth, the logical reasoning of users is also analyzed to understand the impacts of watching videos on their logical reasoning capability. Fifth, a user study is also conducted to further compare the differences in the embodied learning- and conventional learning-centered videos. Such a systematic comparison between embodied and conventional learning videos has provided insights into the overall effectiveness and reception of the embodied learning method compared to the traditional learning approaches, thus helping to identify the areas for improvement and the success of embodied learning. This research study is therefore able to answer the following research questions:

RQ1: Do students feel embodied learning-centered videos are more useful and engaging than conventional learning-centered videos?

RQ2: Are video user comments correlated with video content?

RQ3: Is there any connection between embodied learning content and logical reasoning of students?

The answers to the research questions will help us to design effective robotics instruction to impact robotics workforce training. Specifically, they will help teachers to create the most effective videos to enable engaging robotics instruction so that students are able to follow the lectures to achieve effective workforce training. The use of videos for robotics instruction is popular for both lecture-centered formal robotics instruction and informal experience-driven robotics instruction in STEM and higher eduction. As such, the outcomes of this research will impact both formal and informal robotics instruction for broader STEM education impacts.

The article is organized into the following sections: First, we will outline the methods for the research. This is followed by the presentation of the results of the research. The conclusion of the article is presented along with a discussion of the research findings and insights for robotics education. The future work is also presented, describing our planned work using embodied learning in robotics education.

2. Methods

The methodology of this research was designed to answer research questions through the use of videos. Specifically, through this process, we selected embodied learning and conventional learning videos for the analysis of the sentiment and logical reasoning of users toward the videos. Comprehensive videos have been selected in this study to answer the research questions. Following the collection of video data, large language models were used for the analysis of the user sentiment and logical reasoning and their correlation with the video content to reveal the effectiveness of the embodied learning videos for learning robotics technology.

This research was therefore structured to thoroughly evaluate the comparative effectiveness of the embodied learning versus conventional learning methods in robotics education. The differences between the embodied learning and conventional learning were clearly separated by the presence of hands-on learning components in the videos, including gesture- and body language-focused instruction, gaze, and physical interaction with robotic components, whereas conventional learning was mostly dominated by lecture-style teaching of robotics technology. The contrast in learning elements between these two different types of videos was associated with variation in learning outcomes. User comments on the videos were also analyzed. By systematically analyzing user engagement and sentiment through YouTube video comments, this study aimed to provide a comprehensive understanding of how these different educational approaches impacted user engagement and learning outcomes. Through the correlation analysis of video user comments and video content in terms of users' logical reasoning and sentiments, our research ensures accurate and reliable results that can inform educational practices and policies.

2.1. Video Data Collection

Our study involved the use of videos for the analysis of user preferences for embodied learning or conventional learning techniques when learning robotics technology. A comprehensive data selection process was employed for identifying 100 YouTube videos relevant to robotics education. Both conventional and embodied learning videos were used in this study. We have ensured the selection of a balanced representation of both learning approaches to facilitate a robust comparative analysis of the effectiveness of conventional and embodied learning videos.

The 100 videos were equally categorized into two distinct groups: 50 videos featured embodied learning methods, characterized by hands-on robotics learning activities, gesture-

and body language-focused instruction, gaze, and physical interaction with robotic components, while the other 50 videos featured conventional learning methods, which primarily involved instruction using slides, software coding tutorials with shared screens, and other text- and diagram-centered visual demonstrations.

We have made efforts in ensuring the representativeness of the videos. Diversified videos were studied in our research, which included programming, mathematics, AI, and hardware. Through the process, both formal and informal learning videos have been selected. Specifically, the video contents ranged from introductory to expert-level robotics technology so that the videos were more representative for learning robotics technology. The videos were also comprehensive, covering a significant number of applications including agriculture, manufacturing, healthcare, logistics, construction, energy, and retail. These videos were representative to enable them to be useful for robotics workforce training without suffering from biased video selection.

We have also enforced the random selection of videos by using a keyword search of robotics technology, such as electronics of robotics, navigation of robotics, and mechanical design of robotics. Given the list of the search results of videos on Youtube, we have selected the videos by following the definition of conventional learning and embodied learning. This means that the embodied learning and conventional learning videos were selected randomly based on the learning content rather than the quality and experiences of the teaching and the level of user engagement with the videos. The second criteria of the video selection was to select videos either focusing on conventional learning or embodied learning but not both. We have purposely avoided the biased selection of either overengaged or underengaged video comments for both the embodied learning and conventional learning videos. We have also purposely avoided the selection of videos involving balanced conventional and embodied learning equivalently, thus potentially causing biases in the video selection.

The balanced selection of videos in terms of user comments was also performed. Specifically, videos with balanced comments, such as very few numbers (1–10), somewhat large numbers (10+), and large numbers (100+) of user comments, were selected. Other balance concerns regarding the video comments included the selection of the sentiments toward the videos, where both positive, neutral, and negative comments all needed to be present in the conventional and embodied videos. Through the use of positive, neutral, and negative user comments for the videos, we were able to ensure a fair comparison between the conventional and embodied videos.

The descriptive statistics of the video length and total comments are shown in Table 1. Specifically, it shows the minium, maximum, and average video length and total comments. Specifically, the minimum video length is about 2.21 min with a maximum video length of 96.10 min, and the average video length is 16.07 min for conventional videos and 11.66 min for embodied videos. The number of user comments per video has minimum of 1 comment, a maximum of 2818 comments, and an average of 201 comments for conventional learning videos and 310 comments for embodied learning videos.

The examples of the selected videos are shown in Figures 1 and 2. In these two Figures, the examples of conventional and embodied learning videos are shown, selected respectively based on the conventional and embodied learning selection criteria. In Figures 1 and 2, we show the contrastive differences between conventional and embodied learning. Conventional learning focuses on the use of lectures with charts, coding, mathematics, and diagrams. As shown in the figure, the conventional learning videos heavily focus on the use of text and images through the process. Overall, the lecture-heavy teaching style is emphasized in robotics instruction. While some animations of the robotics technology may be used in these videos, the teaching style of the videos mostly leverages the use of text- and image-based instructional

materials through the process, where hands-on learning is mostly absent. On the other hand, if the hands-on experiential learning of robotics technology becomes dominant, these videos will not be selected as conventional learning videos.

Table 1. Detailed statistical summary of video length and number of user comments for conventional and embodied learning videos.

Metrics	Video Length (min)		Number of Comments	
	Conventional	Embodied	Conventional	Embodied
Minimum	2.40	2.21	3	1
Maximum	96.10	44.33	1324	2818
Average	16.07	11.66	201.54	310.11
Total Count	899.70	862.84	11,286	22,948

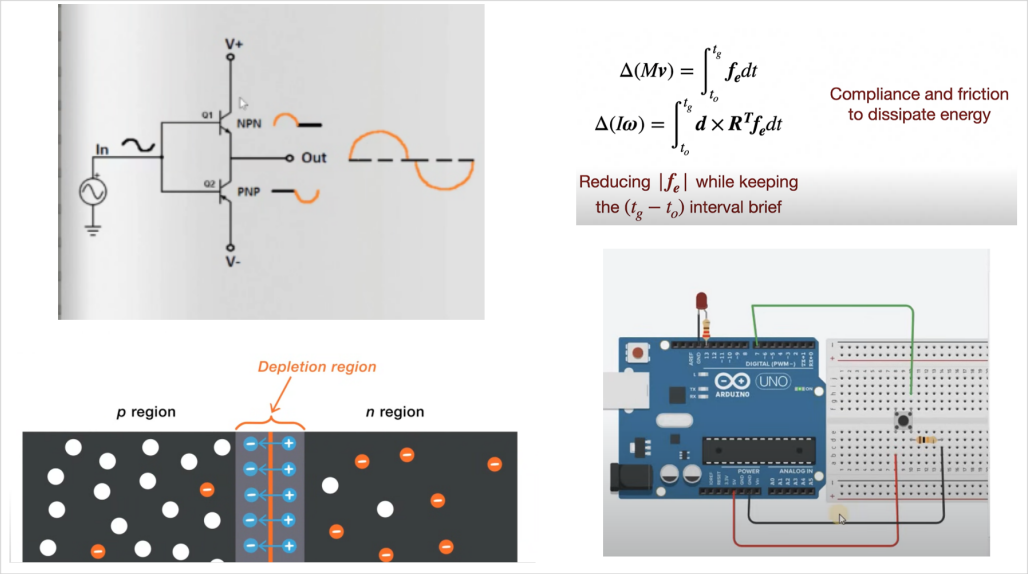


Figure 1. The examples of conventional learning videos. (**Top Left**): The transistor tutorial. (**Top Right**): The friction coefficient calculation tutorial. (**Bottom Left**): The diode tutorial. (**Bottom Right**): The Arduino tutorial.

Similarly, the embodied learning-centered videos are also chosen for the research shown in Figure 2. In contrast to the conventional learning-centered videos, the embodied learning-centered videos focus more on experiential learning involving hands-on skills, kinesthetic-centered learning, and hands-on demonstrations to the learners, where the embodiment elements are emphasized. The embodied learning videos focus on the interaction between learners, software, and hardware. For example, a video can teach learners how to assemble a robotics platform from scratch. In contrast to the conventional learning of robotics, where lecture-based teaching is emphasized, embodied-centered learning involves the use of tangible robotics parts to teach the robotics platform assembling process, as shown in Figure 2.

In contrast to conventional learning methods, embodied learning focuses on the use of hands-on skills through leveraging the actual robotics platform, in-person demonstrations of the electronics, and parts for robots. The major contrast between the two different learning approaches lies in the fact that embodied learning requires the use of embodied components such as gesture, gaze, and hands-on components, while conventional learning

focuses on the delivery of the educational materials primarily by using lecture slides or sharing the screen of the programming software.

Examples of embodied learning are shown in Figure 2. They show that embodied learning videos emphasize the use of gestures to illustrate the robotics technology. Specifically, teachers illustrate the robotics platform construction process through the use of different body gestures. Examples of body gestures include the demonstration of the use of the robot remote controller receiver, where the instructors use the actual remote controller receiver to showcase the use of the device for controlling the robotics platform. Through the process, learners watching the video will learn about the wiring of the remote controller to control robots. Learners watching the video will also learn about the programming of the remote controller receiver through the process to interact with robots. Other examples of the embodied learning of robotics platforms include the hands-on illustration of the DIY process of the robot chassis, as shown in Figure 2. Through the gesture-centered learning of robotics, the user interactions between instructors and learners watching the video are well-utilized.

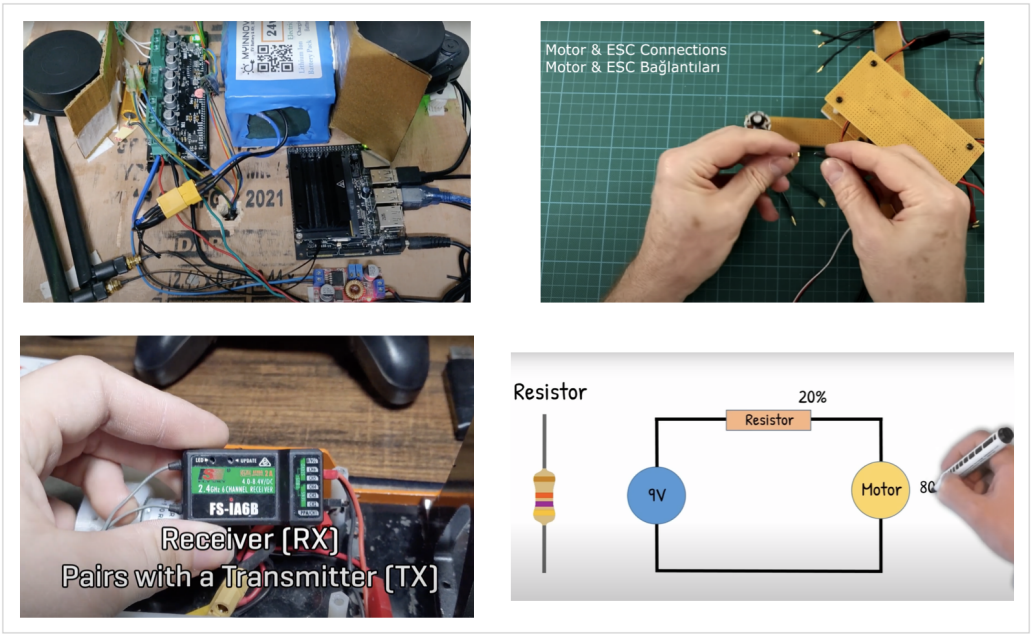


Figure 2. Examples of embodied learning videos. (**Top Left**): The DIY robotics platform. (**Top Right**): The DIY drone platform. (**Bottom Left**): The remote controller receiver. (**Bottom Right**): The circuit tutorial.

2.2. Video User Comment Extraction

For the analysis of the videos, we first retrieved video user comments by specifying the URLs of the videos and the desired output format, as shown in Figure 3. The extraction process was performed by extracting the Youtube comments with Youtube APIs and saving them in a JSON file. This JSON file was then subsequently parsed to access the user comments in the JSON file and iteratively retrieve detailed information about each comment. The resulting JSON output file contained the feedback of users who engaged in commenting on the videos. The video users' commenting data were then parsed for the analysis of user engagement and sentiment with the video. The JSON file was organized with a well-structured format; as such, it could be easily parsed for the subsequent analysis of user sentiment toward the video and we could perform correlation with the transcripts of the video.

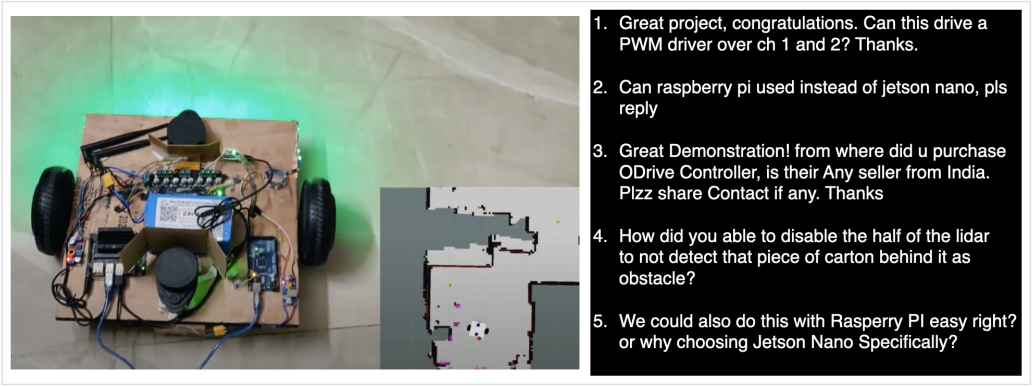


Figure 3. The example of comments extracted from the embodied learning video. (**Left**): The DIY robot tutorial. (**Right**): The comments associated with the video.

2.3. Transcript Extraction

Following the extraction of the user comments on the videos, the Open AI Whisper model was used to transcribe audio data to text to obtain the transcripts of the videos [27]. Through this process, the audio files of the videos were imported. The transcription process involved loading the audio file into the Whisper model and initiating the transcription functionality. The Whisper model utilized its advanced natural language processing (NLP) capabilities to precisely convert the audio input into text output. The model employs a large-scale multilayer transformer architect for the analysis of the spectrogram of the audio [27]. Due to the use of weakly supervised learning, the model is able to learn the features of the audio for precise audio transcription[27]. The transcription results were then saved to an output file for further analysis and documentation. This method was shown to enable automated, accurate, and efficient transcription of the spoken content and facilitate activities that required textual data from audio sources.

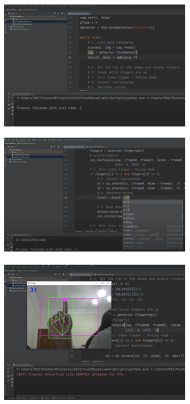
2.4. Summarization

In order to better understand the video content, we performed summarization of the videos, as shown in Figures 4 and 5. Summarization is a key process in natural language processing that condenses lengthy texts into shorter and coherent versions while retaining essential information. It can be extractive, selecting key sentences from the original text, or abstractive, generating new sentences that convey the main ideas [28,29]. OpenAI's GPT is known to be feasible for summarizing videos, with summarization leveraging efficiency and accuracy by processing the video transcripts [30]. As such, the model was used to achieve quality video summarization, making the summaries more concise, understandable, and useful for correlation analysis with user comments.

For summarization, we leveraged OpenAI's GPT-3.5 language model and the Langchain python package V0.3.7. The video summarization process was performed in the following steps: The process started with fetching the video transcripts by extracting the video ID from the provided URL. It also further retrieved the manually created transcript of the video. Following video transcript retrieval, the application summarized the transcript by splitting it into manageable chunks using the recursive character text splitter technique.

The summarization was performed using GPT-3.5 APIs, as shown in Figures 4 and 5. The model was able to understand a wide range of topics and provide high-quality summarization [31]. Following the summarization with the model, the summarized content was presented through a Streamlit Web user interface, which facilitated the accessibility of the summarization of the videos [32]. The results of the summarization were stored in the

JSON file for further analysis of the correlation between user comments and the summary of the video transcript.

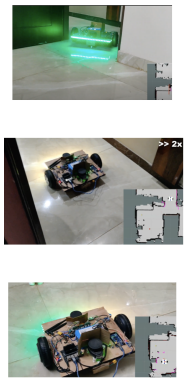


The video tutorial shows how to create an AI-based mouse controller by detecting hand landmarks and tracking and clicking based on these points. The tutorial also explains how to apply smoothing techniques to make the controller more usable. The hand tracking module used in the tutorial was written from scratch and includes methods such as the fingers up method and the fine distance method. The tutorial also introduces the auto-pi library, which allows the user to move and click the mouse using a Python script. The steps involved in creating the AI-based mouse controller include finding the hand landmarks, getting the tip of the index and middle finger, and detecting clicks based on the distance between the fingers. The tutorial provides code snippets for each step and explains how to install the necessary libraries, such as OpenCV and MediaPipe. The tutorial also promotes the creator's premium course, which teaches how to create real-world computer vision apps, such as object detection and augmented reality.

Bullet Summary:

- Tutorial demonstrates creating an AI-based mouse controller using hand tracking
- Smoothing techniques applied to improve usability
- Hand tracking module written from scratch
- Auto-pi library used for mouse movement and clicking
- Steps include finding hand landmarks, getting finger tips, and detecting clicks based on finger distance
- Code snippets provided for each step
- Premium course promoted for learning real-world computer vision apps

Figure 4. Example summary of a video transcript of a conventional learning video. (**Left**): The screenshots of the video. (**Right**): The summary associated with the video.



The creator of an autonomous mobile robot (AMR) has released version 2 with updates including bldc Hub motors with Hall encoders controlled using an O Drive controller, which has improved the accuracy of odometry and localisation. The robot is powered by a 24-volt in ah lithium-ion battery, uses an Arduino to communicate with the Robot Operating System (ROS) and is fitted with an addressable RGB LED strip to change colour based on the state of the robot. The creator has also added a button interface to manually stop and start the robot while moving towards a goal or waiting at a waypoint. To ensure full 360-degree visibility, the robot uses two lidars, one at the front and one at the back. The robot is capable of detecting obstacles in all directions. The video shows the robot being guided through three goal locations, waiting five seconds at each, changing the colour of the LED to green when moving towards a goal, yellow when idle and blue when stopped by a user. The robot detects obstacles and responds by stopping and waiting for a clear path.

Bullet Summary:

- Autonomous mobile robot (AMR) version 2 has been released with updates including motor improvements and lidars at the front and back.
- The robot is controlled using an O Drive controller and a Jetson Nano 4GB variant.
- An Arduino is used to communicate with the Robot Operating System (ROS) and the robot is powered by a 24-volt in ah lithium-ion battery.
- The robot uses an addressable RGB LED strip to change colour based on the state of the robot and has a button interface to manually stop and start the robot while moving towards a goal or waiting at a waypoint.
- The robot is capable of detecting obstacles in all directions and responds by stopping and waiting for a clear path.

Figure 5. Example summary of a video transcript of an embodied learning video. (**Left**): The DIY robot tutorial. (**Right**): The summary associated with the video.

2.5. Correlation Analysis of User Comments with the Transcript and Summary of the Video

The correlation between video user comments and video transcripts was also determined, as shown in Figure 6. The correlation was determined through the use of the video's user comments and the video transcript as well as the summary of the video. The correlation analysis between user comments and the video transcripts along with the video summary began by extracting the comments of the video and the transcript of the video. Each comment was then evaluated using the OpenAI GPT-3.5-turbo model. The model analyzed whether a comment on the video was specifically correlated with the text in the transcript and the summary of the video. Through this process, comments were pre-processed for the removal of duplicates and empty entries. The video transcript was divided into multiple sections with each section being approximately 1000 words. Through the segmentation of the long video transcript into multiple smaller sections, we were able to facilitate a detailed comparison between comments and specific parts of the transcript. Through this process, the correlation between the user comment and the transcript and summary of the video was obtained. The correlation rationale was closely examined to understand the correlation and non-correlation between the user comments and the transcript and summary of the videos. The determination of correlation involved the use of regular expressions to search for the keywords of the correlation results provided by the OpenAI GPT-3.5-turbo model.

The correlation includes both the transcript of the entire video and the summary of the video, which is correlated with the comments of the video. Through the inclusion of both the transcript and summary of the video, the correlation analysis is able to capture the user engagement and its connection with both high-level and low-level details of the video content. The correlation analysis is thereby able to enhance the understanding of the connection between video content and the reaction to the video content of the users. The correlation analysis is done iteratively for each comment of the video and correlated with the summary and transcript of the videos. Through this process, we perform correlation analysis using both embodied and conventional videos to obtain comparative results.

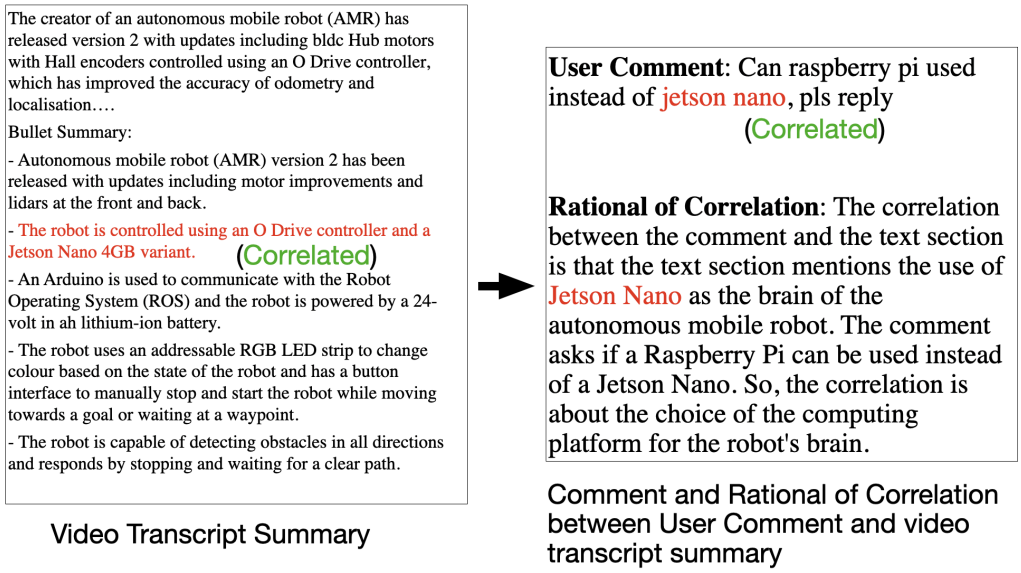


Figure 6. Example summary of a video transcript of an embodied learning video and determination of their correlation. (**Left**): The summary of the video. (**Right**): The comments and the correlation rationale between the comment and the summary of the video.

2.6. Sentiment Analysis

Sentiment analysis was also performed to determine the overall sentiment of the user engagement and the reaction toward the videos using both conventional and embodied learning methods, as shown in Figure 7. It provided analysis of the positive, negative, and neutral sentiments through the video's user comments. The video user comment sentiment analysis algorithms evaluated the comments and assessed the learner engagement. This process helped us to understand the general attitude of users towards the video content. For instance, positive comments indicated their satisfaction regarding the learning process, while negative comments reflected dissatisfaction or criticism and neutral comments suggested indifference or impartiality.

Sentiment analysis of the comments of the videos facilitated the robust analysis of viewer sentiments across conventional and embodied learning approaches. The implementation of sentiment analysis through user video comments was carried out through pysentimiento [33]. This model leverages the use of multiple state-of-the-art sentiment BERT transformer models. Specifically, it includes the use of BERT [34], RoBERTa [35], and ELECTRA [36] for accurate sentiment analysis. Pysentimiento was trained with large text corpora and thus able to understand the contextual information of the text for generalized and effective sentiment analysis.

The sentiment analysis of the video user comments was performed through the extraction of the video comments from the videos. It was further followed by reading the content and splitting the text into paragraphs to facilitate the detailed analysis. Each paragraph

was then evaluated using the sentiment analyzer for the prediction of the sentiment of the user comments. The probabilities of certain sentiments for each comment were analyzed to determine the sentiment and to understand the attitude towards, satisfaction with, and perception of the learning process.

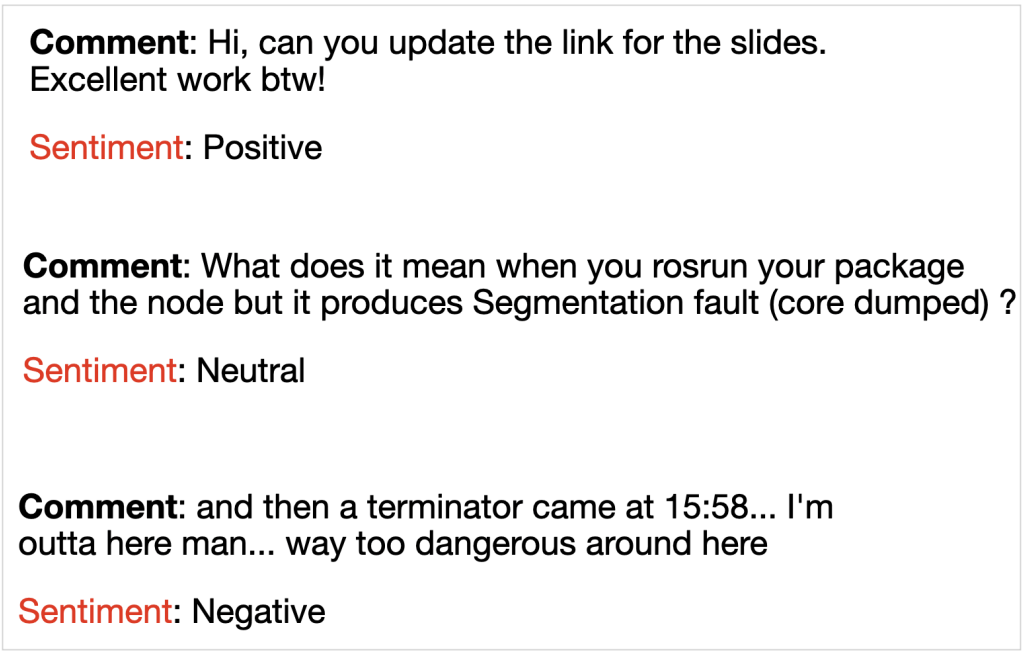
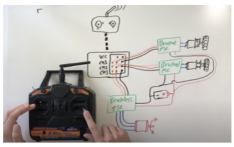


Figure 7. Examples of sentiments in the user comments of the videos. This also shows the sentiment (positive, negative, neutral) analysis of the comments.

2.7. Logical Reasoning in Video Comments

Both conventional and embodied learning videos were involved in the logical reasoning analysis shown in Figure 8. Specifically, we leveraged the use of the user comments of the videos to examine logical reasoning through the comments. Through this process, we examined the logical reasoning which occurred due to watching the videos. It showed the reasoning of the video user comments related to robotics technology. It showed the reasoning and learning from the videos after learners had watched the videos. It also showed the reflections of the learners after watching the videos. The stimulus effects of the videos for learning robotics technology were contrasted between conventional and embodied learning modes.

The logical reasoning due to conventional learning was associated with the learning which occurred in the text and images of the video content. It stimulated users to learn the robotics technology. On the other hand, the contribution of the embodied learning elements to the logical reasoning also occurred due to the hands-on learning components embedded in the videos. Both conventional and embodied learning have shown general effectiveness in assisting with the learning of robotics, as shown in Figure 8. Specifically, the examples of logical reasoning and the contribution of embodied learning toward the logical reasoning process are shown in Figure 8. As shown in the figure, the video demonstrated the use of the remote controller to control robots. After watching the video, users were inspired to ask questions about the use of the remote controller to control the servos of the robotics platform. While the instructor in the video did not specify the details of using the remote controller to control a robotics platform, the embodied learning video content was shown to be able to inspire the video users to raise such questions after watching the video, thus showing the presence of logical reasoning through the learning process.





User Comment on Video: *What if you wanted to build a large complex bot with several servos or linear actuators and you also wanted to use machine learning on it so it has to be able to run python or something? Would respberry pi be the best route or is there something better out there?*

Logical Reasoning	**The user appears to have learned something significant from the video with logical reasoning through their feedback.** The user's comment demonstrates an understanding of the potential uses of a raspberry pi in building a complex bot that requires machine learning and servos or linear actuators. The user is also considering alternative options for controlling the bot beyond just using a raspberry pi, indicating an awareness of different technologies and their applications.

Figure 8. Example of logical reasoning analysis data and output for the 100 videos (50 embodied learning videos and 50 conventional learning videos) in this research. Example of the video and comment on the Youtube video (**Top**) and the related logical reasoning analysis output of the video comment (**Bottom**).

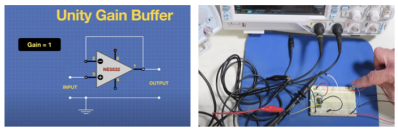
2.8. User Study

A user study was also performed to validate the effectiveness of using embodied learning and conventional learning videos, as shown in Figure 9. For this, we involved 40 students from the Engineering Department at Middle Tennessee State University in the study in watching and performing reflections on both embodied and conventional learning videos, with the study approved by institutional review. The recruited students were from engineering majors, where male students were in greater numbers than female students; as such, 31 male students and nine female students were involved in this study. The age of the students ranged from 20 to 26 years old. All students had basic knowledge of electronics and mechanical design of robots and automation. In total, 60 conventional learning videos and 60 embodied learning videos were used through this research. The user study video selection criteria were identical to the criteria used for the selection in the study of user engagement with online Youtube videos. All the videos were given to students so that they could watch the videos and learn from the videos. Through this process, students had freedom to choose five of the videos which they liked to watch. Students were also asked to reflect and provide their feedback and learning outcomes from watching the videos. Students were advised to carefully watch the details of the videos. In order to facilitate the video-based learning process, students were encouraged to ask for help from the research team throughout the process to resolve their technical challenges in watching videos. Subsequently, students were asked to comment on the learning outcomes of the videos. Specifically, students were asked to comment on the significant learning outcomes from watching the videos. We did not ask students to express strong sentiments, either positive or negative, throughout the feedback process to avoid potential biases in their sentiments toward either embodied learning videos or conventional learning videos.

After watching the videos, the feedback of students was obtained and processed through the extraction of the comments on the videos. Subsequently the logical reasoning and sentiment analysis of the comments toward the videos were conducted to conclude the

logical reasoning learning outcomes along with the sentiments of the participants following watching videos.

The data analysis of the logical reasoning and sentiments was performed using the Llama3 pre-trained model (8B parameters) to obtain the logical reasoning and sentiments from the feedback of students after watching videos. Specifically, the rich written feedback from students enabled us to study the embodied learning element of student's feedback. For this, the embodied learning element in the logical reasoning was also analyzed using the Llama3 model to understand whether the embodied learning elements helped students to improve their logical reasoning capacity through watching the videos.



User Comment on Video: *This video covers the basics of Op-Amps and how they work. An Op-Amp is a high gain differential amplifier, and these devices are extremely cheap, with some being as low as 10 cents in price. Op-amps have high input and low output impedances. Op-amps were first created with vacuum tubes, which makes sense as vacuum tubes were also used to construct instrument amplifiers as well. The same Op-Amp can be used in different circuit designs to create varying results, such as high and low pass filters.*

Logical Reasoning	**Based on the student's paragraph, it seems that there's a clear logical reasoning through the student's feedback:** 1. Simulation: The video covers the basics of Op-Amps, allowing students to see how they work in a visual and auditory format. This helps students understand the concept better by providing an example that they can observe and analyze. 2. Explanation: The student provides a clear explanation of what an Op-Amp is and its functionality as a differential amplifier. This helps students understand the technical terms and concepts related to Op-Amps. 3. Comparison: The student mentions that Op-amps were first created with vacuum tubes, which allows for comparison between the two types of devices. This can help students understand the differences between the two and how they work in different applications. 4. Application: The student provides an example of how the same Op-Amp can be used in different circuit designs to create varying results, such as high and low pass filters. This shows that Op-Amps have multiple applications and can be used in a variety of contexts.

Figure 9. The results of logical reasoning analysis of data and output from the user study of the research. Example of a video and feedback of a student from the user study (**Top**) and the related logical reasoning analysis output of the student's feedback (**Bottom**).

3. Results

3.1. Analysis of Sentiment and Logical Reasoning Regarding Videos

3.1.1. Video Comment Distribution

The analysis of the total number of comments revealed a significant difference between embodied and conventional learning videos. Embodied learning videos received a total of 22,581 comments (81.65%), whereas conventional videos received 5076 comments (18.35%). Since we selected the 100 videos randomly using robotics education keywords from Youtube to avoid the biases in the terms of the video source selection, this result indicates that there is a substantial difference in terms of the user engagement with the two different types of videos, as shown in Figure 10.

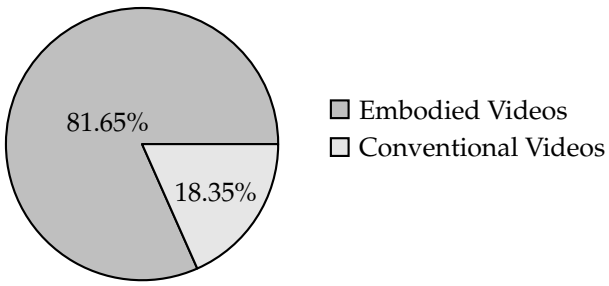


Figure 10. The distribution of the video user comments. Embodied learning videos account for the majority of comments (81.65%) compared to conventional videos (18.35%).

3.1.2. Sentiment Distribution of Video Comments

The sentiment distribution analysis between embodied and conventional videos reveals significant differences across various sentiment categories, as shown in Table 2. For embodied learning videos, positive sentiments account for 35.31% of all comments. In contrast, in conventional learning videos, positive comments account for 40.17% of all comments. Conversely, the negative comments on the embodied learning videos account for 9.57% compared to 12.43% for the conventional learning videos. Neutral sentiments also display a substantial disparity between the embodied learning videos (55.13%) and conventional learning videos (47.4%). These results suggest that the embodied learning videos are not associated with more positive user comments than the conventional learning videos; however, embodied learning videos show a reduction in negative comments and an increase in neutral user comments, which thus shows more balanced learning experiences in robotics.

Table 2. The comparison of user comment sentiments towards conventional and embodied YouTube learning videos.

Learning Type	Positive (%)	Negative (%)	Neutral (%)
Conventional	40.17	12.43	47.40
Embodied	35.31	9.56	55.13

3.1.3. Correlation Between Video Transcription and Video Comments

The analysis of the correlation between video transcription and comments reveals distinct patterns for embodied learning versus conventional learning videos, as shown in Figure 11. The results show that the use of the embodied learning videos yields higher correlations between the the video transcription and summary and video comments. Specifically, 28.1% of user comments are found to be correlated with the video transcription. In contrast, only 20% of user comments are found to be correlated with the input of the model, as shown in Figure 11. These findings suggest that embodied videos tend to generate a higher number of correlated comments compared to conventional videos, indicating potentially higher engagement and relevance of the video content to the viewers' comments.

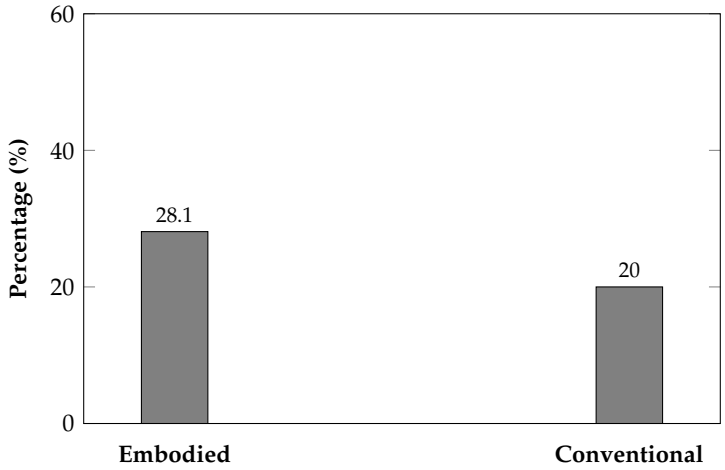


Figure 11. The correlation between video transcription and video comments (in percentages).

3.1.4. Logical Reasoning of Comments on Videos

The logical reasoning of the comments on the videos is shown in Figure 12. It shows that embodied learning videos have a higher percentage of logical reasoning than the conventional learning videos. Specifically, embodied learning videos have 83.57% comments

that are associated with logical reasoning. In contrast, conventional learning videos have 79.76% of comments that are associated with logical reasoning.

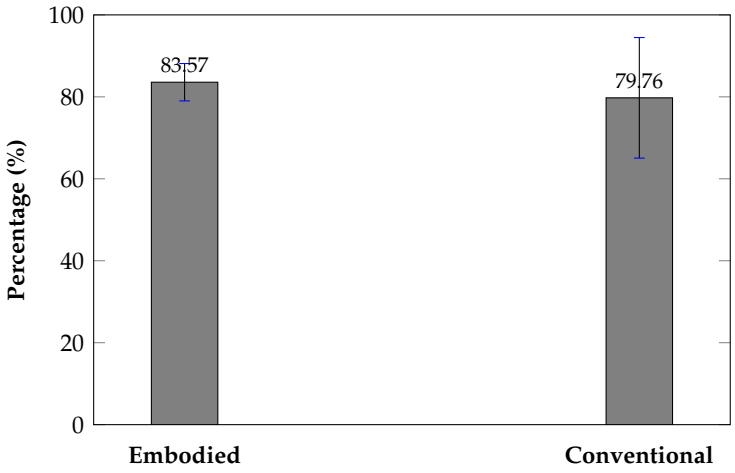


Figure 12. The logical reasoning of the user comments towards the embodied learning and conventional learning videos (in percentages).

3.2. Results of the User Study

3.2.1. Sentiment Distribution of the Feedback on Videos for the User Study

The results of sentiment analysis are shown in Table 3. It shows that there are both positive and neutral sentiment responses to the videos. Specifically, in contrast to the conventional learning videos, students show higher positive sentiment toward the embodied learning videos (22.16% versus 18.46%). However, students show stronger neutral sentiment toward the conventional learning videos (81.54% versus 77.84%). This shows that students have not shown negative sentiment responses toward both embodied learning videos and conventional learning videos. It therefore shows that the videos have not triggered negative sentiment responses from students.

Table 3. Sentiment responses of students toward the embodied learning and conventional learning videos.

Video Type	Positive Sentiment (%)	Neutral Sentiment (%)
Embodied	22.16	77.84
Conventional	18.46	81.54

3.2.2. Logical Reasoning of Videos for the User Study

The logical reasoning analysis of the feedback from students in the user study is shown in Figure 13. It shows that the embodied learning videos (48.27%) are associated with stronger logical reasoning in comparison with conventional learning videos (43.19%). These results are consistent with the results of the logical reasoning analysis based on the comments of students on the Youtube videos, as shown in Figure 12.

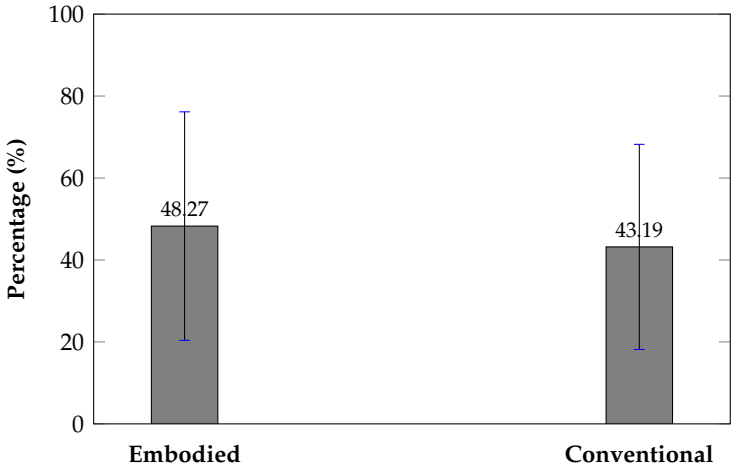


Figure 13. The logical reasoning of the student feedback toward the embodied learning and conventional learning videos (in percentages).

4. Conclusions

The primary goal of this study is to evaluate the effectiveness of user engagement with the embodied learning videos compared to the conventional learning videos in enhancing student engagement and educational outcomes in robotics education. Specifically, our research results are able to answer the three research questions. Overall, the research shows that students have presented engagement with the embodied learning videos. User comments on the videos are related to the video content, consistent with previous research [37]. Similarly, the embodied learning videos are shown to be able to enhance the logical reasoning of users [38,39]. We are able to answer the research questions through the comparative analysis of systematically examining video user comments, sentiments, and correlations with the video transcript and summary. This study utilized advanced data extraction, transcription, sentiment analysis, and statistical techniques to ensure accurate and reliable results, providing valuable insights into the effectiveness of these educational methodologies.

The analysis of the user comments revealed a significant difference in engagement between embodied learning and conventional learning videos. Embodied learning videos received more comments than conventional learning videos. Since we selected videos randomly, this disparity suggests that embodied learning videos maybe generally be more popular than conventional learning videos on Youtube. The higher number of user comments indicates that users are more involved and responsive to content that involves embodied learning activities such as physical and hands-on engagement throughout the learning process. This aligns with the principles of embodied learning, which emphasize active participation and physical interaction as key factors throughout the learning process for user engagement.

The sentiment distribution analysis further highlighted the differences between embodied and conventional learning of robotics. Embodied learning videos attracted more neutral comments than the conventional learning videos. Meanwhile, the embodied learning videos had fewer negative comments than the embodied learning videos. However, embodied learning videos had a lower percentage of significant positive comments than the conventional learning videos. These findings suggest that embodied learning videos do not only engage more viewers but also elicit less negative sentiment toward the videos, indicating that a higher level of physical engagement is overall effective in facilitating the learning of robotics [38,40,41].

The correlation analysis between user comments and video transcripts and summaries provide insights into how users interact with different types of robotics learning content.

Embodied learning videos have a more significant number of correlated comments than conventional learning videos, indicating that viewers are more engaged with specific details in the content. For example, videos involving the use of robotics electronics such as motor drivers have been able to attract a significant amount of user engagement. In contrast, conventional learning videos have fewer correlated comments, suggesting that users are less likely to engage deeply with the learning process. Specifically, it appears that users can better memorize the embodied learning content in the videos and this therefore shows a correlation with the content in the video comments, enabling more correlated video comments.

Users commenting on embodied learning videos show stronger logical reasoning capability based on their comments in contrast to those commenting on the conventional learning videos. This shows that the embodied learning videos may have more engaging learning content for engaging the logical reasoning of the users in comparison with the conventional learning videos. This therefore suggests that more careful embodied leaning activities should be considered in videos to better engage users to enhance their logical reasoning capability.

Our user study has also shown stronger user engagement with embodied learning videos, where students made more positive comments on the embodied learning videos than the conventional learning videos. In addition, the user study showed that students do not exhibit strong negative sentiments for either embodied learning videos or conventional learning videos. Similarly, students show stronger logical reasoning capabilities for the embodied learning videos than the conventional learning videos, therefore suggesting the usefulness of the embodied learning elements in robotics education.

The findings from this study have important implications for the field of robotics education. The higher engagement, positive stronger user engagement, and reaction to the embodied learning instruction suggest that incorporating embodied learning principles can significantly enhance educational outcomes. By involving students in physical activities and hands-on experiences, educators can foster deeper understanding and retention of complex concepts. This approach aligns with educational theories such as Piaget's constructivism and Vygotsky's social constructivism, which emphasize the importance of active engagement and social interaction in the learning process [42–44].

While this study provides valuable insights into the benefits of embodied learning, several challenges need to be addressed for its widespread implementation. These include the need for specialized equipment, teacher training, and curriculum design. Future research should focus on exploring the long-term impacts of embodied learning, identifying best practices for its implementation, and leveraging emerging technologies to enhance its effectiveness. Additionally, studies should investigate the potential applications of embodied learning in other educational fields beyond robotics to determine its broader applicability and benefits.

The results of the user study are comparable to the previous study showing the effectiveness of embodied learning for learning robotics [38,45,46]. While previous research has not studied logical reasoning and its connection with embodied learning, our study has shown that embodied learning is helpful to improve the logical reasoning of learners. The analysis of user comments on Youtube videos has shown that the use of embodied learning is able to reduce negative user experiences through the learning process, which is consistent with the previous studies showing the importance of embodied learning in facilitating the engagement of learners in the underlying topics [47–49].

The results of this study are also insightful for clarifying the benefits of using videos in robotics education. In the user study, students did not voice concerns about using videos for learning robotics technology. As such, students did not provide negative feedback through the learning process. In contrast, students were engaged in the video-based robotics

learning process and requested more videos that we have initially planned. While some neutral comments toward the videos were provided from students, the use of embodied learning videos was more effective in engaging students, obtaininga greater percentage of positive comments than neutral comments. Such results reinforce the previous research findings showing that the use of videos is effective in facilitating the underlying learning subjects [50–52]. As such, our results suggest that the use of videos is important to engage students' interests in learning robotics. In particular, the use of embodied learning is effective in improving user learning experiences.

In conclusion, this research demonstrated that embodied learning, characterized by hands-on activities and physical interaction, significantly enhances student engagement and educational outcomes in robotics education. By fostering active participation and emotional involvement, embodied learning methods can transform educational practices and improve learning experiences. The findings from this study provide a strong foundation for further exploration and implementation of embodied learning in engineering education. This research is limited in the number of participants and videos used through the study. Future work will include the quantification of users' critical thinking and creativity through the use of videos to understand the effectiveness of embodied learning for learning robotics. The future research will also examine gender differences in learning robotics using embodied learning techniques.

Author Contributions: Conceptualization, H.Z. and G.R.; Data Curation, Methodology, and Software, B.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported through the National Science Foundation, Division of Engineering Education and Centers, grant/award number: 2306285.

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki and the protocol was approved by the Ethics Committee of MTSU Institutional Review Board (IRB-FY2023-218).

Informed Consent Statement: Informed consent for participation was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wakefield, E.M.; Goldin-Meadow, S. How gesture helps learning: Exploring the benefits of gesture within an embodied framework. In *The Body, Embodiment, and Education*; Routledge: Abingdon, UK, 2021; pp. 118–135.
2. Chen, N.S.; Fang, W.C. Gesture-based technologies for enhancing learning. In *The New Development of Technology Enhanced Learning: Concept, Research and Best Practices*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 95–112.
3. Zhang, H.; Rajan, R. How Effective Embodied Learning for Robotics Instruction. In Proceedings of the 2025 ASEE Annual Conference & Exposition, Montreal, QC, Canada, 22–25 June 2025.
4. Stolz, S.A. Embodied learning. *Educ. Philos. Theory* **2015**, *47*, 474–487. [CrossRef]
5. Roth, W.M. Gestures: Their role in teaching and learning. *Rev. Educ. Res.* **2001**, *71*, 365–392. [CrossRef]
6. Rho, E.; Chan, K.; Varoy, E.J.; Giacaman, N. An experiential learning approach to learning manual communication through a virtual reality environment. *IEEE Trans. Learn. Technol.* **2020**, *13*, 477–490. [CrossRef]
7. Shakroum, M.A.; Wong, K.W.; Fung, C. The effectiveness of the gesture-based learning system (GBLS) and its impact on learning experience. *J. Inf. Technol. Educ. Res.* **2016**, *15*, 191–210. [CrossRef] [PubMed]
8. Wang, X.-l.; Eberhard, L.; Bernas, R. Gestures can enhance the narrative ability of children with learning disabilities. *Int. J. Literacies* **2014**, *20*, 1. [CrossRef]
9. Cermak, S.A.; Coster, W.; Drake, C. Representational and nonrepresentational gestures in boys with learning disabilities. *Am. J. Occup. Ther.* **1980**, *34*, 19–26. [CrossRef]

10. Ascari, R.E.S.; Silva, L.; Pereira, R. Personalized gestural interaction applied in a gesture interactive game-based approach for people with disabilities. In Proceedings of the 25th International Conference on Intelligent User Interfaces, Cagliari, Italy, 17–20 March 2020; pp. 100–110.
11. Shakroum, M.; Wong, K.W.; Fung, C.C. The influence of gesture-based learning system (GBLS) on learning outcomes. *Comput. Educ.* **2018**, *117*, 75–101. [CrossRef]
12. Freeman, S.; Eddy, S.L.; McDonough, M.; Smith, M.K.; Okoroafor, N.; Jordt, H.; Wenderoth, M.P. Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 8410–8415. [CrossRef]
13. Kontra, C.; Goldin-Meadow, S.; Beilock, S.L. Embodied learning across the life span. *Top. Cogn. Sci.* **2012**, *4*, 731–739. [CrossRef]
14. National Research Council. *How Students Learn*; Donovan, S., Bransford, J., Eds.; National Academies Press: Washington, DC, USA, 2005.
15. Newstead, S.E.; Bradon, P.; Handley, S.J.; Dennis, I.; Evans, J.S.B. Predicting the difficulty of complex logical reasoning problems. *Think. Reason.* **2006**, *12*, 62–90. [CrossRef]
16. Anderson, M.L. Embodied cognition: A field guide. *Artif. Intell.* **2003**, *149*, 91–130. [CrossRef]
17. Barsalou, L.W. Grounded cognition. *Annu. Rev. Psychol.* **2008**, *59*, 617–645. [CrossRef] [PubMed]
18. Michelson, E. Re-membering: The return of the body to experiential learning. *Stud. Contin. Educ.* **1998**, *20*, 217–233. [CrossRef]
19. Giannakos, M.N. Exploring the video-based learning research: A review of the literature. *Br. J. Educ. Technol.* **2013**, *44*, E191–E195. [CrossRef]
20. Yousef, A.M.F.; Chatti, M.A.; Schroeder, U. The state of video-based learning: A review and future perspectives. *Int. J. Adv. Life Sci.* **2014**, *6*, 122–135.
21. Vural, O.F. The impact of a question-embedded video-based learning tool on E-learning. *Educ. Sci. Theory Pract.* **2013**, *13*, 1315–1323.
22. Nadeak, B.; Naibaho, L. Video-based learning on improving students' learning output. *PalArchs J. Archaeol. Egypt/Egyptol.* **2020**, *17*, 44–54.
23. Paivio, A. *Mental Representations: A Dual Coding Approach*; Oxford University Press: Oxford, UK, 1990.
24. Mayer, R.E. Multimedia learning. In *Psychology of Learning and Motivation*; Elsevier: Amsterdam, The Netherlands, 2002; Volume 41, pp. 85–139.
25. Dede, C. *The Role of Digital Technologies in Deeper Learning. Students at the Center: Deeper Learning Research Series*; Jobs for the Future: Boston, MA, USA, 2014.
26. Poria, S.; Cambria, E.; Bajpai, R.; Hussain, A. A review of affective computing: From unimodal analysis to multimodal fusion. *Inf. Fusion* **2017**, *37*, 98–125. [CrossRef]
27. Radford, A.; Kim, J.W.; Xu, T.; Brockman, G.; McLeavey, C.; Sutskever, I. Robust speech recognition via large-scale weak supervision. In Proceedings of the International Conference on Machine Learning, PMLR, Honolulu, HI, USA, 23–29 July 2023; pp. 28492–28518.
28. Schmitz, M.; Soderland, S.; Bart, R.; Etzioni, O. Open language learning for information extraction. In Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, Jeju Island, Republic of Korea, 12–14 July 2012; pp. 523–534.
29. Wu, F.; Weld, D.S. Open information extraction using wikipedia. In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Uppsala, Sweden, 11–16 July 2010; pp. 118–127.
30. Summarizer, Y NoteGPT—YouTube Video Summarizer with AI, 2026. Available online: https://notegpt.io/ (accessed on 5 February 2026).
31. GPT-3.5, O. GPT-3.5 Turbo Fine-Tuning and API Updates, 2023. Available online: https://benched.ai/guides/fine-tuning-gpt-3-5 (accessed on 5 February 2026).
32. Talaviya, A. A Complete Guide to Building Streamlit Data Web Apps with Code Examples, 2022. Available online: https://tinyurl.com/yfumpz3e (accessed on 5 February 2026).
33. Pérez, J.M.; Giudici, J.C.; Luque, F. pysentimiento: A python toolkit for sentiment analysis and socialnlp tasks. *arXiv* **2021**, arXiv:2106.09462v3.
34. Devlin, J. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
35. Tan, K.L.; Lee, C.P.; Anbananthen, K.S.M.; Lim, K.M. RoBERTa-LSTM: A hybrid model for sentiment analysis with transformer and recurrent neural network. *IEEE Access* **2022**, *10*, 21517–21525. [CrossRef]
36. Beno, J.P. ELECTRA and GPT-4o: Cost-Effective Partners for Sentiment Analysis. *arXiv* **2024**, arXiv:2501.00062.
37. Bhuiyan, H.; Ara, J.; Bardhan, R.; Islam, M.R. Retrieving YouTube video by sentiment analysis on user comment. In *Proceedings of the 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*; IEEE: New York, NY, USA, 2017; pp. 474–478.
38. Macedonia, M. Embodied learning: Why at school the mind needs the body. *Front. Psychol.* **2019**, *10*, 2098. [CrossRef] [PubMed]

39. Smyrnaiou, Z.; Sotiriou, M.; Sotiriou, S.; Georgakopoulou, E. Multi-Semiotic systems in STEMS: Embodied Learning and Analogical Reasoning through a Grounded-Theory approach in theatrical performances. *J. Res. STEM Educ.* **2017**, *14*, 99–112.
40. Georgiou, Y.; Ioannou, A. Embodied learning in a digital world: A systematic review of empirical research in K-12 education. In *Learning in a Digital World: Perspective on Interactive Technologies for Formal and Informal Education*; Semantic Scholar: Seattle, WA, USA, 2019; pp. 155–177.
41. Zhong, B.; Su, S.; Liu, X.; Zhan, Z. A literature review on the empirical studies of technology-based embodied learning. *Interact. Learn. Environ.* **2023**, *31*, 5180–5199. [CrossRef]
42. Kirtay, M.; Chevalère, J.; Lazarides, R.; Hafner, V.V. Learning in social interaction: Perspectives from psychology and robotics. In *Proceedings of the 2021 IEEE International Conference on Development and Learning (ICDL)*; IEEE: New York, NY, USA, 2021; pp. 1–8.
43. Daniels, H. Introduction: Psychology in a social world. In *An Introduction to Vygotsky*; Routledge: Abingdon, UK, 2002; pp. 13–39.
44. Utoyo, S.; Ismaniar, I.; Hazizah, N.; Handrianto, C. Validating the kinesthetic play model: A quantitative study on enhancing early mathematical skills in Indonesian preschoolers. *Int. J. Innov. Res. Sci. Stud.* **2025**, *8*, 1440–1449. [CrossRef]
45. Munro, M. Principles for embodied learning approaches. *SATJ S. Afr. Theatre J.* **2018**, *31*, 5–14. [CrossRef]
46. Jusslin, S.; Korpinen, K.; Lilja, N.; Martin, R.; Lehtinen-Schnabel, J.; Anttila, E. Embodied learning and teaching approaches in language education: A mixed studies review. *Educ. Res. Rev.* **2022**, *37*, 100480. [CrossRef]
47. Liu, Z.; Zuo, H.; Zhao, Y.; Lu, Y. The effect of embodied learning on students' learning performance: A meta-analysis. *Front. Psychol.* **2025**, *16*, 1658797. [CrossRef]
48. Dixon, M.; Senior, K. Appearing pedagogy: From embodied learning and teaching to embodied pedagogy. *Pedagog. Cult. Soc.* **2011**, *19*, 473–484. [CrossRef]
49. Ioannou, M.; Ioannou, A. Technology-enhanced embodied learning. *Educ. Technol. Soc.* **2020**, *23*, 81–94.
50. Marsh, B.; Mitchell, N. The role of video in teacher professional development. *Teach. Dev.* **2014**, *18*, 403–417. [CrossRef]
51. Bell, L.; Bull, G. Digital video and teaching. *Contemp. Issues Technol. Teach. Educ.* **2010**, *10*, 1–6.
52. Pisarenko, V.; Krasnoshchekova, G. Video in teaching. In *Proceedings of the 2016 IEEE 10th International Conference on Application of Information and Communication Technologies (AICT)*; IEEE: New York, NY, USA, 2016; pp. 1–6.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.