IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 24 May 2024, accepted 16 July 2024, date of publication 16 August 2024, date of current version 4 October 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3444783

RESEARCH ARTICLE

The Impact of LLM Hallucinations on Motor Skill Learning: A Case Study in Badminton

YEPENG QIU
College of Physical Education and Health Sciences, Chongqing Normal University, Chongqing 401331, China
(e-mail: knyf20@163.com)

This work was supported in part by the Science and Technology Research Program of Chongqing Municipal Education Commission under Grant KJQN202200561; and in part by the Natural Science Foundation of Chongqing, China, under Grant CSTB2023NSCQ-MSX0698.

ABSTRACT The rise of Generative Artificial Intelligence, including Large Language Models (LLMs), has enabled users to engage in self-guided learning of sports skills through conversation-based interactions. However, studies have identified a phenomenon known as "hallucination" in which LLMs generate feedback that is inaccurate or non-existent. While this phenomenon has been observed in various domains, including medicine, academia, and news, its existence and implications in the context of physical exercises, particularly motor skill learning, remain unexplored. This study investigates the presence of LLM hallucinations in badminton skill learning and examines their potential impact on learning outcomes. This study aims to investigate whether LLMs hallucinations exist in the motor skill learning of physical exercises and what impact they may have. Eighty university freshmen with no prior badminton experience participated in a 16-week experiment, with 40 students assigned to the Experimental Group (EG) utilizing LLM-based applications (ChatGPT or New Bing) for self-guided learning, and 40 students in the Control Group (CG) learning under the supervision of 12 university sports teachers and 8 experts that specialized in badminton. Evaluation criteria for badminton skills were established, and assessments were conducted at baseline and 16 weeks using independent sample t-tests and paired-sample t-tests. One-way analysis of variance (One-Way ANCOVA) was employed to compare learning outcomes between the two groups. Interviews were conducted to gain insights into the causes of any observed differences in learning efficiency. Both CG and EG groups demonstrated motor skill improvement (clear: $p < 0.001$; smash: $p < 0.001$; footwork: $p < 0.001$). CG exhibited significantly higher scores in long-distance shots and smashes in the post-test. No significant difference was observed in footwork scores between the two groups. High accordance in specific skill points among students in both groups indicated the common usage of prompts. Interviews with EG students revealed hallucinations in the text generated by LLMs, particularly in the context of "forearm internal rotation swing." LLMs exhibit hallucinations in the context of intricate motor skill learning, such as badminton, where limited corpus data is available. These hallucinations can mislead users and impact learning outcomes. Future research should explore strategies to mitigate LLM hallucinations in physical exercise learning applications.

INDEX TERMS Large language models, hallucination, motor skill learning, badminton skill.

I. INTRODUCTION

The integration of technology and Artificial Intelligence (AI) in physical education has become a burgeoning field of study. Scholars have posited that technology can enhance motor skill acquisition through personalized instruction and immediate feedback mechanisms [60], [61], [62]. AI, in particular, has been recognized for its potential to create immersive learning environments that can simulate real-world sports scenarios, thereby providing students with a rich and interactive learning experience [63].

Our comprehensive review of the literature has underscored the prevalence of technology and AI applications in physical education. Studies have demonstrated that technology can individualize training programs, offering tailored feedback that adapts to the learner's progress in real-time

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Ni.

[64]. Furthermore, the use of AI in physical education has been shown to foster inclusive education by catering to diverse learning styles and abilities [65].

Recent global surveys highlight ''online fitness'' as the predominant fitness trend of 2021, as indicated by Thompson [51]. The advent of cutting-edge artificial intelligence (AI) technologies has empowered educators to leverage media websites and online learning systems, enabling users to learn and exercise independently [24], [28], [43]. Notably, social media platforms such as Facebook, YouTube, and social network websites (SNSs) have evolved into widespread sources for individuals to acquire sports knowledge and skills [23].

Nevertheless, despite the convenience of online learning in terms of time and space, it has demonstrated adverse effects on real-life exercise processes [57]. Fitness videos, for instance, may lack the precision in movements achieved through immediate interactions with coaches, potentially leading to misunderstandings or errors that impede the learning outcomes [22], [33]. Furthermore, resources on effectively utilizing social media for promoting physical activity and learning physical skills are still limited [12].

Identifying an appropriate social media platform for high-quality video content poses challenges. Currently, there is a lack of research and data collection on the quality of online sports videos. Publishers can freely upload videos without undergoing professional audits, and the casual nature of content posting risks providing users with misleading and inaccurate information [3]. Factors such as the identity, quality, and effectiveness of the video's author are critical considerations [2].

Over the decades, search engines like Google or Bing have become indispensable tools for retrieving information. With the advent of Large Language Models (LLMs), there has been a significant breakthrough in enhancing users' learning and comprehension abilities. LLMs, trained on vast amounts of text data from the internet, employ a transformer architecture with hundreds of billions of parameters [14], [30], [45]. These models aim to grasp semantic meaning and intent, delivering structured and organized content based on human feedback and making adjustments accordingly [13], [41], [44], [47] (Sanh et al.; Wei et al.)

Microsoft's Bing Chat, built on LLM technology, represents an advancement over traditional search engines. It serves as both a search engine and a chatbot, offering direct internet access and providing well-informed responses [35]. Bing Chat has expanded its capabilities to support image and video searches, allowing users to input images and videos related to sports skills. The platform then provides web links for users to directly access images and videos for learning sports skills, enabling the integration of theoretical knowledge with practical feedback, such as basketball shooting techniques or deltoid training methods.

The utilization of Large Language Models (LLMs) involves significant risks associated with misinformation, biases, and unethical practices embedded in the vast internet-derived text corpus upon which these models are trained. Existing LLM-based search engines also face challenges in providing accurate feedback, as highlighted by [34]. LLMs exhibit a tendency to fabricate information, a phenomenon often termed as ''hallucinations'' [10]. Concerns have been raised by Weidinger L, Mellor J, Rauh M, et al., regarding the potential dissemination of false or incomplete information by LLMs, leading to incorrect learning outcomes, especially among less informed users. This misinformation poses risks, such as offering erroneous legal or medical advice [54]. In some instances, the generated information may even encourage users to engage in unethical or illegal activities they would not otherwise consider.

The issue of ''hallucination'' in LLMs is evident across various domains. In the medical and health field, for instance, patients may struggle to discern spurious content generated by LLMs, leading to biased outputs that could result in incorrect diagnoses or suboptimal treatment recommendations, potentially harming patients or causing delays in appropriate care. While LLMs can generate credible scientific papers, the data they produce may consist of a blend of authentic and entirely fabricated information, raising concerns about the reliability and accuracy of using LLMs in academic writing. In areas such as news or advertising, LLM-generated misinformation can take the form of fake news or misleading social media posts, contributing to issues of discrimination and potential risks [1], [17], [36].

Beyond Natural Language Processing (NLP), ''hallucination'' is a psychological term referring to a specific type of perception [86], [87]. LLMs like ChatGPT and GPT-4 often generate content that is meaningless or unreal, resembling the characteristics of the psychological phenomenon known as LLM hallucination [7], [35], [58], [59]. Min Zhang and Juntao Li have highlighted significant problems with LLMs, including a lack of consistent understanding of natural language generation and the production of uncontrollable content. GPT-3's training on internet text data containing misinformation and bias often results in the generation of content filled with biased information. Dale suggests that LLM output may contain assertions inconsistent with the truth and produce irrelevant language fragments. For example, ChatGPT and Bard can misrepresent their capabilities, providing entirely fabricated publication lists when asked to cite relevant publications in a review paper [37], [88]. While LLMs may produce information that appears credible and realistic, it can be otherwise incorrect or unverified [58], [89]. Although LLMs for information retrieval (IR) seem promising, they face substantial challenges related to low credibility, providing misleading explanations, incorrect answers, or fabricated information [32].

The integration of AI technology in sports and exercise is experiencing a consistent upward trend, with the postulation that employing AI can significantly enhance students' motivation for learning and improve overall learning efficiency concerning sport-related concepts and skills [15], [52]. These advancements empower students to independently delve into the study and practice of complex motor skills, breaking free from the constraints of traditional course schedules [39].

Nevertheless, it is crucial to acknowledge that mastering sports skills often involves executing a series of intricate body movements within seconds, accompanied by strict technical requirements. Beginners may find it challenging to comprehend the conceptual and technical nuances presented in video content, not to mention the seamless continuity and fluidity of the movements. In such cases, learners often need to seek additional resources that offer more detailed information [48].

The emergence of Large Language Models (LLMs) provides a convenient and straightforward avenue for learning in the sports and exercise domain. However, the outcomes of applying LLMs in this field are still largely unexplored [10]. It remains unknown whether LLMs can potentially induce hallucinations in sports and exercise processes, particularly in scenarios where users are learning targeted skills independently. Comprehensive theoretical research in this regard is currently lacking, highlighting a crucial area for further investigation.

II. THEORETICAL FRAMEWORK

Building upon the extensive literature review, this section delves into the theoretical constructs that underlie our investigation of LLMs' impact on motor skill learning within the context of badminton. The theoretical framework is pivotal in establishing a conceptual roadmap for understanding how AI can be integrated into physical education to enhance learning outcomes.

Cognitive Learning Theory suggests that learners construct knowledge through the interaction of existing knowledge and new information [66]. In the case of motor skill learning, cognitive processes are integral to the initial stages of understanding and replicating complex movements [67].

Constructivism posits that learning is an active process where the learner constructs new ideas based on experience, social interaction, and the existing knowledge framework [68]. This theory is particularly relevant to our study as it emphasizes the role of AI in facilitating a constructivist learning environment where learners can explore and practice badminton skills through interactive AI-based applications.

Social Cognitive Theory, as proposed by [69], extends the constructivist view by highlighting the importance of observation, modeling, and reinforcement. In the context of AI and motor skill learning, this theory underscores the potential of AI to provide models and simulate social interactions that can enhance learners' ability to perform and refine motor skills.

Moreover, the Zone of Proximal Development (ZPD) concept by [70] offers a framework for understanding how AI can support learners in their journey from novice to expert. AI can be tailored to provide instruction and challenges that are just beyond the learner's current capabilities but within reach with guidance [70]. The integration of AI in physical education is further informed by Self-Determination Theory [71], which emphasizes the importance of intrinsic motivation and autonomy in the learning process. AI has the potential to create autonomous learning environments that cater to individual learners' needs and interests, thereby enhancing their intrinsic motivation to learn and master motor skills [72].

In the context of our study, we consider how LLMs, as an AI application, can be leveraged to provide personalized feedback and learning experiences that align with these theoretical principles. We anticipate that AI's ability to adapt to individual learner's paces and styles will contribute positively to motor skill learning in badminton.

This theoretical framework provides a foundation for our methods section, where we detail the design and implementation of our AI-enabled intervention in a badminton learning setting.

III. RESEARCH PURPOSE AND QUESTIONS

The primary purpose of this study is to investigate the impact of Large Language Model (LLM) hallucinations on motor skill learning, with a specific focus on badminton as a case study. This research aims to address several key questions that will guide our investigation and data analysis.

Research Question 1: To what extent do LLM hallucinations occur during the self-guided learning of badminton skills through AI-based applications?

Research Question 2: How do these hallucinations, if present, affect the learning outcomes and motor skill development in beginners?

Research Question 3: Can the presence of hallucinations be mitigated through specific instructional design strategies in AI applications?

Research Question 4: What are the implications of LLM hallucinations for the broader field of physical education and technology-assisted learning?

By posing these research questions, we aim to systematically explore the influence of AI-generated feedback on the learning process and to identify potential solutions to enhance the effectiveness of AI in physical education. Our study contributes to the existing body of knowledge by providing empirical evidence on the intersection of AI, hallucinations, and motor skill acquisition.

This clear statement of purpose and these guiding research questions set the stage for our methods section, where we detail the experimental design and procedures used to address each question and test our hypotheses.

IV. MATERIALS AND METHODS

A. SETTING

The research was conducted within the dynamic badminton sports environment of a university in Southwest China as part of a university-based program in 2023. The program, spearheaded by a Physical Education (PE) professor, is supported by funding from the National Science Foundation.

TABLE 1. Badminton skills learning effect assessment standards.

Skill	Indicator	16pts-20pts	11pts-15pts	6pts-10pts	1pts-5pts
Clear	Contact is made high with a straight elbow in front of body, weight is shifted to front foot as contact is made, racket face is slightly open, follow through is across the body, shuttle lands high and deep inbounds	exemplary	proficient	developing	struggling
	The clear shots consistently drop steeply to the baseline.	exemplary	proficient	developing	struggling
Smash	Smash is taken no further than 3/4 court, mechanics look like the overhead clear, contact is made above the head, in front of the body, with straight elbow. Racket face is pointed downward, follow through is downward across the body.	exemplary	proficient	developing	struggling
	The smash shots travel on a powerful and downward trajectory to the opponent's side The performer hits the shuttle at its highest point and performs a full swing	exemplary	proficient	developing	struggling
Footwork	Start steps and footwork are simple and smooth, fast and rhythmic	exemplary	proficient	developing	struggling

Throughout the duration of this study, the program was actively engaged in providing services to a cohort of 80 university students.

B. PARTICIPANTS

This study encompassed 80 participants who were freshmen from Chongqing Normal University, representing diverse majors across various colleges within the university (56 males and 24 females, averaging 18.06 years of age). Additionally, there were 12 university sports teachers specializing in badminton, comprising 4 associate professors and 8 lecturers, with an average age of 37.75 years and a mean teaching experience of 11.8 years. The expert group, consisting of 8 members, included 2 retired athletes from the university (former champions in men's doubles at the Chongqing City University Games), 2 vice presidents of the Chongqing Badminton Association (with over 20 years of experience in badminton-related activities), and 4 chief referees for the Chongqing Badminton Association's public badminton level tests (possessing extensive experience in evaluating badminton skills). While the sports teachers took on the roles of teaching and supervision, the expert group focused on assessing the questionnaire and conducting related test assessments.

C. SAMPLE SIZE JUSTIFICATION

The sample size was determined based on the desired level of statistical power (1-β), the expected effect size, and the significance level (α). We aimed for a power of 0.80, which is conventionally considered sufficient to detect an effect if it exists [82].

Our sample consisted of 80 university freshmen with no prior badminton experience, which is comparable to similar studies in the field of motor skill learning. This number allowed us to detect medium to large effect sizes, which are common in educational and motor learning research [83]. Additionally, we ensured that our sample size was in line with the recommendations provided by the American Psychological Association for conducting statistical analyses (APA, 2020).

To further validate our sample size, we conducted post-hoc power analyses using G*Power software [83]. The results confirmed that our sample size was adequate to detect the effects observed in our study with a high degree of statistical power.

We acknowledge that while our sample size is sufficient for the current study, future research with larger and more diverse samples may provide additional insights into the impact of LLM hallucinations on motor skill learning across different populations and sports.

D. AI-ENABLED BADMINTON INTERVENTION

To address the need for a detailed account of our AI-enabled badminton intervention, we introduce this subsection. The intervention was designed to integrate AI technology into a 16-week badminton course for university students with no prior experience in the sport. The AI system, developed in partnership with sports education experts and AI specialists, was tailored to provide real-time feedback and personalized guidance to students as they practiced their badminton skills.

The AI-enabled intervention was built upon the principles of adaptive learning and scaffolding, allowing the system to adjust the complexity of tasks and feedback based on each student's performance [72], [73]. The system utilized computer vision to analyze students' movements and provide immediate, constructive feedback, focusing on key skills such as footwork, racket handling, and shot techniques.

A key feature of the AI intervention was its ability to generate customized training modules for each student, based on their progress and areas for improvement. This personalized approach aligns with the principles of differentiated instruction, which suggests that learners benefit from tailored educational experiences that match their readiness and learning profiles [74].

Throughout the intervention, students engaged with the AI system in a virtual training environment that simulated various badminton scenarios. The system's feedback was designed to complement, not replace, the guidance provided by human coaches, ensuring a hybrid learning approach that combined the best of human and artificial intelligence.

We ensured ethical considerations were upheld throughout the intervention, with informed consent obtained from all participants and adherence to data protection regulations [75].

E. INSTRUMENTS

The study's sample estimation primarily relied on the analysis of the "Integrating Play Practice Assessment of Badminton" [26] and "Badminton Skills Assessment Rubrics" (RCampus, 2020) scale models. To enhance the precision of motor skills assessment in badminton, we observed detailed motor skills and identified specific challenging points in skill acquisition. Subsequently, we refined the aforementioned models, creating two more advanced scales with intricate details, enabling a focused examination of motor skills nuances, such as "the snap of the wrist," "full forearm internal rotation," "hit point with body position," and more. To mitigate the potential for random learning outcomes, this study restricted students to using LLM-based applications, specifically ChatGPT and ERNIE Bot. No other LLM-based applications were permitted in this research.

As illustrated in Table 1 and Table 2, the Badminton Skills Learning Effectiveness Assessment Scale (BSLEAS) predominantly evaluates the technique and efficacy of actions like clears and smashes. It employs a four-level scoring system to assess the tester's performance. Conversely, the Badminton Motor Technique Error Analysis Scale (BMTEAS) emphasizes the critical and challenging facets of the badminton skills learning process,including common errors. These technical nuances are frequently overlooked in video resources and explanations. Consequently, subjects are more likely to comprehend and learn badminton skills effectively through textual materials.

F. RESEARCH DESIGN

In this study, students were categorized into two groups: the Control Group (CG, n = 40) and the Experimental Group (EG, n = 40). The Experimental Group students participated in self-directed learning of badminton skills primarily through recommended Large Language Model (LLM)-based applications. Conversely, the Control Group received guidance from university badminton teachers for their badminton skill acquisition.

To accommodate the unique characteristics of badminton skills, we refined and established criteria for evaluating the effectiveness of badminton skill learning and identified key technical points in badminton movements. The experimental procedure commenced with the organization of students into groups during the first week. Initial pre-tests were conducted to ensure no significant differences between groups and that students lacked professional badminton skills.

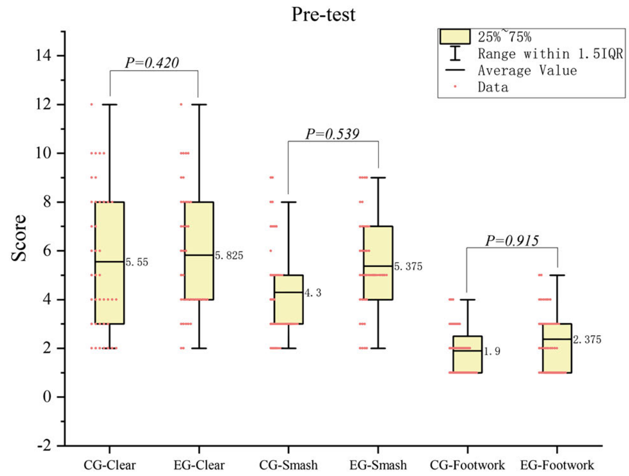

FIGURE 1. Independent sample t-test for Pre-test of badminton skills learning effect.

Starting from the second week, a 14-week badminton skill learning program was implemented, with a focus on clears, smashes, and footwork. In the 16th week, both groups completed post-tests for badminton skills, allowing for the assessment of skill acquisition and comparing the effectiveness of autonomous learning with LLM-based applications against traditional teacher-guided learning.

G. PHASE 1 (GROUPING AND PRE-TEST)

The 80 students were evenly distributed into two groups: a Control Group (CG, N = 40) and an Experimental Group (EG, N = 40). To ensure impartial assignment, both groups underwent the Badminton Skills Learning Effectiveness Assessment Scale (BSLEAS) pre-test. An independent sample t-test was then conducted on the scores of the two groups, as depicted in Figure 1.

H. PHASE 2(BADMINTON SKILLS LEARNING)

The BSLEAS and BMTEAS criteria were assigned to the Control Group (CG) and Experimental Group (EG), respectively, with instructions tailored to their specific evaluation content and key points. In the CG, instruction was delivered by PE teachers. This group was subdivided into 8 smaller groups, each comprising 5 students, and supervised by a PE teacher. CG participants underwent two 100-minute training sessions weekly, totaling 14 consecutive weeks within a semester.

Conversely, the EG utilized a specialized Large Language Model (LLM) platform for acquiring badminton skills. Similar to the CG, the EG was also divided into 8 groups, each led by a designated group leader. Each group leader was responsible for ensuring exclusive access to learning resources via the LLM platform, recording practice times, and monitoring platform usage. For every two groups, there was an appointed supervisor overseeing the group leaders' activities and monitoring the practice sessions of group members.

TABLE 2. Badminton movement technical essential assessment standards.

	Skill	Indicator	16pts-20pts	11pts-15pts	6pts-10pts	1pts-5pts
A	Clear grip	Maintain a correct grip during both preparation and hitting.	exemplary	proficient	developing	struggling
B	Side frame pat posture	The frame posture stretches, the elbows balance with the trunk, forming the chest and back arch posture	exemplary	proficient	developing	struggling
C	Push to turn around	Support foot pedal turn strong, turn smooth and coherent, moderate turn range.	exemplary	proficient	developing	struggling
D	Forearm internal rotation Swing	When swinging, the forearm is fully rotated, and the movement is relaxed and coherent.	exemplary	proficient	developing	struggling
E	Pound point position	The hitting point is above the head, the hitting point is high, and the hitting position is stable.	exemplary	proficient	developing	struggling
F	Whipping force	Whipping force coherent natural, whip strength.	exemplary	proficient	developing	struggling
G	Smash grip	Maintain a correct grip during both preparation and hitting.	exemplary	proficient	developing	struggling
H	The elevation of both elbows	Before the racket and shot, the height of the elbows is high and balanced with the torso.	exemplary	proficient	developing	struggling
I	90 degrees turn to the right/left side	When hitting the ball, the turn movement is coherent and smooth, and the turn range is appropriate, about 90 degrees	exemplary	proficient	developing	struggling
J	The snap of the wrist	When hitting the ball, the finger wrist is very strong to buckle the handle	exemplary	proficient	developing	struggling
K	Full forearm internal rotation, fast whip force	When hitting the ball, the forearm fully internal rotation, and fast whip force, explosive force.	exemplary	proficient	developing	struggling
L	Hit point with body position	The hitting point is above the head and above the front of the body, appropriate with the body position, and the hitting point is stable	exemplary	proficient	developing	struggling

EG participants, like CG, attended two 100-minute sessions weekly, spanning 14 consecutive weeks within one semester. Each session encompassed 20 minutes of theoretical study and 70 minutes dedicated to refining badminton skills.

I. PHASE 3(BADMINTON SKILLS POST-TEST)

After a dedicated period of learning and practice, a post-test will be administered to both the Control Group (CG) and Experimental Group (EG), adhering to the guidelines outlined in the BSLEAS. The post-test sessions will be captured using video equipment, facilitating expert evaluation and scoring of participants' technical skills in alignment with the BMTEAS criteria. Subsequent to the post-test, the recorded data will undergo comprehensive analysis to derive insights and assess the effectiveness of the learning interventions.

J. STATISTICAL METHODS

To assess the advancements in three badminton skills within the Control Group (CG) and Experimental Group (EG), a paired sample t-test was employed to analyze the variations between post-test and pre-test scores. Additionally, a one-way analysis of variance (ANCOVA) was carried out to determine whether there were noteworthy differences in the learning outcomes between the EG and CG. These statistical analyses aim to provide a comprehensive understanding of the impact of the learning interventions on the participants' badminton skills.

K. VARIABILITY ANALYSIS

The potential for variability in outcomes when using distinct LLMs is an important consideration, as it may reflect differences in the algorithms, training data, and design of these applications.

Our analysis involved a comparative study of the feedback provided by several LLM-based applications on the same set of motor skill learning prompts. We found that while there were commonalities in the feedback, there were also significant differences that could impact the learning process. For instance, some LLMs provided more detailed explanations and corrections, while others were less precise and occasionally generated misleading information.

To minimize the randomness in our outcomes, we implemented several strategies. Firstly, we ensured that the LLM applications used in our study were selected based on their demonstrated reliability and accuracy in previous studies. Secondly, we incorporated a human oversight mechanism to review and verify the feedback provided by the LLMs before it was presented to the learners. This process helped to ensure that the information was accurate and consistent, reducing the likelihood of learners receiving misleading feedback.

Furthermore, we suggest that future studies consider the development of standardized guidelines for evaluating and selecting LLM-based applications for use in motor skill learning. This could include criteria such as the application's adherence to educational standards, its ability to provide clear and accurate feedback, and its capacity to adapt to individual learner's needs.

L. ANALYSIS OF VARIABILITY IN LLM-BASED APPLICATIONS

This analysis aimed to identify any systematic differences in the feedback provided for the same motor skill learning prompts.

Comparative Analysis: We selected a set of representative prompts related to badminton skills and submitted them to four different LLM applications. The outputs were evaluated based on their accuracy, consistency, and adherence to established badminton coaching principles. Our results revealed that while minor variations existed in the level of detail and the style of explanation, the core technical advice remained consistent across platforms.

Quality Control Measures: To ensure the reliability of the feedback generated by LLM applications, we implemented a multi-tiered quality control process. This process involved: 1. Cross-referencing the LLM-generated advice with a database of peer-reviewed coaching literature to verify the accuracy of the technical information. 2. Subjecting the LLM outputs to a panel of expert reviewers, including certified badminton coaches and sports scientists, who assessed the applicability and clarity of the advice.

V. RESULTS

After a dedicated practice period, post-tests were administered for both the Control Group (CG) and Experimental Group (EG). Paired-sample t-tests were employed to assess the significance of differences between their pre-test and post-test scores. Following this, an analysis of covariance (ANCOVA) was conducted specifically on the clear test results of the CG and EG, encompassing both the kill and footwork components. This analysis aimed to determine whether a significant difference existed between the EG and CG in terms of badminton skill improvement.

Figure 1 illustrates that there was no statistically significant difference in pre-test performance between the CG and EG, as evidenced by Clear P-value = 0.420, Smash P-value = 0.539, and Footwork P-value = 0.420. This outcome indicates the absence of bias in the allocation of participants to the Control Group and Experimental Group.

Figure 2 succinctly presents the outcomes of the paired-sample t-tests, illustrating noteworthy enhancements in both groups, as evidenced by higher post-test scores compared to pre-test scores. This signifies improved badminton skills

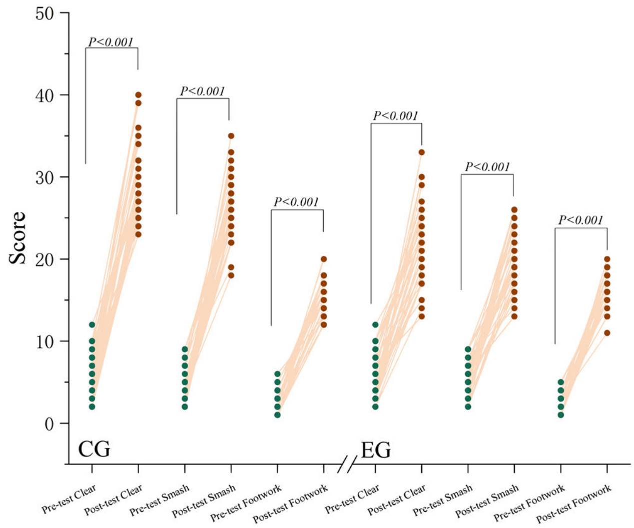

FIGURE 2. Paired-sample t-test results for CG and EG pre-and post-test scores of badminton skills learning effect.

among learners following instruction in various techniques. Before conducting the ANCOVA analysis, a homogeneity test for clear, smash, and footwork was executed. The significance associated with the Group * Pretest scores exceeded 0.05 in the inter-subject effect test, indicating homogeneity of within-group regression coefficients and allowing the progression of the one-way ANCOVA analysis.

Before conducting the one-way ANCOVA, a homogeneity test was executed for clear, smash, and footwork. The significance corresponding to the Group * Pretest scores exceeded 0.05 in the inter-subject effect test, indicating that the homogeneity of within-group regression coefficients was met, allowing the progression of the one-way ANCOVA analysis.

Table 3 elucidates significant differences between the Control Group (CG) and Experimental Group (EG) scores (clear: $F = 57.417$, $p < 0.001$; smash: $F = 56.798$, $p = 0.001$). The average posttest scores for CG in clear and smash were 29.825 and 26.125, respectively, while for EG, the scores in clear and smash were 22.100 and 19.850, respectively. This suggests that the average test score for CG in clear and smash is significantly better than that of EG in clear and smash.

The teaching method implemented by Physical Education (PE) teachers demonstrated a significant improvement in clear and smash compared to students' independent learning and practice method using the Large Language Model (LLM). However, there was no significant difference in posttest scores between CG and EG ($F = 3.410$, $p = 0.069$). The average post-test score for the method taught by PE teachers was 15.525, and the autonomous learning and practice method scored 16.325. This suggests that both the traditional teaching method and the LLM-based autonomous learning approach can effectively enhance footwork skills,

TABLE 3. The one-way ANCOVA table for Clear, Smash and Footwork.

Clear					
	Sum of squares	Degrees of freedom	Mean square	*F*	P
Inter block	1193.512	1	1193.512	57.417	<0.001
Within the group	1621.375	78	20.787		
Aggregate	2814.888	79			
Smash					
	Sum of squares	Degrees of freedom	Mean square	*F*	P
Inter block	787.512	1	787.512	56.798	<0.001
Within the group	1081.475	78	13.865		
Aggregate	1868.987	79			
Footwork					
	Sum of squares	Degrees of freedom	Mean square	*F*	P
Inter block	12.800	1	12.800	3.410	0.069
Within the group	292.750	78	3.753		
Aggregate	305.550	79			

with no significant difference in effectiveness between the two methods.

VI. THE EMERGENCE OF LLM HALLUCINATIONS IN MOTOR SKILL LEARNING

In the post-test, the outcomes of the aforementioned trials reveal a significantly diminished learning effect in the Experimental Group (EG) for clear and smash techniques compared to the Control Group (CG). This implies that the utilization of Large Language Models (LLM) for acquiring badminton sports skills may not yield particularly favorable results. The proficiency in key badminton movements appears to wield a substantial influence on the test scores, thereby impacting the overall efficacy of learning badminton skills. Accurate technical movements are prone to augmenting the effectiveness of shots, whereas imprecise or incorrect technical movements often detrimentally affect shot accuracy. The learning effect of clear and smash techniques in the EG is notably inferior to that observed in the CG. This distinction can be primarily ascribed to the irregularity in badminton skills within the EG and its relatively low proficiency in executing key technical aspects of badminton actions.

Subsequently, we will conduct an in-depth analysis of the technical point scores within the EG to ascertain whether its irregular badminton technical skills and low proficiency in executing key technical aspects are indeed attributable to the LLM method.

Figures that are meant to appear in color, or shades of black/gray. Such figures may include photographs, illustrations, multicolor graphs, and flowcharts.

A. EG BADMINTON MOVEMENT TECHNICAL ESSENTIAL ANALYSIS

In the experimental design, three specific technical points—referred to as D, E, and K—have not been adequately addressed or explained in the video-based learning materials utilized by both researchers and the expert group. When learners encounter these particular technical points, they are likely to resort to relying on text generation by Large Language Models (LLMs) to obtain the necessary information. The learning effect score for corresponding badminton skills is unlikely to be significantly impacted if the text content generated by the large model is accurate or reasonably helpful to the learners. Conversely, if the LLM produces content that is either incorrect or unhelpful, learners may experience a notable decline in their learning effectiveness when attempting to practice badminton skills based on this erroneous information. Importantly, learners might not be able to discern that the information they have received is incorrect or misleading. Therefore, our analysis focuses on the technical points of badminton movement within the Experimental Group (EG) to determine whether an illusion created by the LLM exists concerning the three technical points, D, E, and K.

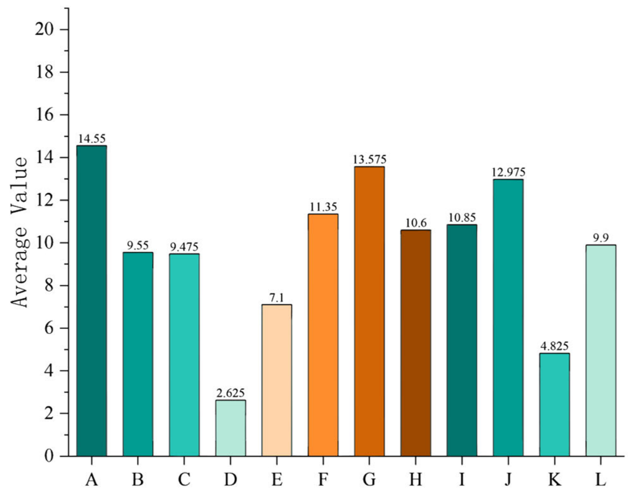

FIGURE 3. EG movement technical essential analysis score average.

Figure 3 illustrates the average technical point scores in the Experimental Group (EG), indicating that the average scores for both D and K technical points are below 5 points

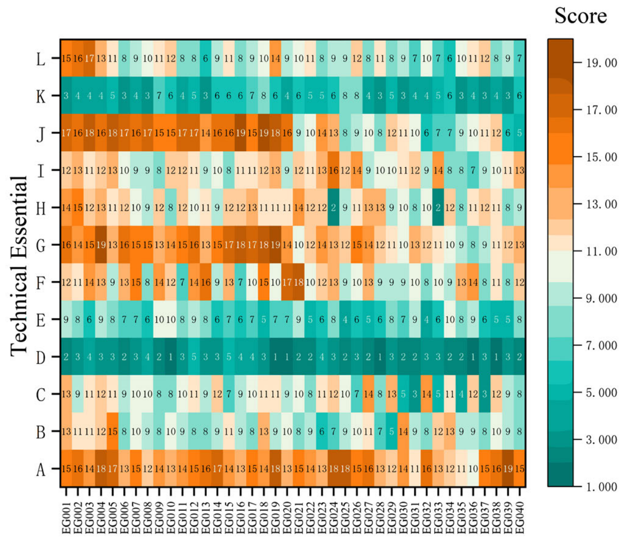

FIGURE 4. Distribution chart of the test scores of badminton movement technical essentials.

(D = 2.625, K = 4.825). This implies that the performance of the EG in categories D and K is unsatisfactory.

Figure 4 depicts the distribution of key points in badminton movement techniques. It is evident that the Experimental Group (EG) performs notably poorly in categories D, E, and K. These specific technical essentials are precisely the areas where learners are highly likely to rely on LLM-based applications for learning, as there is no closely related content readily available through ordinary search engines such as Google or Bing.

B. EXISTENCE OF HALLUCINATION

To further analyze the reasons for the high concordance of scores in the three technical essentials (D, E, and K), face-to-face interviews were conducted with members of the Experimental Group (EG). The interviews were carried out by research team members to gain insight into the learning process of technical essentials D, E, and K. The following are some typical contents recorded during these interviews:

Researcher: Did you have trouble learning DEK skills and did you use LLM after you had trouble?

EG016: I was learning how to whip power and the video asked me to go from a relaxed grip to a tight grip on the racket when hitting the ball, but that didn't give me a good sense of what he was talking about when he said to whip the ball. So, I went ahead and used LLM as requested.

EG023: When I was learning to whip power, the video explained that the feeling is similar to the feeling of whipping with a long rope, so I went home to use a long rope to find the feeling. After I tried it I experienced the feeling of whipping with a long rope, but when I switched back to the badminton racket, the feeling was gone instantly, which was very confusing for me. So, I can only see if I can get some feeling by chatting with LMM.

EG017: I've found very few video explanations of the hit point of a lofted shot, it's hard to find a few, but the videos are all simply one-offs. So, I asked LLM for help, and LLM's explanations are still quite detailed.

EG006: Through this period of study, I think the small arm internal rotation swing to hit the ball is a very important technical movement, it is the most correct swing in line with the human body's physiological structure, but it is also one of the most difficult technical points. I have watched a lot of related videos, but they are not the same, and the explanations are not detailed or even contradictory. So, my turn to LLM to give me a definite content to learn.

EG025: I think the small arm internal rotation swing is so difficult, and the explanations in the videos on the internet are so crude (how to small arm swing, what are the key points and difficulties, why small arm swing, etc., which the videos don't talk about) that I'm confused when I'm learning it, so I'll have to use the LLM to learn it.

All students in the EG group faced significant challenges in learning technical essentials D, E, and K, resulting in consistently low scores distributed in the lower subarea. This phenomenon suggests that students either did not acquire these technical essentials at all or learned them incorrectly. Conversely, as depicted in Table 2, both EG and CG demonstrated overall skill improvement. The plausible explanation for the low and dispersed scores of D, E, and K is that students received inaccurate resources, leading to incorrect learning. During interviews, students expressed difficulty finding corresponding content in videos and websites. Consequently, they had to rely on the text feedback from LLM-based applications for a more comprehensive understanding. The prompts used during interactions with these applications were recorded and analyzed during the interviews to identify any recurring patterns that might indicate "unreal feedback."

Table 4 outlines the prompts and feedback provided for three technical aspects—D, E, and K. An illusionary discrepancy was identified in the output content related to the "Forearm Internal Rotation Swing" technical point. When the input keywords were "how to perform a high badminton ball forearm internal rotation swing," the LLM generated extensive text, including phrases such as "note that this process is the wrist active valgus, so the forearm is internal rotation" and "the wrist over the wrist to drive the arm internal rotation." However, it was unanimously concurred by the experts that no technical action in badminton involves the forearm undergoing internal rotation driven by wrist movement.

Subsequently, when the input keywords were altered to "how to learn the forearm internal rotation swing," the LLM produced the phrase, "the arm internal rotation is mainly achieved by stretching the arm, rotating the arm and wrist." The expert panel opined that, in the application and learning of badminton skills, achieving whipping force necessitates arm relaxation and coordination. Extending the arm induces excessive tension, impeding the effective utilization of whipping force. In summary, it is apparent that LLMs

TABLE 4. Keywords and LLM output text.

prompt	feedback
The hand movement of the badminton forearm	During the backswing, rotate your forearm inward (pronation). This means turning your forearm so that the racket face is pointing downward and the shuttle is behind you. The pronation of the forearm positions the racket face correctly for contact. Initiate the forward swing by extending your arm. Release the wrist pronation during the swing. The forearm should rotate outward as you make contact with the shuttle. (ChatGPT)
Badminton high distance ball how to forearm internal rotation	Initiate the swing by turning your forearm inward (pronation) while keeping the elbow close to your body. Rotate your forearm as you bring the racket head behind you. For a high clear shot, the backswing needs to be more pronounced. Execute the shot by extending your arm forward and releasing the wrist pronation. The forearm should rotate outward during the swing. Focus on making solid and clean contact with the shuttle, sending it high and deep into the opponent's court. (ChatGPT)
how to perform clear forearm internal rotation swing.	At the beginning of the racket, to set the wrist, at the same time, the wrist to drive the forearm internal rotation, note that this process is the wrist active valgus, so the forearm is internal rotation. (ERNIE Bot)
how to learn the forearm internal rotation swing	Arm internal rotation is an important part of badminton's golf technique, which is mainly realized by stretching the arm, rotating the forearm and wrist. During the swing, the strength of the body is transmitted through the arm to the racket, hitting the ball high. (ERNIE Bot)
Badminton high clear hitting point position	The ideal contact point for a high clear is slightly in front of your body, above your head, and close to the center of the racket strings. The shuttle should be hit at its highest point, allowing you to clear the net and send it deep into your opponent's court. (ChatGPT)
Badminton clear hit the ball position which is the most suitable	Strive to make contact with the shuttle in front of your body, slightly above the center of the racket strings. This position allows for better control and a more accurate clear shot. (ChatGPT)
Where is the badminton clear hitting point	The badminton clear shot point is slightly above the right shoulder and about a foot above the ball. This position allows you to hit the ball with full coordination, and can hit the ball to the back of the opposing field, making it difficult for the opponent to return the ball. (ERNIE Bot)
Badminton high distance ball hit the ball point where is the most suitable	The best hitting point for badminton ball is slightly above the right shoulder, about 45 degrees from the ground. This position can help you use the body coordination to hit the ball, and can hit the ball to the back of the opposing field, making the ball back difficult to be threatening. (ERNIE Bot)
Badminton' smash how the forearm internal rotation fully, fast whip force	Keep your wrist cocked during the backswing to store energy for the powerful whip-like motion. As you start the downward swing, fully rotate your forearm internally. This movement is a quick and explosive rotation of the forearm inward. At the point of contact with the shuttle, engage your wrist for a quick snap. The wrist snap is like cracking a whip, and it contributes significantly to the power of your smash. (ChatGPT)
	When waving a racket, you should make full use of the way that the big arm drives the forearm, and pay attention to the inner rotation of the wrist to transfer the power to the racket. At the moment of hitting the ball, the forearm and wrist should be rotated internally quickly, so that the racket surface is perpendicular to the direction of the ball, and use the instantaneous explosive force to hit the ball. Keep your body relaxed during the swinging process. (ERNIE Bot)

generated inaccurate information in the online sports learning process.

Regarding users' awareness of the existence of hallucinations, interviews were conducted with EG students to explore their perception of the authenticity of LLM. Selected excerpts from these interviews are presented below:

Researcher: Has the LLM output been questioned during the learning process?

EG 014: I could comfortably hit the ball out, but it often went sideways, and the ball's trajectory wasn't high or far. I didn't consider it a wrong action.

EG 023: I found it challenging to grasp the feeling of internal rotation through the video, and the generated text made it seem simple. I just wanted to hit the ball, without realizing my actions were incorrect.

EG 033: While playing, my arm felt uncomfortable, and the tension led to muscle soreness.

EG028: Several classmates exhibit proficient badminton skills, consistently delivering high and deep shots with a resonant impact. Intrigued by their performance, I endeavored to emulate their movements during practice sessions.

EG011: The LLM is expected to be accurate, and I am able to hit the shuttlecock as per the instructions, resulting in a satisfying and enjoyable experience. I have not paid particular attention to verifying the accuracy of its statements.

Several participants admitted to unintentionally replicating the movements observed in their peers when faced with challenges in comprehending both video and text content. Regrettably, they were unaware that these replicated actions could be inaccurate. It becomes apparent that LLM generates hallucinations in response to forearm internal rotation, a phenomenon that is not easily discerned. These hallucinations seamlessly blend with authentic content, posing a challenge for learners to distinguish between the two. Moreover, the incorrect actions resulting from these hallucinations may propagate through the mutual communication of learners, eventually establishing a consensus that these actions are correct.

VII. DISCUSSION AND CONCLUSION

In this study, the primary objective is to investigate the occurrence of hallucinations during the utilization of LLM-based applications for sports skills learning. Despite the increasing prevalence of LLMs in various domains, their impact on sports skills acquisition remains underexplored. The central question revolves around the existence of hallucinations induced by LLMs during the learning of motor skills and the potential ramifications of these hallucinations on skill acquisition. The study engaged 80 freshmen with no prior experience in badminton skills, implementing a 16-week experiment that incorporated two modified scales to assess the outcomes of skills learning with LLM-based applications. Statistical methods, including t-tests and one-way ANOVA, were employed to analyze the nuanced technical aspects of skills. Both the trend in test scores and insights from interviews indicate the presence of hallucinations in sports skill learning, particularly in the elaboration of essential components.

Irrespective of the specific skills (e.g., clear, smash, or footwork), students in both the Control Group (CG) and the Experimental Group (EG) demonstrated significant improvement in their post-test scores compared to their initial test scores, indicating that the Experimental Group made progress using LLMs. However, upon conducting a detailed analysis of the technical aspects of badminton actions in the EG, it was observed that students received inaccurate feedback from the LLM when inputting keywords related to the "forearm internal rotation swing" technique, such as "badminton clear how forearm internal rotation swing" and "how to learn high ball arm internal rotation swing." The feedback included phrases like "note that this process is the wrist active valgus, so the forearm is internal rotation" and "the wrist over the wrist to drive the arm internal rotation," which were deemed incorrect. This suggests the occurrence of LLM-induced hallucinations during the process of learning motor skills.

Our investigation into the effects of LLM-induced hallucinations on motor skill learning in badminton presents a critical perspective on the application of AI in physical education. This study contributes to a growing yet underexplored body of research that examines the intersection of advanced technology and motor learning.

Contextualizing Our Findings: The presence of hallucinations in LLMs, as observed in our study, echoes findings from related fields where AI-generated misinformation has been identified. Our results specifically highlight the challenges in motor skill acquisition, corroborating with studies that emphasize the importance of accurate feedback during the learning process [76].

Comparison with Cognitive Load Theory: Drawing from cognitive load theory [77], our findings suggest that the introduction of LLMs may inadvertently increase the cognitive demands on learners, particularly when incorrect information is presented as part of the learning material. This additional cognitive load can hinder the learner's ability to process and integrate new information, especially during the critical initial stages of skill acquisition.

Implications for Motor Learning and Instruction: The implications of our study align with established theories of motor learning, such as [78] model, which underscores the importance of clear instruction and demonstration during the cognitive stage of learning. Our findings indicate that the presence of hallucinations can disrupt this stage, leading to the formation of incorrect motor schemas [79].

Effectiveness of Traditional Teaching Methods: Contrasting the outcomes of the Control Group (CG) and the Experimental Group (EG), our study suggests that traditional teaching methods, which rely on human instructors, may be more effective for imparting complex motor skills like badminton. This finding is consistent with literature that advocates for the nuanced role of human interaction in motor learning [19].

Strategies for Mitigating Hallucinations: Given the potential for hallucinations in AI feedback, our study calls for the development of strategies to address this issue. This includes improving the algorithms that drive LLMs, as well as integrating human oversight to ensure the accuracy of AI-generated feedback [80].

Reliability of LLM-Based Approaches: Our study indicates that while LLMs have the potential to offer personalized and immediate feedback, the presence of hallucinations can significantly undermine their reliability. For beginners, who are at the initial stages of motor skill acquisition, the accuracy of the information provided is paramount. Inaccurate feedback can lead to the formation of incorrect motor patterns, which are challenging to correct later [76]. To enhance the reliability of LLM-based approaches, we suggest several areas for future research. First, there is a need for continuous improvement in the algorithms and data sets used to train LLMs, focusing on domains such as physical education and sports. Second, the development of a verification system to check the accuracy of feedback before it is provided to learners could be beneficial. Third, combining AI feedback with human instructor oversight may offer a hybrid solution that leverages the strengths of both approaches [85]. It is important to note that the reliability of LLM-based approaches is not solely determined by the technology itself but also by how it is integrated into the learning process. As such, educators and instructional designers play a crucial role in curating and overseeing the use of LLMs to ensure that learners receive reliable and valid feedback.

Generalizability to Other Sports: While our study specifically focuses on badminton, the implications of our research may extend to other domains of motor learning. The phenomenon of 'hallucinations' in AI feedback is not unique to badminton and could potentially affect the learning of motor skills across various sports. However, the specific mechanisms through which AI is integrated into training programs and the nature of motor skills required in different sports may moderate the impact of these hallucinations.

Sample Diversity: Our findings provide valuable insights into the impact of LLM-induced hallucinations on motor skill learning in the context of badminton. However, we acknowledge the limitations regarding the sample size and the narrow age range of our study population. To address these limitations and further validate our results, we are planning a follow-up study with the following considerations.

Expanded Age Range: We recognize the importance of age-related factors in motor learning and are committed to including participants from a wider age spectrum. This will enable us to explore potential age-related differences in the susceptibility to LLM-induced hallucinations and the effectiveness of motor skill learning. Larger Sample Size: An increased sample size will not only enhance the statistical power of our study but also allow for a more nuanced analysis of the data. We aim to stratify our sample based on age, experience, and other demographic variables to better understand the interactions between these factors and the learning process.

Alignment with ESG Goals: We have integrated an analysis of our research through the Environmental, Social, and Governance (ESG) criteria. We confirm that our study has minimal environmental impact. The social aspect of our research focuses on democratizing access to motor skill learning through AI, with findings disseminated via open-access channels to ensure inclusivity and knowledge sharing. Governance-wise, we adhere to stringent ethical standards, ensuring informed consent and transparency in our methodology for reproducibility and academic scrutiny. This succinct ESG alignment underscores our commitment to environmental sustainability, social equity, and ethical research integrity.

It is important to note that the transferability of our results should be approached with caution. The complexity and specificity of motor skills in different sports may require tailored AI applications and feedback mechanisms. For instance, team sports with a high degree of strategic and interactive elements may differ significantly from individual, technique-focused sports like badminton [81]. Future research should aim to explore the generalizability of our findings by examining the role of AI in motor skill learning across a range of sports. Comparative studies could provide insights into the commonalities and differences in how AI feedback affects learning outcomes in various athletic contexts.

In conclusion, our research provides empirical evidence of the impact of LLM-induced hallucinations on motor skill learning in badminton. While AI holds great promise for enhancing learning in various domains, our findings underscore the need for caution and critical evaluation of AI tools used in physical education. It is essential to balance the integration of technology with the need for accurate and reliable feedback to support effective motor skill learning.

INSTITUTIONAL REVIEW BOARD STATEMENT

Not applicable.

INFORMED CONSENT STATEMENT

Not applicable.

DATA AVAILABILITY STATEMENT

Not applicable.

CONFLICTS OF INTEREST

The author declare no conflict of interest.

REFERENCES

[1] H. Alkaissi and S. I. McFarlane, "Artificial hallucinations in ChatGPT: Implications in scientific writing," *Cureus*, vol. 15, no. 2, pp. 1–26, Feb. 2023.

[2] A. G. Atci, "Quality and reliability of the information on Youtube videos about botox injection on spasticity," *Romanian Neurosurgery*, vol. 33, pp. 473–477, Dec. 2019.

[3] M. A. Aydin and H. Akyol, "Quality of information available on YouTube videos pertaining to thyroid cancer," *J. Cancer Educ.*, vol. 35, no. 3, pp. 599–605, Jun. 2020, doi: 10.1007/s13187-019-01502-9.

[4] *Badminton Skills Assessment Rubrics*. [Online]. Available: https://www.rcampus.com/rubricshowc.cfm?code=MX55862&sp=yes

[5] J. D. Blom, *A Dictionary of Hallucinations*. Cham, Switzerland: Springer, 2010, p. 168.

[6] *Bing Preview Release Notes: Image and Video Search | Bing Search Blog*. [Online]. Available: https://blogs.bing.com/search/march_2023/TITLE-Bing-Preview-Release-Notes-Image-Video-Search

[7] A. Borji, "A categorical archive of ChatGPT failures," 2023, *arXiv:2302.03494*.

[8] C. Y. Chang, C. H. Chang, K. C. Chen, C. S. Ho, and Y. R. Chen, "The effect of footwork and response capability in elementary school badminton players by six weeks agility training," *Sports Coach.*, vol. 44, pp. 57–66, Jun. 2016.

[9] R. C. Colley, T. Bushnik, and K. Langlois, "Exercise and screen time during the COVID-19 pandemic," *Health Rep.*, vol. 31, pp. 1–11, Jul. 2020.

[10] M. Connor and M. O'Neill, "Large language models in sport science and medicine: Opportunities, risks and considerations," 2023, *arXiv:2305.03851*.

[11] R. Dale, "GPT-3: What's it good for?" *Natural Lang. Eng.*, vol. 27, no. 1, pp. 113–118, Jan. 2021, doi: 10.1017/s1351324920000601.

[12] D. F. Tate, E. J. Lyons, and C. G. Valle, "High-tech tools for exercise motivation: Use and role of technologies such as the Internet, mobile applications, social media, and video games," *Diabetes Spectr.*, vol. 28, no. 1, pp. 45–54, Feb. 2015, doi: 10.2337/diaspect.28.1.45.

[13] J. Deng and Y. Lin, "The benefits and challenges of ChatGPT: An overview," *Frontiers Comput. Intell. Syst.*, vol. 2, no. 2, pp. 81–83, Jan. 2023, doi: 10.54097/fcis.v2i2.4465.

[14] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*.

[15] M. E. Finkenberg, "The Internet in kinesiology and physical education," *Quest*, vol. 49, no. 3, pp. 327–332, 1997.

[16] P. M. Fitts and M. I. Posner. (1967). *Learning and Skilled Performance in Human Performance*. [Online]. Available: https://archive.org/details/humanperformance00fitt

[17] L. Floridi and M. Chiriatti, "GPT-3: Its nature, scope, limits, and consequences," *Minds Mach.*, vol. 30, no. 4, pp. 681–694, Dec. 2020, doi: 10.1007/s11023-020-09548-1.

[18] S. Fullagar, E. Rich, J. Francombe-Webb, and A. Maturo, "Digital ecologies of youth mental health: Apps, therapeutic publics and pedagogy as affective arrangements," *Social Sci.*, vol. 6, no. 4, p. 135, Nov. 2017, doi: 10.3390/socsci6040135.

[19] R. M. Gagne, "Learning outcomes and their effects: Useful categories of human performance.," *Amer. Psychologist*, vol. 39, no. 4, pp. 377–385, 1984.

[20] N. L. Goggin, M. E. Finkenberg, and J. R. Morrow, "Instructional technology in higher education teaching," *Quest*, vol. 49, no. 3, pp. 280–290, 1997.

[21] G. Hall, D. R. Laddu, S. A. Phillips, C. J. Lavie, and R. Arena, "A tale of two pandemics: How will COVID-19 and global trends in physical inactivity and sedentary behavior affect one another?" *Prog. Cardiovascular Diseases*, vol. 64, pp. 108–110, Jan. 2021, doi: 10.1016/j.pcad.2020.04.005.

[22] M. Haseeb, H. I. Hussain, B. Slusarczyk, and K. Jermsittiparsert, "Industry 4.0: A solution towards technology challenges of sustainable Bus. Performance," *Social Sci.*, vol. 8, no. 5, p. 154, May 2019, doi: 10.3390/socsci8050154.

[23] H. J. Oh, C. Lauckner, J. Boehmer, R. Fewins-Bliss, and K. Li, ''Facebooking for health: An examination into the solicitation and effects of health-related social support on social networking sites,'' *Comput. Hum. Behav.*, vol. 29, no. 5, pp. 2072–2080, Sep. 2013, doi: 10.1016/j.chb.2013.04.017.

[24] J. Zou, Q. Liu, and Z. Yang, ''Development of a moodle course for schoolchildren's table tennis learning based on competence motivation theory: Its effectiveness in comparison to traditional training method,'' *Comput. Educ.*, vol. 59, no. 2, pp. 294–303, Sep. 2012, doi: 10.1016/j.compedu.2012.01.008.

[25] M. Kehl, H. Strobl, S. Tittlbach, and J. Loss, '''Der mensch, der handball spielt, braucht den ball, den kontakt und die Gemeinschaft'—Veränderungen im sportangebot durch die COVID-19 pandemie und deren bedeutung fur sportvereine,'' *Das Gesundheitswesen*, vol. 83, no. 3, pp. 159–165, Mar. 2021.

[26] K. Casebolt and P. Zhang, ''An authentic badminton game performance assessment rubric,'' *Strategies*, vol. 33, no. 1, pp. 8–13, Jan. 2020, doi: 10.1080/08924562.2019.1680329.

[27] S. Kumar, V. Balachandran, L. Njoo, A. Anastasopoulos, and Y. Tsvetkov, ''Language generation models can cause harm: So what can we do about it? An actionable survey,'' 2022, *arXiv:2210.07700*.

[28] H. C. Lan, H. C. Wang, and W. C. Chen, ''Investigation of multimedia-assisted instruction on 8-movement Taijiquan's learning effects for the college students,'' *Physical Education Journal*, vol. 43, no. 3, pp. 109–125, 2010.

[29] R. Leser, A. Baca, and J. Uhlig, ''Effectiveness of multimedia-supported education in practical sports courses,'' *J. Sports Sci. Med.*, vol. 10, no. 1, pp. 184–192, 2011.

[30] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer, ''BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension,'' 2019, *arXiv:1910.13461*.

[31] W. Y. Lim and M. Koh, ''E-learning: New opportunities for teaching and learning in gymnastics,'' *Brit. J. Teach. Phys. Educ.*, vol. 37, no. 1, pp. 22–25, 2006.

[32] H. Liu, R. Ning, Z. Teng, J. Liu, Q. Zhou, and Y. Zhang, ''Evaluating the logical reasoning ability of ChatGPT and GPT-4,'' 2023, *arXiv:2304.03439*.

[33] D. Marton, S. Bednar, and J. Modrak, ''Enhancing learning management system through web academic library information systems,'' *Polish J. Manag. Stud.*, vol. 1, pp. 1–23, May 2014.

[34] J. Maynez, S. Narayan, B. Bohnet, and R. McDonald, ''On faithfulness and factuality in abstractive summarization,'' 2020, *arXiv:2005.00661*.

[35] Y. Mehdi. (2023). *Confirmed: The New Bing Runs on OpenAI's GPT-4. Microsoft Bing Blog*. [Online]. Available: https://blogs.bing.com/search/march_2023/Confirmed-the-new-Bing-runson-OpenAI

[36] B. Meskó and E. J. Topol, ''The imperative for regulatory oversight of large language models (or generative AI) in healthcare,'' *Npj Digit. Med.*, vol. 6, no. 1, p. 120, Jul. 2023.

[37] J. G. Meyer, R. J. Urbanowicz, P. C. N. Martin, K. O'Connor, R. Li, P.-C. Peng, T. J. Bright, N. Tatonetti, K. J. Won, G. Gonzalez-Hernandez, and J. H. Moore, ''ChatGPT and large language models in academia: Opportunities and challenges,'' *BioData Mining*, vol. 16, no. 1, p. 20, Jul. 2023, doi: 10.1186/s13040-023-00339-9.

[38] M. Zhang and J. Li, ''A commentary of GPT-3 in MIT technology review 2021,'' *Fundam. Res.*, vol. 1, no. 6, pp. 831–833, Jul. 2021.

[39] B. Mohnsen, *Using Technology in Physical Education*, 6th ed., Vista, CA, USA: Bonnie's Fitware, 2008.

[40] OpenAI et al., ''GPT-4 technical report,'' 2023, *arXiv:2303.08774*.

[41] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. Mishkin, and R. Lowe, ''Training language models to follow instructions with human feedback,'' in *Proc. Adv. Neural Inf. Process. Syst.*, 2022, pp. 27730–27744.

[42] E. Perez, S. Huang, F. Song, T. Cai, R. Ring, J. Aslanides, A. Glaese, N. McAleese, and G. Irving, ''Red teaming language models with language models,'' 2022, *arXiv:2202.03286*.

[43] M. Papastergiou and V. Gerodimos, ''Can learning of basketball be enhanced through a web-based multimedia course? An experimental study,'' *Educ. Inf. Technol.*, vol. 18, no. 3, pp. 459–478, Sep. 2013, doi: 10.1007/s10639-012-9186-z.

[44] C. Qin, A. Zhang, Z. Zhang, J. Chen, M. Yasunaga, and D. Yang, ''Is ChatGPT a general-purpose natural language processing task solver?'' 2023, *arXiv:2302.06476*.

[45] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, ''Improving language understanding by generative pre-training,'' Mike Captain Res. Inst., San Francisco, CA, USA, Tech. Rep. GPT-1, 2018. [Online]. Available: https://www.mikecaptain.com/resources/pdf/GPT-1.pdf

[46] V. Sanh et al., ''Multitask prompted training enables zero-shot task generalization,'' 2021, *arXiv:2110.08207*.

[47] S. Min, M. Lewis, L. Zettlemoyer, and H. Hajishirzi, ''MetaICL: Learning to learn in context,'' in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics*, 2022, pp. 1–20.

[48] T. Shi, L. Ma, and Z. Xing, ''Application of CAI in the teaching of physical education in college,'' in *Proc. 3rd Int. Conf. Teach. Comput. Sci.*, vol. 1, 2012, pp. 429–434.

[49] I. Straw and C. Callison-Burch, ''Artificial intelligence in mental health and the biases of language based models,'' *PLoS One*, vol. 15, no. 12, Dec. 2020, Art. no. e0240376.

[50] Z. Tan and S. Li, ''Multimedia technology in physical education,'' in *Proc. Int. Symp. Comput. Netw. Multimedia Technol.*, Jan. 2009, pp. 1–4.

[51] W. R. Thompson, ''Worldwide survey of fitness trends for 2022,'' *ACSM'S Health Fitness J.*, vol. 26, no. 1, pp. 11–20, Jan. 2022.

[52] R. Thornburg and K. Hill, ''Using Internet assessment tools for health and physical education instruction,'' *TechTrends*, vol. 48, no. 6, pp. 53–55, Nov. 2004, doi: 10.1007/bf02763585.

[53] P. L. Tsai, H. F. Liao, and C. Y. Chen, ''The technical analysis of women's doubles in 2011 London world badminton championships,'' *J. Phys. Educ.*, vol. 31, pp. 1–8, May 2016, doi: 10.6569/NTUJPE.2016.31.01.

[54] L. Weidinger et al., ''Ethical and social risks of harm from language models,'' 2021, *arXiv:2112.04359*.

[55] J. Wei, M. Bosma, V. Y. Zhao, K. Guu, A. Wei Yu, B. Lester, N. Du, A. M. Dai, and Q. V. Le, ''Finetuned language models are zero-shot learners,'' 2021, *arXiv:2109.01652*.

[56] F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. Song, Y. Hu, Z.-W. Tao, J.-H. Tian, Y.-Y. Pei, M.-L. Yuan, Y.-L. Zhang, F.-H. Dai, Y. Liu, Q.-M. Wang, J.-J. Zheng, L. Xu, E. C. Holmes, and Y.-Z. Zhang, ''A new coronavirus associated with human respiratory disease in China,'' *Nature*, vol. 579, no. 7798, pp. 265–269, Mar. 2020, doi: 10.1038/s41586-020-2008-3.

[57] X. Zhou and A. Krishnan, ''What predicts exercise maintenance and well-being? Examining the influence of health-related psychographic factors and social media communication,'' *Health Commun.*, vol. 34, no. 6, pp. 589–597, May 2019, doi: 10.1080/10410236.2018.1428851.

[58] Y. Bang, S. Cahyawijaya, N. Lee, W. Dai, D. Su, B. Wilie, H. Lovenia, Z. Ji, T. Yu, W. Chung, Q. V. Do, Y. Xu, and P. Fung, ''A multitask, multilingual, multimodal evaluation of ChatGPT on reasoning, hallucination, and interactivity,'' 2023, *arXiv:2302.04023*.

[59] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. J. Bang, A. Madotto, and P. Fung, ''Survey of hallucination in natural language generation,'' *ACM Comput. Surveys*, vol. 55, no. 12, pp. 1–38, Dec. 2023, doi: 10.1145/3571730.

[60] M. Pustisek et al., ''The role of technology for accelerated motor learning in sport,'' *Pers. Ubiquitous Comput.*, vol. pp. 969–978, 2021.

[61] R. G. Nowels and J. K. Hewit, ''Improved learning in physical education through immediate video feedback,'' *Strategies*, vol. 31, no. 6, pp. 5–9, 2018.

[62] J. MCNally, ''The effects of immediate supervisory feedback using audiocuing on the interaction between students and preservice physical education teachers,'' CORE, Milton Keynes, U.K., Tech. Rep. 217288400, 2003. [Online]. Available: https://core.ac.uk/reader/217288400

[63] Y. Y. Dyulicheva and A. O. Glazieva, ''Game based learning with artificial intelligence and immersive technologies: An overview,'' in *Proc. CEUR Workshop*, vol. 3077, 2022, pp. 1–23.

[64] O. D. Cote, H. Beverly, and Robert C. Williges, ''Augmented feedback in adaptive motor skill training,'' in *Proc. Hum. Factors Soc. Annu. Meeting*, vol. 22, 1978, pp. 1–26.

[65] J. A. Haegele, ''Inclusion illusion: Questioning the inclusiveness of integrated physical education: 2019 national association for kinesiology in higher education hally beth poindexter young scholar address,'' *Quest*, vol. 71, no. 4, pp. 387–397, 2019.

[66] J. R. Anderson, *The Architecture of Cognition*. London, U.K.: Psychology Press, 2013.

[67] B. Linquist, ''Motor learning concepts and applications,'' *Pediatric Physical Therapy*, vol. 14, no. 1, pp. 57–58, 2002.

[68] J. Piaget, *The Construction of Reality in the Child*. Evanston, IL, USA: Routledge, 2013.

[69] E. A. Locke, *Social Foundations of Thought and Action: A Social-cognitive View*, 1987, pp. 169–171.

[70] V. John-Steiner, *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA, USA: Harvard Univ. Press, 2007.

[71] E. L. Deci and R. M. Ryan, ''What' and 'why' of goal pursuits: Human needs and the self-determination of behavior,'' *Psychological Inquiry*, vol. 11, no. 4, pp. 227–268, 2000.

[72] J. Reeve, *Understanding Motivation and Emotion*. Hoboken, NJ, USA: Wiley, 2018.

[73] C. E. Hmelo-Silver, ''Problem-based learning: What and how do students learn?'' *Educational Psychology Review*, vol. 16, pp. 235–266, Jun. 2004.

[74] C. A. Tomlinson. (2014). *The Differentiated Classroom: Responding to the Needs of All Learners*. [Online]. Available: https://search.worldcat.org/zh-cn/title/1245433099

[75] J. Oates, ''Ethical considerations in psychology research,'' in *Handbook of Research Ethics and Scientific Integrity*, 2020, pp. 783–801.

[76] R. A. Schmidt and C. A. Wrisberg, ''Motor learning and performance: A situation-based learning approach,'' *Hum. Kinetics*, vol. 1, pp. 1–26, May 2008.

[77] F. Paas, A. Renkl, and J. Sweller, ''Cognitive load theory and instructional design: Recent developments,'' *Educ. Psychologist*, vol. 38, no. 1, pp. 1–4, 2003.

[78] P. M. Fitts and M. I. Posner. (1967). *Human Performance*. [Online]. Available: https://psycnet.apa.org/record/1967-35040-000?utm_medium=email&utm_source=transaction

[79] J. A. Adams, ''A closed-loop theory of motor learning,'' *J. Motor Behav.*, vol. 3, no. 2, pp. 111–150, 1971.

[80] C. O'neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New Bremen, OH, USA: Crown, 2017.

[81] D. Farrow and B. Abernethy, ''Can anticipatory skills be learned through implicit video based perceptual training?'' *J. Sports Sci.*, vol. 20, no. 6, pp. 471–485, Jan. 2002.

[82] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. Evanston, IL, USA: Routledge, 2013.

[83] F. Faul, E. Erdfelder, A.-G. Lang, and A. Buchner, ''G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences,'' *Behav. Res. Methods*, vol. 39, no. 2, pp. 175–191, May 2007.

[84] *Publication Manual of the American Psychological Association*, Amer. Psychol. Assoc., Worcester, MA, USA, 2020.

[85] D. R. Moore. (2006). *E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning*. [Online]. Available: https://www.jstor.org/stable/30221321

[86] W. Fish, *Perception, Hallucination, and Illusion*. USA: Oxford Univ. Press, 2009. [Online]. Available: https://books.google.com.sg/books?hl=zh-CN&lr=&id=1cB8Yysp9Y4C&oi=fnd&pg=PR9&dq=(William+Fish+et+al.,+2009&ots=0h-3tJ4xVm&sig=gjU8Y3QuZq6l39QonHJq-NaExX4&redir_esc=y#v=onepage&q=(William%20Fish%20et%20al.%2C%202009&f=false

[87] F. Macpherson and D. Platchias, Eds., *Hallucination: Philosophy and Psychology*. Cambridge, MA, USA: MIT Press, 2013. [Online]. Available: https://books.google.com.sg/books?hl=zh-CN&lr=&id=GcmqAAAAQBAJ&oi=fnd&pg=PR5&dq=Fiona+Macpherson+and+Dimitris+Platchias,+2013&ots=tx1zyaE6F-&sig=yzXgf7FRWi8gIeQj5jCHnVNG90o&redir_esc=y#v=onepage&q=Fiona%20Macpherson%20and%20Dimitris%20Platchias%2C%202013&f=false

[88] R. Zhao et al., ''A practical simulation of carbon sink calculation for urban buildings: A case study of Zhengzhou in China,'' *Sustain. Cities Soc.*, vol. 99, 2023, Art. no. 2104980.

[89] Z. Ji et al., ''LLM internal states reveal hallucination risk faced with a query,'' 2022, *arXiv:2407.03282*.

YEPENG QIU is currently with the College of Physical Education and Health, Chongqing Normal University. His research interests include artificial intelligence and sports management.