International Journal of Organizational Innovation

EXPLORING THE INFLUENCE OF AN AI PIANO TUTORING SYSTEM ON MOTIVATION AND PERFORMANCE ACROSS AGE GROUPS

Ming-Lang Yeh
College of Management, Chung Hua University, Taiwan.
mlyeh@chu.edu.tw

Mei-Ling Hu*
Ph.D. Program of Management, Chung Hua University, Taiwan.
*Corresponding author: d11103001@chu.edu.tw

Wen Pei,
College of Management, Chung Hua University, Hsinchu, Taiwan
wpei@chu.edu.tw

Abstract

Mastering a musical instrument necessitates a prolonged commitment to study and practice, a process particularly critical for children. The cultivation of motivation and sustained interest is vital in facilitating effective learning experiences. This research delves into the influence of Artificial Intelligence (AI) tutoring as a pedagogical approach on students' learning performances and motivation, contrasting it with traditional human piano instruction. Data for this study were acquired through an experimental piano class, incorporating both human and AI tutoring methods. The class structure entailed 30 minutes of conventional human instruction followed by an additional 30 minutes of AI tutoring. To ensure randomization, half of the class commenced with human instruction, while the other half initiated with AI tutoring.

Utilizing a single-group repeated measures research design, this investigation aimed to discern noteworthy variations in student learning performance and motivation between the two teaching methodologies. Results from two-way mixed ANOVAs, based on data from 38 students, revealed a substantial increase in student enthusiasm with the AI teaching method. However, no significant differences in student learning performance

emerged between the two instructional approaches. Furthermore, the interaction between teaching method and independent variables such as age, gender, and self-efficacy did not yield significance.

These findings not only contribute to the discourse on the equitable and sustainable development of music education but also underscore the imperative for tailored AI piano tutoring systems capable of addressing diverse learning needs. The research outcomes furnish valuable insights for AI educational technology researchers and offer guidance to educators in shaping curricular strategies. Significantly, the research underscores the variable impact of teaching models on the motivation and performance of learners across age groups, further substantiating the potential advantages of AI in enhancing student motivation within educational settings.

Keywords: AI-tutoring teaching, Human teaching, piano tutoring system, pedagogical considerations, student learning performance, student learning motivation.

Introduction

Before the rise of Artificial Intelligence (AI), early educational innovations primarily focused on teaching methods utilizing animated characters, referred to as pedagogical agents (Heidig, S., & Clarebout, G., 2011). A comprehensive review by Heidig from 2002 to 2011 revealed a dearth of conclusive evidence regarding the impact of animated characters on students' learning performances and motivation. While most studies suggested no significant disparity between multimedia teaching materials and traditional human methods, Heidig advocated for further investigations examining the influence of multimedia on diverse demographic groups, learning environments, and variations in educational technology designs (Heidig, S., & Clarebout, G., 2011).

Subsequent to Heidig's seminal work, the last decade witnessed a significant breakthrough in educational innovation research. With the proliferation of Internet bandwidth and the abundance of online teaching resources, high-quality teaching videos online were found to outperform conventional teaching. Leddo's exploration of online video teaching, exemplified by Khan Academy, demonstrated an 80% improvement in grades compared to traditional courses (Leddo, et al., 2016). Additionally, students actively engaged with Pearson e-textbooks exhibited three times higher learning performances than those following traditional methods (Leddo et al., 2019). However, Leddo's early research lacked detailed analysis on various factors, prompting subsequent studies to investigate the impact of online teaching videos across subjects, teaching designs, student levels, and gender.

Research by Leddo, J., Boddu, B., Krishnamurthy, S., Yuan, K., & Chippala, S. (2017) expanded the inquiry to include the student level as an intermediary variable, examining the learning of basic web design. The findings suggest-

ed that excellent students performed equally well using online videos or traditional methods, while average students demonstrated better outcomes with human instruction. This nuanced perspective contradicted the previous notion that online videos universally benefit all students. Further advancements in educational technology introduced teaching by robot teachers, with Song, H., Barakova, E. I., Markopoulos, P., & Ham, J. (2021) demonstrating the motivational impact of lively and humorous robot teaching on beginner-level music students.

As technology continued to evolve, AI and machine learning took center stage, surpassing robot teaching and online video self-study. Leddo, J., & Garg, K. (2021) delved into the advantages of AI teaching in mathematics algebra courses, revealing a 37% higher average score for students taught by AI compared to those taught by human teachers. Moreover, AI teaching demonstrated consistent effectiveness across different stages of learning. The emergence of AI teaching marked a paradigm shift in educational technology, particularly in high school mathematics education.

Research Gap and Purpose

This study delves into the less-explored realm of AI application in piano learning within music education. It seeks to understand how AI impacts students' learning performance and motivation, with age, gender, and self-efficacy as independent variables. Research questions examine potential differences between AI tutoring and traditional human teaching, and hypotheses consider interactions with different demographic segments. Aligning with Heidig, S., & Clarebout, G.'s recommendations (2011), age and gender are treated as independent variables, emphasizing self-efficacy's crucial role in music education. Ritchie, L., & Williamon, A.'s findings (2010) highlight self-efficacy nuances among music school and university music department students. Hallam's study (2002) underscores the impact of cognitive processes, self-esteem, and self-efficacy. Hence, self-efficacy is a pivotal variable in our investigation, reflecting its significance in music education literature.

Despite AI-based teaching's decade-long integration into education, its effects on piano learning in music education remain understudied. This research aims to address this gap by exploring the impact of AI tutoring on student learning performance and motivation, considering age, gender, and self-efficacy as independent variables.

Research Questions and Hypotheses

The research questions guiding this study are as follows:

1. Is there a significant difference in student motivation between AI tutoring teaching and human teaching?
2. Is there a significant difference in students' learning performance between AI tutoring teaching and human teaching?
3. Do students' individual demographic segments (age, gender, and efficacy) influence student learning performance and motivation, and

how do these factors interact with AI tutoring and human teaching?

These research questions are translated into 14 null hypotheses, categorized into 4 groups:

Group 1: Pedagogy Comparison

H1a: The means of learning performance scores between AI-tutoring and human pedagogy are equal.

H1b: The means of motivation scores between AI-tutoring and human pedagogy are equal.

Group 2: Age Influence

H2a: The means of learning performance scores between age groups (children, youths, adults) are equal.

H2b: The means of motivation scores between age groups (children, youths, adults) are equal.

H2c: There is no interaction between age groups and pedagogy in terms of learning performance scores.

H2d: There is no interaction between age groups and pedagogy in terms of motivation scores.

Group 3: Gender Influence

H3a: The means of learning performance scores between female and male are equal.

H3b: The means of motivation scores between female and male are equal.

H3c: There is no interaction between gender and pedagogy in terms of learning performance scores.

H3d: There is no interaction between gender and pedagogy in terms of motivation scores.

Group 4: Self-Efficacy Influence

H4a: The means of learning performance scores between low and high efficacy students are equal.

H4b: The means of motivation scores between low and high efficacy students are equal.

H4c: There is no interaction between self-efficacy and pedagogy in terms of learning performance scores.

H4d: There is no interaction between self-efficacy and pedagogy in terms of motivation scores.

Methodology

This study employed a one-group repeated measures design to investigate the impact of AI tutoring on piano learning in the field of music education. It took place in a piano classroom at a music institution in Hsinchu, Taiwan, utilizing AI smart pianos as teaching tools. Due to budgetary constraints, data were collected from a modest sample of 38 participants. The choice of a one-group pilot design was based on the sample size meeting the statistical requirement

of at least 30 but less than 60. Opting for a repeated-measures within-subjects design was considered optimal to mitigate interference from external variables, enhance internal validity, and improve external validity and generalizability.

Experimental Facilities

The study collaborated with a music institution in Hsinchu, Taiwan, utilizing the institution's venue and AI smart pianos for experimental purposes. The facilities, including AI pianos, teachers, and venues, were fully supported by the music institution.

Participant Recruitment

Recruitment was facilitated through personal networks, focusing on individuals who were beginners or had no prior experience in piano learning. To control for the influence of prior knowledge and learning experience, only beginners were accepted into the study.

Experimental Course and Pedagogies

The study adopted a one-group repeated measures design, employing a sequence of human and AI tutoring teaching methods for each participant. The experimental class time consisted of one hour, with each class divided into two parts: 30 minutes of AI tutoring teaching and 30 minutes of human teaching. The classes accommodated 5-7 participants each, with teaching methods randomized to minimize sequence effects. The teaching methods, human and AI tutoring, were conducted in alternating sequences across different classes. The AI tutoring method utilized the AI piano's screen for instructional videos, interactive games, and real-time evaluation, while the human method involved traditional teaching approaches with a small whiteboard, teacher demonstrations, and group interaction. The differences between AI tutoring and human teaching methods were carefully outlined, considering factors such as repertoire, teaching methods, evaluation, questionnaires, equipment, time, and content. The study procedure involved an initial explanation of course content to participants, completion of entry surveys and motivation surveys, and the initiation of teaching methods.

Table 1 is the summary of differences across teaching styles.

The AI tutoring session involved watching instructional videos and interactive games, while the human session utilized traditional teaching approaches. Table 2 lists the contents of different teaching methods.

Randomization Sequence

To uphold the integrity of the study design, a meticulous randomization process was implemented in determining the sequence of teaching methods for each participant. This approach was crucial in minimizing potential biases or order effects that might influence the outcomes, thereby enhancing internal validity and allowing for more reliable conclusions. The process of randomization sequence is summarized in Table 3.

Table 1. Differences in Pedagogies: AI-tutoring and Human

Differences & Learning Tools	*AI-tutoring Teaching Method*	*Human Teaching Method*
Repertoire	Elementary-level piece of music	Elementary-level piece of music
Teaching Methods	Students watch video tutorials, play interactive games and watch on-screen fingering tutorials on their own piano screens.	Listen to the teacher's instructions and watch the teacher's playing demonstration on the teacher's piano.
Evaluation	Using the built-in assessment system of the AI piano for scoring.	Using the built-in assessment system of the AI piano for scoring.
Questionnaires	Entry survey/Motivation survey	Entry survey/Motivation survey
Equipment	AI pianos	AI pianos with the screen off & a small whiteboard
Time	December 2023 to January 2024	December 2023 to January 2024

The AI Piano Evaluation System provided quantitative measurements of learning performance, assessing pitch, rhythm, stability, fluency, and completion. Systematic measurements provide valuable information into participants' progress and proficiency during the learning process. Learning performance was evaluated through the AI Piano Evaluation System, and the randomization of teaching methods aimed to minimize potential biases related to the order of instruction. This approach ensured a comprehensive and unbiased assessment of the impact of different pedagogical approaches.

The study's sampling strategy and participant inclusion criteria focused on a diverse group of beginners with no prior exposure to AI-assisted learning or formal piano lessons. The data collection instruments exhibited reliability and validity, enhancing the robustness of the study's findings.

Table 2. Contents and Procedures of Different Pedagogies

AI-tutoring Teaching Method (15minutes)	*Human Teaching Method (15minutes)*
Students watch built-in instructional videos on their own piano screen to learn scales, time signatures, and the value of each note. *Students watch their own piano screen and play interactive games to learn and memorize scales, time signatures, and the value of each note.* *Students watch interactive instruction on their own piano screen to learn fingerings and corresponding key positions.* *Teachers need to operate the piano tutoring system.*	*The teacher uses a small whiteboard to draw musical notes and staff for teaching.* *The teacher teaches rhythm with clapping hands.* *The teacher demonstrates finger position and playing on the teacher's piano.*
Using AI Piano's built-in evaluation system to score.	*Using AI Piano's built-in evaluation system to score.*

Table 3. Design of the Study

Beginning of the Class Pre-Test	*Treatment 1 (First Class)*	*End of 1 Class Measure 1*	*Treatment 2 (Second Class)*	*End of 2 Class Measure 2*
Entry Survey (Age, Gender, Efficacy) with 19 students	Human teaching	Motivation/learning performance Survey	AI tutoring teaching	Motivation/learning performance Survey
Entry Survey (Age, Gender, Efficacy) with 19 students	AI tutoring teaching	Motivation/learning performance Survey	Human teaching	Motivation/learning performance Survey

Data Analysis

The data analysis employed one-way and two-way factorial mixed ANOVAs, specifically a one-way repeated measures ANOVA for continuous dependent variables (learning performance and motivation). Three two-way mixed ANOVAs were conducted to explore within-subject effects (pedagogies) and between-subject effects (age, gender, and efficacy). These analyses enabled a nuanced examination of main effects and interactions, providing a comprehensive evaluation of research hypotheses.

Variables and Measures

The study involved 6 variables: four independent variables (IVs) and 2 dependent variables (DVs).

Levels of the 4 IVs:

1. Pedagogy (2 levels): Human teaching, AI tutoring teaching.

2. Age (3 levels): Children, Youths, Adults.

3. Gender (2 levels): Male, Female.

4. Self-Efficacy (2 levels): Low, High.

Levels of the 2 DVs:

1. Learning Performance: Assessed on a scale from 0 to 100.

2. Learning Motivation: Evaluated using a scale ranging from 0 to 175.

This framework allowed for a thorough exploration of the impact of instructional methods, age, gender, and self-efficacy on learning performance and motivation. The study's focus on both independent and dependent variables facilitated a nuanced analysis of the research dynamics.

The six variables and their measures are listed in Table 4**.**

Results

Independent Variables: Gender, Age, Efficacy, Pedagogy

The distribution of participants across gender, age, and efficacy levels demonstrated a balanced representation in the study. The gender distribution was 12 males and 26 females, with a normal female-to-male ratio for a piano learning classroom. Age groups were well-represented, with 14 children, 9 youths, and 15 adults. Self-efficacy scores ranged from 36 to 63, and participants were divided into low (22 students) and high (16 students) efficacy groups. Pedagogy was a within-group variable, with all 38 students experiencing both human and AI-tutoring instruction. We presented the frequency distribution of data in Table 5.

Table 4. Variables and Measures

Variables	*Measure*	*Scale of Measure Interval/Ordinal/Ratio/Nominal*
Independent Variables		
Pedagogy: Human and AI.		Nominal
Age: Children, Youths, and Adults	Entry Survey	Ordinal
Gender: Male, Female	Entry Survey	Nominal
Self-efficacy: Low and High	Entry Survey	Ordinal
Dependent Variables		
Students Learning Performance	AI-piano auto-grading	Ratio
Students Motivation	After-Class survey	Ratio (Survey 1.2)

Table 5. Frequency Distribution of Categorical Variables

	N	*Percent*
Gender		
Male	12	31.6%
Female	26	68.4%
Age		
Children (-12)	14	36.8%
Youths (13-20)	9	23.7%
Adults (21-)	15	39.5%
Efficacy		
Low	22	57.9%
High	16	42.1%
Pedagogy		
Human	38	100%
AI tutoring	38	100%

Dependent Variables: Learning Performance and Motivation

Descriptive Statistics: Descriptive statistics revealed that the second test scores for learning performance were

expectedly higher, showing an average increase of 7.85 points (74.36 compared to 66.52). To mitigate the practice effect, the second test scores were adjusted by subtracting 7.85. No practice effect was observed for motivation scores, ensuring the validity of the original scores.

The mean and standard deviation of learning performance and motivation for both human and AI-tutoring methods are summarized in Table 6. After human teaching, the mean motivation score was 77.82, and the learning performance score was 66.52. After AI-tutoring, the mean motivation score increased to 81.82, while the learning performance score remained relatively consistent at 66.95. The data of two dependent variables is presented in Table 6.

Table 6. Descriptive Statistics for Learning Performances and Motivation of the Two Different Pedagogies

	Treatment 1 Human-Based			Treatment 2 AI-Based	
	N	Mean 1	Std. Dev. 1	Mean 2	Std. Dev. 2
Motivation	38	77.82	11.229	81.82	11.622
Learning performance	38	66.521	25.246	66.945	26.197

Descriptive Statistics of All Variables

Table 7 provides the minimum, maximum, mean, and standard deviation for independent and dependent variables. Age, a ratio variable, ranged from 6 to 59, with a mean of 25. Gender was a nominal variable, and efficacy, a ratio variable, had scores ranging from 36 to 63, with a mean of 44.66. Learning performance scores (both human and AI) ranged from 6 to 98, with means of 66.52 and 66.95, respectively. Motivation scores (both human and AI) ranged from 54 to 103, with means of 77.82 and 81.82, respectively.

Two-Way Mixed ANOVA Repeated Measures

Pedagogy Analysis: A one-way ANOVA with repeated measures revealed no statistically significant difference in learning performance between AI-tutoring and human methods ($p = 0.887$). However, a significant difference was observed in motivation scores ($p = 0.001$), with higher scores reported after AI-tutoring.

Age Analysis: Three two-way mixed ANOVAs explored the influence of age, gender, and self-efficacy on learning performances and motivation. Findings indicated a significant difference between age groups for learning

performance ($p = 0.003$). Post hoc tests revealed that youths outperformed adults and children. No significant differences were found in motivation between age groups, and no interactions were observed between age and pedagogy.

Gender Analysis: A significant difference was found between genders for motivation scores ($p = 0.032$), with females demonstrating stronger interest in playing the piano. No significant differences were found in learning performance between genders, and no interactions were observed between gender and pedagogy.

Self-Efficacy Analysis: No significant differences were observed between high and low efficacy groups for both learning performance and motivation. Additionally, no interactions were observed between self-efficacy and pedagogy.

All the results of two-way ANOVAs were shown in Table 8.

Summary of Null Hypotheses

Out of 14 null hypotheses, three were rejected, supporting alternative hypotheses. The rejected null hypotheses are related to motivation scores, age groups influencing learning performance, and gender influencing motivation. Nine null hypotheses were not rejected due to a lack of evidence ($p > 0.05$), suggesting no significant differences in the variables being tested. The complete test results for the null hypotheses are summarized in Table 9.

Table 7. Descriptive Statistics of Variables

	N	*Minimum*	*Maximum*	*Mean*	*Std. Deviation*
Age	38	6	59	25.03	19.88
Gender	38	1	2	1.32	0.471
Self-Efficacy	38	36	63	44.66	5.132
Learning Performance-H	38	6.00	98	66.521	25.246
Learning Performance-Al	38	10.00	98	66.945	26.197
Motivation-H	38	54	96	77.82	11.229
Motivation-AI	38	63	103	81.82	11.6222

Table 8. Results of Repeated Measure ANOVAs, a One Way and Three Two-Way Analyses

Source		*F*	*df*	*Sig.*
Pedagogy				
	Learning performance	0.02	1	0.887
	Learning motivation	13,97	1	0.001[a]
Age				
	Learning performance	6.825	2	0.003*
	Learning motivation	0,710	2	0.499
Pedagogy*Age				
	Learning performance	1.025	2	0.369
	Learning motivation	2.68	2	0.083
Gender				
	Learning performance	0.166	1	0.686
	Learning motivation	4.948	1	0.032*
Pedagogy*Gender				
	Learning performance	0.004	1	0.952[a]
	Learning motivation	0.025	1	0.879
Efficacy				
	Learning performance	0.31	1	0.862
	Learning motivation	0,371	1	0.546
Pedagogy*Efficacy	Learning performance	0.234	1	0.631
	Learning motivation	0.413	1	0.525

* $p<0.05$

Table 9: Test Results for 14 Null Hypotheses

Null Hypothesis	Conclusion
H1a: Learning performance means between AI-tutoring and human are equal.	Not rejected
H1b: Motivation means between AI-tutoring and human are equal.	Rejected (p = 0.001)
H2a: Learning performance means between age groups (children, youths, adults) are equal.	Rejected (p = 0.003)

Null Hypothesis	Conclusion
H2b: Motivation means between age groups are equal.	Not rejected
H2c: No interaction between age groups and pedagogy in learning performance.	Not rejected
H2d: No interaction between age groups and pedagogy in motivation.	Not rejected
H3a: Learning performance means between females and males are equal.	Not rejected
H3b: Motivation means between females and males are equal.	Rejected ($p = 0.032$)
H3c: No interaction between gender and pedagogy in learning performance.	Not rejected
H3d: No interaction between gender and pedagogy in motivation.	Not rejected
H4a: Learning performance means between low and high efficacy students are equal.	Not rejected
H4b: Motivation means between low and high efficacy students are equal.	Not rejected
H4c: No interaction between self-efficacy and pedagogy in learning performance.	Not rejected
H4d: No interaction between self-efficacy and pedagogy in motivation.	Not rejected

Note: $p < 0.05$ indicates statistical significance

Conclusion And Discussion

1. Pedagogy Influence:

The study uncovered a significant impact of AI tutoring on elevating students' motivation compared to traditional human teaching methods. Although inconclusive evidence regarding differences in learning performances was observed, the heightened motivation suggests that AI possesses unique capabilities to engage students and cultivate interest in learning. The study refrained from delving into specific AI functionalities contributing to increased interest, leaving room for future research. The revolutionary implications of low-cost AI-powered education in democratizing access to quality teaching emphasize its transformative potential for education.

2. Age Influence:

Age emerged as a critical factor, with the youth group demonstrating the highest efficiency in acquiring piano skills. The absence of a significant interaction between age and teaching method implies that the positive effect of age on learning performances holds true irrespective of instructional approaches.

This underscores the importance of age-specific considerations in educational strategies. Future research could explore optimal age ranges for AI-enhanced learning in various subjects, ensuring tailored approaches that maximize effectiveness across different developmental stages.

3. Gender Influence:

Gender was identified as a noteworthy factor affecting learning motivation, with females exhibiting higher motivation compared to males. Interestingly, there was no significant interaction between gender and teaching method, suggesting that AI's impact on motivation is consistent across genders. This challenges assumptions about AI-enhanced learning being more appealing to one gender over the other, emphasizing the potential for gender-inclusive AI-powered approaches in addressing educational disparities.

4. AI Functionality and Future Implications:

The study did not explicitly address key AI functionalities, such as a widescreen display for clear piano demonstration, integration of music theory, gamified interactive quizzes for memory reinforcement, real-time assessment, and feedback. Additionally, electronic sheet music features, including automatic page-turning, error highlighting, and the availability of simplified notation and waterfall displays, play a crucial role in enhancing student engagement and interest in learning. These functionalities facilitate easier sight-reading and provide real-time, accurate feedback on playing mistakes. Future research should explore how these features further impact student learning interest and expand their application to other subjects.

In summary, AI tutoring excels in fostering motivation, raising questions about the comparative effectiveness of learning performances against traditional human teaching methods. Age and gender considerations underscore the necessity for personalized, age-sensitive educational approaches. The study sheds light on intricate interactions between variables and teaching methods, urging further research to explore specific AI functionalities influencing motivation, identify optimal age ranges for AI education, and ensure gender-inclusive design.

The discussion reveals comparable effects of AI tutoring and human methods on learning performance, with AI tutoring showing significantly higher motivation scores. The positive impact of age, particularly in youth, highlights the importance of tailoring educational strategies to specific age groups. Gender differences in motivation emphasize the need for gender-inclusive design in AI-powered education. Acknowledging limitations such as a small sample size and brief duration, the study calls for future research to investigate long-term effects and nuanced interactions of age, gender, and self-efficacy in AI-enhanced learning environments. Additionally, the discussion introduces key AI functionalities and their implications, suggesting avenues for future research, including enhancing learning interest and expanding

application to other subjects, providing a comprehensive perspective on AI's transformative potential in education.

References

Hallam, S. (2002). Musical motivation: Towards a model synthesizing the research. *Music Education Research*, *4*(2), 225–244.

Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review, 6(1), 27–54.*

Kindi, R., Leddo, J., Kindi, M., Bhandarkar, S., Chadeva, N., Dasari, S., & Karmacharya, A. (2016). Comparing a-list empire educational software to Khan Academy's: increasing student performance by incorporating artificial intelligence into video-based instruction. *International Journal of Advanced Educational Research, 1(6), 1-6.*

Leddo, J., Boddu, B., Krishnamurthy, S., Yuan, K., & Chippala, S. (2017). The effectiveness of self-directed learning vs. human learning on gifted and talented vs. non-gifted and talented students. *International Journal of Advanced Educational Research, 2(6), 18-21.*

Leddo, J., Guo, Y., Liang, Y., Joshi, R., Liang, I., Guo, W., & Bailey, S. (2019). Artificial intelligence and voice-powered electronic textbooks.