Hsu, T.-C., Huang, H.-L., Hwang, G.-J., & Chen, M.-S. (2023). Effects of Incorporating an Expert Decision-making Mechanism into Chatbots on Students' Achievement, Enjoyment, and Anxiety. *Educational Technology & Society*, *26*(1), 218-231. https://doi.org/10.30191/ETS.202301_26(1).0016

Effects of Incorporating an Expert Decision-making Mechanism into Chatbots on Students' Achievement, Enjoyment, and Anxiety

Ting-Chia Hsu[1*], Hsiu-Ling Huang[2], Gwo-Jen Hwang[2*] and Mu-Sheng Chen[1]

[1]Department of Technology Application and Human Resource Development, National Taiwan Normal University, Taiwan // [2]Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan // ckhsu@ntnu.edu.tw // hsiuling427@gmail.com // gjhwang.academic@gmail.com // mushengchen946@gmail.com

[*]Corresponding author

ABSTRACT: In traditional instruction, teachers generally deliver the content of textbooks to students via lectures, making teaching activities lack vibrancy. Moreover, in such a one-to-many teaching mode, the teacher is usually unable to check on individual students' learning status or to provide immediate feedback to resolve their learning problems. Chatbots provide an opportunity to address this problem. However, conventional chatbots generally serve as information providers (i.e., providing relevant information by matching keywords in a conversation) rather than as decision-making advisors (i.e., using a knowledge-base with a decision-making mechanism to help users solve problems). Thus, this study proposes an expert decision-making-based chatbot to facilitate individual students' construction of knowledge during the learning process. A quasi-experiment was conducted to compare the differences in the performances and perceptions of students using the expert decision-making-based chatbot (EDM-chatbot) and the conventional chatbot (C-chatbot) in the activities of a geography course. One class of 35 students was the experimental group, using the EDM-chatbot. The other class of 35 students was the control group, using the C-chatbot. The results of the study showed that the EDM-chatbot combined with expert decision-making knowledge significantly improved students' learning achievement and learning enjoyment as well as reducing their learning anxiety, showing the value of the proposed approach.

Keywords: Artificial Intelligence in Education, Expert knowledge, Decision tree, Chatbot, Interactive learning system

1. Introduction

In recent years, several studies have reported the benefits of using ICT in traditional instruction, such as the use of multimedia to present learning content. On the other hand, scholars have found that students generally need immediate support to help them address their misconceptions or solve any problems they encounter (Weaver, 2006). However, in a traditional classroom, the teacher may be the only person who can answer students' questions. With dozens of students in a class, it is almost impossible for teachers to provide instant feedback to individual students. Therefore, it is important to encourage students to find answers themselves using information tools. With the increasing use of Artificial Intelligence (AI) technologies in education, the main research topics include intelligent tutoring systems for special education; natural language processing for language education; educational robots for AI education; educational data mining for performance prediction; discourse analysis in computer-supported collaborative learning; neural networks for teaching evaluation; affective computing for learner emotion detection; and recommender systems for personalized learning (Chen et al., 2022). Few studies have considered humanity when employing AI in education. A previous study employed human-centered AI to give students individual responses by analyzing their learning behaviors, learning environments, or strategies (Yang, 2021). Yang (2021) pointed out that AI research in education is encountering new challenges of reshaping the research trend from technology to humanity. The climate unit is one of the most complicated learning topics for students in the discipline of geography because there are numerous conditions and requirements for judging climate classification. Giving students systematic and personalized guidance when learning this topic has become crucial. Therefore human-centered AI should be designed to support the self-learning of geography.

Self-inquiry, that is, making inquiries about questions by oneself, can increase one's learning achievement and is therefore an effective strategy for students to achieve further understanding. In the field of education, chatbots serve as a learning tool where information needed for education can be stored in a database and can be retrieved or supplemented at any time by querying the bot, either orally or through text (Wollny et al., 2021). However, if each learning note in the chatbot is independent and there is no scaffolding option for students to select, they may fall into the loop of the same Q&A cycle or miss some learning notes because they never mention the decision conditions during the conversation.

In this study, the climate unit learning content was organized and constructed so that students could learn by talking to a chatbot with two different mechanisms. Students could acquire knowledge from the chatbot and then organize that knowledge. This study aimed to reduce students' learning anxiety and maintain their learning enjoyment through chatbot learning to promote better learning outcomes. Accordingly, in this study, the control group used a C-chatbot as a teaching assistant to immediately respond to their questions by referring to the database containing each learning note. The experimental group used the EDM-chatbot which incorporated expert knowledge decision making, thus applying Artificial Intelligence in Education (AIED) to achieve adaptive learning. It was expected that the students could increase their learning achievement and enjoyment, while also reducing their learning anxiety through the use of the EDM-chatbot. The research questions in this study are as follows.

(1) Did the students using the EDM-chatbot have better learning achievement than those using the C-chatbot?
(2) Did the students using the EDM-chatbot have lower learning anxiety than those using the C-chatbot?
(3) Did the students using the EDM-chatbot have better learning enjoyment than those using the C-chatbot?

2. Literature review

2.1. Artificial Intelligence in Education (AIED)

AI means the ability of computers to perform tasks by simulating intelligent human behaviors (Duan et al., 2019). AI technologies have been applied in various forms in various fields, such as medical judgment precisely through image recognition via big data (Hulsen et al., 2019), or research on user interfaces that provide personalized feedback to users with voice and gesture recognition and natural language processing, the combination of voice recognition and natural language robots for business models (Okuda & Shoda, 2018), and health management (Nadarzynski et al., 2019).

AIED provides student-centered learning and uses AI to accelerate personalized learning on the one hand, providing students with personalized learning guidance or support based on their learning status, preferences, or personal characteristics (Hwang et al., 2020). Therefore, the role of the teacher changes with the help of AI and robots to provide personalized instruction, shifting to that of a supervisor or facilitator who designs and selects machines to support the students' learning, and who monitors their learning progress (Edwards et al., 2018). Therefore, innovative and productive learning activities have been designed, and better technology-enhanced learning applications have been developed to facilitate teaching, learning, or decision making; in particular, with the help of computer systems that simulate human intelligent reasoning, judgment, or prediction, AI technologies can provide personalized instruction to students (Hwang et al., 2020). For instance, a deep learning-assisted online intelligent English teaching system was proposed to help students improve the efficiency of English teaching based on their knowledge and personality acquisition (Sun et al., 2020), while online learning with social robots was used for assisting curriculum. A previous study attempted to combine the mind-mapping-guided chatbot approach to boost students' English speaking performance. This approach led to better performance than the conventional chatbot approach (Lin & Mubarok, 2021). Based on those successful applications of AIED, one of the AI techniques, supervised machine learning and decision tree, was employed in the interactive learning environment of the current study.

2.2. Chatbots

Chatbots, also known as virtual assistants, are a primitive form of AI software that can mimic human conversations and provide users with a new form of flexibility so as to achieve instant interaction (Dahiya, 2017). For instance, the emergence of chatbots, most notably Apple Siri, Microsoft Cortana, Facebook, and IBM Watson, is becoming a common trend in many fields such as medicine, the product and service industries, and education. Chatbots have a long history of being used as teaching agents in educational settings. The chatbots led to positive learning outcomes and help provide students with better learning and a better personalized learning experience (Vanichvasin, 2021). The use of chatbots in classroom tasks can have motivational effects (Fryer et al., 2017), as well as providing access to multimedia content with portability, flexibility, and immediate searching for information (Gikas & Grant, 2013). Chatbots are not limited to time and place, but can be used for supporting learning anytime and anywhere (Shah et al., 2016). Despite the maturity of chatbot technology, there is still a need to investigate how to properly add value to human practice in education through the use of chatbot technology, including the challenge of designing effective dialogues between humans and robot technology.

Due to the large number of students enrolled in the online course, students solved problems with the support of the instant feedback given by the web bot. There was a study on combining chatbots with a game learning platform to help students enter the game and perform multiple-choice tests through interactive discussions. Nenkov (2015) implemented intelligent agents on the platform IBM Bluemix using the IBM Watson technology. Chatbots have been applied in some courses such as computer science and computer networking fundamentals courses, including for Python learning (Okonkw & Ade-Ibijola, 2020). In another study, by working with a chatbot, post-secondary writers developed a thesis statement for their argumentative essay outlines, and the chatbot helped them refine their peer review feedback (Lin & Chang, 2020). A knowledge-based chatbot system was integrated into the teaching activities of a physical examination course in nursing education, using smartphones as learning devices to guide students in practicing their anatomy knowledge and analyzing the effectiveness and enjoyment of their learning (Chang et al., 2022). The impact of a teaching simulation activity using chatbots on pre-service teacher effectiveness was studied by Song et al. (2022). Accordingly, the chatbots have been used in language learning (Fryer et al., 2017), writing skills (Lin & Chang, 2020). Accordingly, the chatbot in the current study is a task-based chatbot designed to achieve learning goals by obtaining the intention and entities in the user's messages with natural language processing (NLP), adopting a free-form textual dialogue model that does not constrain the user's choices, and allows the user to interact more naturally with the robot.

2.3. Expert systems

Expert systems research has been one of the longest running and most successful areas of AI (Wagner, 2017). An expert system is a knowledge-based program that can be used to solve problems in a specific domain and provide "professional level" answers like human experts. The methodologies used in the domain can provide much help to geographers as a means of presenting geographic knowledge in a form that is accessible to many people (Fisher, 1989). Early research, based on domain knowledge provided by experienced teachers, proposed an expert system-based instructional approach to effective context-aware ubiquitous science learning (Wu et al., 2013). Using AI technologies to simulate teachers' knowledge and experience to provide individual students with personalized supports or guidance has been recognized as a potential solution (Pai et al., 2020).

A decision tree is a classification of knowledge and the relations of the concept nodes. Concepts shown as nodes and the relationships between the tree are connected with lines, like a concept map of learning material according to the classification of expert knowledge. In this study, an EDM-based chatbot was constructed based on the learner's prior knowledge measured against the results of a pre-assessment test, and a decision tree was generated based on the prior knowledge of the learner and similar former learners who had previously completed the course. The learning path was then recommended to the learner as a personalized learning tree. Decision tree classification is an important data classification technique which represents a mapping relationship between object attributes and object values. In order to employ the expert's knowledge in the application, the expert knowledge decision tree uses decision tables and decision trees to retrieve expert knowledge. The decision tables are used to confirm the completeness and correctness of the knowledge retrieval and to present the retrieved knowledge in a rule-based manner.

2.4. The current study

Effective classroom questioning is crucial for effective teaching and learning. Student questioning is an important self-regulatory strategy with multiple benefits for teaching and learning science (Van der Meij, 1994). Questioning is important for knowledge construction, discussion, self-assessment, and cognitive curiosity, and is also useful for enhancing learning achievement. For example, mutual rhetorical strategies in reading lessons were found to improve reading comprehension (Ersianawati et al., 2018). In addition, questioning strategies enhance the memory of text details in second language learning, and the comprehension of main ideas (Liu, 2021). A previous study explored the benefits of repetitive practice of short-answer questions which could enhance students' long-term memory for subsequent improvements in learning performance (Lu et al., 2021). However, it is rare to see students asking questions in conventional classes; meanwhile, teachers do not always have enough time to answer all of the students' questions in one class with the pressure of instructional progress. Therefore, this study attempted to develop an EDM-chatbot with a decision tree by using the expert system architecture and features to optimize the conversation path between the chatbot and the students, and to help students concentrate on the learning goals and focus on the interaction.

2.5. Learning enjoyment and anxiety

Learning anxiety refers to the negative emotions that students experience during the learning process; they may feel anxious at different stages of learning (Alnuzaili & Uddin, 2020), which is a common negative emotional response of learners during the learning process. Learners with higher levels of anxiety are more burdened with learning, resulting in lower learning efficiency; however, the learning process cannot be completely free of anxiety, meaning that learners with the right level of anxiety can perform better. Andrade and Williams (2009) suggested that this anxiety, called "facilitative anxiety," can make learners work harder and pursue better performance on tasks in class.

Enjoyment of learning is an affective orientation that stems from the pleasure and happiness that learners derive from learning activities (Shumow et al., 2013). By enhancing students' enjoyment of learning, they may develop a high level of interest in the learning goal, which will then allow them to sustain their learning and enhance their learning experience (Jack & Lin, 2018). In this study, a chatbot was used to help students learn about climate concepts. The chatbot acted as a teacher to guide students, and it was hoped that its use would enhance students' learning enjoyment.

3. Development of the Expert Decision Making (EDM)-based chatbot

This study used IBM Watson to build a chatbot for the geographical climate unit of a science course. Climate change is a complex environmental problem that can be used to examine students' understanding, gained through classroom communication, of climate change and its interaction. Jakobsson et al. (2009) found in a study conducted through a written test that students' understanding of climate change was poor. They pointed out, however, that a written test does not explicitly reveal students' knowledge. Therefore, in the present study, it was considered that students' understanding or meaning making of complicated issues such as climate change would be better if a communicative approach was used.

Table 1 shows examples of the expert knowledge for building the ID3 decision tree (Quinlan, 1983). There are 16 classifications (i.e., C_1, C_2…C_{16}) of weather, composed of nine constructs (i.e., elevation, cold in winter and cool in summer, latitude, rainfall, dry season, summer dry, stationary front, needle forests, snow (no rain)) which have their own different critical feature values, as shown in Table 1.

Table 1. Illustration of examples

Features									Class
Elevation[m] (A)	Cold in winter and cool in summer (B)	Latitude (L)	Rainfall [mm] (D)	Dry season (E)	Summer Dry (F)	Stationary front (G)	Needle Forests (H)	Snow and no rain (I)	
A2	Yes	L1	D2	Yes	No	No	Yes	Yes	C_1
A2	No	L1	D2	Yes	No	No	Yes	Yes	C_2
A1	Yes	L2	D1	Yes	No	No	No	No	C_3
A1	Yes	L2	D1	No	No	No	No	No	C_4
A1	Yes	L2	D2	Yes	No	Yes	No	No	C_5
A1	Yes	L2	D2	Yes	No	No	No	No	C_6
A1	Yes	L2	D4	Yes	No	No	No	No	C_7
A1	Yes	L3	D2	No	Yes	No	No	No	C_8
A1	Yes	L3	D2	Yes	No	Yes	No	No	C_9
A1	Yes	L3	D2	Yes	Yes	No	No	No	C_{10}
A1	Yes	L3	D2	Yes	No	No	No	No	C_{11}
A1	Yes	L3	D3	Yes	No	No	No	No	C_{12}
A1	Yes	L3	D4	Yes	No	No	No	No	C_{13}
A1	No	L4	D3	Yes	No	No	No	Yes	C_{14}
A1	No	L4	D3	Yes	No	No	Yes	No	C_{15}
A1	No	L4	D3	Yes	No	No	No	No	C_{16}

Entropy is used to determine the importance of the construct which is used for classification, so as to form an effective decision tree. We can calculate the gained information of each feature shown in the following based on the training data.

- Feature A. Elevation: A1 < 3000; A2 ≥ 3000
- Feature B. Cold in winter and cool in summer? Yes; No
- Feature L. Latitude: L1 = None, L2 < 30, 30 ≤ L3 < 60, L4 ≥ 60
- Feature D. Rainfall: D1 ≥ 1500, 500 ≤ D2 < 1500, 250 ≤ D3 < 500, D4 < 250
- Feature E. Dry season: Yes; No
- Feature F. Summer Dry: Yes; No
- Feature G. Stationary front: Yes; No
- Feature H Needle Forests: Yes; No
- Feature I. Snow and no rain: Yes; No

Example 1:
Gain(S,A)= Entropy(S)- Entropy(A) =
$-p_{C1} \times log_2(p_{C1}) - p_{C2} \times log_2(p_{C2}) - p_{C3} \times log_2(p_{C3}) \dots - p_{C16} \times log_2(p_{C16}) - (-p_{A1} \times log_2(p_{A1}) - p_{A2} \times log_2(p_{A2})) = \left(-\left(\frac{1}{16}\right) \times log_2\left(\frac{1}{16}\right) \times 16\right) - \left(-\left(\frac{2}{16}\right) \times log_2\left(\frac{2}{16}\right) - \left(\frac{14}{16}\right) \times log_2\left(\frac{14}{16}\right)\right) \cong 3.46$

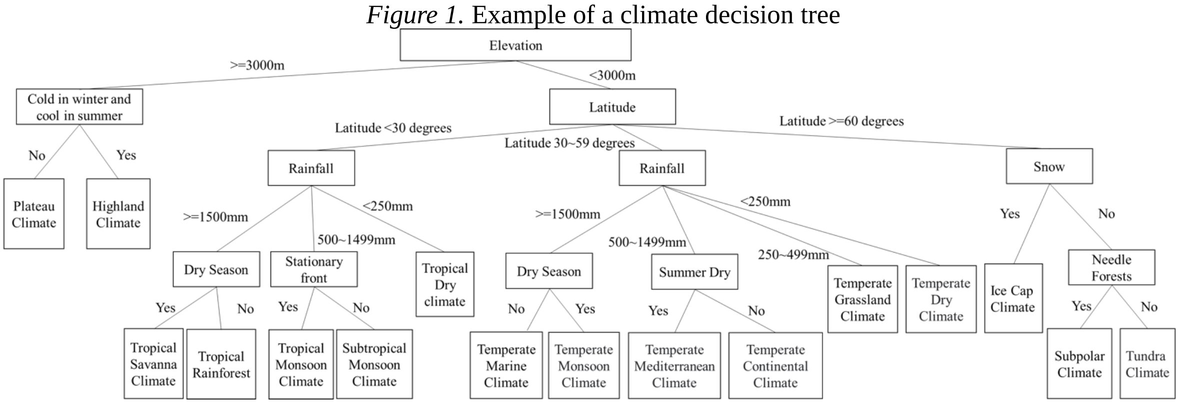
To develop a decision rule for correctly classifying training examples, ID3 performs feature tests by first selecting a feature, and then using the selected feature to classify the examples into subclasses. Next, it calculates the information entropy to determine the importance of the feature based on the following Formula 1.

Formula 1:
$Entropy(I) = \sum_{i=1}^{C} -\frac{Ni}{N} log_2\left[\frac{Ni}{N}\right]$

In this formula, N, N_i, and C represent the total number of training examples, the number of examples that belong to class i, and the number of classes, respectively. Entropy can be used as an indicator of the messiness of the information quantity. The calculation of Gain (S, A) indicates the profit of using attribute A (elevation in Table 1) to partition the data set S. The larger the value of Gain, the less messy the data in attribute A, and the better A can be used to classify data; the smaller the value of Gain, the greater the confusion of data in attribute A, and the worse the classification of data will be. Therefore, the information gain (S, A) represents the degree of reduction of the information complexity under the specific condition of using attribute A, equal to the information gain value of feature A. The result is calculated to be 3.46 in example 1. For example, when testing the feature "elevation," the 16 samples are divided into two subclasses, "≥ 3000" and "< 3000." Then, the sum of information entropy of each subcategory can be calculated. By subtracting the information entropy of these subclasses from the information entropy of the original training example set, ID3 deduces the information gain of the feature "elevation" as the root note at the present stage. In a similar way, the information gain for each feature can be obtained separately for testing.

When ID3 searches for features that provide the greatest information gain, the maximum information gain is obtained by comparing the gain of each feature. Next, other features are tested and the decision tree is expanded until all leaf nodes contain examples falling into a single class, as shown in Figure 1. Five vegetation groups can be distinguished as the equatorial zone, arid zone, temperate zone, cool temperate zone, and polar region. The second letter of the classification is precipitation (weather or names of climate types), and the third letter is the temperature of the location (Kottek et al., 2006).

Figure 1. Example of a climate decision tree



This chatbot has the function of learning, and adopts fuzzy matching in IBM Watson as a technique to make the conversation with students smoother. Fuzzy matching enables the system to deal with stemming, misspelling, or partial matches. For instance, the term “running” could also be interpreted as “run,” and “bananas” could be interpreted as “banana” when dealing with the “stemming” status. Such a stemming problem occurs more in English than in Chinese. On the other hand, misspelling and partial matches more frequently occur in Chinese interactions. For example, dealing with misspelling means that even if the order of words in a phrase is incorrectly located or reversed, the original sentence can still be interpreted. “Partial match” refers to the function whereby the system is able to judge the meaning of the statement as long as certain attributes are detected in that statement. The system architecture is shown in Figure 2. The C-chatbot is shown in Figure 3. The system will search for examples and rules when it receives any questions.

Figure 2. The system architecture diagram of the EDM-chatbot

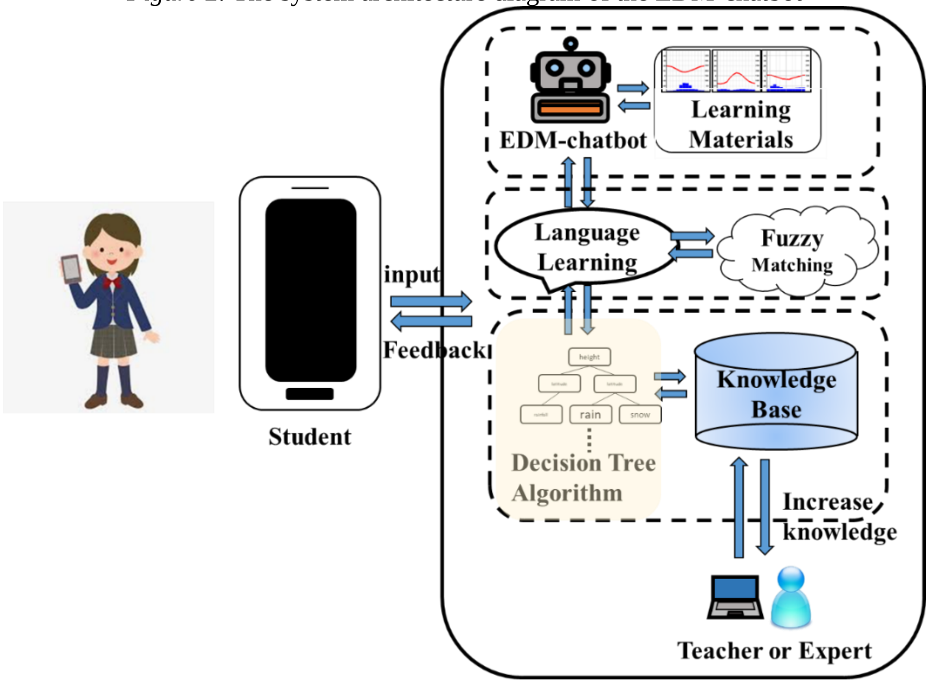
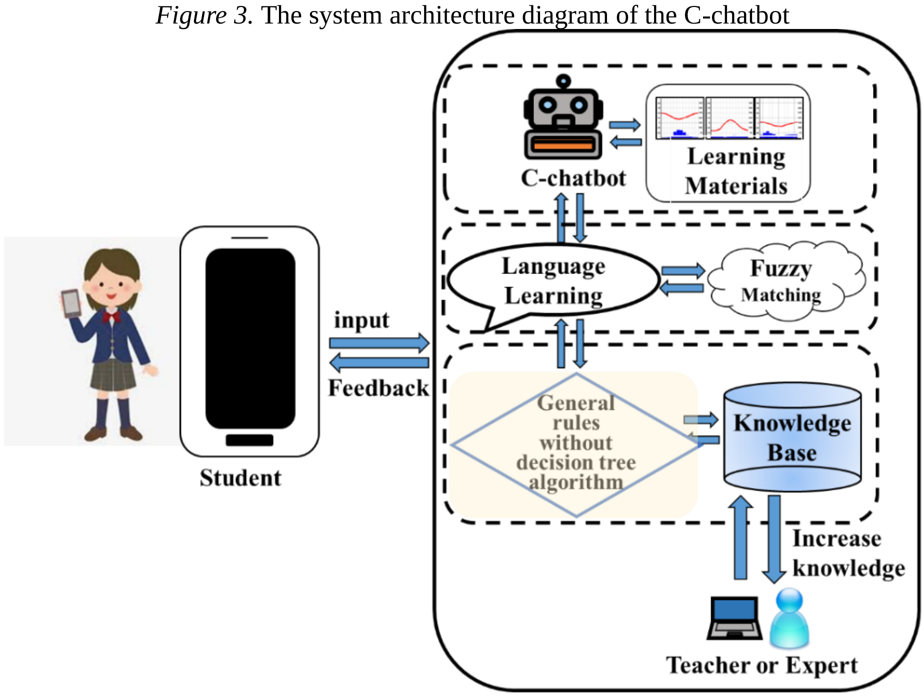
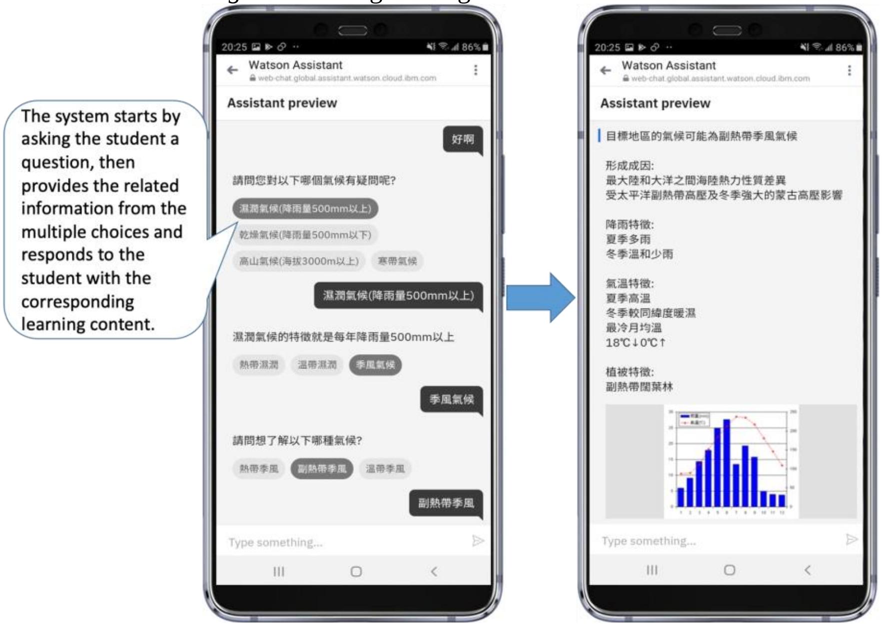


Figure 3. The system architecture diagram of the C-chatbot



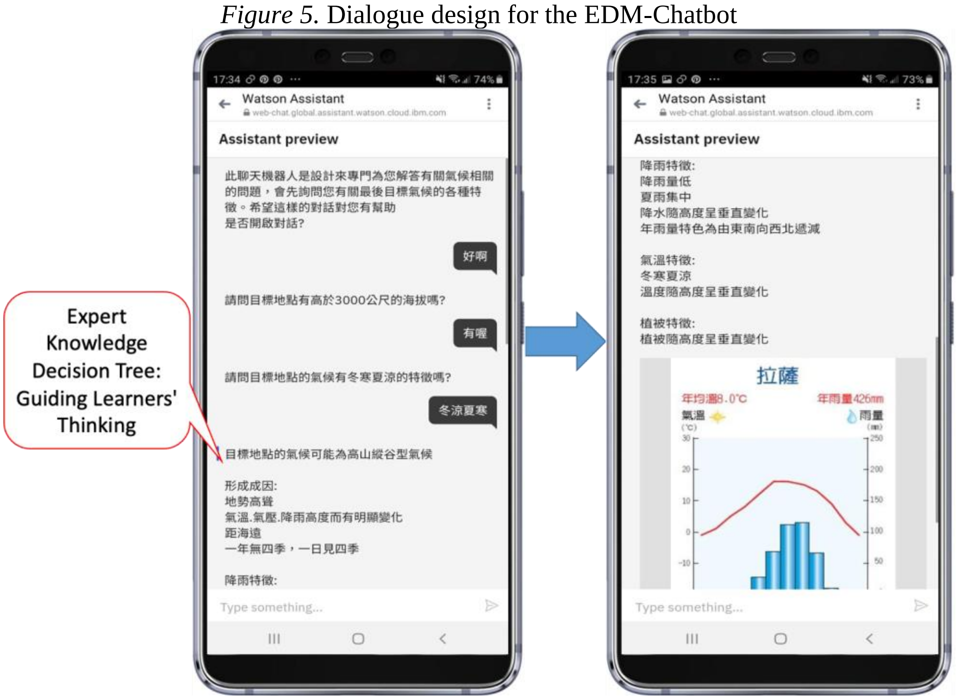
The C-chatbot conversations were arranged according to the same climate feature sequences, and the dialogue replies were designed using the IBM Watson technology which can recognize similar semantics said by the students. For example, in Figure 4, the system starts by asking the student a question, then provides related information from the multiple choices, and responds to the student with the corresponding learning content in the database. This is a so-called conventional chatbot. Because the C-chatbot easily falls into the same conversation loop, the example provides the conversation options for students to choose by clicking the dialogue items when they want to interact with the C-chatbot. Meanwhile, the students can also directly reply with the words they want to say if they do not just want to click the options.

Figure 4. Dialogue design for the C-Chatbot



The EDM-chatbot conversations were processed by an algorithm, so their conversations were more streamlined based on the expert knowledge and decision tree, and students were able to organize their knowledge and find their learning goals more easily. Examples comparing the two systems are shown in Figures 4 and 5.

Figure 5. Dialogue design for the EDM-Chatbot



4. Experimental design

The geographical climate expert system was designed to be used as a reference for many natural ecological studies and human activities. Each climate variable was analyzed separately for climate patterns, or data could be aggregated by using climate classifications. These classifications usually correspond to vegetation distributions, in the sense that each climate type is dominated by a vegetation zone or an ecological region (Belda et al., 2014). Köppen was trained as a plant physiologist and believed that plants are indicators of many aspects of climate change (Belda et al., 2014). Köppen's climate classification is based on two climate elements, temperature and precipitation, and is confirmed by the distribution of natural vegetation.

4.1. Participants

In order to examine the effects of the chatbots on enhancing the learning performance of the geographical climate unit, two classes of high school students were recruited. Their average age was 17 years old. One class (N = 35) was the experimental group using the EDM-chatbot, while the other (N = 35) was the control group applying the C-chatbot. The same teacher taught both groups. The study was approved by the Research Ethics Committee of the Graduate Institute of Digital Learning and Education (approval number REA-2020-0705A). Subjects were informed that participating in the experiment was voluntary and they could withdraw from the study at any stage.

4.2. Measuring tools

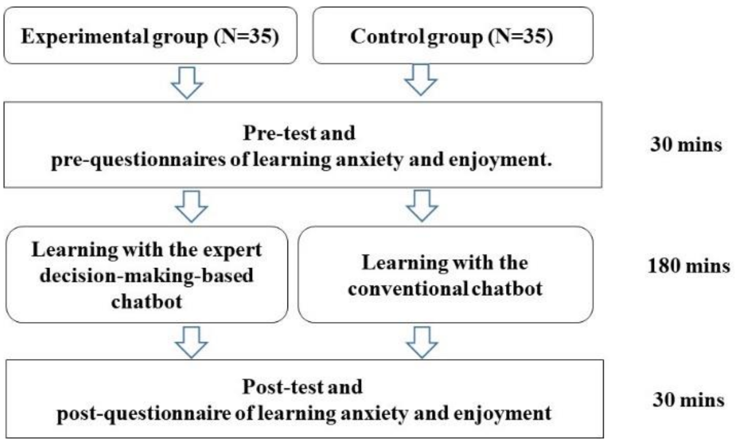
For this study of applying a chatbot to the climate unit, two master students with teaching experience of 2 years on average were involved in the content development. The test of the content was jointly reviewed by two experts, and corresponded to the learning content of the chatbot. The test comprised 10 multiple-choice items, with a perfect score of 100 points in total.

The scale of learning anxiety and enjoyment used in this study was selected from the Learning Anxiety and Engagement Questionnaires (Hsu & Hwang, 2021). There are nine items in the scale of learning anxiety, which was assessed on a 5-point scale with an internal consistency reliability of 0.91. An example item is: *Learning with the chatbot makes me nervous.* There are three items in the scale of enjoyment, which was evaluated on a 5-point scale with an internal consistency reliability of 0.90. An example item is: "*The actual process of learning with the chatbot is pleasant.*"

4.3. Experimental procedure

The experimental process is shown in Figure 6. Before the chatbot-based learning activity, the students took a pre-test to examine their basic knowledge related to geographical climate and filled out the learning anxiety and enjoyment questionnaires.

Figure 6. Research process



During the learning activity, each group spent three hours in total. The students were first guided to install the chatbot on their mobile phones and use it to complete their individual learning tasks by answering a set of questions on learning sheets prepared by the teacher. All the students in the same group used their personalized chatbot in the same classroom. Individual students needed to interact with the chatbot to get hints for the self-learning tasks during the three periods, where each period was 50 minutes with a 10-minute break between. They could talk to the chatbot via audio or text input. It should be noted that both groups were asked to complete the same geography learning unit on climate. The only difference between the two groups was that the experimental group used the EDM-chatbot, while the control group used the C-chatbot. Both groups completed the experiment in half a day, but the two experiments were conducted on different days.

After the learning activity, the students took a post-test during which they could not use the chatbot. The learning achievement post-test comprised 10 multiple-choice items related to the knowledge of geographical climate in the chatbot. The students also completed the learning anxiety and enjoyment post-questionnaires.

After the experiment, the statistical analysis was performed. The results are presented in the next section.

5. Experimental results

The normality test was firstly carried out using the Kolmogorov-Smirnov test according to the research data; it was found that all data of each group did not conform to the normal distribution (i.e., all the p values of Shapiro-Wilk were smaller than 0.05). Therefore, the statistical methods of non-parametric analysis were conducted.

5.1. Learning achievement

First, the Wilcoxon signed-rank test was performed to compare the learning achievement pre-test and post-test of each group, as shown in Table 2. The results revealed that the learning achievement of the post-test (M = 57.429, SD = 11.464) was significantly higher than that of the pre-test (M = 53.143, SD = 12.071) in the control group (Z = -2.044^{*}, $p < .05$). Meanwhile, the learning achievement of the post-test (M = 65.714, SD = 15.202) was remarkably higher than that of the pre-test (M = 57.143, SD = 23.082) in the experimental group (Z = -2.736^{**}, $p < .01$). Consequently, both systems were helpful for self-learning.

Table 2. Results of the Wilcoxon signed-rank test on learning achievement for the two groups

Group	N	Pre-test		Post-test		Z
		Mean	SD	Mean	SD	
Experiment	35	57.143	23.082	65.714	15.202	-2.736^{**}
Control	35	53.143	12.071	57.429	11.464	-2.044^{*}

Note. $^{**}p < .01$; $^{*}p < .05$.

Next, the Mann-Whitney U test was performed for comparing the pre-test of the two groups. The results confirmed that there was no significant difference between the prior knowledge of the students (U = 485.500; Z = -1.527; p = .127 > .05). Finally, the Mann-Whitney U test was performed again for comparing the post-test of the two groups. The results found that the learning achievement (M = 65.714, SD = 15.202) of the experimental group outperformed the learning achievement (M = 57.429, SD = 11.464) of the control group significantly (U = 416.500, $p < .05$), as shown in Table 3.

Table 3. Results of the Mann-Whitney U test on learning achievement for the two groups

Group	N	Mean	SD	Average Rank	Rank Sum	U	W	Z
Experiment	35	65.714	15.202	41.10	1438.50	416.500^{*}	1046.500^{*}	-2.364^{*}
Control	35	57.429	11.464	29.90	1046.50			

Note. $^{*}p < .05$.

5.2. Learning anxiety

First, the Wilcoxon signed-rank test was performed to compare the learning anxiety pre-test and post-test of each group, as shown in Table 4. The results revealed that there was no significant difference between the anxiety pre-test (M = 3.083, SD = 0.439) and the anxiety post-test (M = 3.117, SD = 0.279) of the control group (Z = -0.432, $p > .05$). On the contrary, there was a significant difference between the anxiety pre-test (M = 2.844, SD = 0.490) and the anxiety post-test (M = 2.390, SD = 0.611) in the experimental group (Z = -2.893^{**}, $p < .01$). It was found that the EDM-chatbot was helpful for significantly decreasing the students' learning anxiety.

Next, the Mann-Whitney U test was performed for comparing the learning anxiety pre-test of the two groups. The results confirmed that there was no significant difference between the prior learning anxiety of the students (U = 454.500; Z = -1.883; $p > .05$). Finally, the Mann-Whitney U test was performed again for comparing the learning anxiety post-test of the two groups. The results found that the learning anxiety (*M* = 2.390, *SD* = 0.611) of the experimental group was lower than the learning anxiety (*M* =3.117, *SD* = 0.279) of the control group, significantly (U =216.500***, p < .001), as shown in Table 5.

Table 4. Results of the Wilcoxon signed-rank test on learning anxiety for the two groups

Group	*N*	Pre-test		Post-test		Z
		Mean	*SD*	Mean	*SD*	
Experiment	35	2.844	0.490	2.390	0.611	-2.893**
Control	35	3.083	0.439	3.117	0.279	-0.432

Note. **p < .01.

Table 5. Results of the Mann-Whitney U test on learning anxiety for the two groups

Group	*N*	Mean	*SD*	Average Rank	Rank Sum	U	W	Z
Experiment	35	2.390	0.611	24.19	846.50	216.500***	846.500***	-4.691***
Control	35	3.117	0.279	46.81	1638.50			

Note. ***p < .001.

5.3. Learning enjoyment

First, the Wilcoxon signed-rank test was performed to compare the learning enjoyment pre-test and post-test of each group, as shown in Table 6. The results revealed that the enjoyment post-test (*M* = 2.790, *SD* = 0.801) was lower than the enjoyment pre-test (*M* = 3.419, *SD* = 0.711) in the control group, significantly (Z = -3.105**, p < .01). This finding revealed that the students perceived lower learning enjoyment when they carried out self-learning with the C-chatbot. On the contrary, there was no significant difference between the enjoyment pre-test (*M* = 3.324, *SD* = 0.810) and the enjoyment post-test (*M* = 3.343, *SD* = 0.865) in the experimental group (Z = -0.082, p > .05).

Table 6. Results of the Wilcoxon signed-rank test on learning enjoyment for the two groups

Group	*N*	Pre-test		Post-test		Z
		Mean	*SD*	Mean	*SD*	
Experiment	35	3.324	0.810	3.343	0.865	-0.082
Control	35	3.419	0.711	2.790	0.801	-3.105**

Note. **p < .01.

Next, the Mann-Whitney U test was performed for comparing the learning enjoyment pre-test of the two groups. The results confirmed that there was no significant difference between the prior learning enjoyment of the students (U = 570.000; Z = -0.524; p>*.05*). Finally, the Mann-Whitney U test was performed again for comparing the learning enjoyment post-test of the two groups. The results found that the learning enjoyment (*M* = 3.343, *SD* = 0.865) of the experimental group was higher than the learning enjoyment (*M* = 2.790, *SD* = 0.801) of the control group, significantly (U = 404.000*, p < .05), as shown in Table 7.

Table 7. Results of the Mann-Whitney U test on learning enjoyment for the two groups

Group	*N*	Mean	*SD*	Average Rank	Rank Sum	U	W	Z
Experiment	35	3.343	0.865	41.46	1451.00	404.000	1034.000	-2.566*
Control	35	2.790	0.801	29.54	1034.00			

Note. *p < .05.

6. Discussion

The learning discipline in the current study, geography, is one of the humanities learning subjects. This study adopted an AI chatbot as an interactive mentor for self-learning students and compared two different chatbot designs for smart phones so as to determine the contributions of expert-based decision tree chatbots with human-centered AI to the humanities learning subjects. The EDM-chatbot can provide different levels of responses from a decision tree according to students' answers. Precision education is very similar to precision medicine in that precision medicine must be tailored to each individual difference, including genes, living environment, and lifestyle (Lin et al., 2021); in the same way, each student will face different difficulties and obstacles in learning which can be addressed by precision education. Rus et al. (2013) found that the effectiveness of teaching and learning can be improved by using an intelligent assistance system with conversational capabilities or in the form of a chatbot. The current study also proved that the chatbot used in self-learning of humanities subjects is a good means of application to promote the learning achievement of self-learning.

The C-chatbot is a passive way to perform conversation with students, although it can recognize most of the students' semantics. However, each learning note is separately stored in the database. The conversation starts from the same sequence for every student so that the students' anxiety cannot significantly decrease. They have to pay attention so as not to miss any key point or fall into the loop of the problem. The current study provided the students with the EDM-chatbot with embedded expert decisions underpinning the system so as to provide appropriate guidance for individual students and to check each learning note based on the decision tree during conversation. Thus, the application of human-centered AI could be achieved. With such a form of self-inquiry underpinned by expert decision tree scaffolding for individuals, students can systematically and actively gain relevant concepts for knowledge construction. From the perspective of meaningful learning, connecting information from different sources in an attempt to combine what they have learned is intended to reinforce meaning and enable learners to construct knowledge effectively (Dahiya, 2017). By constructing learning nodes through expert knowledge, meaningful learning is constructed, and appropriate learning paths are selected for learners to proceed in a sequential manner.

In this study, the EDM chatbot played the role of an interactive knowledge map that provided learners with learning paths, learning support for different learners, and self-adjustment. Students using the EDM chatbot to learn could make adjustments according to their needs. For example, if the student was already familiar with the classification of highland climates, he or she would then skip this classification result according to the chatting interaction and be guided to the next type of result. This is why the students showed better academic performance after self-learning with the EDM chatbot than those who used the C-chatbot, because the application of the decision tree checking during conversation became an automatic mind tool for students or scaffolding of learning nodes. In traditional education, teachers may be discriminatory in their conversations with students, even if they are unaware of it. In chatbot learning, discriminatory language is removed during the process of setting up the chatbot. If teachers pass on the wrong knowledge and do not correct it in time, it may cause learning difficulties for students. With the chatbot approach to learning, this problem can be solved by making sure that the chatbot is built to be free of knowledge errors and guidance. In sum, the EDM-chatbot group showed lower learning anxiety than the C-chatbot group because they did not need to be afraid of the level of questions they asked, and they could get the required learning responses from the robots (Babel et al., 2021). Simplifying the chatbot conversation process by means of decision trees allows students to find adaptive learning content or answers more quickly, so they will not always be in the same dialogue loop. Therefore, the EDM-chatbot can not only reduce students' learning anxiety, but can also maintain their learning enjoyment.

7. Conclusions

The core of the human-centered AIED research is to support students' learning by designing instruments which address students' learning dilemmas and provide them with equitable access to learning opportunities. In this study, an EDM-chatbot was constructed using IBM Watson, and expert decision making was incorporated into a multi-round dialogue mechanism to provide students with adaptive learning. In AI algorithmic systems, biased words related to culture, religion, and gender are avoided, providing learners with a level playing field, and new algorithms can achieve closer to human performance with intelligent analysis, diagnosis, prediction, treatment and prevention, providing adaptive learning for students (Yang, 2021). Personalizing instruction to the unique needs of learners, developing teaching strategies (Tempelaar et al., 2021), and creating human-centered learning technologies achieved the standards of precision education (Luan & Tsai, 2021). The experimental results showed that the EDM-chatbot was more effective than the C-chatbot in terms of promoting students' learning achievement, reducing their learning anxiety, and increasing their learning enjoyment. The chatbots use natural language processing to judge the focus of the students' conversation. They will not respond to students using any biased or discriminatory language, but will converse fluently and answer the climate issue first. The conversations of the chatbots in this study were centered on the learning content and were verified to contain no discriminatory language. The learning content was designed based on the textbook content and was verified by the instructor to be explanatory and reliable. Teaching requires interaction, and chatbots provide students with immediate guidance and answers, thereby increasing learning achievement and interest, and enhancing students' enjoyment of learning (Fryer et al., 2019).

Shneiderman (2020) described human-centered AI as a promising direction for designing AI systems that support human self-efficacy, promote creativity, clarify responsibility, and facilitate social participation. This study used a chatbot to help students learn knowledge about the climate. Chatbots can solve the problems of conventional education. It is difficult for teachers to deal with the problems encountered by each student or to spend too much time on specific learning content. Students can use a chatbot to find answers on their own and to study the content they are not familiar with at any time. However, chatbots have some limitations. Chatbots are more suitable for structured or rule-based learning content. The process of building chatbots for unstructured learning content will be very complicated, and it is also difficult for students using general chatbots to organize their knowledge structure. The chatbot does not know the student's ability in advance or their learning situation during the conversation, so it may be necessary to confirm with a pop quiz, or as in this study, options to hint and guide the students' direction can be used, as in C-chatbot, or a decision tree to structure and check the learning nodes of each student can be used, as in EDM-chatbot.

Despite the positive findings, there are some limitations to the present study that should be noted. First, if the students' answers are irrelevant to the question at hand, the chatbot might have to start the conversation from the beginning, which may make the students feel impatient. In addition to system stability and accuracy adjustment, future studies are encouraged to include a machine learning mechanism to refine the chatbot's natural language processing ability by analyzing the behavioral patterns and feedback of the students using the chatbots. It would also be valuable for future research to track students' learning emotions, or to compare the difference in the effects that voice chatbots and physically human-like chatbots have on students' learning. It is recommended that future studies first collect the learning achievement and engagement of students in traditional lectures, so that the performance of the students using e-learning combined with an AI mechanism for self-learning can be compared with the performance of students taught by a teacher in a traditional lecture class which cannot take any personalized responses into consideration. Because this study compared two mechanisms under the precondition of self-learning, teachers did not intervene in students' learning in this study. Research has identified teachers' intentions to adopt AI tools in the classroom as a factor that influences the integration of AI technologies or applications into educational curriculum design (Wang et al., 2021). Therefore, teachers' perspectives on chatbots can also be explored in future studies. Future studies are encouraged to propose other research objectives and hypotheses which are different from those in this study. In other words, it is suggested that teachers become an independent variable in further studies. Another limitation of this study is that it employed chatbots in a geographical climate unit only with limited self-learning time, so it is suggested that future studies try the highly interactive design of chatbots for different disciplines and courses for a longer period of time.

Acknowledgement

This study is supported in part by the National Science and Technology Council in Taiwan under contract numbers NSTC 111-2410-H-003-168-MY3 and 111-2410-H-011-007-MY3.

References

Alnuzaili, E. S., & Uddin, N. (2020). Dealing with anxiety in foreign language learning classroom. *Journal of Language Teaching and Research, 11*(2), 269-273. http://dx.doi.org/10.17507/jltr.1102.15

Andrade, M., & Williams, K. (2009). Foreign language learning anxiety in Japanese EFL university classes: Physical, emotional, expressive, and verbal reactions. *Sophia Junior College Faculty Journal, 29*(1), 1-24.

Babel, F., Kraus, J., Miller, L., Kraus, M., Wagner, N., Minker, W., & Baumann, M. (2021). Small talk with a robot? The Impact of dialog content, talk initiative, and gaze behavior of a social robot on trust, acceptance, and proximity. *International Journal of Social Robotics*, *13*, 1485-1498. https://doi.org/10.1007/s12369-020-00730-0

Belda, M., Holtanová, E., Halenka, T., & Kalvová, J. (2014). Climate classification revisited: From Köppen to Trewartha. *Climate research, 59*(1), 1-13. https://doi.org/10.3354/cr01204

Chang, C. Y., Kuo, S. Y., & Hwang, G. H. (2022). Chatbot-facilitated nursing education. *Educational Technology & Society, 25*(1), 15-27.

Chen, C. H., Koong, C. S., & Liao, C. (2022). Influences of integrating dynamic assessment into a speech recognition learning design to support students' English speaking skills, learning anxiety and cognitive load. *Educational Technology & Society, 25*(1), 1-14.

Dahiya, M. (2017). A Tool of conversation: Chatbot. *International Journal of Computer Sciences and Engineering, 5*(5), 158-161.

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data–evolution, challenges and research agenda. *International Journal of Information Management, 48*, 63-71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021

Edwards, C., Edwards, A., Spence, P. R., & Lin, X. (2018). I, teacher: Using artificial intelligence (AI) and social robots in communication and instruction. *Communication Education, 67*(4), 473-480. https://doi.org/10.1080/03634523.2018.1502459

Ersianawati, N. L., Santosa, M. H., & Suprianti, G. (2018). Incorporating reciprocal questioning strategy and numbered heads together in reading class. *International Journal of Language and Literature, 2*(1), 19-29. http://dx.doi.org/10.23887/ijll.v2i1.16090

Fisher, P. F. (1989). Expert system applications in geography. *Area, 31*(3), 279-287.

Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An Experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior, 75*, 461-468. Elsevier. https://doi.org/10.1016/j.chb.2017.05.045

Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in Human Behavior, 93,* 279-289. https://doi.org/10.1016/j.chb.2018.12.023

Gikas, J., & Grant, M. M. (2013). Mobile computing devices in higher education: Student perspectives on learning with cellphones, smartphones & social media. *The Internet and Higher Education, 19*, 18-26. https://doi.org/10.1016/j.iheduc.2013.06.002

Hsu, T. C., & Hwang, G. J. (2021). Interaction of visual interface and academic levels with young students' anxiety, playfulness, and enjoyment in programming for robot control. *Universal Access in the Information Society*. https://doi.org/10.1007/s10209-021-00821-3

Hulsen, T., Jamuar, S. S., Moody, A. R., Karnes, J. H., Varga, O., Hedensted, S., & McKinney, E. F. (2019). From big data to precision medicine. *Frontiers in Medicine, 6*, 34. https://doi.org/10.3389/fmed.2019.00034

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, 100001, 1-5. https://doi.org/10.1016/j.caeai.2020.100001

Jack, B. M., & Lin, H.-S. (2018). Warning! Increases in interest without enjoyment may not be trend predictive of genuine interest in learning science. *International Journal of Educational Development, 62*, 136-147. https://doi.org/10.1016/j.ijedudev.2018.03.005

Jakobsson, A., Mäkitalo, Å., & Säljö, R. (2009). Conceptions of knowledge in research on students' understanding of the greenhouse effect: methodological positions and their consequences for representations of knowing. *Science Education, 93*(6), 978-995. https://doi.org/10.1002/sce.20341

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift, 15*(3), 25-263. https://doi.org/10.1127/0941-2948/2006/0130

Lin, C. J., & Mubarok, H. (2021). Learning analytics for investigating the mind map-guided AI Chatbot approach in an EFL flipped speaking classroom. *Educational Technology & Society, 24*(4), 16-35.

Lin, H. C., Tu, Y. F., Hwang, G. J., & Huang, H. (2021). From precision education to precision medicine. *Educational Technology & Society, 24*(1), 123-137.

Lin, M. P. C., & Chang, D. (2020). Enhancing post-secondary writers' writing skills with a chatbot. *Educational Technology & Society, 23*(1), 78-92.

Liu, H. (2021). Does questioning strategy facilitate second language reading comprehension? The Effects of comprehension measures and insights from reader perception. *Journal of Research in Reading, 44*(2), 339-359. https://doi.org/10.1111/1467-9817.12339

Lu, O. H., Huang, A. Y., Tsai, D. C., & Yang, S. J. (2021). Expert-authored and machine-generated short-answer questions for assessing students learning performance. *Educational Technology & Society, 24*(3), 159-173.

Luan, H., & Tsai, C. C. (2021). A Review of using machine learning approaches for precision education. *Educational Technology & Society, 24*(1), 250-266.

Nadarzynski, T., Miles, O., Cowie, A., & Ridge, D. (2019). Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A Mixed-methods study. *Digital Health, 5*, 1-12. https://doi.org/10.1177/2055207619871808

Nenkov, N. (2015). Implementation of a course in "Artificial Intelligence and Expert Systems" on top of a distance-learning platform. *Computer Modelling & New Technologies, 19*, 34-36.

Novak, J. D. (1990). Concept mapping: A Useful tool for science education. *Journal of Research in Science Teaching, 27*(10), 937-949. https://doi.org/10.1002/tea.3660271003

Okonkwo, C. W., & Ade-Ibijola, A. (2020). Python-bot: A Chatbot for teaching Python programming. *Engineering Letters, 29*(1), 25-34.

Okuda, T., & Shoda, S. (2018). AI-based chatbot service for financial industry. *Fujitsu Scientific and Technical Journal, 54*(2), 4-8.

Pai, C. K., Liu, Y., Kang, S., & Dai, A. (2020). The Role of perceived smart tourism technology experience for tourist satisfaction, happiness and revisit intention. *Sustainability, 12*(16), 6592. https://doi.org/10.3390/su12166592

Quinlan, J. R. (1983). Learning efficient classification procedures and their application to chess end-games. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine Learning: An Artificial Intelligence Approach* (pp. 463–482). Morgan Kaufmann. Springer. https://doi.org/10.1007/978-3-662-12405-5_15

Rus, V., D'Mello, S., Hu, X., & Graesser, A. (2013). Recent advances in conversational intelligent tutoring systems. *AI Magazine, 34*(3), 42-54. https://doi.org/10.1609/aimag.v34i3.2485

Shah, H., Warwick, K., Vallverdú, J., & Wu, D. (2016). Can machines talk? Comparison of Eliza with modern dialogue systems. *Computers in Human Behavior, 58*, 278-295. Elsevier. https://doi.org/10.1016/j.chb.2016.01.004

Shneiderman, B. (2020). Human-centered artificial intelligence: Three fresh ideas. *AIS Transactions on Human-Computer Interaction, 12*(3), 109-124. https://doi.org/10.17705/1thci.00131

Shumow, L., Schmidt, J. A., & Zaleski, D. J. (2013). Multiple perspectives on student learning, engagement, and motivation in high school biology labs. *The High School Journal, 96*(3), 232-252.

Song, D., Oh, E. Y., & Hong, H. (2022). The Impact of teaching simulation using student chatbots with different attitudes on preservice teachers' efficacy. *Educational Technology & Society, 25*(3), 46-59.

Sun Z., Anbarasan M., & Praveen Kumar, D. (2020). Design of online intelligent English teaching platform based on artificial intelligence techniques. *Computational Intelligence, 37*(3), 1166-1180. https://doi.org/10.1111/coin.12351

Tempelaar, D., Rienties, B., & Nguyen, Q. (2021). The Contribution of dispositional learning analytics to precision education. *Educational Technology & Society, 24*(1), 109-122.

Van Der Meij, H. (1994). Student questioning: A Componential analysis. *Learning and Individual Differences, 6*(2), 137-161. https://doi.org/10.1016/1041-6080(94)90007-8

Vanichvasin, P. (2021). Chatbot development as a digital learning tool to increase students' research knowledge. *International Education Studies, 14*(2), 44-53. https://doi.org/10.5539/ies.v14n2p44

Wang, Y., Liu, C., & Tu, Y. F. (2021). Factors affecting the adoption of AI-Based applications in higher education: An Analysis of teachers' perspectives using structural equation modeling. *Educational Technology & Society, 24*(3), 116-130.

Wagner, W. P. (2017). Trends in expert system development: A Longitudinal content analysis of over thirty years of expert system case studies. *Expert Systems with Applications, 76*, 85-96. https://doi.org/10.1016/j.eswa.2017.01.028

Weaver, M. R. (2006). Do students value feedback? Student perceptions of tutors' written responses. *Assessment & Evaluation in Higher Education, 31*(3), 379-394. https://doi.org/10.1080/02602930500353061

Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are we there yet?-A systematic literature review on chatbots in education. *Frontiers in Artificial Intelligence, 4*, 654924. https://doi.org/10.3389/frai.2021.654924

Wu, P. H., Hwang, G. J., & Tsai, W. H. (2013). An Expert system-based context-aware ubiquitous learning approach for conducting science learning activities. *Educational Technology & Society, 16*(4), 217-230.

Yang, S. J. (2021). Guest editorial: Precision education-a new challenge for AI in education. *Educational Technology & Society, 24*(1), 105-108.