A Project entitled

Instant Messenger Chatbot for Intelligent Tutoring and Quiz System

An Exploratory Study

Submitted by

Tse Nok Hang

submitted to The Education University of Hong Kong

for the degree of *Bachelor of Education (Honours) (Secondary) - Information and Communication Technology*

in *May 2018*

Declaration

I, *Tse Nok Hang* , declare that this research report represents my own work under the supervision of *Dr Lai Yiu Chi*, and that it has not been submitted previously for examination to any tertiary institution.

Signed ___

Tse Nok Hang

11 May 2018

Abstract

The emergence of mobile devices has facilitated the development of instant messaging, which is a common means to communicate with people in this word nowadays. People are familiar with these technologies as the they have already entered their lives. Therefore, people especially youngsters, should have enough knowledge in Information and Communication Technology in order to prepare themselves to survive in this technological world. A 3-interation design-based research was carried out. In which, an intelligent tutoring and quiz system was built on an instant messenger. It makes use of the technologies they are familiar with to enhance the learning effectiveness. Experiments was carried out to evaluate their performance before and after using the system. The usability of the system (usefulness and ease of use) was also examined using a set of survey. Detailed interviews were conducted to inspect the possible improvements to the system. Then, the design of the system was revised after each iteration according to the opinions provided by the participants. For the results, the tests results show that the system was effective as the subjects generally performed better in the post-test. The survey results also show positive trend for usefulness across 3 iterations. Both usefulness and ease of use of the system received satisfactory mean scores. Despite the interview results show that there were room for improvements to the system, some of which were improved and the subjects generally agreed that the system had more advantages and disadvantages. Therefore, the system is considered to be effective and for interactive tutoring and delivering ICT course.

Table of contents

List of Figures

List of Tables

Instant Messenger Chatbot for Intelligent Tutoring and Quiz System

An Exploratory Study

Introduction

The prevalence of mobile devices has changed people's daily lives. It made the popularisation of instant messengers possible, which has become a common method of communication these days. In Hong Kong, more than 99% of youngsters (aged from 15 to 24) own smartphones (Census and Statistics Department, 2017). This figure reflects that youngsters in Hong Kong rely heavily on technologies. It is important to educate this group of people about information and communication technology (ICT), so that they can prepare themselves in this technological world. The reasons above inspired me to develop a tutoring and quiz system on an instant messenger application. After the development, a research about its effectiveness in teaching ICT and its usability was conducted. The detailed procedures of the development and the research will be discussed in this article.

Research Aim & Questions

The research aim is to investigate the learning effectiveness of using instant messenger as a tutoring and quiz system, which gives rise to the following research questions.

1. Can chatbot be an effective tool for interactive tutoring?
2. Can chatbot be used to deliver ICT course?

Literature Review

In this part, the previous works done by different researchers will be reviewed so as to investigate the feasibility and effectiveness of using chatbot and intelligent tutoring system on instant messenger.

Instant Messaging in Education

Instant messaging in education usually involves real people. The most iconic example is the communication between teachers and students. There are several reasons for students to communicate with their instructors, including resolving urgent matters, dealing with administrative problems and asking for academic guidance (Lauricella & Kay, 2013). Another example in education is the communication between students themselves. Also, students agreed that instant messaging should be implemented in their learning environment (De Bakker, Sloep , & Jochems, 2007). The above statements show that the demand for instant messaging in education should not be ignored.

Intelligent Tutoring System

Intelligent tutoring system (ITS) is a broad term because any type of computer programs that includes intelligence can be regarded as ITS (Freedman, Ali, & Mcroy, 2000). ITS is a branch of computer-aided instruction (CAI), which means teaching with the help of computer. It can be as simple as an educational website that shows all the information without actively filtering

the information that the learners do not need. With the advancement of technology, ITS branched from CAI. It has a distinct element comparing to CAI, which is intelligence. For example, it allows customisation for different learners so as to offer a unique learning experience for every learner. Chatbot system can also be a type of ITS as it also involves intelligence.

Chatbot in Education

An example of chatbot in education is AutoTutor, which was developed by the University of Memphis in 1997 (Graesser, 2016). It was designed to teach computer literacy at first. From that time, many variants were developed to teach different fields and skills, such as physics, reading comprehension and biology. The user interfaces (UI) of these systems are similar. The UI generally has an animated avatar as the teacher with a question and a dialogue box that allows the learner to input the answer to the question raised by the 'teacher'. During the tutorial, the 'teacher' initialises the conversation by conveying the knowledge in a specific field and then elicits answers from the learners. The learner can also ask definitional questions (what is the definition of X) by inputting natural language (Jackson, 2008).

Effectiveness of Chatbot in Education

A Previous research shows that intelligent tutoring chatbot is feasible and effective (Abbasi & Kazi, 2014). The researchers developed a standalone chatbot system for teaching programming.

After the development, the effectiveness of using the chatbot system and conventional search engines in terms of memory retention and learning outcomes was tested. It was concluded that the subjects who used the chatbot system performed significantly better than those who used conventional search engines in both aspects (memory retention and learning outcomes).

The above literatures reveal 2 main problems of most chatbots in education. The first one is that the learners cannot start a conversation if they do not know the terminology they want to ask. For example, the learner knows that a type of malwares will steal personal data, but he/she does not know its name, which is Trojan Horse. He/she cannot look for this topic as he/she cannot input the keyword. It can be attributed to the chatbot for not providing any input for the learners. To be specific, the learners can only look for what they want under the premise that they know the exact terminologies. The second problem is unfamiliar user interface. Those chatbots are often built in standalone applications/websites, which have different user interfaces. Unlike instant messengers, learners must learn how to use the chatbot again if new chatbots are presented to them. Apart from the problems of most chatbots in education, the literatures indicate that using the chatbot function on an instant messenger for intelligent tutoring and quiz system might be feasible as the learners found it useful and have the need in communicating with the instructors for various matters, especially for seeking academic guidance. For this reason, the learners at least will not be averse to the idea of using instant messaging for learning. Also, the effectiveness of using chatbot in education was tested. It provides a base for the

intelligent tutoring and quiz system on instant messenger in this research as the system does not only include chatbot function but also contents and quiz. Thus, the combination of instant messaging, chatbot and the contents might be effective and useful for teaching and learning ICT. Therefore, the system was developed and the research was done to find out the answer.

Research Design

This is a design-based research, which involves 4 components, namely (1) analysis, (2) design, (3) development, and (4) implementation (Wang & Hannafin, 2005). The stages of this design-based research were simplified to only 2 stages for clearer explanation, the development stage and experiment stage. The development stage includes design and development, while the experiment stage includes implementation and analysis. This kind of research is expected to be run for several iterations for more optimised design. Therefore, 3 iterations were undertaken. Below are the descriptions of both stages.

Procedures of Development

This research started with the development of the system. It went through 5 steps. The first step was choosing teaching materials. The contents were chosen from the teaching materials of 2 introductory courses of ICT in the Education University of Hong Kong, Foundations of Information & Communication Technologies and Web Authoring & Internet Media. The selected materials cover Social Issues of ICT – Cybercrime & Information, and Multimedia –

Video & Animation. The second step was shortening the selected materials. They should be adjusted to an optimal length for the sake of carrying out the experiment in a short duration. The third step was developing the system architecture. It allowed the users to navigate the system in a logical and smooth way. The fourth step was inputting contents into the database. The data and the teaching materials were inputted into the database so that the system could run as it was intended. Below shows the architecture of the system.

Technological Aspect

The technological aspect, for example, what technologies and services were used will be explained briefly in the following. 6 technologies or services were used in this development mainly. They were (1) Telegram API, a chatbot function based on an instant messenger, (2) NodeJS, a server-side JavaScript programming language for coding the system, (3) MySQL, a relational Database Management System (DBMS) for storing data, such as the teaching materials, test results, customised contents, in a complicated data structure, (4) Dialogflow, a natural language conversation engine that allows the chatbot to hold human-like conversations with the users, (5) Amazon Web Services, offering a text-to-speech (TTS) engine for read aloud function, and (6) Google Cloud Platform, offering a voice recognition engine.

System Architecture

This system requires the users to complete a pre-test when they press ‘start’. Then, they need

to input the command. If the users input command by just pressing an on-screen button (select topic), they will be brought to the next level, which is selecting subtopic. The users will finally be brought to the content page as they continue. Otherwise, if the users input a command by typing a natural language command (e.g. what are the differences between X and Y), the system will analyse the intent of the command and then perform relevant operation. There are more functions included in this system. However, they are yet to be covered because it is not the main focus of this article.

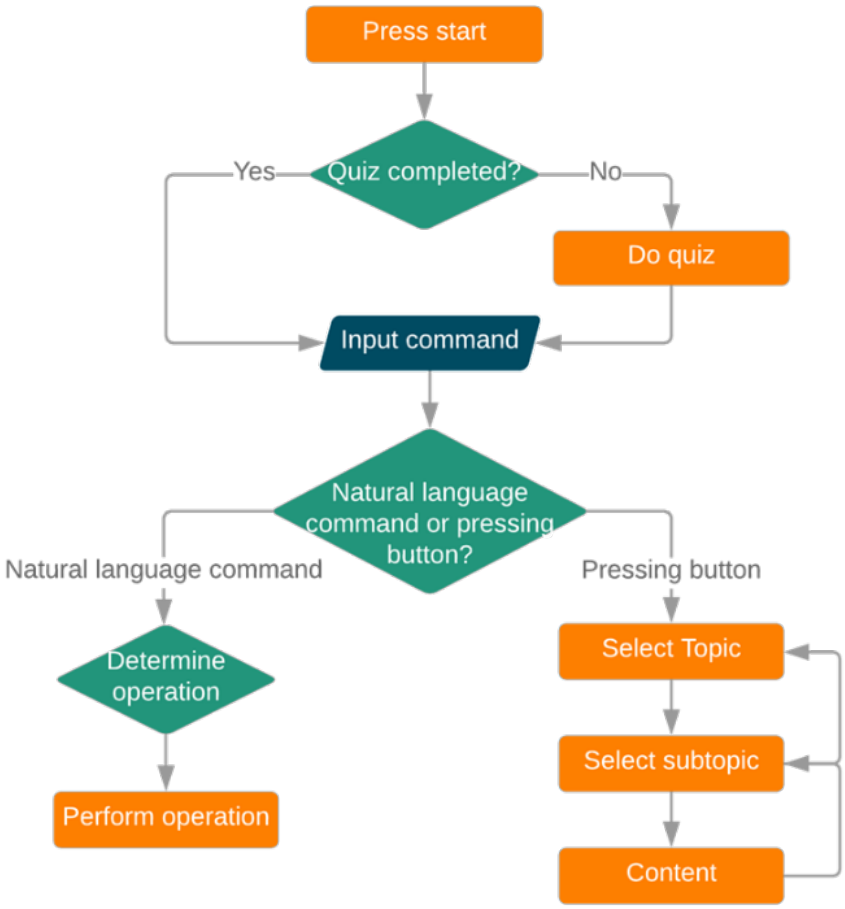


Figure 1. The basic system architecture that illustrates the flow of using the system

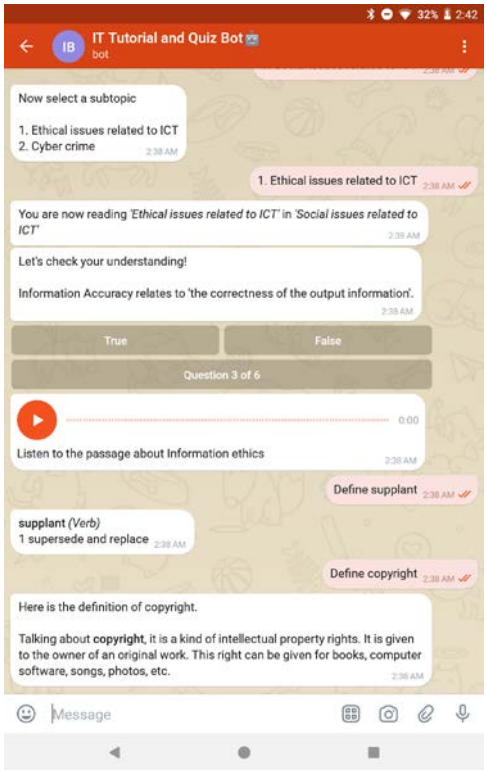


Figure 2. A screenshot of the system that demonstrates several features (read aloud function, natural language command, and checking questions)

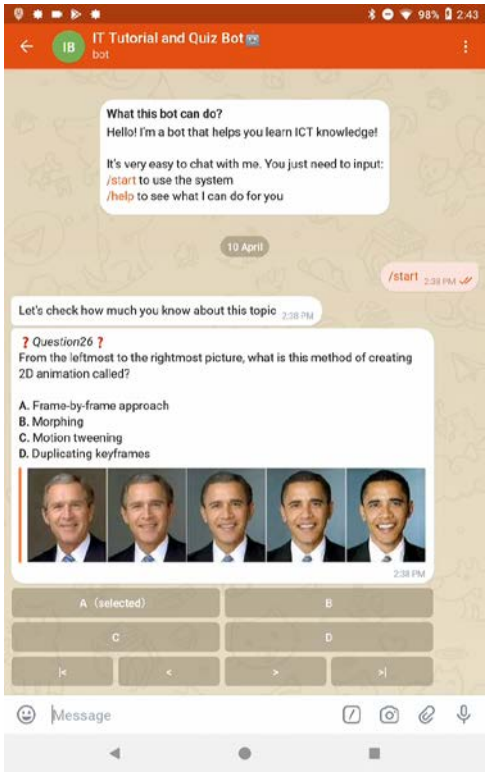


Figure 3. A screenshot of the system during pre-test/post-test

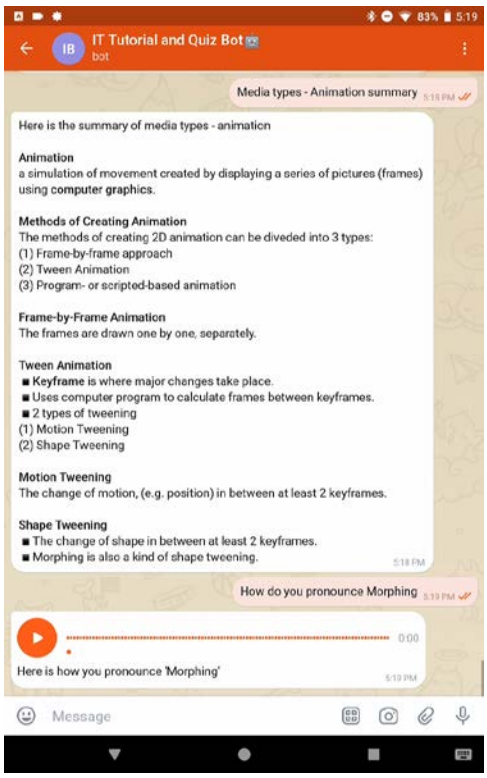


Figure 4. A screenshot of the system during self-study section

Methodology

This part aims to illustrate the methodology of conducting the experiment and receiving the feedback about the design. This research was meant to be run for 3 iterations including the part of experiment. It involved 10 participants. These participants were chosen randomly at EdUHK campus. The first and the second iteration had 3 participants for each, while the third iteration had 4 participants. All participants should be ongoing higher education students or graduates. The reason for targeting this group of subjects is that they generally rely on ICT more heavily. For example, they need to access online resources for academic purposes frequently. Apart from the participants, the duration of the experiment was at least 80 minutes. The whole experiment was done on a tablet computer installed with Telegram to access the system. Each subject went

through the following sections individually. Firstly, the subjects did a 15-minute pre-test. Then, they were required to complete a 30-minute self-study section. In which, they would not be restricted to read the contents in a specific order. Instead, they could jump to different topics and pages freely. After the self-study section, a 15-minute post-test was done. Finally, the subjects were asked to answer a survey and some questions in an interview section.

Experiment

Both pre-test and post-test consist of the same 28 questions. There are 2 topics and each topic has 2 subtopics, which make 4 subtopics in total as stated in the section of *Procedures of Development*. Each subtopic occupies 7 questions. As a result, 28 questions from the 4 subtopics were asked. These questions were designed according to a revised version of Bloom's Taxonomy. The cognitive process of it consists of 6 dimensions (Krathwohl, 2002). The higher the dimension, the higher the order of thinking. Back to the questions design, about a half of the questions are associated with the lowest dimension, remembering. That is, recalling the definition from long-term memory. The other half are associated with the higher dimensions like evaluating, which means judging the opinions based on criteria. Also, these 4 subtopics comprise the contents of the self-study materials. Each subtopic has similar length.

Survey

After the experiment, a survey was done by the subjects without the intervention by the

researcher. It covers mainly 2 parts. The first part is about the personal information of the subjects, while the second part is about the opinions on the system. The part of personal information asks for the gender, education details (education level & field of study and the experience in studying senior secondary ICT), usage of instant messengers, and self-perception of computer literacy and motivation in learning ICT. When it comes to the part about the opinions on the system, it mainly covers the system usefulness (6 statements) and the ease of use (7 statements). Besides, there are 3 statements which cannot be categorised into the 2 parts above. They are about the comparison between the system and the paper-based materials/self-learning website, and the suitability for primary/secondary school students to learn ICT. All statements except for those about personal information are rated on a scale from 1 to 4, where 1 represents ‘strongly disagree’ and 4 represents ‘strongly agree’.

Interview

This is the final section and is about the opinions on the system in different aspects. The interview is semi-structured and contains 6 open-end questions. It follows a framework yet allows follow-up questions. It starts with the overall perception on the system. The questions include (1) Which are the features that help you learn ICT? (2) What can be improved? (3) What features do you want to add? Then, it asks for the improvement on the key feature of this system and the navigation. Finally, it asks for the strengths and the weaknesses of this system comparing to the paper-based/self-learning website.

Approaches to Analyse the Data

This research adopts a hybrid approach. In which, both quantitative and qualitative approach will be used. The scores of the tests and the survey will be analysed using quantitative approach, while the results of the interview will be analysed qualitatively. The mean scores of the tests will first be calculated and then they will be compared to see if there is any difference and how much they are different from each other. On the other hand, the mean scores of the all 3 iterations will be compared against each other to see if the improvements between the iterations are consistent or not. Statistically, a paired T-test will be done to examine the significance between the overall mean scores of both tests. A nonparametric test will also be done due to the small sample size of this research. Similarly, it aims to find out if the mean scores are significantly different or not. The analysis of the survey results will be done through the comparison of the overall mean scores and the mean scores of iterations. These scores will be broken down into 2 parts, (1) usefulness of the system, and (2) ease of use. In which, the mean scores of some questions which are worth discussing will be discussed with the support of interview results. Finally, the interview results will be a part of qualitative analysis which focuses on the detailed explanations of the opinions on the system.

Results and Discussion

This part will start with the quantitative analysis first. It will be followed by the qualitative analysis.

Scores of Tests

The following is the bar chart that shows the mean scores of pre-test and post-test.

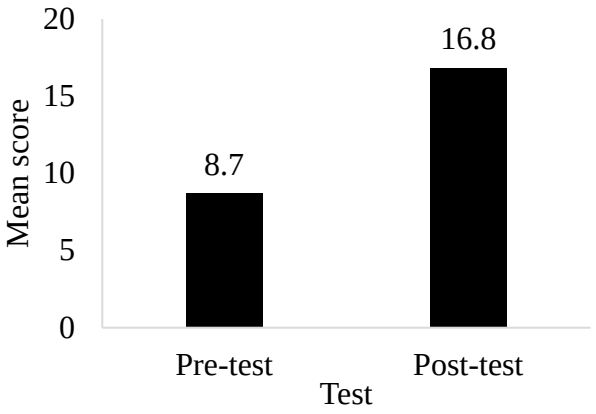


Figure 5. The mean scores of pre-test and post-test

The total marks of both pre-test and post-test are 28. The mean score of pre-test is 8.7, while that of post-test is 16.8. It shows that there is an overall improvement after reading the self-study materials. As for the difference, the increments of absolute value and percentage are 8.1 and 93.1% respectively. It is assumed that the subjects are in a normal distribution as they were chosen from a normally distributed population. Also, the mean scores show that the subjects performed better after reading the materials generally. It is expected that the system has a positive effect and the post-test scores are higher than the pre-test score. Therefore, the alternative hypothesis is that the post-test mean score is larger than the pre-test mean score (H_1: $\mu > 8.7$) and so the null hypothesis is that the post-test mean score is equal to the pre-test mean score (H_0: $\mu = 8.7$). A 1-tailed paired T-test was done using the scores of both tests for all subjects. Below is the result.

Table 1

Paired Samples Test

	Paired Differences						
				95% Confidence Interval of the Difference			
	Mean	Std. Deviation	Std. Error Mean	Upper	t	df	Sig. (1-tailed)
Post-test - Pre-test	8.10000	4.79467	1.51621	10.87938	5.342	9	0.00023

The result shows that a 1-tailed *p*-value of 0.00023 and the significance level is 0.05 in this case. The *p*-value is smaller than the significance level (i.e. $0.00023 \leq 0.05$). Thus, the null hypothesis can be rejected because the result is significantly different. The subjects performed significantly better in the post-test after reading the self-study materials. Apart from 1-tailed paired T-test, the significance of improvement was examined using 1-tailed Wilcoxon Signed-Rank Test, which is a nonparametric test, due to the small sample size. Below is the result.

Table 2

Wilcoxon Signed Ranks Test

	Post-test – Pre-test
Z	-2.708[a]
Asymp. Sig. (1-tailed)	0.00338

a. Based on negative ranks.

The result of this nonparametric test shows a *p*-value of 0.00338. The significance level of this test is 0.05. The *p*-value is less than the significance level (i.e. 0.00388 ≤ 0.05). Therefore, it can be said that the improvement is significant. To conclude, all statistical data yield the same outcome that the result is significant. In other words, the improvement between pre-test and post-test is significant overall.

When it comes to the tests results of the 3 iterations, they show a similar trend to the overall statistical data mentioned above.

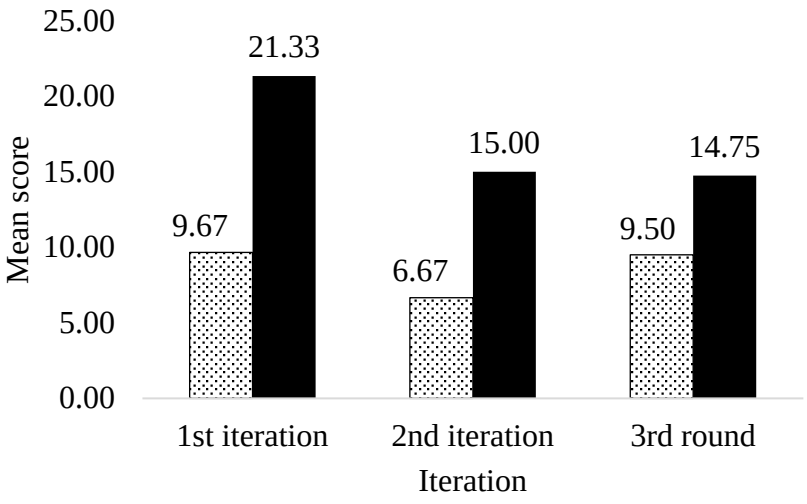


Figure 6. Pre-test and post-test mean scores of 3 iterations

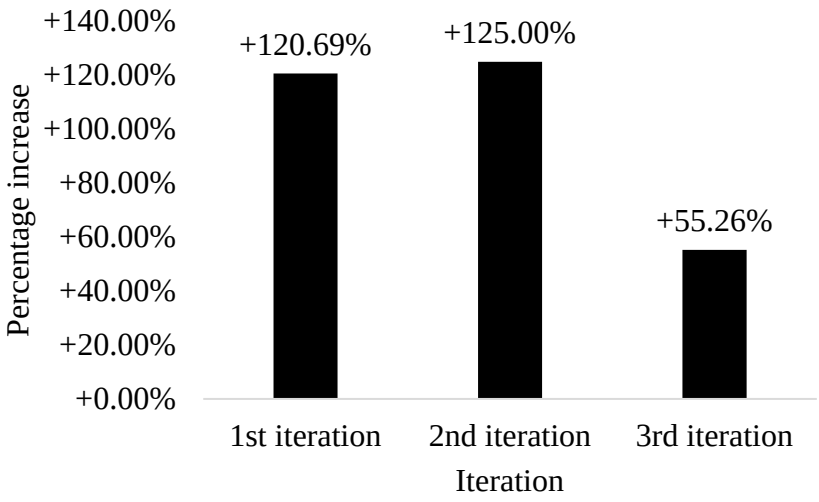


Figure 7. Percentage change in mean scores of 3 iterations

That is, the subjects of each iteration performed better in the post-test. In the first 2 iterations, the mean scores of pre-test increase 11.66 for 1st iteration and 8.33 for 2nd iteration. Both iterations show consistent increments in percentage, which are 120.69% and 125.00%. However, the 3rd iteration increases only 55.26% (5.25 in absolute value). This is because a subject performed worse after studying the materials in the 3rd iteration. His score dropped from 15 in pre-test to 14 in post-test. If this record is excluded, the scores of the post-test of all subjects are higher than those of the pre-test.

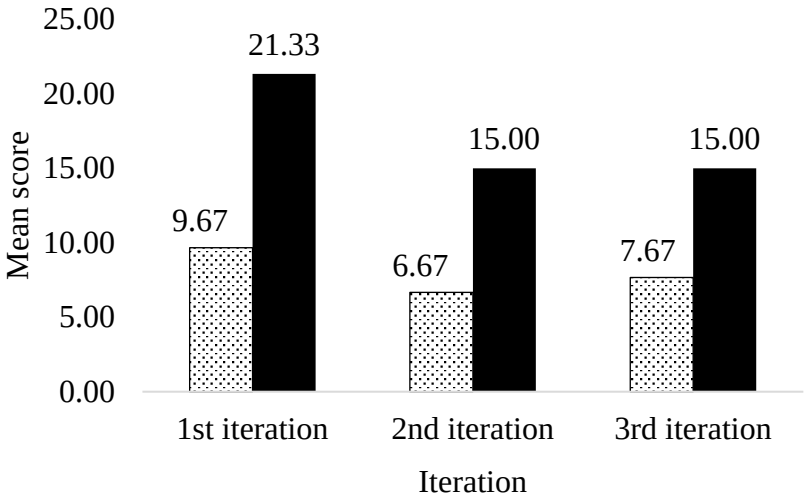


Figure 8. Pre-test and post-test mean scores of 3 iterations (Subject No.8 excluded)

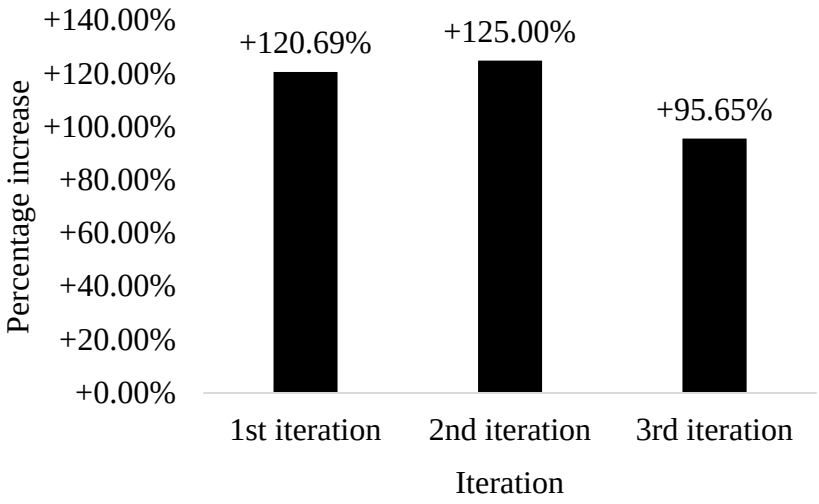


Figure 9. Percentage change in mean scores of 3 iterations (Subject No.8 excluded)

The increment in percentage becomes 95.65% (7.33 in absolute value). It shows a more consistent percentage increase.

The reason for the subject performing worse in the post-test is worth discussing. The design and the composition of questions are stated in the part of *Experiment* in **Methodology** that a half of the questions are of lowest dimension, while another half are of higher dimensions. For

this subject, all the correct answers in the pre-test that became wrong in the post-test are to the questions in the lowest dimension (e.g. What is the definition of X). On the contrary, some questions in the higher dimensions were answered wrongly in the pre-test, but they were corrected in the post-test. It is believed that the subject is weak in recognising and remembering as the corrected answers do show an improvement in higher dimensions. Moreover, the system does not provide a review for indicating wrong answers after doing the pre-test, it forces the subjects to study all materials including the materials for the questions they have answered correctly. Due to the characteristic of the subject, he wrongly remembered the definitions and thus finally gave the wrong answers to the questions he originally answered correctly. After all, the special case might be caused by personal characteristic. To conclude, the result answers the research questions that the chatbot can be an effective tool for interactive learning and the chatbot can be used to deliver ICT courses because the subjects scored higher after using the chatbot to learn ICT.

Scores of Survey

The part that asks for the opinions on the system is divided into 2 categories with 3 questions which cannot be categorised. The first category asks for the usefulness of the system. The overall mean score of all 6 statements for usefulness is 3.1 out of 4. In which, the statements that have the highest mean scores are statement 1 and 2. Both of them have 3.2 as they mean scores. They can be combined to become one statement, which is ‘the system helps me learn

ICT more effectively and efficiently'. The subjects found the system useful as they could actually learn ICT from the system (effective) in a fast and easy way (efficient). The following are the notable quotes from the subjects that help illustrate the point. It should be noted that these quotes are translated directly from Cantonese to English. "I could ask the questions and get the answers directly" said Subject Number 3. The quote shows that instead of searching through the contents, the subject could just ask for what she wanted and elicited the answers from the system successfully. Also, Subject Number 7 said "It is convenient. I can use the system as long as I have a mobile device". The system is based on Telegram. To use the system, the users only need mobile devices (e.g. smartphone) that can connect to the Internet and have the application installed. Once the said steps are done, they can use the system anywhere and anytime they want. It indicates that the system makes learning ICT more efficient. The lowest overall mean score in usefulness is statement 5, which has a mean score of 2.9. The statement is "The natural language command feature helps me get the answer". Some of the subjects did not agree to the statement. One of them, Subject Number 2 said "the system could not answer me the symptoms of infecting Trojan Horse". The natural language command feature did not give the subject a proper answer to what she asked because the knowledge base did not have such entry. It has a negative impact on the effectiveness. As a result, this statement received the lowest mean score in usefulness.

The next category is the ease of use. Similarly, the highest and the lowest overall mean score will be presented. The statements that received the highest mean scores of 3.3 are statement 12 and 13. They are combined to become one statement, that is “the dialogue style user interface and natural language command feature help me make sense of the ongoing event at the moment”. There are 2 quotes that support the point. Subject Number 9 said “The conversations between me and the system are human-like” and Subject Number 7 said “It is like asking teacher questions”. The natural conversation between humans can be illustrated using a communication model. It involves at least 2 interlocutors, the sender and the receiver (Mcquail & Windahl, 2015). They share a context in which communication occurs. To initiate a communication, the sender encodes a message and the receiver decodes it. Then, the receiver makes a feedback. Instant messaging makes use of this principle combining with the natural language conversation engine to achieve human-like communication. That is the reason for the highest scores for these 2 statements. Statement 8 has the lowest overall mean score in ease of use, which is 2.8. The statement is “the system is responsive”. Some of the subjects disagreed to this statement. For example, Subject Number 7 said “I needed to wait after I selected the answers in quiz”. The subject was not satisfied with the speed of the system as she needed to wait for the system to perform a specific task. It made the system less responsive, and therefore the mean score is the lowest in this category.

Besides, the mean scores of the 3 iterations were calculated for both category. Below show the trends for both categories.

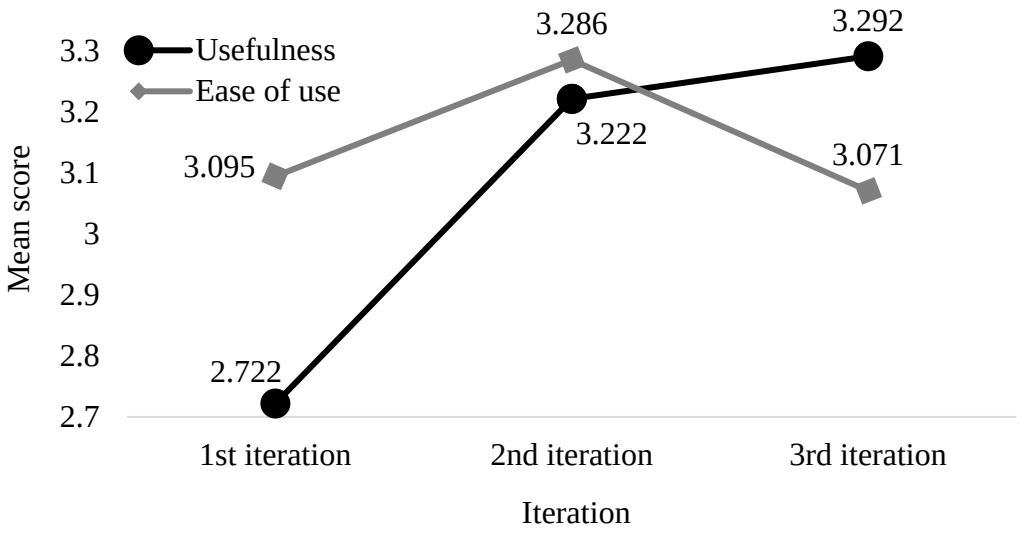


Figure 10. Trends of mean scores of usefulness and ease of use

The category of usefulness has an increasing trend. It increases from 2.722 in the 1st iteration to 3.222 in the 2nd iteration. Finally, it goes slightly higher to 3.292 in the 3rd iteration. It indicates that the system became more useful after each iteration. On the other hand, the mean scores of 3 different iterations for the ease of use increases then drops. It first rises from 3.095 in the 1st iteration to 3.286 in the 2nd iteration. However, it drops to 3.071 in the 3rd iteration. It means that the system was easy to use at the beginning, but some amendments were done in the development stage of the 3rd iteration which made the system difficult to use. The interview result helps explain this phenomenon. There were 4 subjects involved in the 3rd iteration. All of them expressed that the highlight function being added in the development stage of the same

iteration was difficult to use as the steps to highlight were cumbersome and complicated. “Highlight is a bit complicated. It would be more convenient if I could highlight the keywords directly” said Subject Number 10. Despite the fluctuation in this category, the overall mean scores (3.1 for usefulness and 3.143 for ease of use) and the general trends of the 3 iterations (both scored at least 3 in the final iteration) for both categories give the positive answers for the research questions.

Interview

This part will first explain the overall results of the 3 iterations combined. Then, it will focus on the separated results of each iteration, since some of the results depend heavily on the features updated and added in the development stage of each iteration. The results will be comprehensive and explainable only if they are explained iteration-by-iteration. Below are the overall results where the numbers in the parentheses represent the numbers of subjects who agreed to the opinions. For the overall results involving 10 subjects, regarding the features that help them learn ICT the most, most of them agreed that the natural language command (9) and the checking questions (7) were the most helpful features for them to learn ICT. The following question asks for the advantages and the disadvantages of the system comparing to paper-based materials. Most subjects thought that the system had the advantages of allowing interactions (6), more direct question & answer (6), innovative learning method (5), and all-in-one features (5). In contrast, most of them thought that the lack of highlight function (3) and bad for eyes

when using the system on a mobile device (2) were the disadvantages comparing to paper-based materials. Similarly, The following question ask for the advantages and disadvantages, but those of the system comparing to self-learning websites. The most outstanding advantages comparing to self-learning websites are the same as those in Question 6a, which are allowing interaction (7), more direct question & answer (7), and all-in-one features (4). For the disadvantages, most subjects thought that the content of the system was not enough comparing to self-learning websites (2). In short, the subjects generally agreed that the system was superior than both paper-based materials and self-learning websites. There are still more questions and they will be discussed at below paragraph, which cover the improvements, functions and improvements to navigation suggested by the subjects.

The following table shows the overall improvements to the system suggested by the participants. It also indicates whether the suggestions were accepted or denied and in which iteration they were adopted or the reasons they were denied. The denied suggestions or the suggestions that are especially meaningful will be discussed.

Table 3

Suggested improvements and the follow-up actions

Suggested improvements	Status	Remark / Justification
More intelligent	Accepted	Continuously adding data to knowledge base
Show related page for wrong answer for checking question	Accepted	Added in 3rd round development
Checking question after each page	Denied	Excessive drilling
More checking questions	Accepted	Added in 3rd round development
Dictionary function shows derivatives directly	Denied	Illogical
Integrate definition and pronunciation functions	Accepted	Will add in future development
Dictionary function shows examples and pictures	Denied	Messy and lengthy user interface
Pre-test shows wrongly answered question	Denied	Result in biased self-study experience

The participants in the 1st iteration thought the natural language command feature could be more intelligent as the system sometimes failed to answer them. This suggested improvement was accepted and more data is being added to the database continuously. Apart from that, the participants in the 2nd iteration wanted to have a checking question after each page for checking understanding. However, it was denied due to excessive drilling, which may also annoy the users. The participants in the 3rd iteration stated that the dictionary function should show derivatives of words directly instead of telling the users to look up the words manually.

However, the suggestion was seen as illogical because the users should get the definition of word X but not word Y if they are looking up word X. Moreover, the subjects wanted the dictionary to show other information such as examples and pictures. It was denied as excessive information makes the user interface messy and lengthy. Finally, the subjects suggested to show wrongly answered question after doing the pre-test. It was also denied because it results in biased self-study and therefore unfair post-test, which is undesirable for research.

Besides the improvements, the functions suggested by the subjects are as follow, which is a table having the similar structure as the table for suggested improvements.

Table 4

Suggested functions and the follow-up actions

Suggested functions	Status	Remark / Justification
Summary	Accepted	Added in 2nd round development
Checking questions	Accepted	Added in 2nd round development
Pronunciation	Accepted	Added in 2nd round development
Dictionary	Accepted	Added in 2nd round development
Highlight	Denied then accepted	Deemed unfeasible in 2nd round development, but in great demand
Passage read aloud	Denied then accepted	Deemed unnecessary in 2nd round development, but in great demand
Voice input	Accepted	Added in 3rd round development

Translation	Denied	English is supposed to be the medium
Note taking	Denied	Restriction of instant messenger

In the 1st iteration, the subjects suggested 6 functions to be added, which are summary and checking questions after each subtopic, pronunciation, dictionary, highlight, and passage read aloud function. The subjects wanted summary and checking questions to be added because they thought that summary could help them to review what they had learned and checking questions could check their understanding to see how much they had learned. These 2 functions were added in the 2nd iteration. According to the overall interview results, most of the subjects found these functions useful. For highlight function, some subjects said that they had a learning habit of highlighting the key points in the passage. However, this function was deemed unfeasible and denied in the 2nd iteration, because the simplest method for highlighting is long pressing on the keywords on mobile devices, which is impossible to be implemented as instant messengers generally do not provide such feature. The last function that was suggested to be added is passage read aloud function. This feature was denied because it was mentioned by only one subject and was not the main focus suggested by the subject. This suggestion was made when the subject was commenting the necessity of pronunciation, “It is not quite necessary to have pronunciation function. Unless it reads aloud the whole passage” said Subject Number 3. Therefore, it was deemed unnecessary and denied. Then, in the 2nd iteration, there were 3 suggested functions. They were highlight, passage read aloud, and voice input function. The

former 2 functions suggested in the previous iteration were still on the list in this iteration. Therefore, after consideration, they were accepted and added in the development stage of 3rd iteration due to the great demands. It is proven that learners have different learning styles, such as visual, auditory, and kinesthetic (Kanar, 2014). Passage read aloud function can fulfil the learners with auditory learning style. As a result, it was added in the final iteration. Also, highlight function was added to cater for learners' needs. However, the steps to highlight keywords are complicated because of the restriction of instant messenger. It resulted in the negative comment about this feature as seen in both the results of survey and interview. Finally, the translation and note taking function were required to be added in the 3rd iteration and both were denied. English is supposed to be the medium of instruction, it is unnecessary to be translated to another language. For note taking, it is restricted by the functionality of instant messenger. The steps to take note will be as complicated as highlighting. Therefore, it can be foreseen that the function will degrade the user experience if it is added.

The following table shows the improvements for system navigation and the follow-up actions taken.

Table 5

Suggested improvements for navigation and the follow-up actions

Suggested improvements	Status	Remark / Justification
Shorter loading time	Denied	Depends on internet connection speed
Larger button	Accepted	Fixed in 2nd round development
Show chapter page	Accepted	Added in 2nd round development
Jump to specific test question directly	Accepted	Will add in future development
Flag questions in test	Accepted	Will add in future development
More obvious notification	Denied	Restriction of mobile OS and instant messenger design
Remove redundant buttons	Accepted	Will fix in future development
More obvious topic dialog box	Accepted	Will fix in future development
Checking question jump to related page if wrongly answered	Denied	Users may lose focus

In the 1st iteration, some subjects suggested that the loading time could be shorter. However, it was denied because of technical reason. The loading time of the system depends on the internet connection speed and other uncontrollable factors (e.g. server performance of third party web services). In the 3rd iteration, the subjects suggested more obvious notification. They complained that the system did not respond to button pressing. This happened because they did not notice the notification at the top of the screen. However, it should be attributed to the

application design of Telegram. Therefore, the suggestion was denied. Also, the subjects suggested that the system should jump to the related page when the checking question was wrongly answered. However, this may cause the users to lose focus as the quiz is ongoing and the users still have more checking questions to complete.

To conclude, the interview can answer the research questions. The subjects generally agreed that the system had more advantages than disadvantages comparing to both self-learning websites and paper-based materials. Therefore, it is an effective tool for interactive learning and it can be used to deliver ICT course.

Future Application

This research focuses on exploring the feasibility to use this system. Therefore, the all-in-one installation package is not yet developed. One should have intermediate computer knowledge to deploy the system at this moment. Below are the software requirements.

Table 6

Software requirements

Software needed	Detailed description
Database	MySQL
Server OS	Preferred Microsoft Windows
Web server	NodeJS, Apache & PHP (XAMPP)
NodeJS packages	node-telegram-bot-api, uuid/v4, mysql, apiai, aws-sdk, fs, request, @google-cloud/speech

The integrated all-in-one executable file is yet to be developed. However, the components can be installed manually. Firstly, the users who want to host the system can install XAMPP. It includes MySQL database, Apache & PHP server. MySQL is setup during the installation. The database structure can be built using the backup file (hp.sql). Then, install NodeJS. The NodeJS packages can be installed using a batch (install_package.bat) file. Then, go to Telegram and register for an API key. Replace the API key and MySQL settings in the main file (app.js) with their own key and settings. Due to the restriction of trial version, the web services like Dialogflow, Oxford Dictionary API, Amazon Web Services, and Google Cloud only have limited bandwidth. One must register their own accounts for them to function normally. In short, this system can only be installed manually with relatively complicated steps at this moment. However, it will be easy even for beginners if an executable file is developed and the paid versions of web services are used. It will be the future development of this system.

Limitations

For development, the knowledge base of the system was not large enough. It can be larger, so that the system can answer more questions. For research, the sample size and variations were not large enough. The research only involved 10 participants, which lowered the accuracy of the statistical data. If the sample and variations were larger, more statistical analysis could be carried out (e.g. the correlation of motivation and test result) and the data would be more representative. Also, manpower was not sufficient. Developing a software and conducting the research by only one person led to smaller scale of both the development and the research. If more manpower could be involved, the scale would be larger and it would result in more in-depth research. Finally, time for the experiment was not enough. The experiment was designed to last a week. However, it was impossible to achieve due to the limited time. If more time was allowed, the experiment could simulate the use of the system instead of forcing the participants to use all functions in only half an hour.

Conclusion

Despite the limitations and the weaknesses of the system, the research gives positive answers to the research questions. Chatbot can be an effective tool for interactive tutoring. The subjects performed significantly better in the post-test after reading the self-study materials on the system. Also, the results of the usability survey generally show positive trends and the subjects were satisfied with the usefulness and the ease of use of the system. Besides, chatbot is suitable

to deliver ICT course. It can be seen from the interview results that the subjects were willing to use the system. In addition, the interview results show that there are more advantages than disadvantages comparing to both paper-based materials and self-learning websites. To conclude, the system is considered to be successful in being an effective tool for interactive tutoring that delivers ICT course.

(6227 words)

References

Abbasi, S., & Kazi, H. (2014). Measuring effectiveness of learning chatbot systems on student's learning outcome and memory retention. *Asian Journal of Applied Science and Engineering, 3*(2), pp. 251-260.

Census and Statistics Department. (2017). *Thematic Household Survey Report No. 62.* Hong Kong: Census and Statistics Department.

De Bakker, G., Sloep , P., & Jochems, W. (2007). Students and instant messaging: a survey of current use and demands for higher education. *ALT-J, 15*(2), pp. 143-153.

Freedman, R., Ali, S. S., & Mcroy, S. (2000). Links: what is an intelligent tutoring system? *Intelligence, 11*(3), pp. 15-16.

Graesser, A. C. (2016). Conversations with AutoTutor help students learn. *International Journal of Artificial Intelligence in Education, 26*(1), pp. 124-132.

Jackson, T. G. (2008). Learning through natural language dialogue that adapts to the cognitive

and affective states of the learner. In D. H. Robinson, & G. Schraw, *Recent Innovations in Educational Technology That Facilitate Student Learning* (pp. 95-125). Greenwich, United States: Information Age Publishing.

Kanar, C. C. (2014). *The Confident Student.* Boston, MA: Wadsworth.

Krathwohl, D. R. (2002). A revision of Bloom's Taxonomy: An overview. *Theory into practice, 41*(4), pp. 212-218.

Lauricella, S., & Kay, R. (2013). Exploring the use of text and instant messaging in higher education classrooms. *Research in Learning Technology, 21*(1), pp. 1-17.

Mcquail, D., & Windahl, S. (2015). *Communication Models for the Study of Mass Communications.* Abingdon, Oxon: Routledge.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development, 53*(4), pp. 5-23.

Appendix

Visit the link (https://goo.gl/W166mq) for the following documents

Appendix I. Information sheet & consent form

THE EDUCATION UNIVERSITY OF HONG KONG
The Department of Mathematics and Information Technology

CONSENT TO PARTICIPATE IN RESEARCH

Instant Messenger Chatbot for Intelligent Tutorial and Quiz System

I _____ hereby consent to participate in the captioned research supervised by Dr LAI Yiu Chi and conducted by TSE Nok Hang.

I understand that information obtained from this research may be used in future research and may be published. However, my right to privacy will be retained, i.e., my personal details will not be revealed.

The procedure as set out in the **attached** information sheet has been fully explained. I understand the benefits and risks involved. My participation in the project is voluntary.

I acknowledge that I have the right to question any part of the procedure and can withdraw at any time without negative consequences.

Name of participant _____
Signature of participant _____
Date _____

INFORMATION SHEET

Instant Messenger Chatbot for Intelligent Tutorial and Quiz System

You are invited to participate in a project supervised by Dr LAI Yiu Chi and conducted by TSE Nok Hang, who are staff / students of the Department of Mathematics and Information Technology in The Education University of Hong Kong.

The introduction of the research

This research aims to find out the possibility and usability of using instant messaging to enhance students' learning effectiveness and efficiency by implementing intelligent tutoring system on an instant messaging mobile application.

You were chosen for this research because you have/had been a higher education student using ICT in learning extensively. Your opinion on this research will be useful for improvement.

The methodology of the research

This study will include at least 10 participants.

You will be asked to finish a 60-minute section, which includes studying the course materials and completing the quizzes about the materials if prompted. In which, a 15-minute pretest consists of 28 questions will be conducted. You will then be required to study the course materials in 30 minutes. The materials are related to ICT, specifically (1) Multimedia elements – video and animation and, (2) Social issues – cyber crime and information ethics. After finishing the course materials, you will be asked to complete the posttest. The questions of the pretest and time required are the same as the pretest, which also consists of 28 questions and costs 15 minutes.

If you agree, you can choose to participate in an extra part, which will be using paper-based version of the learning materials after using the intelligent tutorial and quiz system (the part stated in the previous paragraph). The contents of the paper-based version are exactly the same as that of the system. It takes 30 minutes. This part is optional. However, if you can complete this part, you will be rewarded an extra 50HKD supermarket cash coupon as reimbursement.

Then, you will be asked to complete a survey about user experience (e.g. the usefulness and ease of use). The survey will last approximately 10 minutes.

Finally, an interview will be conducted, which is an extension of the survey. You will be asked some open-ended questions about how the system can be improved. which will last 10 minutes.

If you can complete the whole experiment (excluding the extra part), which in other words, you can finish the tasks stated above, you will be rewarded a 50HKD supermarket cash coupon as reimbursement.

The potential risks of the research

This research has no potential risk.

Your participation in the project is voluntary. You have every right to withdraw from the study at any time without negative consequences. All information related to you will remain confidential, and will be identifiable by codes known only to the researcher.

How results will be potentially disseminated

The results obtained in this research will be disseminated for academic purposes only. The results will be published in the form of thesis and oral presentation. All personal data will not be published by any means.

If you would like to obtain more information about this study, please contact TSE Nok Hang at telephone number or their supervisor Dr LAI Yiu Chi at telephone number

If you have any concerns about the conduct of this research study, please do not hesitate to contact the Human Research Ethics Committee by email at hrec@eduhk.hk or by mail to Research and Development Office, The Education University of Hong Kong.

Thank you for your interest in participating in this study.

TSE Nok Hang
Principal Investigator

香港教育大學

數學與資訊科技學系

參與研究同意書

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

本人_____同意參加由黎耀志博士負責監督,謝諾行執行的研究項目。

本人理解此研究所獲得的資料可用於未來的研究和學術發表。然而本人有權保護自己的隱私，本人的個人資料將不能洩漏。

研究者已將所附資料的有關步驟向本人作了充分的解釋。本人理解可能會出現的風險。本人是自願參與這項研究。

本人理解我有權在研究過程中提出問題,並在任何時候決定退出研究，更不會因此而對研究工作產生的影響負有任何責任。

參加者姓名: _____

參加者簽名: _____

日期: _____

有關資料

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

誠邀閣下參加黎耀志博士負責監督，謝諾行負責執行的研究計劃。她/他們是香港教育大學學生/教員。

研究計劃簡介

本研究旨在找出使用以即時通訊軟件作為平台的智能教學系統對學習效果及效率的可行性及可用性。

閣下現正／曾經為高等教育學生，並廣泛地使用資訊及通訊科技作為學習用途，有效為本研究提供有用的意見，所以閣下被揀選作為本次研究的對象。

研究方法

本研究涉及最少 10 名參與者。

是次研究將會要求閣下完成 60 分鐘的部份，包括自學教材以及系統所提供的測驗。研究中，閣下首先會被要求進行一個為時 15 分鐘的前測，前測包括 28 條題目。其後你需要完成 30 分鐘的教材。教材將會與資訊及通訊科技有關，具體而言包括(1) 多媒體元素－視像及動畫，及(2)社會議題－網絡罪案與資訊倫理。完成此部份後，閣下將被邀請完成後測，後測與前測的內容及耗時一樣，即同樣包含 28 條題目及需時 15 分鐘。

如果閣下同意，你可選擇參加額外部份。此部份要求閣下於使用智能教學及評估系統後（即上一段所載部份），完成紙本教材。紙本教材的內容與智能教學及評估系統的內容完全一致，需時 30 分鐘。閣下可自由選擇是否參與此部份，但是閣下如果能夠完成此部份，將可獲得一張額外港幣 50 元的超級市場現金券作報酬

其後，閣下將會被邀請填寫一個關於使用者體驗的問卷（如：有用性及易用性）。此部份將會花費大約 10 分鐘。

最後，閣下將會被邀請進行一個訪問，作為問卷的延伸。閣下將會回應一些關於如何改善智能教學及評估系統的開放式問題。本部份需時 10 分鐘。

如閣下能夠完成整個研究，即完成上述的任務，將可獲得一張港幣 50 元的超級市場現金券作報酬。

風險

本研究沒有任何潛在風險。

閣下的參與純屬自願性質。閣下享有充分的權利在任何時候決定退出這項研究,更不會因此引致任何不良後果。凡有關閣下的資料將會保密,一切資料的編碼只有研究人員得悉。

描述將如何發佈研究結果

本研究所獲得的結果只用作學術用途。研究結果將會以論文及口頭報告發佈。所有個人資料將不會以任何方式公開。

如閣下想獲得更多有關這項研究的資料,請與謝諾行聯絡,電話
或聯絡她/他們的導師黎耀志博士,電話

如閣下對這項研究的操守有任何意見,可隨時與香港教育大學人類實驗對象操守委員會聯絡(電郵:hrec@eduhk.hk; 地址:香港教育大院研究與發展事務處)。

謝謝閣下有興趣參與這項研究。

謝諾行
首席研究員

Appendix II. Pre-test and post-test questions

1. Which of the following item is **NOT** included in PAPA Model?
   A. Privacy
   B. Availability
   C. Property
   D. Accessibility

2. Which of the following action violates the 6 Data Protection Principles?
   A. Allow the data owner to review the personal data he/she has inputted for registering a lucky draw.
   B. Notify the data owners why their data are collected.
   C. Safely Keep the personal data of a former customer as long as possible.
   D. Collect only the names and contact numbers of customers for product warranty.

3. What is information accuracy concerned with?
   A. The correctness of the output information.
   B. Whether the information is recent or not.
   C. Whether the information is mentioned by large amount of people or not.
   D. Whether the information is readily available to people.

4. What is the relationship between copyright and intellectual property right?
   A. Copyright is a kind of intellectual property rights.
   B. Copyright is another name of intellectual property right.
   C. Intellectual property right is a kind of copyrights.
   D. Copyright is protected by laws while intellectual property right is not.

5. Which one is **NOT** protected by copyright law?
   A. Student writing submitted to the teacher.
   B. Emojis on a smartphone.
   C. Instagram photo posted by a user.
   D. The symbol of the Apple of Apple Inc., a technology company.

6. What is digital divide?
   A. A phenomenon caused by insufficient production of digital devices.
   B. The economic and social gap between different groups of people in accessing, using ICT.
   C. The unacceptance of technology of different groups of people.
   D. The difference between proficiency of using ICT of different groups of people.

7. Which one is an example of digital divide?
   (1) A man who lives in distant area cannot connect his computer to the Internet.
   (2) An elder person who does not know how to use a smartphone cannot access online news.
   (3) An elder person refuses to read books on a digital device.

   A. (1) and (2)
   B. (1) and (3)
   C. (2) and (3)
   D. All of the above

8. What is the difference between hackers and crackers?
   A. Hackers is more skilful than crackers in terms of hacking.
   B. Hackers has great passions in computer programming, system and network, while crackers do not.
   C. Hackers write the script for hacking by himself while crackers use the script written by hackers.
   D. Hackers aim to find out loopholes of the system and fix or report to the systems' owners, while crackers aim to attack the systems with malicious intentions.

9. Which of the following is the examples of hacking?
   (1) An online game player uses bugs to win a game.
   (2) A criminal breaks into an e-commerce database to retrieve credit card information.
   (3) A student breaks into the school's computer system, but reports it to the teacher.

   A. (1) and (2)
   B. (1) and (3)
   C. (2) and (3)
   D. All of the above

10. Why do criminals target the Internet for fraud?
   A. The people on the Internet are usually richer as they can afford electronic devices and the Internet.
   B. There are more businesses on the Internet. The people who run businesses are usually richer.
   C. Internet is a global environment. The fraudulent messages have higher chance in getting contact with all people.
   D. People on the Internet have lower alertness to fraud.

11. Which one is the most possible example of phishing?
    A. A cinema website requests the users to input contact number when buying movie ticket.
    B. A website found on Google that asks for your email for a lucky draw.
    C. A software that encrypted your files and asks for money to unlock.
    D. A software that steals your data without being noticed.

12. What is the difference between scam and hoax emails?
    A. Scam emails are usually sent by the people the recipients know, while hoax emails are sent by unknown people.
    B. Scam emails usually ask for money, while hoax emails usually ask for sensitive information.
    C. Scam emails usually include incredible offer, while hoax emails pretend to be trusted entity.
    D. Scam emails usually target companies, while hoax emails usually target individual users.

13. What is the difference between virus and worm?
    A. Virus can start itself, while worm cannot.
    B. Virus usually just consumes system resources, while worm usually damages the system.
    C. Virus can transmit through the Internet, while worm generally does not replicate itself.
    D. Virus is usually attached to a normal program, while worm is usually a standalone program.

14. Which of the following is most likely to be the symptom of Trojan Horse infected?
    A. No symptoms.
    B. Computer is becoming slower.
    C. Some files disappear suddenly.
    D. Computer crashes suddenly.

15. Which description about digital video is correct?
    A. Videos are usually generated using computer.
    B. Videos must contain sound.
    C. Videos are usually real-world motion.
    D. A video is a sequence of drawing.

16. What is the most common unit of frame rate.
    A. RPM
    B. BPS
    C. FPS
    D. Hz

17. Scan type does **NOT** determine how the frames of a video:
    A. Stored
    B. Produced
    C. Transmitted
    D. Displayed

18. What are the characteristics of interlaced scan?
    (1) The moving object appears to be combing.
    (2) A frame is displayed at a time.
    (3) Odd lines will be displayed after even lines.

    A. (1) and (2)
    B. (1) and (3)
    C. (2) and (3)
    D. All of the above

19. Below is a zoomed frame captured from a video. Which scan type is the video?

    A. Progressive scan
    B. Ingressive scan
    C. Interlaced scan
    D. Full-frame scan

20. How many frames at least for the human visual system to consider it as motion?
    A. 10 to 12
    B. 22 to 24
    C. 28 to 30
    D. 50 to 60

21. How many frames in a second for a video of 50i?
    A. 50
    B. 49.97
    C. 25
    D. 30

22. Cathy and Ivan are discussing about keyframes. Cathy says, “Each keyframe illustrates a key event in the timeline of the animation.” Ivan says, “The minimum number of keyframes required in a tweening process is three.”
    A. Both Cathy and Ivan are correct.
    B. Both Cathy and Ivan are wrong.
    C. Cathy is correct while Ivan is wrong.
    D. Cathy is wrong while Ivan is correct.

23. What are the differences between digital video and digital animation?
    A. Video has sound, while animation does not
    B. Video is measured in frames per second, while animation is measured in Pictures per second
    C. Video is a sequence of frame, while animation is a sequence of pictures.
    D. Video is produced using video camera, while animation is produced using computer graphics.

24. Peter is watching Anime (日本動畫) after watching a basketball match broadcasted on TV. He noticed that the basketball match is smoother than the Anime. What is the most possible reason?
    A. The animation is interlaced scan, while the basketball match is progressive scan.
    B. The animation has lower FPS than that of the basketball match.
    C. He did not turn on the deinterlacing feature.
    D. The animation has less moving objects than that in a basketball match.

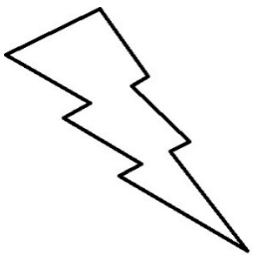
25. Which method of creating 2D animation did Miyazaki Hayao (宮崎 駿) use in his films?
    A. Frame-by-frame approach
    B. Morphing
    C. Tweening
    D. Duplicating keyframes

26. From the leftmost to the rightmost picture, what is this method of creating 2D animation called?

    A. Frame-by-frame approach
    B. Morphing
    C. Tweening
    D. Duplicating keyframes

27. What is the use of keyframe?
    A. Mark the point of interest on the timeline for later investigation.
    B. The frame that generated by the computer program.
    C. Indicate where a major change of animation takes place.
    D. Acts as a unique identification of the animation.

28. How many keyframes at least should there be in order to draw the following pattern?



    A. 12
    B. 8
    C. 6
    D. 3

Appendix III. Blank survey

可用性問卷

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 個人資料

性別

□男 □女

年齡

□18-23 □24-29 □30 歲以上

是否在學大專生

□在學 □已畢業

課程種類

□電腦科學及資訊科技 □與醫學及衛生有關的學科 □生物科學 □物理科學 □數學 □工程及科技 □建築學及城市規劃 □工商管理 □社會科學 □法律 □大眾傳播 □語言及相關科目 □人文學科 □藝術、設計及演藝 □教育（請選擇主修科目）

□其他：_____

是否於高中時選修資訊及通訊科技

□是 □否

是否曾經使用即時通訊軟件（如：WhatsApp）

□是 □否（請跳至第 2 部份）

使用過的即時通訊軟件

□微訊 □Telegram □QQ □Skype □Facebook Messenger □Line □其他：_____

有否使用過即時通訊軟件中的機械人功能（bot）

□有 □沒有

你認為自己的電腦知識水平為（1 為最低，4 為最高）:

□1 □2 □3 □4

你認為自己學習電腦知識的動機為（1 為最低，4 為最高）:

□1 □2 □3 □4

2. 系統有用性

	非常不認同	不認同	認同	非常認同
使用本系統使我更有效學習電腦知識	☐	☐	☐	☐
使用本系統使我學習電腦知識更快捷	☐	☐	☐	☐
使用本系統令學習電腦知識變得更加容易	☐	☐	☐	☐
本系統中的自然語言指令功能能夠解答我的問題	☐	☐	☐	☐
本系統中的自然語言指令功能令我更容易得到想要的答案	☐	☐	☐	☐
本系統中的自然語言指令功能有助我理解課程內容	☐	☐	☐	☐

3. 系統易用性

	非常不認同	不認同	認同	非常認同
使用本系統對我來說是容易的	☐	☐	☐	☐
本系統的反應是迅速的	☐	☐	☐	☐
我與本系統的互動是流暢的	☐	☐	☐	☐
本系統的操作流程是合理的	☐	☐	☐	☐
本系統的介面是簡單直覺的	☐	☐	☐	☐
本系統採用的對話形式使我更易理解當前發生的事件	☐	☐	☐	☐
本系統中的自然語言指令功能使我更易理解當前發生的事件	☐	☐	☐	☐

4. 其他

	非常不認同	不認同	認同	非常認同
就學習電腦知識而言，本系統比				
紙本自學教材優勝	☐	☐	☐	☐
就學習電腦知識而言，本系統比自學網站優勝	☐	☐	☐	☐
本系統適合中小學生學習電腦知識	☐	☐	☐	☐

Appendix IV. Interview questions

訪問稿

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 你認為本系統哪一功能對於學習電腦知識最有用？

2. 你認為本系統哪一功能最需要改善？

3. 你希望本系統新增哪一功能？

4. 對於自然語言辨識系統，你最想其新增哪一種題型（例如：如何、為何）？

5. 你認為本系統的導航可以有甚麼改善？

6. 作為自學平台，你認為使用聊天機械人比使用自學網站及紙本材料有何好處？

Appendix V. Raw survey results

Subject No.: 1

可用性問卷

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 個人資料

性別
☐男 ☑女

年齡
☑18-23 ☐24-29 ☐30 歲以上

是否在學大專生
☑在學 ☐已畢業

課程種類
☐電腦科學及資訊科技 ☐與醫學及衛生有關的學科 ☐生物科學 ☐物理科學 ☑數學 ☐工程及科技 ☐建築學及城市規劃 ☐工商管理 ☐社會科學 ☐法律 ☐大眾傳播 ☐語言及相關科目 ☐人文學科 ☐藝術、設計及演藝 ☐教育（請選擇主修科目）
☐其他：_____

是否於高中時選修資訊及通訊科技
☐是 ☑否

是否曾經使用即時通訊軟件（如：WhatsApp）
☑是 ☐否（請跳至第 2 部份）

使用過的即時通訊軟件
☑微訊 ☑Telegram ☐QQ ☑Skype ☑Facebook Messenger ☑Line ☐其他：_____

有否使用過即時通訊軟件中的機械人功能（bot）
☐有 ☑沒有

2. 系統有用性

	非常不認同	不認同	認同	非常認同
使用本系統使我更有效學習電腦知識	☐	☐	☑	☐
使用本系統使我學習電腦知識更快捷	☐	☐	☑	☐
使用本系統令學習電腦知識變得更加容易	☐	☑	☐	☐
本系統中的自然語言指令功能能夠解答我的問題	☐	☑	☐	☐
本系統中的自然語言指令功能令我更容易得到想要的答案	☐	☑	☐	☐
本系統中的自然語言指令功能有助我理解課程內容	☐	☐	☑	☐

3. 系統易用性

	非常不認同	不認同	認同	非常認同
使用本系統對我來說是容易的	☐	☐	☑	☐
本系統的反應是迅速的	☐	☑	☐	☐
我與本系統的互動是流暢的	☐	☐	☑	☐
本系統的操作流程是合理的	☐	☐	☑	☐
本系統的介面先簡單直覺的	☐	☐	☑	☐
本系統採用的對話形式使我更易理解當前發生的事件	☐	☐	☑	☑
本系統中的自然語言指令功能使我更易理解當前發生的事件	☐	☐	☑	☐

4. 其他

	非常不認同	不認同	認同	非常認同
本系統適合中小學生學習電腦知識	☐	☑	☐	☐

Subject No.: 2 Date: 25 Nov 2018 Time: 15:48

可用性問卷

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 個人資料

性別
□男 ☑女

年齡
☑18-23 □24-29 □30 歲以上

是否在學大專生
☑在學 □已畢業

課程種類
□電腦科學及資訊科技 □與醫學及衛生有關的學科 □生物科學 □物理科學 □數學 □工程及科技 □建築學及城市規劃 □工商管理 □社會科學 □法律 ☑大眾傳播 □語言及相關科目 □人文學科 □藝術、設計及演藝 □教育（請選擇主修科目）
□其他：_____

是否於高中時選修資訊及通訊科技
□是 ☑否

是否曾經使用即時通訊軟件（如：WhatsApp）
☑是 □否（請跳至第 2 部份）

使用過的即時通訊軟件
☑微訊 ☑Telegram □QQ ☑Skype ☑Facebook Messenger ☑Line □其他：Whatsapp

有否使用過即時通訊軟件中的機械人功能（bot）
□有 ☑沒有

你認為自己的電腦知識水平為（1 為最低，4 為最高）：
□1 ☑2 □3 □4

你認為自己學習電腦知識的動機為（1 為最低，4 為最高）：
□1 ☑2 □3 □4

2. 系統有用性

	非常不認同	不認同	認同	非常認同
使用本系統使我更有效學習電腦知識	□	□	☑	□
使用本系統使我學習電腦知識更快捷	□	□	☑	□
使用本系統令學習電腦知識變得更加容易 無法回答問題	□	☑	□	□
本系統中的自然語言指令功能能夠解答我的問題	□	☑	□	□
本系統中的自然語言指令功能令我更容易得到想要的答案	□	☑	□	□
本系統中的自然語言指令功能有助我理解課程內容	□	☑	□	□

3. 系統易用性

	非常不認同	不認同	認同	非常認同
使用本系統對我來說是容易的	□	□	☑	□
本系統的反應是迅速的	□	□	☑	□
我與本系統的互動是流暢的	□	□	☑	□
本系統的操作流程是合理的	□	□	☑	□
本系統的介面是簡單直覺的	□	□	☑	□
本系統採用的對話形式使我更易理解當前發生的事件	□	□	☑	□
本系統中的自然語言指令功能使我更易理解當前發生的事件	□	□	☑	□

4. 其他

	非常不認同	不認同	認同	非常認同
就學習電腦知識而言，本系統比紙本自學教材優勝	□	□	☑	□
就學習電腦知識而言，本系統比自學網站優勝	□	□	☑	□
本系統適合中小學生學習電腦知識	□	□	☑	□

Subject No.: 3 Date: 25 May 2018 Time: 17:30

可用性問卷

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 個人資料

性別
☐男 ☑女

年齡
☑18-23 ☐24-29 ☐30 歲以上

是否在學大專生
☐在學 ☑已畢業

課程種類
☐電腦科學及資訊科技 ☐與醫學及衛生有關的學科 ☐生物科學 ☐物理科學 ☐數學 ☐工程及科技 ☐建築學及城市規劃 ☑工商管理 ☐社會科學 ☐法律 ☐大眾傳播 ☐語言及相關科目 ☐人文學科 ☐藝術、設計及演藝 ☐教育（請選擇主修科目）
☐其他：_____

是否於高中時選修資訊及通訊科技
☐是 ☑否

是否曾經使用即時通訊軟件（如：WhatsApp）
☑是 ☐否（請跳至第 2 部份）

使用過的即時通訊軟件
☐微訊 ☐Telegram ☐QQ ☑Skype ☑Facebook Messenger ☑Line ☑其他：Whats app

有否使用過即時通訊軟件中的機械人功能（bot）
☐有 ☑沒有

你認為自己的電腦知識水平為（1 為最低，4 為最高）：
☐1 ☐2 ☑3 ☐4

你認為自己學習電腦知識的動機為（1 為最低，4 為最高）：
☐1 ☐2 ☐3 ☑4

2. 系統有用性

	非常不認同	不認同	認同	非常認同
使用本系統使我更有效學習電腦知識	☐	☐	☑	☐
使用本系統使我學習電腦知識更快捷	☐	☐	☑	☐
使用本系統令學習電腦知識變得更加容易	☐	☐	☐	☑
本系統中的自然語言指令功能能夠解答我的問題	☐	☐	☐	☑
本系統中的自然語言指令功能令我更容易得到想要的答案	☐	☐	☑	☐
本系統中的自然語言指令功能有助我理解課程內容	☐	☐	☑	☐

3. 系統易用性

	非常不認同	不認同	認同	非常認同
使用本系統對我來說是容易的	☐	☐	☐	☑
本系統的反應是迅速的	☐	☑	☐	☐
我與本系統的互動是流暢的	☐	☐	☑	☐
本系統的操作流程是合理的	☐	☐	☑	☐
本系統的介面是簡單直覺的	☐	☐	☐	☑
本系統採用的對話形式使我更易理解當前發生的事件	☐	☐	☑	☐
本系統中的自然語言指令功能使我更易理解當前發生的事件	☐	☐	☐	☑

4. 其他

	非常不認同	不認同	認同	非常認同
就學習電腦知識而言，本系統比紙本自學教材優勝	☐	☐	☑	☐
就學習電腦知識而言，本系統比自學網站優勝	☐	☐	☑	☐
本系統適合中小學生學習電腦知識 簡單易用，易上手	☐	☐	☐	☑

Subject No.: 4 Date: 21/1/2018 Time: 12:30

可用性問卷

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 個人資料

性別
□男 ☑女

年齡
☑18-23 □24-29 □30 歲以上

是否在學大專生
☑在學 □已畢業

課程種類
□電腦科學及資訊科技 □與醫學及衛生有關的學科 □生物科學 □物理科學 □數學 □工程及科技 □建築學及城市規劃 □工商管理 ☑社會科學 □法律 □大眾傳播 □語言及相關科目 □人文學科 □藝術、設計及演藝 □教育（請選擇主修科目）
□其他：_____

是否於高中時選修資訊及通訊科技
□是 ☑否

是否曾經使用即時通訊軟件（如：WhatsApp）
☑是 □否（請跳至第 2 部份）

使用過的即時通訊軟件
☑微訊 □Telegram □QQ ☑Skype ☑Facebook Messenger ☑Line ☑其他：Whatsapp ~~Beetalk~~

有否使用過即時通訊軟件中的機械人功能（bot）
☑有 □沒有 Google Form

你認為自己的電腦知識水平為（1 為最低，4 為最高）：
□1 ☑2 □3 □4

你認為自己學習電腦知識的動機為（1 為最低，4 為最高）：
□1 □2 ☑3 □4

2. 系統有用性

	非常不認同	不認同	認同	非常認同
使用本系統使我更有效學習電腦知識	□	□	☑	□
使用本系統使我學習電腦知識更快捷	□	□	☑	□
使用本系統令學習電腦知識變得更加容易	□	□	☑	□
本系統中的自然語言指令功能能夠解答我的問題	□	□	☑	□
本系統中的自然語言指令功能令我更容易得到想要的答案	□	□	☑	□
本系統中的自然語言指令功能有助我理解課程內容	□	□	☑	□

3. 系統易用性

	非常不認同	不認同	認同	非常認同
使用本系統對我來說是容易的	□	□	☑	□
本系統的反應是迅速的	□	□	☑	□
我與本系統的互動是流暢的	□	□	☑	□
本系統的操作流程是合理的	□	□	☑	□
本系統的介面是簡單直覺的	□	□	☑	□
本系統採用的對話形式使我更易理解當前發生的事件	□	□	☑	□
本系統中的自然語言指令功能使我更易理解當前發生的事件	□	□	☑	□

4. 其他

	非常不認同	不認同	認同	非常認同
就學習電腦知識而言，本系統比紙本自學教材優勝	□	□	☑	□
就學習電腦知識而言，本系統比自學網站優勝	□	□	☑	□
本系統適合中小學生學習電腦知識	□	□	□	☑

Subject No.: 5 Date: 2/4/2018 Time: 14:38

可用性問卷

使用即時通訊軟件的鄰天機械人功能作智能教學及評估系統

1. 個人資料

性別
□男 ☑女

年齡
☑18-23 □24-29 □30 歲以上

是否在學大專生
☑在學 □已畢業

課程種類
□電腦科學及資訊科技 □與醫學及衛生有關的學科 □生物科學 □物理科學 □數學 □工程及科技 □建築學及城市規劃 □工商管理 □社會科學 □法律 □大眾傳播 □語言及相關科目 □人文學科 □藝術、設計及演藝 ☑教育（請選擇主修科目）
□其他：_____

是否於高中時選修資訊及通訊科技
□是 ☑否

是否曾經使用即時通訊軟件（如：WhatsApp）
☑是 □否（請跳至第 2 部份）

使用過的即時通訊軟件
☑微訊 □Telegram ☑QQ □Skype □Facebook Messenger □Line □其他：_____

有否使用過即時通訊軟件中的機械人功能（bot）
□有 ☑沒有

你認為自己的電腦知識水平為（1 為最低，4 為最高）：
□1 ☑2 □3 □4

你認為自己學習電腦知識的動機為（1 為最低，4 為最高）：
□1 □2 □3 ☑4

2. 系統有用性

	非常不認同	不認同	認同	非常認同
使用本系統使我更有效學習電腦知識	□	□	☑	□
使用本系統使我學習電腦知識更快捷	□	□	□	☑
使用本系統令學習電腦知識變得更加容易	□	□	□	☑
本系統中的自然語言指令功能能夠解答我的問題	□	□	☑	□
本系統中的自然語言指令功能令我更容易得到想要的答案	□	□	☑	□
本系統中的自然語言指令功能有助我理解課程內容	□	□	☑	□

3. 系統易用性

	非常不認同	不認同	認同	非常認同
使用本系統對我來說是容易的	□	□	□	☑
本系統的反應是迅速的	□	□	☑	□
我與本系統的互動是流暢的	□	□	□	☑
本系統的操作流程是合理的	□	□	□	☑
本系統的介面是簡單直覺的	□	□	□	☑
本系統採用的對話形式使我更易理解當前發生的事件	□	□	□	☑
本系統中的自然語言指令功能使我更易理解當前發生的事件	□	□	□	☑

4. 其他

	非常不認同	不認同	認同	非常認同
就學習電腦知識而言，本系統比紙本自學教材優勝	□	□	□	☑
就學習電腦知識而言，本系統比自學網站優勝	□	□	□	☑
本系統適合中小學生學習電腦知識	□	□	□	☑

Subject No.: 6 Date: 4 April 2018 Time: 18:27

可用性問卷

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 個人資料

性別
□男 ☑女

年齡
☑18-23 □24-29 □30 歲以上

是否在學大專生
☑在學 □已畢業

課程種類
□電腦科學及資訊科技 □與醫學及衛生有關的學科 ☑生物科學 □物理科學 □數學 □工程及科技 □建築學及城市規劃 □工商管理 □社會科學 □法律 □大眾傳播 □語言及相關科目 □人文學科 □藝術、設計及演藝 □教育（請選擇主修科目）
□其他：_____

是否於高中時選修資訊及通訊科技
□是 ☑否

是否曾經使用即時通訊軟件（如：WhatsApp）
☑是 □否（請跳至第 2 部份）

使用過的即時通訊軟件
□微訊 ☑Telegram □QQ ☑Skype ☑Facebook Messenger ☑Line ☑其他：Whatsapp

有否使用過即時通訊軟件中的機械人功能（bot）
☑有 □沒有

你認為自己的電腦知識水平為（1 為最低，4 為最高）：
□1 □2 ☑3 □4

你認為自己學習電腦知識的動機為（1 為最低，4 為最高）：
□1 ☑2 □3 □4

2. 系統有用性

	非常不認同	不認同	認同	非常認同
使用本系統使我更有效學習電腦知識	□	□	□	☑
使用本系統使我學習電腦知識更快捷	□	□	☑	□
使用本系統令學習電腦知識變得更加容易	□	□	☑	□
本系統中的自然語言指令功能能夠解答我的問題	□	□	□	☑
本系統中的自然語言指令功能令我更容易得到想要的答案	□	□	☑	□
本系統中的自然語言指令功能有助我理解課程內容	□	□	☑	□

3. 系統易用性

	非常不認同	不認同	認同	非常認同
使用本系統對我來說是容易的	□	□	☑	□
本系統的反應是迅速的	□	□	☑	□
我與本系統的互動是流暢的	□	□	☑	□
本系統的操作流程是合理的	□	□	☑	□
本系統的介面是簡單直覺的	□	□	☑	□
本系統採用的對話形式使我更易理解當前發生的事件	□	□	☑	□
本系統中的自然語言指令功能使我更易理解當前發生的事件	□	□	☑	□

4. 其他

	非常不認同	不認同	認同	非常認同
就學習電腦知識而言，本系統比紙本自學教材優勝	□	□	□	☑
就學習電腦知識而言，本系統比自學網站優勝	□	□	□	☑
本系統適合中小學生學習電腦知識	□	□	□	☑

Subject No.: 7 Date: 10/4/2018 Time: 10:35

可用性問卷

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 個人資料

性別
□男 ☑女

年齡
☑18-23 □24-29 □30 歲以上

是否在學大專生
☑在學 □已畢業

課程種類
□電腦科學及資訊科技 □與醫學及衛生有關的學科 □生物科學 □物理科學 □數學 □工程及科技 □建築學及城市規劃 □工商管理 □社會科學 □法律 □大眾傳播 ☑語言及相關科目 □人文學科 □藝術、設計及演藝 □教育（請選擇主修科目）
□其他：_____

是否於高中時選修資訊及通訊科技
☑是 □否

是否曾經使用即時通訊軟件（如：WhatsApp）
☑是 □否（請跳至第 2 部份）

使用過的即時通訊軟件
☑微訊 □Telegram ☑QQ □Skype ☑Facebook Messenger □Line □其他：_____

有否使用過即時通訊軟件中的機械人功能（bot）
□有 ☑沒有

你認為自己的電腦知識水平為（1 為最低，4 為最高）：
□1 ☑2 □3 □4

你認為自己學習電腦知識的動機為（1 為最低，4 為最高）：
□1 □2 □3 ☑4

2. 系統有用性

	非常不認同	不認同	認同	非常認同
使用本系統使我更有效學習電腦知識	□	□	☑	□
使用本系統使我學習電腦知識更快捷	□	□	☑	□
使用本系統令學習電腦知識變得更加容易	□	□	☑	□
本系統中的自然語言指令功能能夠解答我的問題	□	□	☑	□
本系統中的自然語言指令功能令我更容易得到想要的答案	□	□	☑	□
本系統中的自然語言指令功能有助我理解課程內容	□	□	☑	□

3. 系統易用性

	非常不認同	不認同	認同	非常認同
使用本系統對我來說是容易的	□	□	☑	□
本系統的反應是迅速的	□	☑	□	□
我與本系統的互動是流暢的	□	□	☑	□
本系統的操作流程是合理的	□	□	☑	□
本系統的介面是簡單直覺的	□	□	☑	□
本系統採用的對話形式使我更易理解當前發生的事件	□	□	☑	□
本系統中的自然語言指令功能使我更易理解當前發生的事件	□	□	☑	□

4. 其他

	非常不認同	不認同	認同	非常認同
就學習電腦知識而言，本系統比紙本自學教材優勝	□	□	☑	□
就學習電腦知識而言，本系統比自學網站優勝	□	□	☑	□
本系統適合中小學生學習電腦知識	□	□	☑	□

Subject No.: 8 Date: 10/4/2019 Time: 12:30

可用性問卷

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 個人資料

性別
☑男 ☐女

年齡
☑18-23 ☐24-29 ☐30 歲以上

是否在學大專生
☑在學 ☐已畢業

課程種類
☐電腦科學及資訊科技 ☐與醫學及衛生有關的學科 ☐生物科學 ☐物理科學 ☐數學 ☐工程及科技 ☐建築學及城市規劃 ☐工商管理 ☑社會科學 ☐法律 ☐大眾傳播 ☐語言及相關科目 ☐人文學科 ☐藝術、設計及演藝 ☐教育（請選擇主修科目）
☐其他：_____

是否於高中時選修資訊及通訊科技
☐是 ☑否

是否曾經使用即時通訊軟件（如：WhatsApp）
☑是 ☐否（請跳至第 2 部份）

使用過的即時通訊軟件
☑微訊 ☐Telegram ☐QQ ☑Skype ☑Facebook Messenger ☑Line ☐其他：_____

有否使用過即時通訊軟件中的機械人功能（bot）
☐有 ☑沒有

你認為自己的電腦知識水平為（1 為最低，4 為最高）：
☐1 ☐2 ☑3 ☐4

你認為自己學習電腦知識的動機為（1 為最低，4 為最高）：
☐1 ☐2 ☑3 ☐4

2. 系統有用性

	非常不認同	不認同	認同	非常認同
使用本系統使我更有效學習電腦知識	☐	☐	☑	☐
使用本系統使我學習電腦知識更快捷	☐	☐	☑	☐
使用本系統令學習電腦知識變得更加容易	☐	☐	☑	☐
本系統中的自然語言指令功能能夠解答我的問題	☐	☐	☐	☑
本系統中的自然語言指令功能令我更容易得到想要的答案	☐	☐	☐	☑
本系統中的自然語言指令功能有助我理解課程內容	☐	☐	☐	☑

3. 系統易用性

	非常不認同	不認同	認同	非常認同
使用本系統對我來說是容易的	☐	☐	☑	☐
本系統的反應是迅速的	☐	☐	☑	☐
我與本系統的互動是流暢的	☐	☐	☑	☐
本系統的操作流程是合理的	☐	☐	☑	☐
本系統的介面是簡單直覺的	☐	☐	☑	☐
本系統採用的對話形式使我更易理解當前發生的事件	☐	☐	☑	☐
本系統中的自然語言指令功能使我更易理解當前發生的事件	☐	☐	☑	☐

4. 其他

	非常不認同	不認同	認同	非常認同
就學習電腦知識而言，本系統比紙本自學教材優勝	☐	☐	☑	☐
就學習電腦知識而言，本系統比自學網站優勝	☐	☐	☑	☐
本系統適合中小學生學習電腦知識	☐	☐	☑	☐

Subject No.: 9 Date: 10/4/18 Time: 14:30

可用性問卷

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 個人資料

性別
□男 ☑女

年齡
☑18-23 □24-29 □30 歲以上

是否在學大專生
☑在學 □已畢業

課程種類
□電腦科學及資訊科技 □與醫學及衛生有關的學科 □生物科學 □物理科學 □數學 □工程及科技 □建築學及城市規劃 □工商管理 □社會科學 □法律 □大眾傳播 □語言及相關科目 □人文學科 □藝術、設計及演藝 ☑教育（請選擇主修科目）
□其他：P.E

是否於高中時選修資訊及通訊科技
□是 ☑否

是否曾經使用即時通訊軟件（如：WhatsApp）
☑是 □否（請跳至第 2 部份）

使用過的即時通訊軟件
□微訊 □Telegram □QQ ☑Skype ☑Facebook Messenger □Line □其他：_____

有否使用過即時通訊軟件中的機械人功能（bot）
□有 ☑沒有

你認為自己的電腦知識水平為（1 為最低，4 為最高）：
□1 □2 ☑3 □4

你認為自己學習電腦知識的動機為（1 為最低，4 為最高）：
□1 □2 ☑3 □4

2. 系統有用性

	非常不認同	不認同	認同	非常認同
使用本系統使我更有效學習電腦知識	□	□	☑	□
使用本系統使我學習電腦知識更快捷	□	□	☑	□
使用本系統令學習電腦知識變得更加容易	□	□	☑	□
本系統中的自然語言指令功能能夠解答我的問題	□	□	☑	□
本系統中的自然語言指令功能令我更容易得到想要的答案	□	□	☑	□
本系統中的自然語言指令功能有助我理解課程內容	□	□	☑	□

3. 系統易用性

	非常不認同	不認同	認同	非常認同
使用本系統對我來說是容易的	□	□	☑	□
本系統的反應是迅速的	□	□	☑	□
我與本系統的互動是流暢的	□	□	☑	□
本系統的操作流程是合理的	□	□	☑	□
本系統的介面是簡單直覺的	□	☑	□	□
本系統採用的對話形式使我更易理解當前發生的事件	□	□	☑	□
本系統中的自然語言指令功能使我更易理解當前發生的事件	□	□	☑	□

4. 其他

	非常不認同	不認同	認同	非常認同
就學習電腦知識而言，本系統比紙本自學教材優勝	□	□	☑	□
就學習電腦知識而言，本系統比自學網站優勝	□	□	☑	□
本系統適合中小學生學習電腦知識	□	☑	□	□

Subject No.: 10 Date: 10/04/2018 Time: 1630

可用性問卷

使用即時通訊軟件的鄧天機械人功能作智能教學及評估系統

1. 個人資料

性別
☐男 ☑女

年齡
☑18-23 ☐24-29 ☐30 歲以上

是否在學大專生
☑在學 ☐已畢業

課程種類
☐電腦科學及資訊科技 ☐與醫學及衛生有關的學科 ☐生物科學 ☐物理科學 ☐數學 ☐工程及科技 ☐建築學及城市規劃 ☐工商管理 ☐社會科學 ☐法律 ☐大眾傳播 ☐語言及相關科目 ☐人文學科 ☐藝術、設計及演藝 ☑教育（請選擇主修科目）
☐其他：_____

是否於高中時選修資訊及通訊科技
☐是 ☑否

是否曾經使用即時通訊軟件（如：WhatsApp）
☑是 ☐否（請跳至第 2 部份）

使用過的即時通訊軟件
☑微訊 ☐Telegram ☑QQ ☐Skype ☑Facebook Messenger ☐Line ☐其他：_____

有否使用過即時通訊軟件中的機械人功能（bot）
☑有 ☐沒有

你認為自己的電腦知識水平為（1 為最低，4 為最高）：
☐1 ☑2 ☐3 ☐4

你認為自己學習電腦知識的動機為（1 為最低，4 為最高）：
☐1 ☐2 ☑3 ☐4

2. 系統有用性

	非常不認同	不認同	認同	非常認同
使用本系統使我更有效學習電腦知識	☐	☐	☐	☑
使用本系統使我學習電腦知識更快捷	☐	☐	☐	☑
使用本系統令學習電腦知識變得更加容易	☐	☐	☐	☑
本系統中的自然語言指令功能能夠解答我的問題	☐	☐	☑	☐
本系統中的自然語言指令功能令我更容易得到想要的答案	☐	☐	☑	☐
本系統中的自然語言指令功能有助我理解課程內容	☐	☐	☐	☑

3. 系統易用性

	非常不認同	不認同	認同	非常認同
使用本系統對我來說是容易的	☐	☐	☑	☐
本系統的反應是迅速的	☐	☐	☐	☑
我與本系統的互動是流暢的	☐	☐	☐	☑
本系統的操作流程是合理的	☐	☐	☑	☐
本系統的介面是簡單直覺的	☐	☐	☑	☐
本系統採用的對話形式使我更易理解當前發生的事件	☐	☐	☐	☑
本系統中的自然語言指令功能使我更易理解當前發生的事件	☐	☐	☐	☑

4. 其他

	非常不認同	不認同	認同	非常認同
就學習電腦知識而言，本系統比紙本自學教材優勝	☐	☐	☐	☑
就學習電腦知識而言，本系統比自學網站優勝	☐	☐	☐	☑
本系統適合中小學生學習電腦知識	☐	☐	☐	☑

Appendix VI. Raw interview results

Subject No.: 1

訪問稿

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 你認為本系統哪一功能對於學習電腦知識最有用？ ~~對答~~ IM形式
2. 你認為本系統哪一功能最需要改善？ 對答 ⇒ 可以更AI化
3. 你希望本系統新增哪一功能？ Summary after chapter + T/F Qs
4. 對於自然語言辨識系統，你最想其新增哪一種題型（例如：如何、為何）？ Pronunciation ✓
5. 你認為本系統的導航可以有甚麼改善？ button can be more pronounced ✓
6. 作為自學平台，你認為本系統與 Moodle 比較如何？ More direct ✓✓✓ step-by-step

Low motivation in ~~as~~ asking questions.

Back vs previous page confusing

Ugly UI

Long Passage

Video long loading time ~~technical problem~~

Pronunciation ✓

Subject No.: 2 Date: 25 Mar 2018 Time: 15:48

訪問稿

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 你認為本系統哪一功能對於學習電腦知識最有用？

如可回答問題，自然語言指令最有用。
Test，因為test到是否學習到，即時

2. 你認為本系統哪一功能最需要改善？

NLC，symptoms未能回答

3. 你希望本系統新增哪一功能？

需要完善多啲新增
for自學就係需要check understanding，如果是formal test，就要有summary & checking.

4. 對於自然語言辨識系統，你最想其新增哪一種題型（例如：如何、為何）？

Yes/No，check interpretation of content.

5. 你認為本系統的導航可以有甚麼改善？

Show chapter page
Bucle Confuse problem ~~does~~ does not exist

6. 作為自學平台，你認為使用聊天機械人比使用自學網站(及紙本教材)有何好處？

紙本：有互問，學生比較喜愛，有interaction (e.g. ask question)，較易整理materials；
自學網站：interaction有用，如可完善對話可令學習更有效。
└ drawback：navigation唔夠清晰，唔能夠by category去選擇。

Subject No.: 3 Date: 25 Mar 2018 Time: 17:30

訪問稿

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 你認為本系統哪一功能對於學習電腦知識最有用？

問答（NLU）→ 可以即時問，不需尋找相關章節
分段 → 清晰
Test → Check understanding

2. 你認為本系統哪一功能最需要改善？

做 test 較慢，選擇答案時要 load
完 topic 時要有 indication，eg. end of this topic

3. 你希望本系統新增哪一功能？

字典 (Dictionary)
不太需要 pronunciation，除非成段讀
Summary + T/F Question，check understanding
highlight 功能

4. 對於自然語言辨識系統，你最想其新增哪一種題型（例如：如何、為何）？

Yes/No, check understanding and interpretation of content.

5. 你認為本系統的導航可以有甚麼改善？

Back confusing problem does not exist

整體 OK，局上手

6. 作為自學平台，你認為使用聊天機械人比使用自學網站及紙本材料有何好處？

紙本：方便，不需翻查，問問題可直接知答案，有 device 就 OK。
drawback：不可 highlight 重點，要睇成段先搵到重點
自學網站：較快捷，不需翻查，可直接問問題
↳ 相似，無需一連串動作方可使用

Subject No.: 4 Date: 2/9/2018 Time: 12:30

訪問稿

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 你認為本系統哪一功能對於學習電腦知識最有用？

Summary + checking Qs，可總結，Qs可知道是否真的汲收，
自然語言指令功能（define / what is），字眼示清楚可查詢
導航合邏輯及簡單易用

2. 你認為本系統哪一功能最需要改善？

速度：OK long passage
keyboard
較難上手
→ 整體而言，好處較多

3. 你希望本系統新增哪一功能？

highlight，強調main point
voice input

4. 對於自然語言辨識系統，你最想其新增哪一種題型（例如：如何、為何）？

為何
relationship

5. 你認為本系統的導航可以有甚麼改善？

loading notification

6. 作為自學平台，你認為使用聊天機械人比使用自學網站及紙本材料有何好處？

紙本：NLC，可立即知道答，有multimedia elements (e.g. video)，珠係，比較有趣 ↗容易整理
└不可highlight，technical problem (loss of data)
自學網站：有checking Qs，pre/posttest，loading faster，夠統一（materials + quiz）

interaction + multimedia

Subject No.: 5 Date: 2/4/2018 Time: 14:38

訪問稿

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 你認為本系統哪一功能對於學習電腦知識最有用？

Checking Qs
向系統問問題（字典）
pretest/posttest，看到進步幅度，有目的找尋前測中不足的範疇
Multimedia examples（直接）

2. 你認為本系統哪一功能最需要改善？

問答不一定答到　Long passage → no problem
問答反應較慢
Checking Qs 可提示有關內容 / checking Qs after each page + 可以繼續做

3. 你希望本系統新增哪一功能？

中文翻譯
Highlight
~~Voice input~~

4. 對於自然語言辨識系統，你最想其新增哪一種題型（例如：如何、為何）？

需求不大
inflection

5. 你認為本系統的導航可以有甚麼改善？

最後頁改成'做題'
~~inflect~~

6. 作為自學平台，你認為使用聊天機械人比使用自學網站及紙本材料有何好處？

紙本：比較有趣，互動性比較強 / 紙本無聊，無法做 checking Qs，不能查字典/問問題，無法即時知道答案
　└ 對眼睛不好
自學網站：互動性較強，直覺 / 有條理，學習次序較合理
　└ ~~眼花繚亂~~，介面較漂亮，

Subject No.: 6 Date: 2/9/2018 Time: 18:27

訪問稿

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 你認為本系統哪一功能對於學習電腦知識最有用？

直接問(不用揸書)
Pronunciation (~~Google translate~~)
T/F Qs (可以即時test)
分類較清楚
Multimedia elements

2. 你認為本系統哪一功能最需要改善？

① T/F Qs 可以更多
上手容易
Long Passage X
速度OK

3. 你希望本系統新增哪一功能？

Voice input 可有可冇 (~~Lack~~ detection accuracy)
Highlight (長遠學習，更加容易尋找)
中文翻譯可有可有 MOLE
成段朗讀 (較容易滿足不同學習需要)

4. 對於自然語言辨識系統，你最想其新增哪一種題型（例如：如何、為何）？

Inflection
Example → more
Summary (不用回去 chapter's last page)

5. 你認為本系統的導航可以有甚麼改善？

loading notification 已有
最後頁次成'欲返'，但已有 page no.
Back及 previous page confusion does not exist

6. 作為自學平台，你認為使用聊天機械人比使用自學網站及紙本材料有何好處？

互動性
紙本：可以主動知道答案，可直接去到特定頁面，不同功能要分開做，統一(因有得查字典問問題)
L X
自學網站：直覺，可以直接去到特定頁面，字典、讀音
L

Subject No.: 7 Date: 10/4/2018 Time: 10:35

訪問稿

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 你認為本系統哪一功能對於學習電腦知識最有用？

因為人人都有 devices，快捷，很少看書，用 app 較高效率

對答 (define / difference / summary)，

└ ~~查字~~ 另外查字比較不便

Read aloud (不同 channel 更有動力，文字冇咁吸引)，比較似真人授課

分段，列點 (清楚，短時間學習更多)

Checking Qs (check understanding 可糾正觀念)

Multimedia elements (內容豐富，吸引)

2. 你認為本系統哪一功能最需要改善？

速度慢 (Quiz，數題答案要 load 耐 / define 要等)

~~Define 慢~~ Define show derivatives

Highlight 太複雜 (以致不會使用)

3. 你希望本系統新增哪一功能？

自己寫 notes 功能

4. 對於自然語言辨識系統，你最想其新增哪一種題型（例如：如何、為何）？

句式較長，大不方便

可變成按鈕，可 long press 字 → define / example ...

5. 你認為本系統的導航可以有甚麼改善？

Quiz → jump to

→ flag question

6. 作為自學平台，你認為使用聊天機械人比使用自學網站及紙本材料有何好處？

自學網站：不用轉換頁面 (效率高)，可揀記，查字 (可靠近電腦組拔)，可行度高，不會分心，

└ 對眼睛不好，分心 (notification)

紙本：很方便，不用 ... ，多功能，新穎，更像跟老師問問題 (更智能更好)，

└ 筆記較快

不用眼花繚亂，不會不見，有條理

Subject No.: 8 Date: 10/4/2018 Time: 12:30

訪問稿

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 你認為本系統哪一功能對於學習電腦知識最有用？

對答（立即解答），如有要自己搵
Highlight（溫書可 mark main point），否則花時間
讀音 Read aloud（溫書，audio channel 比較有效），否則有咁深刻

Drilling 可以不停做，可 check 自己是否 understand

2. 你認為本系統哪一功能最需要改善？

Pretest → ~~建議~~ show wrong answered Qs ⇒ 提示頁數
Passage length OK
上手 OK
Checking Qs 數目 OK

3. 你希望本系統新增哪一功能？

可以 ~~marktdo~~ 直接 highlight

4. 對於自然語言辨識系統，你最想其新增哪一種題型（例如：如何、為何）？

Keywords 不明白，就可以直接問，對比起 Google，有直接且有效的答案
Example 可更詳盡（morphing example）更多

5. 你認為本系統的導航可以有甚麼改善？

Chapter page 每可~~都~~更容易尋找某一 page 所提及的內容

6. 作為自學平台，你認為使用聊天機械人比使用自學網站及紙本材料有何好處？

自學網站：資料量適中，容易尋找課題，不用時時 Google，比較直覺（有指示）
　└ 有真人影片，更多生活例子（內容）
紙本：多功能（read aloud, 對答），純文有單一功能
　└ 比較方便，~~手冊~~一拿出來便可，但如有此 system，都 prefer 用 system（功能）

對答形式

如有 device，會使用此 system（對答）比上網方便（多功能）

Subject No.: 9 Date: 10/4/2018 Time: 14:30

訪問稿

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 你認為本系統哪一功能對於學習電腦知識最有用？

字典（即時查，不用網上查）
Highlight 位
對話（似 Whatsapp，似真人對話）
Read aloud（可查讀音，讀音舒服）
多功能（All in one）
Checking question + summary（無需回翻查）

2. 你認為本系統哪一功能最需要改善？

Highlight 較麻煩（多工序）

3. 你希望本系統新增哪一功能？

單字讀音 → 已有（可符合期望）
讀音加 definition

4. 對於自然語言辨識系統，你最想其新增哪一種題型（例如：如何、為何）？

Examples 不足
Definition 可加圖片

5. 你認為本系統的導航可以有甚麼改善？

速度較慢（e.g. quiz miss Qs 要按好多次先可以去返嗰題）
Notif 較不明顯
需要教學及指示（T/F Qs 有反應）

6. 作為自學平台，你認為使用聊天機械人比使用自學網站及紙本材料有何好處？

自學網站：似 whatsapp（比較肯去用，有趣味）似真人對答），多功能，直接問
└ 內容較充實
紙本：All-in-one 不用再帶紙張出街，
└ 較少接觸紙張，直接 highlight（不用多工序），一眼睇晒

Subject No.: 10 Date: 10/4/2018 Time: 16:30

訪問稿

使用即時通訊軟件的聊天機械人功能作智能教學及評估系統

1. 你認為本系統哪一功能對於學習電腦知識最有用？

NLC (Summary，一問就直接彈出，即時接受知識，可以更流暢地學習)
Highlight (可以在原本對話出現個人化內容，製造內容)
Checking Qs (加深印象)

2. 你認為本系統哪一功能最需要改善？

Highlight 操作太複雜
Passage length OK
速度 OK (除了 highlight)

Checking Qs 一答錯可選跳到有關頁面

3. 你希望本系統新增哪一功能？

自己加 notes → 手寫
d 可以加埋讀音及例子

4. 對於自然語言辨識系統，你最想其新增哪一種題型（例如：如何、為何）？

Define program-scripted (Not in database)
句式較長

5. 你認為本系統的導航可以有甚麼改善？

Back disappears
Checking Qs 'Skip' 唔知做乜
每個 subtopic 嘅 dialog box 可更明顯

6. 作為自學平台，你認為使用聊天機械人比使用自學網站及紙本材料有何好處？

自學網站：知道自己的學習痕跡 (對話好處)，介面較熟悉 (較舒服及易於使用，更專心)，網站=作業，System=不會職悉 對功課 係
└ 需要學習操作
紙本：Multimedia elements，可修改，可直接得到答案 (不用揭)，環保
└ 較容易操作 (新手無需學習)，如果內容較少，prefer 紙本，vice versa

Appendix VII. Spreadsheets of tests, survey and interview results

Pretest

Subject No	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28	Score
1	C	C	C	D	C	B	D	D	C	C	C	C	C	D	C	C	C	C	C	C	C	D	C	B	A	D	A	D	10
2	A	C	D	C	B	B	A	D	C	D	C	B	D	C	B	A	B	A	B	B	B	A	C	B	D	B	C	C	9
3	C	C	A	C	C	D	A	C	A	C	D	C	D	D	C	D	D	D	B	C	C	C	C	A	B	D	D	A	10
4	C	B	B	C	B	C	A	D	D	C	C	B	A	B	D	C	B	D	C	B	D	C	A	A	A	C	A	B	8
5	B	D	D	C	B	C	C	C	A	C	D	B	C	C	A	D	C	D	A	B	C	D	B	A	D	A	D	D	4
6	C	D	A	C	B	A	A	D	A	C	A	A	D	D	A	D	C	C	B	B	A	C	B	A	A	B	A	D	8
7	B	D	A	C	C	C	C	D	A	C	C	D	C	B	A	B	B	B	A	D	B	D	C	A	A	A	B	A	9
8	B	A	A	A	A	D	A	C	C	C	B	C	D	D	A	C	B	D	C	B	C	C	D	B	A	D	A	B	15
9	C	B	D	C	A	C	D	A	D	B	A	A	C	D	D	C	A	D	B	B	B	A	C	A	A	A	B	B	3
10	C	A	A	A	B	B	B	D	C	B	D	C	B	C	D	C	C	D	C	B	B	C	D	B	D	D	C	A	11

Posttest

Subject No	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28	Score
1	B	C	B	A	C	B	A	D	B	C	B	C	D	D	C	C	A	B	C	A	C	C	D	B	D	B	C	C	20
2	B	C	A	A	A	B	A	D	B	C	B	C	D	A	C	C	B	B	C	A	B	C	D	B	A	B	C	D	23
3	B	C	A	A	B	B	A	D	D	C	B	B	D	B	C	C	A	B	A	B	C	B	D	B	A	B	C	D	21
4				A	B	B	A	D	D	C	D	B	C	B	C	C	B	A	A	A	B	A	C	A	B	C	A	B	11
5	B	D	A	A	B	B	A	D	B	C	D	C	B	C	C	C	C	B	A	A	A	C	D	B	A	B	C	B	19
6	B	D	A	C	B	A	A	D	B	D	B	C	C	A	A	C	C	B	A	A	A	C	D	A	C	B	C	D	15
7	B	C	A	A	B	B	D	D	A	B	A	C	A	B	C	C	B	B	B	B	C	C	D	A	A	A	C	C	16
8	B	A	A	A	A	B	A	A	D	B	B	C	B	C	C	C	A	A	C	B	C	C	C	B	A	C	A	C	14
9	B	A	B	B	A	B	B	D	D	C	D	D	C	D	D	D	B	D	D	D	B	C	C	A	B	B	C	A	10
10	B	B	A	A	A	B	D	D	B	D	B	C	D	C	C	C	B	B	A	A	A	C	C	B	A	B	D	A	19

Overall		1st round		2nd round		3rd round		3rd round (Subject No 8 excluded)	
Pretest mean	8.7	Pretest mean	9.66666667	Pretest mean	6.66666667	Pretest mean	9.5	Pretest mean	7.66666667
Posttest mean	16.8	Posttest mean	21.3333333	Posttest mean	15	Posttest mean	14.75	Posttest mean	15
		Difference	11.6666667	Difference	8.33333333	Difference	5.25	Difference	7.33333333
		Percentage change	120.69%	Percentage change	125.00%	Percentage change	55.26%	Percentage change	95.65%

Subject No	Gender	Age Group	Is U Student	Education Programme	Programme Type	Studied ICT
1	F	18-23	1	1	Mathematics	0
2	F	18-23	1	0	Communication	0
3	F	18-23	0	0	BBA	0
4	F	18-23	1	0	Socal science	0
5	F	18-23	1	1	Chinese	0
6	F	18-23	1	0	Bioscience	0
7	F	18-23	1	0	Linguistics	1
8	M	18-23	1	0	Socal science	0
9	F	18-23	1	1	Sports	0
10	F	18-23	1	1	IT	0

Subject No	Used IM	WeChat	Telegram	QQ	Skype	Facebook Messenger	Line	WhatsApp	Used Bot	Bot Example	Literacy Level	ICT Learning Motivation
1	1	1	1	0	1	1	1	1	0		3	1
2	1	1	1	0	1	1	1	1	0		2	2
3	1	0	0	0	1	1	1	1	0		3	4
4	1	1	0	0	1	1	1	1	1	Google Form	2	3
5	1	1	0	1	0	0	0	0	0		2	4
6	1	0	1	0	1	1	1	1	1	Telegram	3	2
7	1	1	0	1	0	1	0	0	0		2	4
8	1	1	0	0	1	1	1	1	0		3	3
9	1	0	0	0	1	1	0	1	0		3	3
10	1	1	0	1	0	1	0	0	1		2	3
										Mean	2.500	2.900

Subject No	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16
1	3	3	2	2	2	3	3	2	3	3	3	4	3	4	4	2
2	3	3	2	2	2	2	3	3	3	3	3	3	3	3	3	3
3	3	3	4	4	3	3	4	2	3	3	4	3	4	3	3	4
4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4
5	3	4	4	3	3	3	4	3	4	4	4	4	4	4	4	4
6	4	3	3	4	3	3	3	3	3	3	3	3	3	4	4	4
7	3	3	3	3	3	3	3	2	3	3	3	3	3	3	3	3
8	3	3	3	4	4	4	3	3	3	3	3	3	3	3	3	3
9	3	3	3	3	3	3	3	3	3	3	2	3	3	3	3	2
10	4	4	4	3	3	4	3	4	4	3	3	4	4	4	4	4
Mean	3.200	3.200	3.100	3.100	2.900	3.100	3.200	2.800	3.200	3.100	3.100	3.300	3.300	3.400	3.400	3.300

	Overall mean	3.169	Usefulness mean	3.100	Ease of use mean	3.143
	1st round	2.98	1st round	2.722	1st round	3.095
	2nd round	3.35	2nd round	3.222	2nd round	3.286
	3rd round	3.17	3rd round	3.292	3rd round	3.071

Q1. Features that help learning ICT

	IM style	NLC	Test	Pagination	Summary	Checking Qs	Multimedia	Pronunciation	Category	Read aloud	Mobility	Highlight	Dictionary	All-in-one multi-functioned
1	Y	Y												
2		Y	Y											
3		Y	Y	Y										
4		Y			Y	Y								
5		Y	Y			Y	Y							
6		Y				Y	Y	Y	Y					
7		Y		Y		Y	Y		Y	Y	Y			
8		Y				Y				Y		Y		
9	Y				Y	Y				Y		Y	Y	Y
10		Y				Y						Y		
Count	2	9	3	2	2	7	3	1	2	3	1	3	1	1

Q2. Functions to be improved

	More intelligent AI	Loading time	Physical keyboard	Long passage	Qs prompt related topics	Qs after each page	Qs continue	More Qs	Complicated highlight	Show derivatives directly	Pretest show wrong answer	Qs jump to related pages	Define + pronunciation	Define + picture
1	Y	Y		Y										
2	Y													
3		Y												
4		N	Y	Y										
5	Y	Y		N	Y	Y	Y							
6		N		N				Y						
7		Y							Y	Y				
8				N				N	Y		Y			
9									Y				Y	Y
10	Y	N		N					Y			Y	Y	
Agree	4	4	1	2	1	1	1	1	4	1	1	1	2	1
Disagree	0	3	0	4	0	0	0	1	0	0	0	0	0	0

Q3. Functions to be added

	Summary	Checking Qs	Pronunciation	Dictionary	Highlight	Voice input	Translation	Read aloud	Notes
1	Y	Y	Y						
2	Y	Y							
3	Y	Y	N (unless whole para)	Y	Y			Y	
4					Y	Y			
5					Y	N	Y		
6					Y	N	N	Y	
7									Y
8									
9									
10									Y
Agree	3	3	1	1	4	1	1	2	2
Disagree	0	0	0	0	0	2	1	0	0

Q4. NLC functions improvement

	Yes/No	Why	Relationship	Inflection problem	More example	Summary	Long command	Long press --> button
1								
2	Y							
3	Y							
4		Y	Y					
5				Y				
6				Y	Y	Y		
7							Y	Y
8					Y			
9					Y			
10							Y	
Count	2	1	1	2	3	1	2	1

Q5. Navigation improvement

	Larger buttons	Back/previous page confusing	Show chapter page	End topic indicator	Loading notification	next page --> do Qs	Quiz jump to	Quiz flag question	Unobvious notification	Need training	Back button hidden	Unknown button usage	More obvious topic dialog box
1	Y												
2		N	Y										
3		N		Y									
4					Y								
5						Y							
6		N			N	N							
7							Y	Y					
8													
9									Y	Y			
10											Y	Y	Y
Count	1	0	1	1	1	1	1	1	1	1	1	1	1

Q6. Pros: System vs paper-based

	Interesting (innovative learning method)	Interaction (NLC)	Materials more organised	Mobility	More direct	Multimedia	Checking Qs	All-in-one	Won't lose data	Editable
1										
2	Y	Y	Y							
3		Y	Y	Y	Y					
4	Y	Y	Y		Y	Y				
5	Y	Y			Y		Y	Y		
6		Y			Y			Y		
7	Y	Y	Y	Y	Y			Y	Y	
8								Y		
9	Y			Y				Y		
10					Y	Y				Y
Count	5	6	4	3	6	2	1	5	1	1

	More direct	Step-by-step	Interaction (NLC)	Materials more organised	Checking Qs	Pretest/ Posttest	Faster loading	All-in-one	Highlight	Trust-worthy	Concentration on system (all in one)	Familiar UI	Relaxing UI
1	Y	Y											
2			Y	Y									
3	Y		Y										
4					Y	Y	Y	Y					
5	Y		Y	Y									
6	Y		Y					Y					
7	Y	Y						Y	Y	Y	Y		
8	Y		Y	Y				Y				Y	Y
9	Y		Y										
10			Y										
Count	7	2	7	3	1	1	1	4	1	1	1	1	1

Q7. Pros: System vs Self-learning website

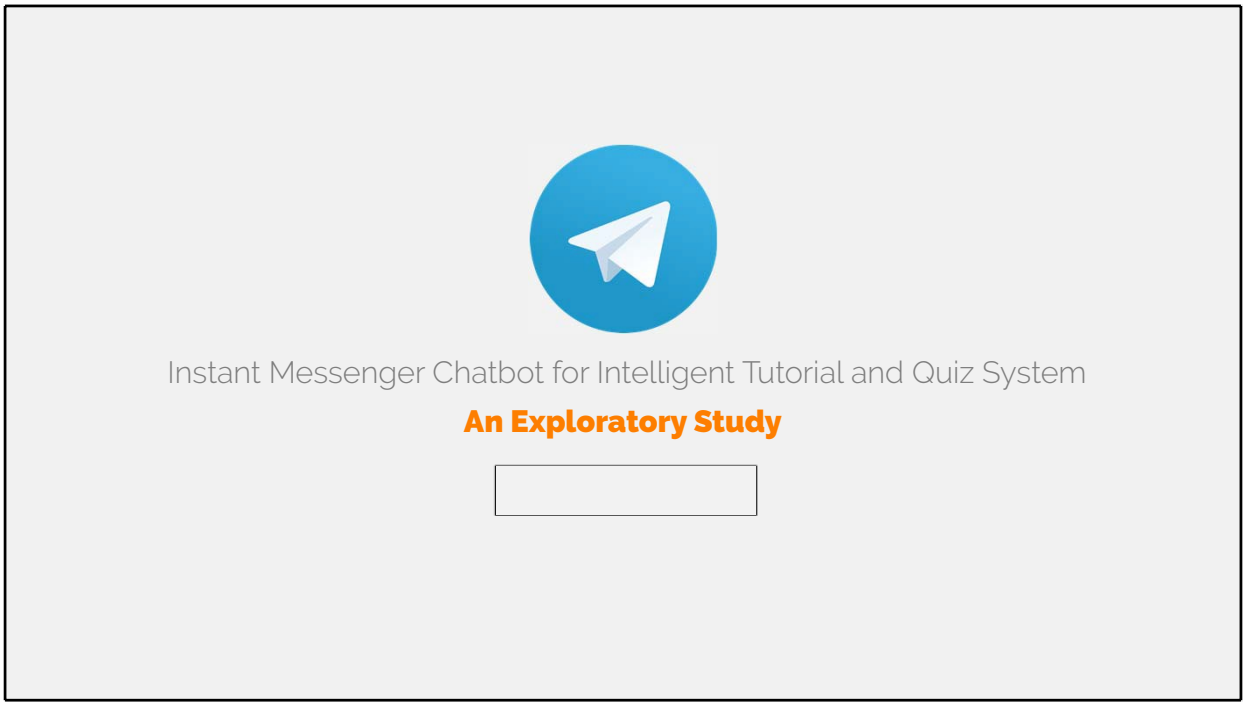
Q8. Cons: System vs paper-based

	No highlight	Loss of data	Bad for eyes	Distraction	Lower mobility	Unable to present at once	Need learning
1							
2							
3	Y						
4	Y	Y					
5			Y				
6							
7			Y	Y			
8					Y		
9	Y					Y	
10							Y
Count	3	1	2	1	1	1	1

Q9. Cons: System vs Self-learning website

	Navigation not clear	No highlight	Less attractive UI	Slower jotting notes	More content	More multimedia	Need learning
1							
2	Y						
3		Y					
4			Y				
5							
6							
7				Y			
8					Y	Y	
9					Y		
10							Y
Count	1	1	1	1	2	1	1

Appendix VIII. Presentation slides



3

Introduction

WHAT THIS RESEARCH IS ABOUT

REASONS FOR DOING THIS PROJECT

Nowadaysinstantmessengersare popular Youngsters are familiar with this kind of technology Most of us have mobile devices Youngsters rely heavily on ICT → Know more about ICT

4

Research Aim & Questions

WHAT THIS RESEARCH IS ABOUT

RESEARCH AIM

To investigate the learning effectiveness of using instant messenger as a tutoring and quiz system

RESEARCH QUESTION 1

Can chatbot be an effective tool for interactive tutoring?

RESEARCH QUESTION 2

Can chatbot be used to deliver ICT course?

5

Literature Review

6

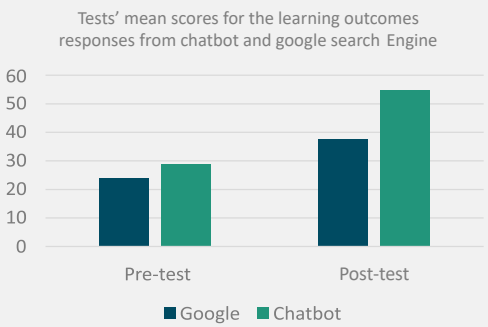
Literature Review

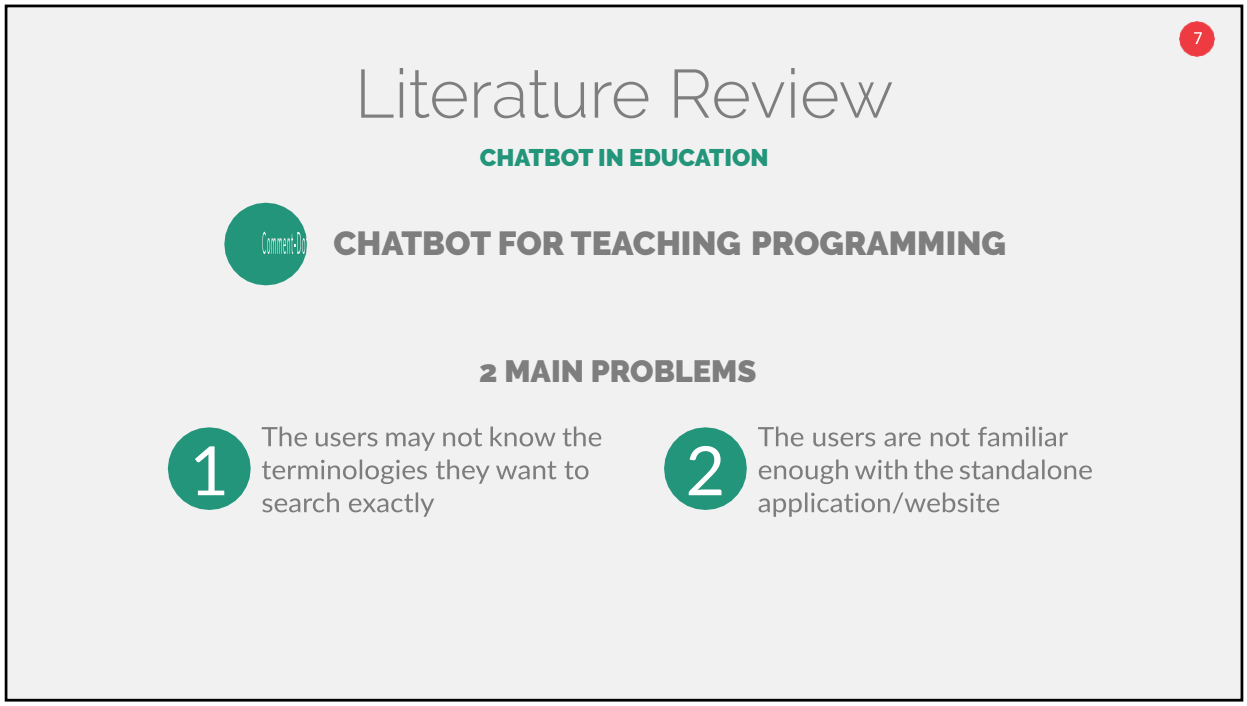
CHATBOT IN EDUCATION

CHATBOT FOR TEACHING PROGRAMMING

"From the overall evaluations, it is concluded that students learning in terms of memory retention and learning outcomes by using the chatbot system is significantly high in terms of quality as well as quantity as compares to the learning through the conventional search engines."

-(Abbasi & Kazi, 2014, p. 65)



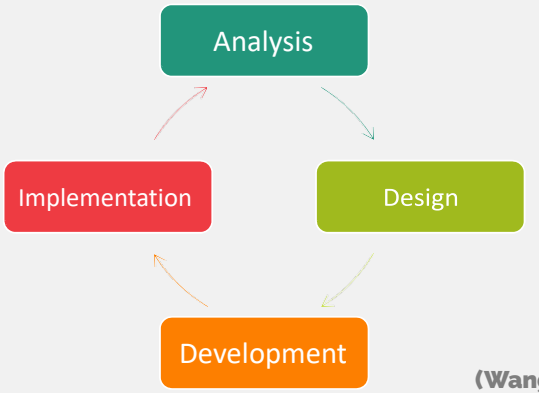


9

Research Design

HOW THIS RESEARCH WAS CONDUCTED

DESIGN-BASED RESEARCH

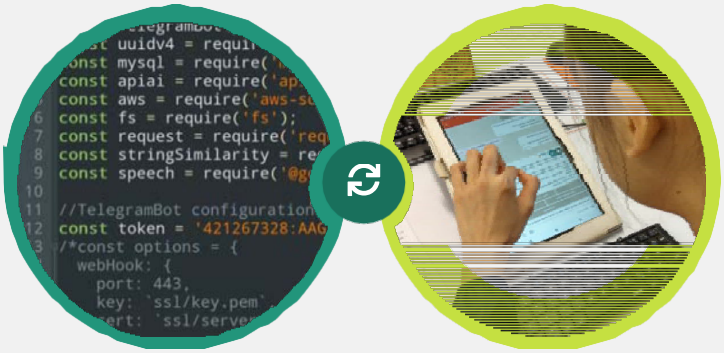


(Wang & Hannafin, 2005)

10

Research Design

HOW THIS RESEARCH WAS CONDUCTED



DEVELOPMENT **EXPERIMENT**

11

Development

HOW THIS PLATFORM WAS MADE

TECHNICAL ASPECT

The following services and technologies are used





12

Development

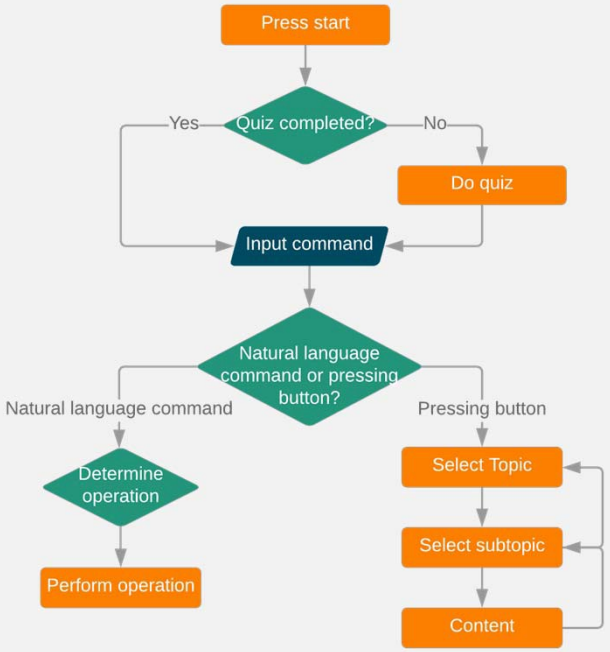
HOW THIS PLATFORM WAS MADE

1. **CHOOSE TEACHING MATERIALS**
2. **ADJUST THE LENGTH OF SELECTED MATERIALS**
3. **DEVELOP THE STRUCTURE OF THE SYSTEM**
4. **INPUT CONTENT INTO DATABASE**
5. **DEVELOP OTHER FEATURES**

13

Flowchart of the System

WHAT THE STRUCTURE IS & HOW THE USERS USE IT



14

Demonstration

HOW IT IS USED

01 **Structure of the system**

02 **Read aloud function**

03 **Natural language command**

04 **Checking questions**

And more...

15

Methodology - Experiment

THE CRITERIA AND HOW IT IS CONDUCTED

PARTICIPANTS

Number: 10

Education: University students

DURATION

At least 80 minutes

- Pre-test: 15 minutes
- Self-study: 30 minutes
- Post-test: 15 minutes
- Survey & interview: 20+ minutes

TESTS ANDSELF-STUDY CONTENTS

Both tests have 28 questions

Topic:

Social issues

- Cybercrime
- Information ethics

Multimedia

- Video
- Animation

16

Methodology - Survey

WHAT QUESTIONS ARE COVERED IN THE SURVEY

PERSONAL INFORMATION

Age

group

Education

- Programme

Studied DSE ICT?

Have used Instant Messengers (IM)?

- What IM have you used?
- Have used bot function in IM?

Computer literacy

USEFULNESS

6 Questions to assess usefulness

EASE OF USE

7 Questions to assess ease of use

MISCELLANEOUS

Superior than paper-based materials / self-learning website?

Suitable for primary / secondary school students to learn ICT?

17

Methodology - Interview

WHAT QUESTIONS ARE COVERED IN THE INTERVIEW

OVERALL

Which are the features that help you learn ICT?

What can be improved?

What features do you want to add?

COMMAND

What can be improved?

NAVIGATION

What can be improved?

COMPARISON

Strengths and weaknesses

18

Methodology

WHAT METHODS ARE USED TO ANALYSE THE DATA

QUANTITATIVE

Analysis of the scores of pre-test & post-test in both descriptive and statistical way

- Mean comparison
- T-test for 2 independent sample
- Mann-Whitney-Wilcoxon test

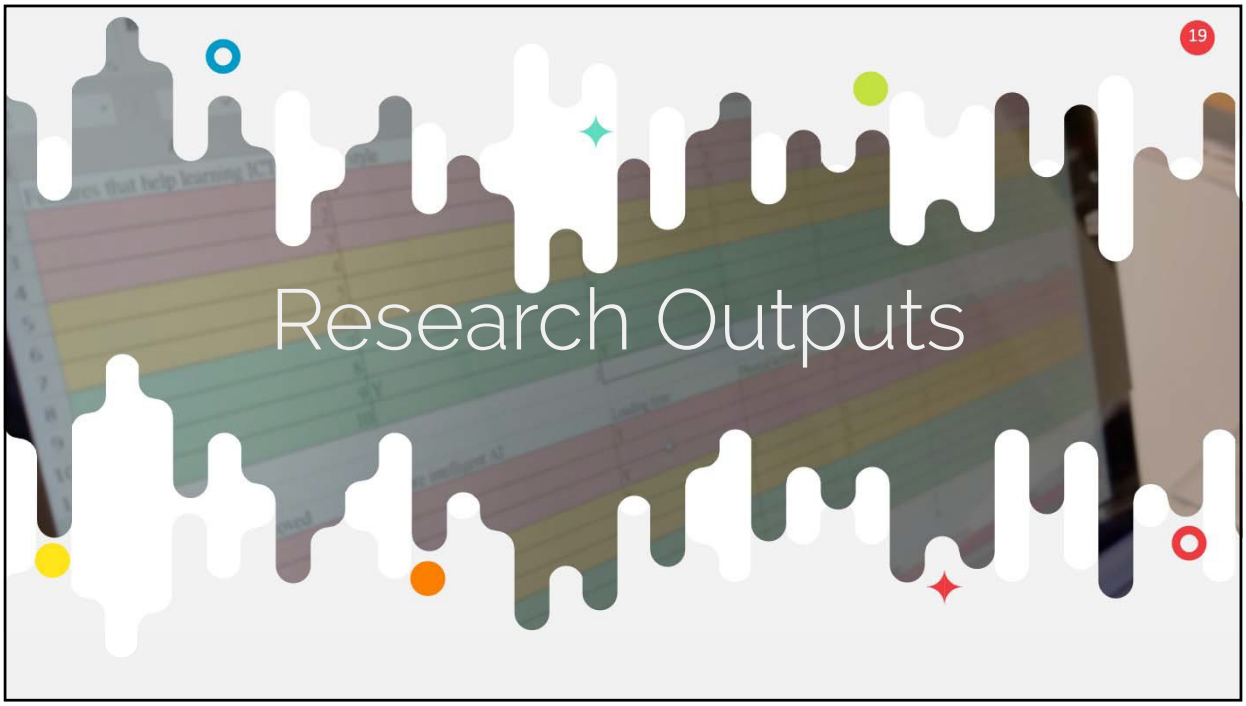
Analysis of the survey for usability (from Technology Acceptance Model) in both descriptive and statistical way

- Mean comparison

QUALITATIVE

Interview for different aspects of the system

- Detailed
- Reasons behind will be investigated



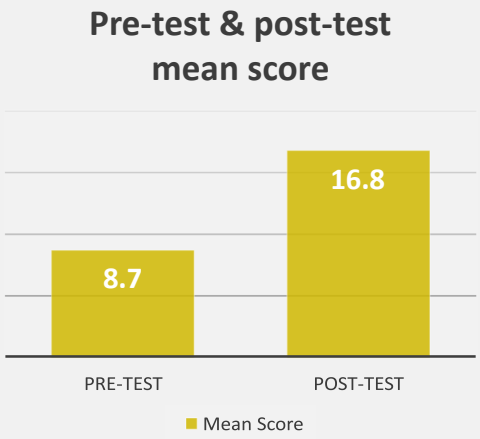
20

Research Outputs (Tests)

SCORES OF TESTS

Mean
- Pre-test: 8.7 / 28
- Post-test: 16.8 / 28

Difference
- Absolute value: +8.1
- Percentage change:+93.1%



21

Research Outputs (Tests)

SCORES OF TESTS

Subject No.	1	2	3	4	5	6	7	8	9	10
Pre-test	10	9	10	8	4	8	9	15	3	11
Post-test	20	23	21	11	19	15	16	14	10	19

Subjects were chosen from a normally distributed population

Paired Samples Test (one-tailed)

- *p*-value = 0.00023
- Significance level = 0.05

p-value < significance level → The result is significant

22

Research Outputs (Tests)

SCORES OF TESTS

Subject No.	1	2	3	4	5	6	7	8	9	10
Pre-test	10	9	10	8	4	8	9	15	3	11
Post-test	20	23	21	11	19	15	16	14	10	19

Wilcoxon Signed Ranks Test (nonparametric test) (one-tailed)

- *p*-value = 0.00338
- Significance level = 0.05

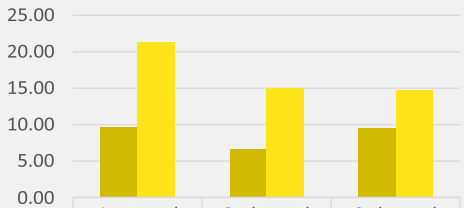
p-value < significance level → The result is significant

23

Research Outputs (Tests)

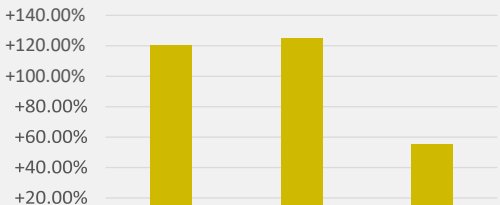
SCORES OF TESTS

Pre-test and post-test mean scores of 3 rounds



	1st round	2nd round	3rd round
Pre-test	9.67	6.67	9.50
Post-test	21.33	15.00	14.75

Percentage change in mean scores of 3 rounds



	1st round	2nd round	3rd round
Change	+120.69%	+125.00%	+55.26%

Show positive changes in 3 rounds
All subjects performed better after studying the materials
Except Subject No. 8

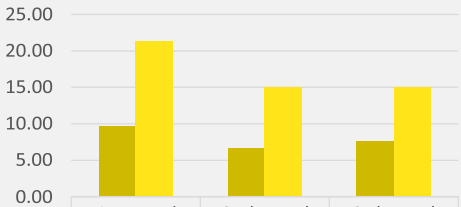
- Pre-test: 15 → post-test: 14

24

Research Outputs (Tests)

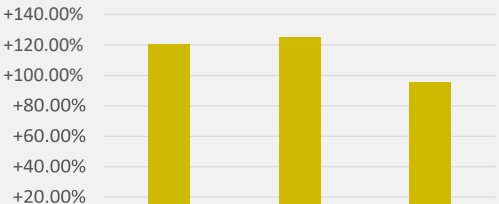
SCORES OF TESTS

Pre-test and post-test mean scores of 3 rounds



	1st round	2nd round	3rd round
Pre-test	9.67	6.67	7.67
Post-test	21.33	15.00	15.00

Percentage change in mean scores of 3 rounds



	1st round	2nd round	3rd round
Change	+120.69%	+125.00%	95.65%

If this entry is removed → More consistent percentage increase

25

Research Outputs (Tests)

DOES IT ANSWER THE RESEARCH QUESTIONS?

RESEARCH QUESTION 1

Can chatbot be an effective tool for interactive tutoring?

RESEARCH QUESTION 2

Can chatbot be used to deliver ICT course?

YES

The subjects scored higher after using the chatbot to learn ICT

26

Research Outputs (Survey)

SCORES OF SURVEY

USEFULNESS (6 QUESTIONS)

Mean (full marks: 4)

- 1st round (3 subjects): 2.722
- 2nd round (3 subjects): 3.222
- 3rd round (4 subjects): 3.292
- Overall: 3.100

EASE OF USE (7 QUESTIONS)

Mean (full marks: 4)

- 1st round (3 subjects): 3.095
- 2nd round (3 subjects): 3.286
- 3rd round (4 subjects): 3.071
- Overall: 3.143

27

Research Outputs (Survey)

SCORES OF SURVEY

Highest overall mean in usefulness (score: 3.2)

Question 1 & 2:
The system helps me learn ICT more effectively and efficiently.

"I could ask the questions and get the answers directly."

-Subject No. 3

"It is convenient. I can use the system as long as I have a mobile device."

-Subject No. 7

28

Research Outputs (Survey)

SCORES OF SURVEY

Lowest overall mean in usefulness (score: 2.9)

Question 5:
The natural language command feature helps me get the answer.

"The system could not answer me the symptoms of infecting Trojan Horse."

-Subject No. 2

29

Research Outputs (Survey)

SCORES OF SURVEY

Highest overall mean in ease of use (score: 3.3)

Question 12 & 13:
The dialog style user interface and natural language command feature help me make sense of the ongoing event at the moment.

"The conversations between me and the system are human-like."

-Subject No. 9

"It is like asking teacher questions."

-Subject No. 7

30

Research Outputs (Survey)

SCORES OF SURVEY

Lowest overall mean in ease of use (score: 2.8)

Question 8:
The system is responsive.

"I needed to wait after I selected the answer in quiz."

-Subject No. 7

31

Research Outputs (Interview)

THE MOST PROMINENT OPTIONS

Question 1

Which are the features that help you learn ICT?

- Natural language command (8)
- Checking questions (7)

Question 2

What functions can be improved?

- Can be more intelligent (4)
- Loading time can be shorter (4)
- Highlight steps can be minimised (4)

32

Research Outputs (Interview)

THE MOST PROMINENT OPINIONS

Question 3

What functions do you want to add?

- Highlight (4)
- Summary (3)
- Checking questions (3)

Question 4

What functions can be improved for natural language command?

- More example (3)
- Yes/No questions (2)
- Inflection problem (2)

Question 5

No prominent option

33

Research Outputs (Interview)

THE MOST PROMINENT OPINIONS

Question 6a

What are the advantages of this system comparing to paper-based materials?

- Allow interaction (6)
- More direct (6)
- Innovation learning method (5)
- All-in-one (5)

Question 6b

What are the disadvantages of this system comparing to paper-based materials?

- No highlight function (3)
- Bad for eyes (2)

34

Research Outputs (Interview)

THE MOST PROMINENT OPINIONS

Question 7a

What are the advantages of this system comparing to self-learning website?

- Allow interaction (7)
- More direct (7)
- All-in-one (4)

Question 7b

What are the disadvantages of this system comparing to self-learning website?

- Content is not rich enough (2)

35

Research Outputs (Survey + Interview)

DOES IT ANSWER THE RESEARCH QUESTIONS?

RESEARCH QUESTION 1

Can chatbot be an effective tool for interactive tutoring?

RESEARCH QUESTION 2

Can chatbot be used to deliver ICT course?

YES

High overall mean scores of usefulness (3.100) and ease of use (3.143) give a positive answer for both questions. Generally, the subjects expressed that the system had more advantages than

36

Research Outputs (Interview)

HOW THE SYSTEM CAN BE IMPROVED & ARE THEY ACCEPTED OR DENIED?

Suggested improvements	Status	Remark / Justification
More intelligent	Accepted	Knowledge base needs more data, continuously adding data
Show related page for wrong answer	Accepted	Added in 2nd round development
Checking question after each page	Denied	Excessive drilling
More checking questions	Accepted	Added in 3rd round development
Dictionary function shows derivatives directly	Denied	Illogical
Integrate definition and pronunciation functions	Accepted	Will add in future development
Dictionary function shows examples and pictures	Denied	Messy and lengthy user interface
Pre-test shows wrongly answered question	Denied	Result in biased self-study experience

37

Research Outputs (Interview)

HOW THE SYSTEM CAN BE IMPROVED & ARE THEY ACCEPTED OR DENIED?

Suggested functions	Status	Remark / Justification
Summary	Accepted	Added in 2nd round development
Checking	Accepted	Added in 2nd round development
Pronunciation	Accepted	Added in 2nd round development
Dictionary	Accepted	Added in 2nd round development
Highlight	Accepted	Deemed unfeasible in 2nd round development, but in great demand
Voice input	Accepted	Added in 3rd round development
Translation	Denied	English is supposed to be the medium
Paragraph read aloud	Accepted	Deemed unnecessary in 2nd round development, but in great demand
Note taking	Denied	Restriction of instant messenger

38

Research Outputs (Interview)

HOW THE SYSTEM CAN BE IMPROVED & ARE THEY ACCEPTED OR DENIED?

Suggested improvements for navigation	Status	Remark / Justification
Shorter loading time	Denied	Depends on internet connection speed
Checking question jump to related page if wrongly answered	Denied	Users may lose focus
Larger button	Accepted	Fixed in 2nd round development
Show chapter page	Accepted	Added in 2nd round development
Jump to specific test question directly	Accepted	Will add in future development
Flag questions in test	Accepted	Will add in future development
More obvious notification	Denied	Restriction of mobile OS and instant messenger design
Remove redundant buttons	Accepted	Will fix in future development
More obvious topic dialog box	Accepted	Will fix in future development

39

Limitation

WHAT CAN BE DONE BETTER

Development

- Knowledge base of the system can be larger

Research

- Accuracy will be higher if sample size and variations can be larger
- More iterations can be conducted
- The scale can be larger if manpower is enough
- Accuracy will be higher if the experiment can last for a longer period of time

40

Conclusion

FINALLY WE ARE HERE

Chatbot can be an effective tool for interactive tutoring

- Subjects performed better
- Results of usability survey show positive trend

Chatbot is suitable to deliver ICT course

- Interview results show that subjects are willing to use
- Interview results show that it has more pros and cons comparing to paper-based materials and self-learning website

41

References

Abbasi, S., & Kazi, H. (2014, 7 16). Measuring effectiveness of learning chatbot systems on student's learning outcome and memory retention. *Asian Journal of Applied Science and Engineering, 3*(2), pp. 251-260.

Wang, F., & Hannafin, M. J. (2005, 12). Design-based research and technology-enhanced learning environments. Educational Technology Research and Development, 53(4), pp. 5-23.

42

THANK YOU

HOPE YOU ENJOY