Do AI-generated Field-specific Vocabulary Quizzes Provide Consistent Results? Measuring Learners' Performance and their Perceptions

Amr Selim Wannas[1] and Inas Hussein Hassan[2]

[1] Magdi Yacoub Heart Foundation (MYF), Aswan Heart Centre (AHC), Egypt.
[2] College of Language and Communication (Alexandria), Arab Academy for Science, Technology and Maritime Transport, Egypt.

E-Mails: amrsilem8842@outlook.com, inas.hussein@aast.edu

ABSTRACT

Artificial Intelligence (AI) is increasingly becoming an incredibly beneficial instrument for language education. There is a variety of AI-powered educational instruments that are embedded in miscellaneous platforms whether they are educational or non-educational. These AI-powered platforms are said to have the ability to improve English language learners' communication skills through the provision of interactive and personalized learning experiences. Moreover, learning English for Specific Purposes (ESP) vocabulary is sometimes problematic for many learners and the search for new effective learning strategies is not an option anymore. In addition, there is a need to create and generate reliable, authentic, specific and valid ESP tests. Thus, the principal aim of this study is to examine the reliability of AI generated field-specific text-based activities and the attitude of cardiac care nurses towards these AI-generated activities. In this endeavour, four AI-composed quizzes were disseminated to cardiac care nurses (n=70) in various departments (e.g., Intensive Care Unit, Cardiac Care Unit, Adult Ward, Operating Room) at a cardiac care Centre and a questionnaire, based on a five-point Likert scale, was administered following the application of the quizzes to identify the attitude of the participants towards these quizzes. To confirm the results, interviews were also conducted with a subset of the questionnaire participants (n=8). The results revealed that the AI-composed vocabulary quizzes pleaded consistent results in terms of learners' level of proficiency. Additionally, the nurses' attitude was positive towards using these quizzes as an assessment tool along with raising their interest and motivation to learn more nursing vocabulary. Furthermore, a model for improving ESP vocabulary through field-specific, contextualized, authentic and level-suitable texts and activities has been constructed. The model is based on the findings of the questionnaire.

Keywords: AI-generated quizzes; AI-composed quizzes for ESP Learners; Artificial Intelligence

1. INTRODUCTION

Artificial Intelligence (AI) is increasingly becoming an incredibly beneficial instrument for language education. There is a variety of AI-powered educational instruments that are embedded in miscellaneous platforms whether they are educational or non-educational. AI-powered platforms are said to have the ability to improve English language learners' communication skills through the provision of interactive and personalized learning experiences (Kushmar et al. 2022; Rusmiyanto et al. 2023).

1.1 Overview of Artificial Intelligence

The evolution of computers in the last eight decades becoming an integral part of humans' daily life has brought about several sequential developments (Roser 2023). The advent of AI occurred because of the emergence of big data, cloud computing, AI neural networks machine learning which, in turn, enthused engineers to construct a human-intelligence simulating machine (Zhai et al. 2021). John McCarthy first used the phrase "Artificial Intelligence" in 1956 at the Dartmouth Workshop, which was a significant occasion that signaled the official beginning of AI as a field of study (Wikipedia 2023). King and ChatGPT

(2023) traced the origin of AI back to the 1950s stating that scientists commenced the experimentation and exploration of this realm of knowledge to finally create the first version of an AI chatbot called ELIZA which mimicked human conversations. Today, several AI technologies have been developed by scientists. These technologies involve Natural Language Processing (NLP), speech recognition, image recognition and processing, Autonomous agents, affect recognition, data mining for prediction and, finally, artificial creativity (Artificial Intelligence in Education 2023).

West and Allen (2018) identified some of the qualities of AI which included intentionality since it can take decisions, intelligence as it can discern information to reach accurate results and adaptability because it can adapt and learn as it takes decisions. This makes it an incredibly useful tool in education as it might be able to overcome some of its most pressing issues, adopt innovative teaching and learning strategies, and get closer to achieving Sustainable Development Goal 4 (Artificial Intelligence in Education 2023). Nowadays, language teachers have a myriad of technological alternative solutions to be utilized for language learning enhancement (Kessler 2016). That is why Kessler (2018) argues that professional development opportunities are required for teachers to recognize the potential of these established and quickly emerging technologies, considering the participatory nature of contemporary society. Numerous AI tools have been proven to improve learners' skills and knowledge after using them in language classes, as stated by Woo and Choi (2021). Similarly, Mageira et al. (2022) argues that AI-powered chatbots improve students' language interactive skills and cultural content. Additionally, Chen et al. (2021) noted that students' language proficiency had increased and that they had a positive attitude towards AI technology and language learning and even recommended using them in language classes.

1.2   *English for Specific Purposes (ESP)*

English for Specific Purposes (ESP) is a type of English language education that caters for learners' needs for communicating in English within a particular field (Hutchinson and Waters 1987; Işık-Taş and Kenny 2019). Basturkmen (2015) added that ESP courses focus either on academic or occupational purposes where the learner is the focus. According to Hutchinson and Waters (1987), the 20th century witnessed the emergence of English as a global lingua franca, facilitating and streamlining international communication. Furthermore, Widodo (2016) added that the occurrence of this phenomenon is a consequence closely associated with the inception of ESP. Due to the United States of America's post-World War II economic, scientific and technical

advancements as well as its advantageous position in the global economic and commercial system, experts from a variety of fields knead in the international marketplace for professionally communicating employees who could function effectively in an English-mediating environment.

Medical terminology originates in specialized language that non-professionals may not be familiar with; yet it is frequently used in the context of the medical profession and overlaps with general English vocabulary to some extent (Panocova 2017). Hsu (2013) stated that medical students significantly depend on their ability to read specialized textbooks in the field, the majority of which are published in English. Still, Muller (2011) noted some issues that foreign students have while picking up medical jargon in English. Their incapacity to read or scan technical writings, their need for some time to talk, their annoyance with communication breakdowns, and their difficulty pronouncing or understanding the meaning of particular terms are all part of these issues. Thus, the first step in comprehending medical English is knowing medical terminology (Hsu 2013).

1.3   *Previous Studies on AI and English Language Learning (ELL)*

There is a dearth of literature regarding using AI in ELL. Zhuomin Sun, Anbarasan and Kumar (2020) concluded that online intelligent English learning platforms with AI support can increase student learning effectiveness and personalize the curriculum. Fenglin Jia et al. (2022) added that AI-powered English language learning applications can assist in learning vocabulary and grammar in English as a second language, enabling widespread and real-world learning. In another study concerned with writing, Xin Zhao (2022) discovered that writing aids driven by artificial intelligence, such as Wordtune, can assist EFL authors in writing consistently and in learning practical methods to convey their thoughts in written English. In addition, Hwang et al. (2023) concluded that writing in English as a Foreign Language (EFL) may be improved with the use of AI-generated example sentences and customized writing feedback, which will increase coherence and consistency. Finally, Gore et al. (2023) found that AI can produce random clausal phrases, which is helpful for online English learners since it offers a range of practice challenges. However, none of the previous studies investigated the utilization of AI in creating vocabulary exercises or quizzes relevant to the medical or the nursing field which is a gap that needs to be filled with proper and accurate scientific research. Also, none of the studies explored the use of AI in the field of ESP which is a relatively new field with some recent expansion, especially English for occupational purposes.

1.4   *Research Problem and Aim*

Teaching ESP presents several challenges such as a shortage of qualified instructors and instructional materials, alongside ineffective teaching methodologies (Khalil and Kholofel 2020). Many learners struggle with ESP vocabulary, necessitating the exploration of new effective learning strategies (Zamfir 2022). Furthermore, there is an urgent need to develop reliable, authentic, specific and valid ESP assessments (Bucur and Neagu 2015). To this end, the study aims to examine cardiac care nurses' consistency of results as provided by AI generated field-specific text-based vocabulary activities and their attitude towards these AI-generated activities and quizzes. The study purports to find practical answers to the following questions:

Question 1: To what extent are the descriptive results of AI-generated quizzes consistent in evaluating the English language performance of cardiac care nurses?

Question 2: What is the attitude of Egyptian cardiac care nurses towards AI-composed ESP quizzes?

2.   METHODOLOGY

The researchers employed a mixed-method approach, quantitative and qualitative, utilizing three data collection methods for the sake of triangulation: AI-generated quizzes, a questionnaire constructed by the researchers and interviews conducted by one of the researchers. The study is exploratory as it attempts to discover a new realm of knowledge that could be of much effect in the field of teaching English as a Foreign Language (TEFL).

The participants of the study were professional cardiac care nurses working at one of the renowned Egyptian cardiac care centers offering free-of-charge treatment for patients with various cardiac diseases, congenital and developed. The numbers of participants in the quizzes were: quiz 1= 77, quiz 2= 78, quiz 3= 72 and quiz 4= 72. The questionnaire was administered to most of the participants (n=70) and the interviews were conducted with a subset of the participants (n=10). All quizzes consisted of 20 items with one mark allocated to each.

All the AI-generated quizzes contained a passage with highlighted words, nursing information about a specific topic and six skimming, scanning and inference questions so that learners would identify vocabulary in context and almost authentic text. The topics of the quizzes included Cardiology and Catheterization (quiz 1), Antibiotics (quiz 2), Types of Medications for Cardiac Patients (quiz 3) and Pain Management (quiz 4). Following the reading passage and the comprehension questions, each of the quizzes included two vocabulary activities. The first one was a matching activity that targeted identifying the meanings of the seven highlighted words in the reading passage. The second activity was a contextualized gap-filling activity including seven gaps so that learners would use the words they had known from the text. These quizzes have been administered through google form with some strict rules to prevent cheating or any type of malpractice and ensure the accuracy of results.
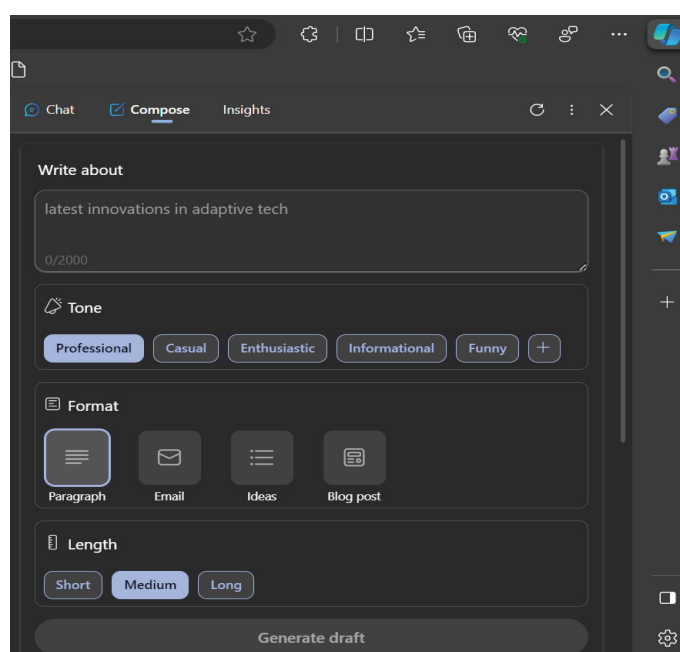

Figure 1:Bing AI (Copilot now) as displayed in Microsoft Edge Note. Researchers' own source.

All quizzes were generated by Microsoft Bing which is a technological AI instrument incorporated in the Microsoft Edge browser. The Bing search engine's AI chatbot from Microsoft relies on the technology of ChatGPT 4 to provide more accurate search results and perform further tasks. The underlying GPT technology enables this chatbot, which is driven by artificial intelligence (AI), to replicate typical human dialogue. The Bing AI chatbot may respond to queries in a more natural and verbose manner as opposed to merely delivering pages of straightforward links and page descriptions interspersed with knowledge boxes, which is similar to what other search engines do. Compose is the feature used to generate field-specific reading texts and vocabulary texts.

Table 1. Distribution of nurses according to demographic information (n =70)

Demographic Information	No.	%
Gender		
Male	40	57.1
Female	30	42.9
Age		
Between 20 and 25	30	42.9
Between 26 and 30	28	40.0
Between 31 and 35	10	14.3
Over 35	2	2.9
Years of Experience in nursing		
Less than a year	12	17.1
1 to 3 years	23	32.9
4 to 6 years	21	30.0
7 to 10 years	11	15.7
More than 10 years	3	4.3

The demographic data in the questionnaire is distributed according to nurses' gender, age and years of experience. As demonstrated in table 1, in terms of gender, 40 male nurses representing 57% and 30 female nurses constituting almost 43% took the questionnaire following their participation in the AI-generated quizzes. Regarding their age, 58 nurses collectively were between 20 to 30 years old whereas 12 nurses were above 30. Moreover, 35 nurses had from less than a year to 3 years of experience in nursing, 32 nurses had from 4 to 10 years of experience and only 3 had more than 10 years of nursing experience. Data were fed into the SPSS and the following statistical tests were used:

Participants' responses in percentages towards the items included in the questionnaire.

Arithmetic mean rank to identify the weight of responses towards the questionnaire items.

Pearson Correlation (r) to ensure the validity of the questionnaire items.

Cronbach's alpha coefficient (α) to measure the reliability of the study's questionnaire.

Following the analysis of the frequencies of the questionnaire items by the SPSS program, a validity test was conducted. The questionnaire administered to the nurses at AHC included 19 statements. All statements had a significant correlation at the ($P \leq 0.01$) level and none of the item's correlation was significant at the level ($P \leq 0.05$) which means that the questionnaire is valid. In addition, the reliability test of

Cronbach Alpha for the 19 items of the questionnaire (α = .94) revealed that the questionnaire is highly reliable and applicable.

The interview participants were 8 cardiac care nurses. They were selected from miscellaneous departments in the cardiac care Centre. These departments included Pediatric Cardiac Intensive Care Unit (PCICU), Adult Cardiac Intensive Care Unit (ACICU), Adult Ward (AW), New Ward (NW). In reference to the analysis, the interview responses were analyzed thematically and coded. Participants were given alternative names apart from their real names such as (MN 1) and (FN 2). The M letter refers to the word male whereas the F letter refers to the word female and the N letter refers to the word nurse. The number at the end of the acronym refers to the number of the interviewee in the interview. These names were given to hide the identity of nurses for confidentiality. Additionally, all interviews were conducted in a comfortable room by one of the researchers. In general, participants were given the opportunity to express themselves freely and openly without any kind of peer or organizational pressure. Interviews were conducted in Arabic and responses were translated into English using ChatGPT
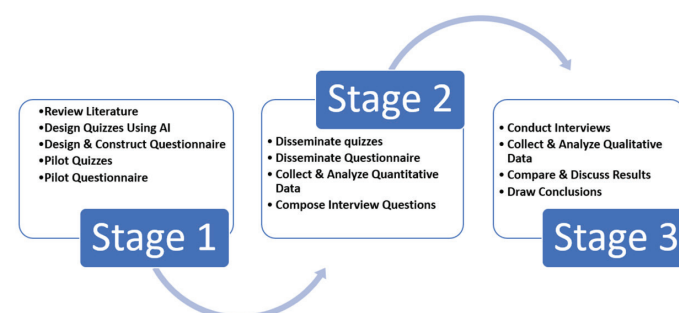

Figure 2: Stages of the study
Note. Researchers' own work.

The research study passed through 3 stages and 10 steps. They are as follows:

1)   Though there was scarcity in the literature about Artificial Intelligence in education, the researchers strived to find out about the gap of the study and its possible utilization in language education.

2)   The researchers then started to know more about AI tools and figure out how they can be used in language education in general and ESP in particular. For the latter one, the researchers selected 4 quizzes out 6 generated by AI to see if they can be used to measure the performance of cardiac care nurses, especially in the point of consistency of results.

3) To measure the perceptions of cardiac care nurses, the researchers constructed a questionnaire to be disseminated after participants take the quizzes.

4) The quizzes and the questionnaire were piloted prior to the experiment to have preliminary results. The results were promising and encouraging to continue the experiment.

5)   The researchers disseminated the AI-generated quizzes during the nursing English sessions and the questionnaire was disseminated at the end of the experiment.

6) Following that, the quantitative collected data were analyzed using the Statistical package for Social Sciences (SPSS) program.

7) To ensure the results of the quantitative data, interview questions were composed and prepared and the interviews were conducted by one of the researchers to have an in-depth insight into the experiences of nurses regarding the AI-generated vocabulary quizzes.

8) Interview data were collected, analyzed thematically and coded.

9) A comparison was implemented to ensure the similarities and discover the differences between the quantitative and qualitative data.

10) Finally, a conclusion was drawn based on this comparison following the discussion of the results.

3. FINDINGS

This part presents the results of the AI-generated quizzes, the questionnaires and the conducted interviews. In addition, the results are interpreted in accordance with the participants' responses.

3.1 *Results of the AI-generated Quizzes*

A total number of 4 AI-generated quizzes have been administered to cardiac care nurses in different English language sessions and they have been given 30 minutes to finish each quiz. The descriptive data demonstrated in figure 3 shows the exploratory results calculated by google forms. These descriptive data included the mean score (referred to as the average), the median and the range. All quizzes were out of 20 points with one point for each question. There were no inferential data calculated by the Google forms.
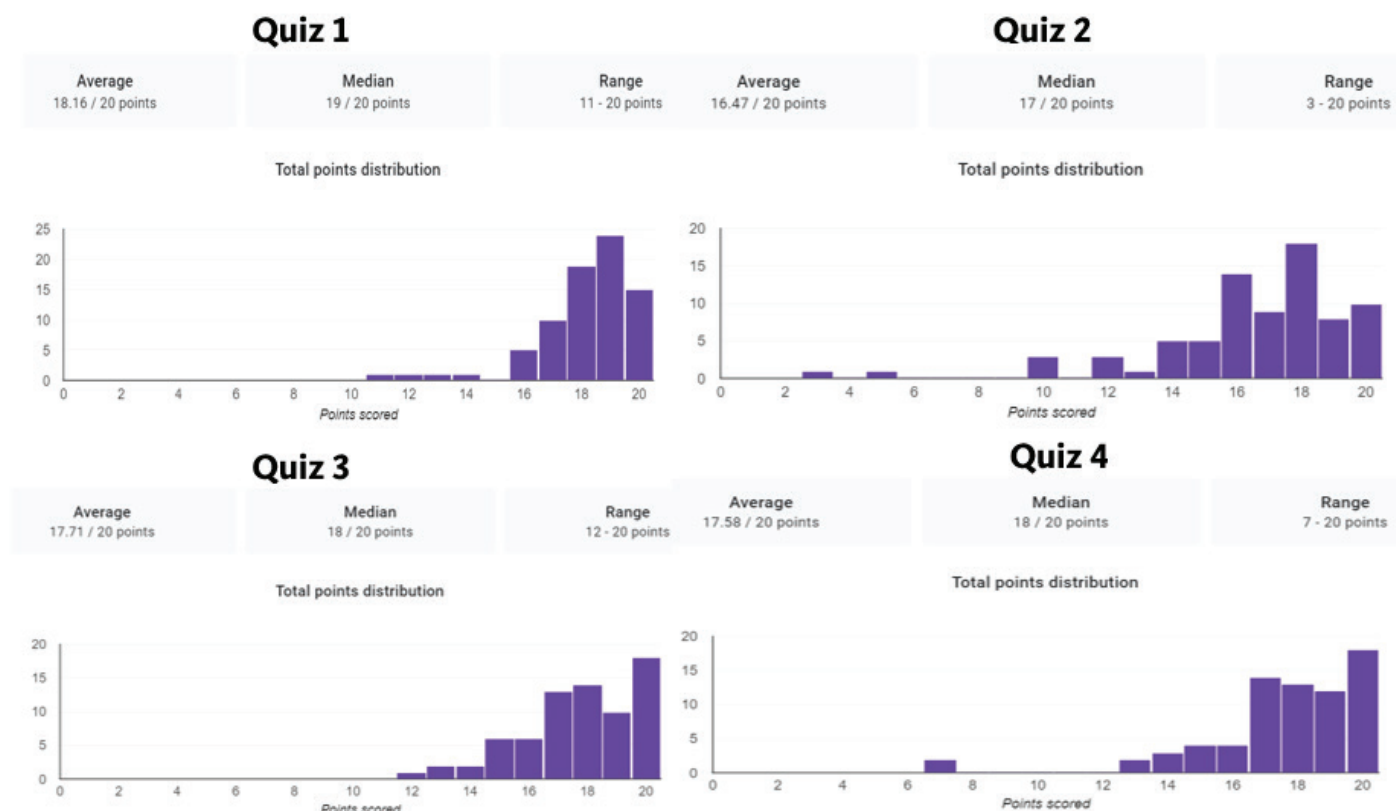

Figure 3: Quizzes results as calculated by Google forms

The vertical axis refers to the number of nurses whereas the horizontal axis refers to the quiz points obtained by a nurse. In Quiz 1, the mean score was (M=18.16), the median was 19 points, and the range was from 11 to 20 points. Quiz 2 was quite lesser than quiz one in terms of the mean score (M=16.47), the median was 17 points, and the range was from 3 to 20. In quiz 3, the mean score was (M=17.71), the median was 18 and range was from 12 to 20 points. Quiz 4 was the closest to quiz 3 in terms of the mean score (M=17.58). The median of quiz 4 was 18 points and the points ranged from 7 to 20 points. What one can deduce from the descriptive data collected and calculated by google forms is that all mean scores and medians are very close to one another whereas the range of points has significant variations among the 4 quizzes. This shows consistency of results in the performance of cardiac care nurses in the AI-composed field-specific quizzes.

3.2 Frequencies of the Questionnaire Items

Tables 2 and 3 present the data collected by the researchers through the previously constructed questionnaire. Aside from the demographic items, the questionnaire contained 19 items eliciting information regarding the perceptions of cardiac care nurses on the AI-composed quizzes taken during their English sessions.

Table 2. Participants' responses towards AI-composed texts and vocabulary quizzes

	Items		Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Mean Rank	Direction
1	The AI-composed texts include useful information about cardiac care nursing	no.	0	0	8	39	23	4.21	High
		%	0.0	0.0	11.4	55.7	32.9		
2	The AI-composed texts include useful vocabulary about cardiac care nursing	No.	0	1	7	41	21	4.17	High
		%	0.0	1.4	10.0	58.6	30.0		
3	I used the information and the vocabulary I learned from the AI-composed texts orally and in the written form at work.	No.	0	2	23	30	15	3.83	High
		%	0.0	2.9	32.9	42.9	21.4		
4	The AI-composed texts carrying nursing-specific vocabulary are informative and useful for me	No.	0	0	13	38	19	4.09	High
		%	0.0	0.0	18.6	54.3	27.1		
5	The AI-composed texts carrying nursing-specific vocabulary are not informative and useful for me	No.	16	44	9	0	1	0.94	Low
		%	22.9	62.9	12.9	0.0	1.4		
6	I tried to guess the meaning of the highlighted vocabulary in the AI-composed texts	No.	1	5	18	37	9	3.69	High
		%	1.4	7.1	25.7	52.9	12.9		
7	I used nursing-specific vocabulary in gap-filling contextualized activities designed by AI	No.	0	3	14	39	14	3.91	High
		%	0.0	4.3	20.0	55.7	20.0		
8	The AI-composed texts and quizzes were interesting because they were field-specific	No.	0	0	9	35	26	4.24	High
		%	0.0	0.0	12.9	50.0	37.1		
9	The AI-composed texts and quizzes motivated me to learn more vocabulary because they are related to nursing	No.	0	0	14	32	24	4.14	High
		%	0.0	0.0	20.0	45.7	34.3		
10	I understood most of the vocabulary in the text because the AI-composed text was appropriate for my level of English language	No.	0	0	9	40	21	4.17	High
		%	0.0	0.0	12.9	57.1	30.0		

In table 2, out of 70 nurses, 62 nurses representing 88.6% either agreed or strongly agreed that the AI-composed texts include useful information about cardiac care nursing whereas only

were neutral. The same proportion reported that the AI-composed texts include useful vocabulary about cardiac care nursing, 7 nurses remained neutral and only one disagreed. Moreover, 45 nurses constituting 64.3% confirmed that they used the information and

the vocabulary they learned from the AI-composed texts orally and in written form at work whereas 23 remained neutral and 2 disapproved. In general, 57 nurses comprising 81.4% agreed that the AI-composed texts carrying nursing-specific vocabulary are informative and useful for them, 13 remained neutral and none disagreed. In reverse, 60 nurses representing 85.8% either disagreed or strongly disagreed that the AI-composed texts carrying

nursing-specific vocabulary are not informative and useful for them whereas only one strongly agreed and 9 remained neutral. In terms of vocabulary learning, 46 nurses constituting 65.8% stated that they tried to guess the meaning of the highlighted vocabulary in the AI-composed texts while 6 disagreed and 18 remained neutral. Also, 53 nurses representing 75.7% reported that they used nursing-specific vocabulary in gap-filling contextualized activities designed by AI. 61 nurses representing 87.1% stated that the AI-

composed texts and quizzes were interesting because they were field-specific and 9 remained neutral. A total number of 56 nurses comprising 80% confirmed that the AI-composed texts and quizzes motivated them to learn more vocabulary because they are related to nursing. Also, 61 nurses representing 87.1% reported that they understood most of the vocabulary in the text because the AI-composed texts were appropriate for their level of English language.

Table 3. Participants' responses towards AI-composed texts and vocabulary quizzes

	Items		Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Mean Rank	Direction
11	I understood the information in the text because the AI-composed text was appropriate for my level of English language	No.	0	1	7	40	22	4.19	High
		%	0.0	1.4	10.0	57.1	31.4		
12	I learned the parts of speech (noun, verb, adj., adv., etc.) of the ESP vocabulary highlighted in the AI-composed texts	No.	0	3	19	35	13	3.83	High
		%	0.0	4.3	27.1	50.0	18.6		
13	I learned the register (formal, semi-formal, informal) of the vocabulary highlighted in the AI-composed texts or quizzes	No.	0	6	27	30	7	3.54	High
		%	0.0	8.6	38.6	42.9	10.0		
14	I learned the meanings of the highlighted words from the AI-composed vocabulary-definition matching activity	No.	0	2	16	37	15	3.93	High
		%	0.0	2.9	22.9	52.9	21.4		
15	I gained some vocabulary knowledge from the AI-composed texts and quizzes related to the nursing profession.	No.	0	3	7	42	18	4.07	High
		%	0.0	4.3	10.0	60.0	25.7		
16	I did not gain any vocabulary knowledge from the AI-composed texts and quizzes related to the nursing profession.	No.	18	44	4	4	0	0.91	Low
		%	25.7	62.9	5.7	5.7	0.0		
17	The AI-composed texts and quizzes aroused my interest in knowing more about cardiac care nursing and nursing-related vocabulary.	No.	0	0	14	39	17	4.04	High
		%	0.0	0.0	20.0	55.7	24.3		
18	The AI-composed texts and vocabulary quizzes improved my vocabulary knowledge (meaning, parts of speech and degree of formality) leading to improving my nursing-related vocabulary skills.	No.	0	4	14	42	10	3.83	High
		%	0.0	5.7	20.0	60.0	14.3		
19	The AI-composed texts and vocabulary quizzes improved my vocabulary knowledge because they are appropriate for my English language proficiency level leading eventually to improve my nursing-related vocabulary.	No.	0	0	11	45	14	4.04	High
		%	0.0	0.0	15.7	64.3	20.0		

In table 3, a total number of 62 nurses comprising 85.5% revealed that they understood the information in the text because the AI-composed texts were appropriate for their level of English language. Also, 48 nurses representing 68.6% approved that they learned the parts of speech (noun, verb, adj., adv., etc.) of the ESP vocabulary highlighted in the AI-composed texts, whereas 3 nurses disagreed and 19 remained neutral. 37 nurses constituting 52.9 stated that they learned the register (formal, semi-formal, informal) of the vocabulary highlighted in the AI-composed texts or quizzes whereas 6 disagreed and 27 nurses stayed neutral. Additionally, 52 nurses representing 74.3%

approved that they learned the meanings of the highlighted words from the AI-composed vocabulary-definition matching activity. In addition, 60 nurses constituting 85.7% approved that they gained some vocabulary knowledge from the AI-composed texts and quizzes related to the nursing profession. The same number either disagreed or strongly disagreed that they did not gain any vocabulary knowledge from the AI-composed texts and quizzes related to the nursing profession whereas only 4 agreed and 4 nurses remained neutral. Also, 56 nurses comprising 80% confirmed that the AI-composed texts and quizzes aroused their interest in knowing more about

cardiac care nursing and nursing-related vocabulary. Moreover, 52 nurses comprising 74.3% approved that the AI-composed texts and vocabulary quizzes improved their vocabulary knowledge (meaning, parts of speech and degree of formality) leading to the improvement of their nursing-related vocabulary skills. Finally, 59 nurses representing 84.3% reported that the AI-composed texts and vocabulary quizzes improved their vocabulary knowledge because they are appropriate for their English language proficiency level leading eventually to improve their nursing-related vocabulary.

3.3 Interviews Responses

One of the researchers conducted interviews with eight cardiac care nurses from different departments. They were only a subset of the questionnaire participants. The codes extracted from the collected data were field-specific information, vocabulary contextuality and authenticity, and level of proficiency. Two themes were identified: interest in learning and vocabulary knowledge.

Participants were asked to provide information whether the quizzes were informative or not. In this respect, FN1 showed signs of surprise knowing that the texts were composed by AI stating, *"I thought that a doctor or a nurse wrote the texts"*, which indicates the accuracy of the AI-composed texts. Also, FN 2 and FN 3 who are staff nurses in the New Ward of the Cardiac Care Centre stated, *"They were informative regarding clinical information, especially the idioms used in the texts"* and *"They were informative in a scientific way, and I gained some knowledge about nursing. They were interesting and excellent because they were about anticoagulants and coronary syndrome"*. Also, MN 2 who works in the same department mentioned that *"some topics were new for me, and I searched for them"*. Another nurse in the Pediatric Ward department acknowledged the beneficial information she gained from the texts stating, *"I refreshed my knowledge about medications especially because there was some missed information from me"*. The perceived impact of these texts and quizzes was obvious through the comment by the same nurse: *"The most interesting one was the medications quiz because I wanted to know more about them"*. The same comment was mentioned by MN 3 who works in the Pediatric Intensive Care Unit stating, *"After the quizzes I found out that I need to read more about the topics"*. Based on the last comment, these quizzes, may contribute to nurses' knowledge enhancement and thereby to patient safety and the provision of the best possible cardiac care.

In terms of the contextualized and authentic vocabulary,

medical vocabulary is sometimes problematic since *"we have many acronyms and abbreviations and according to the sentence and the text, we know what it refers to"* (FN 1). She also stated, *"Sometimes I know the meaning of a word, but I don't know how to use it as word meanings differ from one sentence or a text to another"*. FN 3 continued stating, *"Words are not helpful unless I use them in sentences and then use them in spoken English"*. Also, FN 4 reported, *"I preferred to guess the meaning of a word rather than go to translate it"*. FN 5 also stated, *"I knew most of the words mentioned in the texts but there were a few that I searched for their meanings and others that I understood from the context after reading their sentences a few times"*. These comments confirm the importance of contextualizing field-specific words using AI-generated texts.

In addition, it is important to consider learners' level of English language proficiency, as it is a key factor in capturing their interest and motivating them to learn. All quizzes were suitable for the learners' level of English language proficiency, and this was apparent in their comments. As MN 3 stated, *"I didn't find any difficulty as it was all related to the medical field. The vocabulary was not frustrating"*. In addition, FN 4 and FN 5 also respectively stated, *"I did not find any difficulty in answering the questions"* and *"There were a few difficult parts, but they were understandable through the context"*. Furthermore, FN1 reported, *"I was surprised when I found them in nursing and suitable to my level of English because this made me more interested in learning about nursing"*. Applying Krashen's i+1 hypothesis in designing the quizzes through AI was very helpful to the cardiac care nurses and it led them to be interested in learning the field-specific vocabulary and increase their vocabulary knowledge as well. According to Krashen's i+1 idea (Krashen 1985), language acquisition occurs most effectively when learners are exposed to input that is just a little bit beyond their present competency level. The significance of understandable input for successful language acquisition is highlighted by this concept.

The themes of interest in learning and vocabulary knowledge were evident in the comments of the cardiac care nurses. For instance, MN 3 stated, *"The quizzes were interesting, precise and detailed pieces of information"*. Another comment that shows the relationship between vocabulary knowledge and interest is FN 1's as she stated, *"The more I know, the more I want to know, like knowing the meaning, the part of speech of a word and whether it is formal or informal"*. Moreover, MN 1 was interested in knowing more about the vocabulary he was not familiar with

stating that *"I searched for the meaning of some vocabulary"*. He also continued stating *"When I know a word, I use its synonyms to explain it, as knowing the keyword helps me find related words"*. Similarly, FN 2 was also interested in knowing more about the vocabulary she gained reporting that *"in the second vocabulary exercise I was seeking to know more about the vocabulary"*. All these comments lead us to seriously consider using AI in generating texts and vocabulary exercises in teaching field-specific vocabulary.

4.   DISCUSSION

Analysis of the mean and median scores from the four vocabulary quizzes in section 3.1 reveals that AI-generated quizzes in field-specific contexts can assist instructors achieve improved learning outcomes through offering consistent results and potentially alleviating workload. In terms of the attitude and perceptions of learners regarding these quizzes, almost 80% of participants asserted that the texts were informative. Moreover, almost 75% confirmed that they have used nursing vocabulary in the gap-filling activities which were highlighted in the texts. In addition, their level of interest has been increased as 87% of participants acknowledged. The same number of participants acknowledged that the quizzes were convenient to their level of English proficiency which made them even more interested and motivated to continue and finish these tests together with looking up words and information mentioned in the texts. Also, 85% of participants approved that they have gained vocabulary knowledge from the AI composed texts and vocabulary exercises.
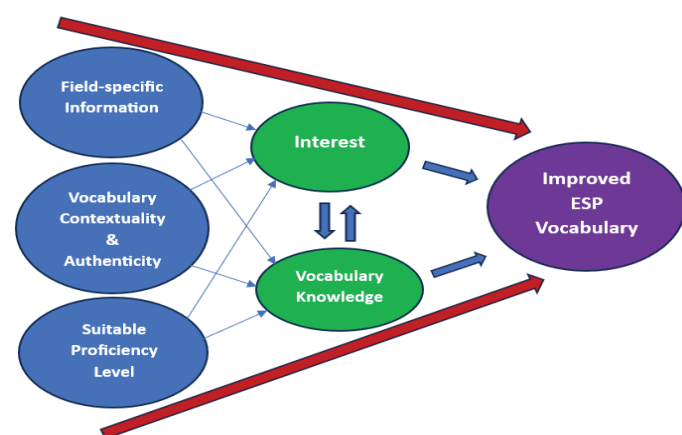

Figure 4: Model constructed based on questionnaire, quizzes and interview results

The four AI-generated vocabulary quizzes administered to cardiac care nurses are a prime example of the potentiality of using AI to generate ESP exercises, quizzes and even full exams under the review of ESP instructors. These quizzes fulfilled certain criteria to teach ESP vocabulary. These criteria included encompassing information about the nursing and medical field so as to grasp the interest of nurses to take part in these quizzes actively. Contextualized and authentic vocabulary is the second criterion, and it is to be guessed and comprehended from the text first then after finishing the quiz (if there is a confusing word), they may go to look it up in a dictionary (Taboada and Rutherford 2011). The third criterion is the suitability of the quizzes level to the learners' abilities (Glušac and Milić 2021). All the three criteria are conducive to two major components: interest and vocabulary knowledge. Both components have a reciprocal relationship that is conducive to improved ESP knowledge and vocabulary (see Figure 4). This is because when nurses acquire vocabulary knowledge, their interest in learning additional language items increases, leading to a greater desire to understand the words they have already learned (Santi et al. 2021).

5.   CONCLUSION

The aim of this study was to measure the reliability of AI-composed field-specific vocabulary quizzes in terms of their consistency of results and the attitude of cardiac care nurses. The findings extended the investigations of previous studies (Anbarasan and Kumar, 2020; Fenglin Jia et al. 2022; Xin Zhao 2022; Wu-Yuin Hwang et al. 2023; Ankita Gore et al. 2023). They were demonstrated in figure 4 and they are as follows:

1)   Artificial Intelligence (AI) can generate field-specific texts to be used in reading quizzes and exams for ESP learners.

2)   The texts generated by AI can be highly similar to authentic texts and can embed the teaching-intended vocabulary in them.

3)   These texts can be generated according to the level of ESP learners.

4)   AI-generated quizzes are reliable in terms of consistency of results.

5)   AI-generated texts can raise the interest of learners through the provision of field-specific useful information and vocabulary.

6)   AI-generated texts can improve learners' field-specific vocabulary knowledge.

7)   Improved field-specific vocabulary knowledge

can raise the interest of learners and vice versa thereby improving their ESP vocabulary in general.

The results imply that using texts and vocabulary activities created by AI is a good way to engage ESP students and improve their learning outcomes. By streamlining the resource search and saving time, this approach not only lightens the job for ESP educators but also gives them the ability to design personalized learning materials that highlight terminology unique to their profession and particular educational topics.

1.   REFERENCES

Basturkmen, Helen. *Developing Courses in English for Specific Purposes*. Springer., 2015.

Bucur, Norica-Felicia, and Camelia Neagu. "THE LIMITS OF ESP TESTS." *Challenges of the Knowledge Society*, 2015, 898–901. https://www.proquest.com/scholarly-journals/limits-esp-tests/docview/1698605013/se-2.

Chen, Xieling, Di Zou, Gary Cheng, and Haoran Xie. "Artificial Intelligence-Assisted Personalized Language Learning: Systematic Review and Co-Citation Analysis." In *2021 International Conference on Advanced Learning Technologies (ICALT)*, 241–45. IEEE, 2021. https://doi.org/10.1109/ICALT52272.2021.00079.

Contributors, Wikipedia. "History of Artificial Intelligence." Wikimedia Foundation, January 2023.

Glušac, Tatjana, and Mira Milić. "Quality of Written Instructions in Teacher-Made Tests of English as a Foreign Language." *English Teaching & Learning* 46, no. 1 (March 9, 2022): 39–57. https://doi.org/10.1007/s42321-021-00079-1.

Gore, Ankita, Vanshika Bajaj, Preeti Yadav, Vaishnavi Chouhan, Madhuri Tayal, and M. Sasi Kumar. "Sentence Generator for English Language Using Formal Semantics." *International Journal of Next-Generation Computing*, February 15, 2023. https://doi.org/10.47164/ijngc.v14i1.1090.

Hsu, Wenhua. "Bridging the Vocabulary Gap for EFL Medical Undergraduates: The Establishment of a Medical Word List." *Language Teaching Research* 17, no. 4 (October 9, 2013): 454–84. https://doi.org/10.1177/1362168813494121.

Hutchinson, Tom, and Alan Waters. *English for Specific Purposes*. Cambridge University Press., 1987.

Hwang, Wu-Yuin, Rio Nurtantyana, Siska Wati Dewi Purba, Uun Hariyanti, Yuniar Indrihapsari, and Herman Dwi Surjono. "AI and Recognition Technologies to Facilitate English as Foreign Language Writing for Supporting Personalization and Contextualization in Authentic Contexts." *Journal of Educational Computing Research* 61, no. 5 (September 19, 2023): 1008–35. https://doi.org/10.1177/07356331221137253.

Işik-Taş, Elvan Eda, and Nalan Kenny. "Current Practices, Challenges, and Innovations in English for Specific Purposes Instruction and Research." In *English for Specific Purposes Instruction and Research: Current Practices, Challenges and Innovations*, 2019. https://doi.org/10.1007/978-3-030-32914-3_1.

Jia, Fenglin, Daner Sun, Qing Ma, and Chee-Kit Looi. "Developing an AI-Based Learning System for L2 Learners' Authentic and Ubiquitous Learning in English Language." *Sustainability* 14, no. 23 (November 22, 2022): 15527. https://doi.org/10.3390/su142315527.

Kessler, Greg. "Technology and the Future of Language Teaching." *Foreign Language Annals* 51, no. 1 (March 19, 2018): 205–18. https://doi.org/10.1111/flan.12318.

———. "Technology Standards for Language Teacher Preparation." In *The Routledge Handbook of Language Learning and Technology*, 2016. https://doi.org/10.4324/9781315657899-16.

Khalil, Lubna, and Brunhielda Kholofelo Semono-Eke. "Appropriate Teaching Methods for General English and English for Specific Purposes from Teachers' Perspectives." *Arab World English Journal* 11, no. 1 (2020). https://doi.org/10.24093/awej/vol11no1.19.

King, Michael R. "A Conversation on Artificial Intelligence, Chatbots, and Plagiarism in Higher Education." *Cellular and Molecular Bioengineering*, 2023. https://doi.org/10.1007/s12195-022-00754-8.

Krashen, Stephen D. *The Input Hypothesis: Issues and Implications*. Longman., 1985.

Lesia Viktorivna, Kushmar, Vornachev Andrii Oleksandrovych, Korobova Iryna Oleksandrivna, and

Kaida Nadia Oleksandrivna. "Artificial Intelligence in Language Learning: What Are We Afraid Of." *Arab World English Journal*, no. 8 (July 25, 2022): 262–73. https://doi.org/10.24093/awej/call8.18.

Mageira, Kleopatra, Dimitra Pittou, Andreas Papasalouros, Konstantinos Kotis, Paraskevi Zangogianni, and Athanasios Daradoumis. "Educational AI Chatbots for Content and Language Integrated Learning." *Applied Sciences* 12, no. 7 (March 22, 2022): 3239. https://doi.org/10.3390/app12073239.

Muller, Amanda. "Addressing the English Language Needs of International Nursing Students." *Journal of Academic Language and Learning* 5 (January 2011): A14–A22. https://journal.aall.org.au/index.php/jall/article/view/145.

Panocová, Renáta. *The Vocabulary of Medical English: A Corpus-Based Study*. England: Cambridge Scholars Publishing, 2017.

Roser, Max. "The Brief History of Artificial Intelligence: The World Has Changed Fast – What Might Be Next?," January 2023. https://ourworldindata.org/brief-history-of-ai.

Rusmiyanto, Rusmiyanto, Nining Huriati, Nining Fitriani, Novita Kusumaning Tyas, Agus Rofi'i, and Mike Nurmalia Sari. "The Role Of Artificial Intelligence (AI) In Developing English Language Learner's Communication Skills." *Journal on Education* 6, no. 1 (May 25, 2023): 750–57. https://doi.org/10.31004/joe.v6i1.2990.

Santi, Eka, Rochmawati Kholipa, Marina Gratciana Putri, and Mujiono. "Reading Interest Strength and Vocabulary Acquisition of EFL Learners: A Meta-Analysis." *Journal of Language and Linguistic Studies*, 2021. https://doi.org/10.52462/jlls.87.

Sun, Zhuomin, M. Anbarasan, and D. Praveen

Kumar. "Design of Online Intelligent English Teaching Platform Based on Artificial Intelligence Techniques." *Computational Intelligence* 37, no. 3 (August 29, 2021): 1166–80. https://doi.org/10.1111/coin.12351.

Taboada, Ana, and Vanessa Rutherford. "Developing Reading Comprehension and Academic Vocabulary for English Language Learners Through Science Content: A Formative Experiment." *Reading Psychology* 32, no. 2 (March 24, 2011): 113–57. https://doi.org/10.1080/02702711003604468.

UNESCO. "Artificial Intelligence in Education | UNESCO," 2023. https://www.unesco.org/en/digital-education/artificial-intelligence.

Widodo, Handoyo Puji. "Teaching English for Specific Purposes (ESP): English for Vocational Purposes (EVP)." In *English Language Education*, Vol. 5, 2016. https://doi.org/10.1007/978-3-319-38834-2_19.

Woo, Jin Ha, and Heeyoul Choi. "Systematic Review for AI-Based Language Learning Tools." *Journal of Digital Contents Society* 22, no. 11 (2021). https://doi.org/10.9728/dcs.2021.22.11.1783.

Zamfir, Tania. "On Vocabulary Learning Strategies in ESP: A Students' Perspective." *SERIES V - ECONOMIC SCIENCES*, June 20, 2022, 97–104. https://doi.org/10.31926/but.es.2022.15.64.1.10.

Zhai, Xuesong, Xiaoyan Chu, Ching Sing Chai, Morris Siu Yung Jong, Andreja Istenic, Michael Spector, Jia Bao Liu, Jing Yuan, and Yan Li. "A Review of Artificial Intelligence (AI) in Education from 2010 to 2020." *Complexity*, 2021. https://doi.org/10.1155/2021/8812542.

Zhao, Xin. "Leveraging Artificial Intelligence (AI) Technology for English Writing: Introducing Wordtune as a Digital Writing Assistant for EFL Writers." *RELC Journal* 54, no. 3 (December 2, 2023): 890–94. https://doi.org/10.1177/00336882221094089.

Appendix

Interview Questions

1- How did you find AI-composed texts? **(Informative – non-informative)**

2- Did these AI-composed texts arouse your interest in knowing more vocabulary and information in your field? **(Yes - Maybe – No)** please explain why?

3- Did putting this vocabulary in context help you to use it effectively? **(Yes - Maybe – No)** please explain how?

4- Does knowing more about the vocabulary you learn (part of speech, register, etc.) arouse your interest in knowing more vocabulary?

5- Would you like to add anything else?