The effects of a theory-based summary writing tool on students' summary writing

Chiou Sheng Chew1 | Norisma Idris1 | Er Fu Loh2 | Wen-Chi Vivian Wu3,4 | Yan Piaw Chua5 | Andrew Thomas Bimba1

1 Department of Artificial Intelligence, Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia
2 Academy of Language Studies, Universiti Teknologi MARA, Melaka, Malaysia
3 Department of Foreign Languages and Literature, Asia University, Taichung City, Taiwan
4 Department of Medical Research, China Medical University Hospital, China Medical University, Taichung, Taiwan
5 Institute of Educational Leadership, University of Malaya, Kuala Lumpur, Malaysia

Correspondence
Norisma Idris, Department of Artificial Intelligence, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia.
Email: norisma@um.edu.my

Abstract
This paper focuses on the design and evaluation of a theory-based computer-assisted summary writing learning environment called Summary Writing-PAL (SW-PAL). The SW-PAL was developed based on four aspects: summarizing strategies, learning theories, prior knowledge, and cognitive load. A quasi-experiment that involved 58 undergraduates majoring in Computer Science was conducted to examine the effectiveness of SW-PAL in writing summaries. Two intact classes were selected with 28 and 30 students in control and experimental groups, respectively. The conventional teaching approach was employed in the control group, whereas the SW-PAL was introduced to the experimental group. Pretest and posttest were administrated to both groups. The findings indicated that SW-PAL improved students' summary writing performance. A significant variance was noted between intrinsic and extraneous cognitive load among students with varying levels of English proficiency in the experimental group, signifying that the SW-PAL is more suitable for students with lower language proficiency.

1 | INTRODUCTION

Secondary and tertiary institutions throughout ASEAN countries have adopted summary writing as a measurement to assess the ability of students in comprehending texts in the English language. Although a substantial number of studies have investigated summary writing, only a handful have initiated innovative techniques that have led to summary writing enhancement (Abdi, Idris, Alguliyev, & Aliguliyev, 2016; Cho, 2012; Friend, 2001; Idris, Baba, & Abdullah, 2011; Ke & Hoey, 2014; Marzec-Stawiar ska, 2016; McDonough, Crawford, & De Vleeschauwer, 2014; Sung, Liao, Chang, Chen, & Chang, 2016; Wichadee, 2014; Yang, 2015). Obviously, some students have poor summary writing skill, particularly among English as a second language (ESL) students (Hosseinpur, 2015; Idris et al., 2011; Kim, 2001; McDonough et al., 2014; Wichadee, 2014). Summary writing reflects repetition of sections from the original text with absence of comprehension. Apprehension among students towards summary writing is partly due to lack of acquired summary writing skills (Wichadee, 2014). Prior studies have mostly emphasized on the assessment of summary, while dismissing the aspect of making students comprehend the source text. In fact, a correlation exists between good understanding of the source text and good summaries (Mokeddem & Houcine, 2016).

In order to acquire good summary writing skills, reading comprehension, without doubt, is indeed significant. Reading comprehension is defined as the process of simultaneously extracting and constructing meaning of what is read (Nandhini & Balasundaram, 2013). Students need to decode what they have read, apart from making connections between what they have read and their prior knowledge (Ausubel, 1963). Prior knowledge is formed from the experiences one has had, seen, and heard in the past. More often, students tend to forget a previously learnt concept, thus requiring some form of simulation to trigger prior knowledge in the long-term memory.

Apart from prior knowledge, one key area related to summary writing skills refers to identifying commonly used summarizing...
strategies. Several researchers have identified a number of effective summarizing strategies applied by both students and teachers (Abdi et al., 2016; Idris et al., 2011), such as topic sentence selection, deletion, sentence combination, paraphrasing, generalization, and invention. The outcomes from these studies signify that mastery of summarizing strategies aids in producing good summaries. Nevertheless, most of the summarizing strategies identified failed to completely incorporate summary writing instructions in schools at all levels. Even though some basic strategies have been applied in teaching indirectly, some students still fail to write good summaries (Idris et al., 2011; McDonough et al., 2014; Wichadee, 2014). Poor knowledge of summarizing strategies appear to be the main reason for the deteriorating summary writing (Hosseinpur, 2015).

An important consideration in summary writing is the cognitive load exacted by the task on the students. Summarizing is a highly complex, recursive reading–writing activity that can impose an overwhelming cognitive load on students, thereby adversely affecting performance (Kirkland & Saunders, 1991). Thus, examining students’ cognitive load during the process of summary writing would provide useful insights into the role that cognitive load plays in mediating their performance in relation to their levels of English language proficiency in summary writing.

Recent interest of researches concerning summary writing has embedded computer technology in the education thread (Abdi et al., 2016; Idris et al., 2011; Sung et al., 2016; Wade-Stein & Kintsch, 2004). Vast research outcomes have typically proven that learners achieve learner autonomy in language learning and significant learning gains via computer-assisted learning (CAL) environment (Mukama, 2009; Ridgway, 1986; Wang, Shang, & Briody, 2013). Therefore, this study has designed and developed a theory-based CAL environment for summary writing called Summary Writing-PAL (SW-PAL) that can be used to assist students in learning summarizing strategies, as well as to support activation of students’ prior knowledge while reading source texts. In addition, a suitable measurement of cognitive load was carried out to determine the variations in the effectiveness of learning environments as a function of instructional formats and learner characteristics.

The purpose of this study is threefold. First, it investigated if the SW-PAL learning environment could enhance the performance of ESL students’ summary writing. Second, it assessed the extent to which the SW-PAL learning environment affected the ESL students’ cognitive load based on their English language proficiency levels. Third, this study evaluated the students’ perceptions pertaining to the use of concept mapping, worked examples, and feedback.

The approach of this study was guided by the following four research questions:

1. Is there a significant difference in ESL students’ summary writing performance after using the SW-PAL?
2. Is there a significant difference in intrinsic cognitive load between students with modest (Band 3) and satisfactory (Band 4) English language proficiency levels after using the SW-PAL?
3. Is there a significant difference in extraneous cognitive load between students with modest (Band 3) and satisfactory (Band 4) English language proficiency levels after using the SW-PAL?
4. What are the students’ perceptions towards the use of concept mapping, worked examples, and feedback?

2 | LITERATURE REVIEW

Summary writing is denoted as the process of producing a more concise text by selecting important information and deleting other supporting and explanatory details from the source text (Abdi et al., 2016; Idris et al., 2011; Kim, 2001; Kirkland & Saunders, 1991). Summarizing skills appear to be integral in comprehending a text, hence, have been employed to evaluate students’ understanding in secondary and tertiary education in most ASEAN nations (Abdi et al., 2016; Idris et al., 2011). In order to produce a good quality summary, effective summarizing strategies are essential (Abdi et al., 2016; Cho, 2012; Friend, 2001; Hosseinpur, 2015; Idris et al., 2011; Ke & Hoey, 2014; McDonough et al., 2014).

The effective summarizing strategies listed by Idris et al. (2011) and Abdi et al. (2016) are topic sentences selection, deletion, sentences combination, paraphrasing, generalization, and invention. The strategy of “topic sentence selection” determines the main idea within a paragraph. Next, “deletion” discards insignificant lexical and phrases from a selected sentence. Meanwhile, “sentence combination” amalgamates a couple or more sentences to arrive at a summarized sentence. “Paraphrasing” refers to sentence rewriting using one’s own words while retaining the meaning. In “generalization,” a word, a phrase, or a sentence is constructed to substitute lexical or a list of selected ideas (Idris et al., 2011). Lastly, “invention” reflects sentence construction to assert an implicit main idea in relation to the topic at hand.

Choy and Lee (2012), Hosseinpur (2015), Ke and Hoey (2014), and McDonough et al. (2014) concluded that instructions related to summarizing strategies (e.g., selection of main ideas and topic sentence, syntactic transformation, abstraction, copy/paste, paraphrasing, deletion, generalization, invention, and sentence combination) displayed substantial enhancement and positive impacts on performance of ESL summary writing. Nevertheless, these summarizing strategies have yet to be embedded in the related instructions at schools. Meanwhile, the conventional approach of writing summary seems to be inefficient because of limited summarizing strategies implementation.

Studies concerning CAL have escalated in line with technology advancement so as to enhance summary writing skills among students (Abdi et al., 2016; Chiu, 2015; He, Hui, & Quan, 2009; Idris et al., 2011; Sung et al., 2016; Wade-Stein & Kintsch, 2004). CAL instills confidence amidst teachers and students to develop both pedagogical and technological skills (Jeong, 2017; Wang et al., 2013). In fact, CAL has been proven time and again in its efficacy of promoting enhanced learning performances, especially when compared with conventional methods (Chiu, 2015; Jeong, 2017; Wade-Stein & Kintsch, 2004;
Wang et al., 2013). An article entitled “The No Significant Difference Phenomenon” by Thomas Russell (1999) asserted the insignificance of technology in classrooms that concludes similar effects of both conventional and technological approaches upon teaching. Similarly, Zhao, Lei, Lai, and Tan (2005) revealed insignificant variance in student outcomes from comparing courses that embedded technology and conventional classroom settings. In a more recent study, Rajab (2018) compared the effectiveness of using CAL environment and conventional classroom teaching by assessing e-learning experience and yielded similar results with Russell (1999) and Zhao et al. (2005). Such inconclusive results in CAL studies call for more researches to provide empirical evidence pertaining to the effectiveness of CAL in general and of summary writing in CAL environment in particular.

Since the previous decade, a number of CAL tools were developed for summary writing so as to enhance summarizing skills among students, such as summary street (Wade-Stein & Kintsch, 2004), summary assessment system (He et al., 2009), summary sentence decomposition tool (Idris et al., 2011), online summary assessment and feedback system (Sung et al., 2016), and relevance detection summarizing strategies identification tool (Abdi et al., 2016). Nonetheless, it seemed that these tools happen to omit the prewriting comprehension phase. This is because the very phase of writing the summary excludes the phase where source text needs to be comprehended. According to Ambrose, Bridges, DiPietro, Lovett, and Norman (2010), one with activated prior knowledge would be able to better understand the source text and produce flawless summaries, in comparison with one without any prior knowledge. Besides, most of the developed summary writing tools only focus on summary quality, content coverage, and strategies of writing a summary but omit one vital aspect of using the tool, which is cognitive load.

Leppink, Paas, Van der Vleuten, Van Gog, and Van Merrienboer (2013) categorized cognitive load into three types: intrinsic load (IL), extraneous load (EL), and germane load (GL). IL is defined as a combination of the innate difficulty of the material being learned, as well as the learner’s characteristics. EL refers to the load placed on working memory that does not contribute directly towards material comprehension. GL is the instructional features that are necessary to learn a material (Leppink et al., 2013). In order to tap into one’s prior knowledge effectively, one’s EL should be at a minimum level, while the IL at maximum during the learning process (Leppink, Paas, Van der Vleuten, Van Gog, & Van Merrienboer, 2014). When EL is higher while learning with limited prior knowledge, more resources are required by the working memory and redundant information processes occur (Chandler & Sweller, 1991), thus disabling effective acquisition. As such, measuring both aspects of IL and EL is essential in learning. Upon low EL and optimum IL, acquisition is bound to take place, such as the GL imposed via self-elaboration (Leppink et al., 2014). Leppink and van den Heuvel (2015) claimed that measurement of GL can never be accurate because it is challenging to determine the assimilation between GL and one’s working memory, mainly due to its limited capacity.

Apart from cognitive load, incorporation of learning theories is greatly lacking in the existing summary writing tools (Abdi et al., 2016; He et al., 2009; Idris et al., 2011; Sung et al., 2016; Wade-Stein & Kintsch, 2004). According to Alzaghoul (2012), incorporating learning theories into educational tools enhances one’s learning experience. Hence, there is a pressing need to develop a theory-based summary writing learning environment.

3 | A THEORY-BASED CAL ENVIRONMENT: SW-PAL

The learning theories that underpin the SW-PAL are Rosenblatt’s transactional theory (1988), Ausubel’s theory of meaningful learning (1963), Sweller’s theory of cognitive load (1988), and Vygotsky’s theory of the zone of proximal development (Vygotsky, 1978). According to Rosenblatt (1988), transactional theory links reading with prior knowledge to effectively understand source text. New meaning is established after associating prior knowledge with text (Anson, 1989). Nevertheless, in order to meaningfully organize information, one should link past knowledge with the newly acquired information (Ausubel, 1963).

Ausubel initiated the meaningful learning theory in 1963, which upholds the notion that the prior and new knowledge link leads to meaningful learning. Besides, the application of advance organizer enhances one’s comprehension due to the process that relates existing and new knowledge by activating data processing, which assists learners to learn actively (Rohman, 2017). Regardless of the numerous advance organizers available, concept mapping appears to be the best for summary writing (Desoiza, 2011; Yang, 2015) due to its graphical features that embed multiconcepts for grasping essential gist in summary writing (Tseng, Chang, Lou, Tan, & Chiu, 2012). Farshi and Tavakoli (2014) reported that students appeared to better organize and relate their ideas using concept mapping, apart from linking past and new knowledge. Hence, concept mapping was integrated in SW-PAL as a vital feature.

Sweller (1988) developed the cognitive load theory that asserts effective learning takes place best upon aligned instructional conditions and cognitive architecture, which is composed of working, long-term, and sensory memories. In fact, the theory claims that learning happens at unlimited long-term memory and limited working memory (handling new knowledge with restricted duration and capacity). Thus, it is integral to weigh in cognitive load offered by instructional design because of restricted working memory (Sweller, Van Merrienboer, & Paas, 1998). Besides, investigations that have adopted this theory highlight the implementation of instructional method (e.g., worked examples) to reduce cognitive load and enhance learning process (McLaren & Isotani, 2011).

According to Sweller et al., (1998), worked examples offer students instances of problems and solutions in a step-by-step manner, which has been widely applied amidst domains that are well structured (e.g., mathematics, computer programming, and physics (Sweller, Ayres, & Kalyuga, 2011). Although worked examples have been proven effective for the less-structured domain (e.g., language learning, social studies, and history; Kyun, Kalyuga, & Sweller, 2013; Yu,
2015), a number of subjective dimensions have been noted as the outcomes heavily rely on comprehension of texts and proficiency, inclusive of writing summary task (Roscoe & McNamara, 2013). Because the technique of worked examples seemed to be an effective solution to learn writing summary, it was integrated in the SW-PAL as the instructional approach to teach summarizing strategies.

Vygotsky promulgated the zone of proximal development theory in year 1978, which defines the difference of student attainment with or without assistance. He added that a learner who follows an example closely would eventually accomplish tasks without aid. Scaffolding is an effective channel provided in zone of development towards accomplishing tasks, wherein initial support at high levels is withdrawn slowly to promote independent learning. Therefore, scaffolding reflects feedback offered to students in acquiring summarizing strategies with positive encouragement to independently produce summaries. Here, feedback is the basic means that allows learners to assess and enhance writing summary skills (Roscoe & McNamara, 2013) in an independent manner without much reliance on the instructor. As such, feedback was integrated in SW-PAL as a tool for scaffolding.

The significant aspects extracted from the above-mentioned theories of learning were prior knowledge, advance organizer, instructional approach, and scaffolding. Additionally, concept mapping, worked examples, and feedback were embedded in SW-PAL. Figure 1 illustrates the conceptual framework of SW-PAL, along with the integrated features that are further described in the following sections.

4 | THE DEVELOPMENT OF SW-PAL

The three main components developed in SW-PAL are prior knowledge activation, summarizing strategies instruction, and scaffolding. The concept mapping technique was employed in prior knowledge activation component, which was applied as advance organizer in activating schemata among learners during the reading phase. The concept mapping, which relates new and existing knowledge to understand source texts, adopted the design guidelines outlined by Novak and Canas (2008) based on two features: (a) concepts that are hierarchically arranged with most inclusive and general notions at the top and the most specific at the bottom depending on the knowledge domain and (b) cross links that bridge concepts in varied domains found in the concept map.

Based on the design guidelines recommended by Novak and Canas, the learners need to identify three primary elements: (a) main topic, (b) main idea or key concepts, (c) supporting details for each main idea. Figure 2 portrays the concept mapping with main general topic, followed by main idea as key concept (refers to node) and supporting details (lines) to construct meaningful idea.

Figure 3 displays the concept mapping tool adapted from an open source mapping tool built by Richard (2017), which had offline and online applications based on HTML5 technology. Delete, save, add, and image export were several functions embedded in the concept mapping, which resembled a tool for drawing for learners to outline concept maps at the reading phase so as to associate new and existing knowledge.

Next, worked examples adapted from Sweller’s et al. (2011) cognitive load theory had been applied in the component of summarizing strategies instruction as the instructional approach. In order to keep the cognitive load at a minimal level, worked examples presentation features, as well as redundancy and split-attention effects, had been weighed in. The integration of varied sources of information (e.g., images and explanatory texts) into a schema leads to split-attention effect (Chandler & Sweller, 1991; Sweller et al., 2011). Upon temporal or spatial distribution of sources, the learners would need to constantly shift their attention, thus defeating the purpose of worked examples. This is because worked examples demand learners to go through information derived from multiple sources (images, audio clips, or texts) that becomes cognitively demanding; therefore, interfering with learning instead of facilitating it (Tarmizi & Sweller, 1988). Hence, multiple information resources were discarded from the worked examples design for SW-PAL.

Second, multiple sources with similar information can lead to redundancy effect that may hinder learning from taking place (Chandler & Sweller, 1991). For example, a written text (visual
presentation) that is presented concurrently with reading aloud (auditory presentation) results in lower learning output, when compared with presentation of the material either orally or visually (van Mierlo, Jarodzka, Kirschner, & Kirschner, 2012). The integrated presentation may burden the cognitive capacity of a learner, hence leading to ineffective learning process. As such, redundant information was excluded from SW-PAL by only providing visual presentation of the instructional design of worked examples.

Third, multiple instances allow comparisons, wherein learners are able to identify and comprehend the variances between elements, thus distinguishing relevant information from that irrelevant. Besides, similarities also should be comprehended in deciding the relevant
aspects. As such, learners acquire the skills of abstracting from irrelevant features, which happens to be a criterion to induce schema (Atkinson, Derry, Renkl, & Wortham, 2000). Therefore, the tool of multiple worked examples was embedded in SW-PAL for every single text, as illustrated in Figure 4. As such, the learners can assess various worked examples, as summarized by the instructor to comprehend the application of the strategies in writing summaries.

Lastly, the third component, scaffolding (see Figure 5), refers to strategies identification feedback tool that was developed for students to check the strategies they employed in writing summaries. This tool was constructed by adopting Abdi et al.’s (2016) algorithm called "Identification Summarizing Strategies Linguistic Knowledge." The summarizing strategies embedded in the feedback tool were deletion, sentence combination, copy-verbatim, topic sentence selection, and paraphrasing. Determining the correct strategies for summary writing is essential. Hence, this feedback tool allows students to learn and practice summary writing in an independent manner, without adding to the excessive load of teachers.

5 | METHODS

5.1 | Subjects and research design

A total of 58 Malaysian undergraduate students, which composed of 15 males and 43 females, participated in this study. The subjects were from the Department of Computer Science at a local university located in Malaysia. They have learnt the ESL for more than 14 years, from primary school right up to the present university. These subjects were from two nonequivalent intact classes, wherein 28 students

FIGURE 4  Worked examples tool [Colour figure can be viewed at wileyonlinelibrary.com]
were selected for the control group and 30 students in the experimental group. The average age of the subjects in both the control and experimental groups was 21.

A nonequivalent pretest–posttest control group quasi-experimental design was carried out to analyse the impact of SW-PAL on the performance of ESL students in writing summaries. The study was conducted over a 5-week period for the two groups of students: 1 week for pretest, 3 weeks of practice session, and the last week for posttest. The control group was taught summary writing using the conventional approach, whereas the experimental group was exposed to SW-PAL. Two measures were employed in the analysis: pretest and posttest scores, in order to determine if the subjects had improved in their summary writing performance.

Assessing the cognitive load among the subjects involved students with varying levels of English language proficiency based on their Malaysia University English Test (MUET) results. MUET is a standardized exam that measures the level of English competency (Council, 2015) among undergraduate students in Malaysia. The MUET is composed of all four language skills of listening, speaking, reading, and writing. Based on the MUET description, the level of proficiency among the students can be classified based on bands, as depicted in Table 1. The scores are graded in six bands, with Band 6 reflecting the highest language proficiency.
proficiency and Band 1 as the lowest language proficiency. In this study, the subjects only scored Bands 3 and 4 in the MUET exam. As indicated in Table 1, students with Bands 3 and 4 in the MUET exam display modest and satisfactory levels of the English proficiency, respectively. The subjects in the experimental group with modest and satisfactory levels of the English language proficiency were 16 and 14, respectively, whereas the subjects in the control group with modest and satisfactory levels of the English language proficiency were 16 and 12, respectively.

5.2 | Materials

Flesch Reading Ease readability index was applied to select the sample text. Seven expository sample texts with Flesch Reading Ease readability index between 30 and 50, which were pitched at the undergraduate students’ level (Fielding, 2006), had been selected. One of the expository texts was used for pretest and posttest, whereas the other six were applied during the practice session. Two worked examples of the summary were employed for each text using various types of summarizing strategies, as prepared by an instructor using SW-PAL.

5.3 | Procedures

Every subject in the experimental group underwent a pretest, a training session with SW-PAL, six SW-PAL practice sessions, and a posttest. As for the control group, the subjects were exposed to a pretest, a conventional instruction session, six conventional summary writing practice sessions, and a posttest. Upon completion of posttest, the cognitive load psychometric test was administered to the subjects in both groups. The summaries gathered from both the pretest and the posttest were evaluated by two experienced ESL lecturers. After the cognitive load psychometric test, six voluntary students from the experimental group were selected for a follow-up focus-group interview, with a purpose to explore the participants’ perceptions pertaining to their learning experience using concept mapping, worked examples, and feedback techniques. Figure 6 shows the entire instructional and data collection process.

A 1-hour pretest was administered to subjects in both experimental and control groups in the first week. Prior to the pretest, the instructor explained to the subjects the meaning of summarizing a text. The pretest required the subjects to read a text and thereafter provide a summary of the text (see Appendix A).

After the pretest, to familiarize the students with the SW-PAL tool before starting the practice sections, the students in the experimental...
group went through a 90-min SW-PAL induction training session where they were taught about the summarizing process, including concept mapping, worked examples, and feedback. The instructor acted as a facilitator to train the students how to use the SW-PAL tool. However, as for the conventional instruction, the students in the control group were taught how to summarize by the instructor using the traditional paper and pencil-based method, adopted from McWhorter (2015). The strategies we employed to guide our students step by step in the conventional instruction are as follows:

1. Highlight or write brief notes on the material.
2. Write one sentence that addresses the writer’s overall concern or most important idea of the material. This sentence is the topic sentence of student summary.
3. Rewrite in their own words (paraphrasing).
4. Review the major supporting information.
5. Check the amount of detail included, if any, depending on the purpose of writing the summary.
6. Present ideas in the summary in the same order as they appeared in the original material.

Six practice sessions were held for both the experimental and control groups from Week 2 until Week 4. All the subjects from both groups practiced only one text per session, and each session took approximately 90 min. The texts prepared for both groups were the same. In the experimental group practice sessions, the subjects interacted with the SW-PAL by engaging themselves in activities such as drawing concept maps, understanding the instructor’s worked examples to acquire summarizing strategies and applied them when summarizing the text. Besides, the subjects were free to request strategies identification feedback from the tool to check the summarizing strategies applied. As for the control group, the conventional method was carried out, whereby the subjects were taught summary writing by an instructor during the practice sessions.

The posttest session was conducted in Week 5. All the subjects took the posttest following completion of the practice phase. All procedures for posttest for experimental and control groups were the same as those applied in the pretest. The text used was also same to that employed for the pretest. The rationale of using the same text for both premeasurement and postmeasurement was to assure objectively comparable texts, thus avoiding the problem of equating varied surface linguistic forms of the pre- and post-texts. The 4-week interval between the tests administration had been considered fairly adequate to control short-term memory effects. This is because the subjects were not provided with the correct summary or versions of the correct summary after the pretest; hence, they were clueless if their summary was correct or they were unlikely to remember how they had summarized the text initially. Hence, the interval was deemed sufficient to control significant learning, except for that due to the training (Lin, 2010).

The summaries collected from pretest and posttest were evaluated and graded by two experienced ESL lecturers using the Grading Rubric for summaries (Desoiza, 2011), consisting of five criteria, including main ideas, accuracy, words and style, organization, and length. The marking scheme consisted of four grading levels that ranged from exemplary (4 points), proficient (3 points), adequate (2 points), and needs to improve (1 point). When the score for each criterion was summed up, a cumulative summarization performance score (ranging between 5 and 20 points) was obtained. Prior to using the rubrics for scoring, the content validity of the instrument was determined by two experienced ESL lecturers who have taught summary writing to university students at a local university for at least a decade. The students’ pretest and posttest were graded by two raters. The interrater reliability scores of the pretest and posttest were established by using the Pearson’s product–moment correlation, which yielded 0.92 and 0.90, respectively.

After the posttest, the experimental and control groups were required to answer a cognitive load psychometric questionnaire, which was adopted from the instrument developed by Leppink and van den Heuvel (2015). The items in the instrument were measured by using a 10-point Likert scale that ranged between Strongly disagree and Strongly agree. Two types of cognitive load were tested: IL and EL. Items 1–3 in the questionnaire reflected IL, whereas Items 4–6 denoted EL. This test determined the efficacy of the types of cognitive load exerted by the subjects after learning summary writing via SW-PAL. Determination of cognitive load leads to identification of a learning method efficacy, which is bound to vary among subjects and teaching styles.

The content validity of the instrument was verified by two instructional design experts from the Faculty of Education at a local university with more than a decade of experience in the instructional design arena. The Cronbach’s α values for the reliability of the rating scale applied in IL and EL were 0.82 and 0.84, respectively, as presented in Table 2.

After the cognitive load test, a semistructured focus-group interview with six participants from experimental group using interview protocols developed by the researchers (see Appendix B) was conducted to further explore their overall learning experiences and their

### TABLE 2 Cognitive load psychometric instrument and reliability analysis

<table>
<thead>
<tr>
<th>No.</th>
<th>Item</th>
<th>Domain</th>
<th>Cronbach α</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The content of this activity was very complex.</td>
<td>Intrinsic load</td>
<td>0.82</td>
</tr>
<tr>
<td>2</td>
<td>The problem/s covered in this activity was/were very complex.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>In this activity, very complex terms were mentioned.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>The explanations and instructions in this activity were very unclear.</td>
<td>Extraneous load</td>
<td>0.84</td>
</tr>
<tr>
<td>5</td>
<td>The explanations and instructions in this activity were full of unclear language.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>The explanations and instructions in this activity were, in terms of learning, very ineffective.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

perceptions with regard to SW-PAL features, for instance, concept mapping, worked examples, and feedback. The interview questions asked the participants to (a) describe their learning experiences in using the SW-PAL, and (b) express if they liked using the features presented in SW-PAL as a learning tool, and make comments regarding the features in SW-PAL that could be useful and helpful in terms of summary writing, and lastly, (c) provide suggestions on future improvement of the SW-PAL. The researcher recorded the interview, transcribed it, and then analysed the interview contents as qualitative data of this study.

6 | RESULTS

6.1 | Normality of data

Normality test was performed prior to the analysis of pretest and posttest in order to select an appropriate statistical method. The skewness and kurtosis values of the pretest scores for the control group were 1.069 and 1.155, respectively, whereas the values of the posttest scores were 0.61 and 0.564, respectively. Next, the skewness and kurtosis values of the pretest scores for the experimental group were 0.028 and −0.866, respectively, whereas the values of the posttest scores in the experimental group were −0.903 and 0.312, respectively. All the above values ranged between −1.96 and +1.96, signifying that the data were normally distributed for all variables (Chua, 2009); hence, the data were indeed appropriate for analysis using parametric tests.

6.2 | Effectiveness of SW-PAL

Research question 1 looked into the effectiveness of SW-PAL in enhancing the performance of ESL students in summary writing. The effectiveness of SW-PAL had been determined by using the split-plot analysis of variances. The split-plot analysis of variance test assesses the effect of a treatment in an experimental study by comparing between-subject groups (control and experimental) and within-subject measurements (pretest and posttest) together in an analysis to obtain the interaction effect (or treatment effect). The significant outcome in the Multivariate Pillai’s Trace test displays significant treatment effect (Chua & Don, 2013).

The results tabulated in Table 3 show that the mean scores of pretest and posttest for the experimental group were 10.22 and 13.07, respectively, indicating an increase in the mean scores. Meanwhile, the mean scores of pretest and posttest for the control group were 9.79 and 11.04, respectively, which also indicated increment in the mean score. Nevertheless, the variance of mean scores between pretest and posttest scores for the experimental group (2.85) exceeded the variance of mean scores between pretest and posttest scores for the control group (1.25), hence significant variance, \( F(1, 56) = 40.45, p < 0.001 \), with a large effect size (\( \eta_p^2 = 0.42 \)). Therefore, the SW-PAL contributed 42% of the variance in the performance of the subjects in summary writing. In summary, the outcomes portrayed a positive impact of SW-PAL upon subjects' summary writing performance.

### Table 3

<table>
<thead>
<tr>
<th>Group</th>
<th>Within-subject</th>
<th>M (SD)</th>
<th>Mean difference</th>
<th>Multivariate Pillai’s Trace test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( df1, df2 ) ( F ) ( p ) Effect size (( \eta_p^2 ))</td>
</tr>
<tr>
<td>Control</td>
<td>Pretest</td>
<td>9.79 (1.49)</td>
<td>1.25</td>
<td>1, 56</td>
</tr>
<tr>
<td></td>
<td>Posttest</td>
<td>11.04 (1.41)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>Pretest</td>
<td>10.22 (2.12)</td>
<td>2.85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Posttest</td>
<td>13.07 (1.69)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Effect size (\( \eta_p^2 \)): small = 0.01; moderate = 0.06; large = 0.14 (Cohen, 1988).

*Significant at \( p < 0.001 \).
English language proficiency, displaying the efficacy of SW PAL on EL of subjects with modest and satisfactory levels of English language proficiency. The mean scores for EL for subjects with modest language proficiency were 3.38 and 4.70, respectively, indicating a significant variance of 1.32. Hence, it can be deduced that with the assistance of SW-PAL, the subjects with modest level of English language proficiency in the experimental group recorded less EL, when compared with those with satisfactory level of English language proficiency. The mean scores for EL for subjects with modest and satisfactory levels of English language proficiency were 3.38 and 4.70, respectively, indicating a significant variance of 1.32. Hence, it can be deduced that with the assistance of SW-PAL, the subjects with modest level of English language proficiency in the experimental group recorded less EL, when compared with those with satisfactory level of English language proficiency. As for the control group, no difference was noted for subjects with modest and satisfactory levels of English language proficiency at the level of significance $p < 0.05$.

6.4 Effect of using SW-PAL on EL of subjects with varied levels of English language proficiency

Research question 3 looked into the effect of using SW-PAL on EL of subjects with varied levels of English language proficiency after they had learnt summary writing using SW-PAL. The independent sample $t$-test outcomes, as presented in Table 5 ($t = -2.20$, $df = 28$, and $p < 0.05$), presented a significant variance in the mean scores for EL between subjects with modest and satisfactory levels of English language proficiency. The mean scores for EL for subjects with modest and satisfactory levels of English language proficiency were 3.38 and 4.70, respectively, indicating a significant variance of 1.32. Hence, it can be deduced that with the assistance of SW-PAL, the subjects with modest level of English language proficiency in the experimental group recorded less EL, when compared with those with satisfactory level of English language proficiency, displaying the efficacy of SW-PAL among subjects with lower level of English language proficiency. As for the control group, however, no variance was observed for subjects with modest and satisfactory levels of English language proficiency at the level of significance $p < 0.05$.

6.5 Students’ perceptions towards the use of concept mapping, worked examples, and feedback

Research question 4 concerns the students’ perceptions towards features installed in SW-PAL, such as concept mapping, worked examples, and feedback. The students’ overall summary writing learning experiences on these features were gathered via semistructured focus-group interviews (see Appendix B). Most of the students were motivated in using the tool and stated that the SW-PAL assisted them significantly with, and enhanced their summary writing performance. They also felt that concept mapping was beneficial in organizing concepts and thoughts in a logical manner prior to summarizing, apart from activating their prior knowledge when reading. Nonetheless, several students found the concept mapping features challenging or troublesome to use as they had never learnt them in school to produce a summary. All the students stated their preference towards using the features presented in worked examples to learn the summarizing strategies. A student responded that the use of worked examples increased his participation as he preferred learning summary writing using computers/tools, instead of learning it via conventional way in classrooms. A number of students responded by stating that using the worked examples tool did help them to learn multiple types of summarizing strategies, instead of the sole single strategy of paraphrasing commonly practiced at the university at present time. The students also stated that worked examples allowed them to practice repetitively before they were able to master the skillset of summarizing strategies. Lastly, all the students responded that feedback was indeed useful in helping them to check their summarizing strategies at the task. One constructive comment made by two students pertaining to improvement in the feedback function, “the feedback should be able to detect specific mistakes in writing the summary sentences, thus enabling them to improve their overall language proficiency.”

7 DISCUSSION AND CONCLUSION

This study has successfully designed and developed a theory-based CAL environment for summary writing (SW-PAL) that composed of concept mapping, worked examples, and feedback features. Additionally, it investigated the efficacy of SW-PAL in improving the performance of ESL students in summary writing, cognitive load imposed on the use of SW-PAL, and students’ perceptions towards the features embedded in SW-PAL.
In this study, an experiment was conducted to assess the SW-PAL by involving 58 ESL undergraduate students who were requested to summarize a text in pretest and posttest with SW-PAL (experimental group) and without the developed tool (control group). The study outcomes showed that learning summary writing via CAL environment appeared to be more effective, in comparison to the conventional method. In fact, the results appear to concur with those reported in prior investigations (see Chiu, 2015; Jeong, 2017; Wade-Stein & Kintsch, 2004; Wang et al., 2013), wherein CAL environment enhanced and facilitated language learning performance.

Generally, the study outcomes indicate that the synergistic operations of concept mapping, worked examples, and feedback in SW-PAL facilitated the performance exerted by ESL students in writing summary by first, activating their prior knowledge to enhance their comprehension towards source text via concept mapping, second, modelling summarizing strategies through worked examples, and finally, providing scaffolding in the form of feedback. The worked example instructional approach applied in SW-PAL displayed a positive effect and enhanced the students' skills in summary writing within the writing domain—an ill-structured domain. A significant variance was noted between IL and EL among the respondents with varying levels of English language proficiency after learning via SW-PAL. Those with lower English language proficiency (modest level) demonstrated higher IL and lower EL, compared with students with higher English language proficiency (satisfactory level). Higher English language proficiency students exhibited higher EL and were easily distracted, which made the learning process challenging. With increasing expertise, the worked examples may become ineffective and eventually redundant, hence resulting in expertise reversal effect (Kalyuga, Chandler, Tuovinen, & Sweller, 2001). Therefore, the worked examples instructional approach implemented in SW-PAL was more suitable for students with lower English language proficiency, in comparison with those with higher English language proficiency.

The outcomes of cognitive load are in agreement with the findings reported by Kyun et al. (2013), which signified that the worked example instructional approach is effective in ill-defined domain. Moreover, the subjects involved in the experiments had been similar, ESL. In fact, the only variance between this and prior studies appears to be the instruments used to test cognitive load, wherein Kyun et al. (2013) adopted Paas and Van Merriënboer's (1994) 9-point mental effort subjective rating questionnaire, whereas this present study tested IL and EL by applying Leppink and van den Heuvel's (2015) latest cognitive load psychometric instrument. Apart from that, the participants in Kyun et al.'s (2013) study were undergraduate students with high-level English language proficiency, whereas this present study selected undergraduate students with moderate-level of English language proficiency. Thus, the methodology applied in these two studies differed, thus the varying outcomes. For future study, students with higher level English language proficiency could be selected for both instruments to verify the variances.

Generally, the SW-PAL was favoured by the students, especially for the use of worked examples as a scaffolding technique in learning how to use the summarizing strategies to facilitate their summary writing, despite of some initial anxiety concerning the use of concept mapping as an advance organizer in activating their prior knowledge, as well as the desired function for the inclusion of specific feedback to enhance their overall language proficiency. The researcher's intent is to make conscientious efforts to perfect this online tool in the near future.

The present study has several limitations. First, the sample size of 58 participants seemed to be small to generalize the effectiveness of SW-PAL to other populations. Second, the participants were ESL learners majoring in computer science at a university. Their English proficiency level was considered average for freshmen amidst Malaysian universities. The study outcomes regarding the effects of SW-PAL can only be generalized to ESL populations with similar characteristics and cannot necessarily be generalized to other ESL populations, such as young ESL learners or college students with lower or higher level of English proficiency. Finally, the yields of this study might not be sufficiently valid or reliable due to the duration of the experiment, which was carried out for only 5 weeks, too short a period.

The findings of this study need to be interpreted with caution in terms of the prolonged engagement of the experimental group. Several avenues do exist for future research and enhancement of the SW-PAL tool. The author may take more effort in assessing the extent to which ESL students with varying levels of language proficiency benefit from SW-PAL in their summary writing task. The primary focus of this present study had been on the features of worked examples; thus, future studies may look into the effects of concept mapping and feedback components separately or in combination on the performance of ESL students in writing summary. Hence, further investigation is required to analyse the provision of intelligent feedback in SW-PAL in the attempt of enhancing summary writing performance.

ORCID

Chiou Sheng Chew https://orcid.org/0000-0002-6718-4663
Norisma Idris https://orcid.org/0000-0002-8006-7496
Andrew Thomas Bimba https://orcid.org/0000-0001-5682-5931

REFERENCES

in Saudi Arabia. IEEE Access, 6(c), 6783–6794. https://doi.org/10.1109/ACCESS.2018.2800164


APPENDIX A


Instruction: Please read a passage on the next page very carefully and then write a summary in English within 60 minutes. The word limit of the summary is 200 words.

Title: Research Method

A student at a school or university will be no stranger to research. Any projects, dissertations or theses will likely involve an effort to collect facts or to find out new information. Beyond academic institutions, there is still a heavy emphasis on research in many fields. Businesses may carry out research to develop or market products, to improve services or to keep ahead of their competitors. It is therefore important to understand the various methods of research and how research is applied.

Every part of research—identifying a subject to investigate, gathering information, analysing information—presents a researcher with a range of practical problems that need to be solved. In addition to the work they already have, researchers, then, need to be busy finding solutions to these problems. Fortunately, there is a body of work they can draw on. For hundreds of years, solutions have been developed and tested. The ideas behind research techniques have been evolving for as long as there have been thinkers, dating back to the philosophers of ancient Greece. The debate about knowledge and how it is acquired were central in the curious minds of Plato and Aristotle.

Research methods involve techniques used to collect information. These are the tools of the trade. In addition to collecting data, research methods allow us to sort and analyse the information. Once this is done, we can start drawing conclusions about the topic we are studying. Different research methods are used for different types of enquiry. When carrying out research, therefore, it is necessary to know the correct method to use and also how to use it to best effect. If we use the wrong tool for the job, our analysis may not be valid and it will be difficult to convince other people of our conclusions. The rigour with which the research is carried out will also be reflected in the quality of our results.

Research broadens our body of knowledge, and many kinds or research are employed to accomplish this. Descriptive research is dependent on observation. It examines situations and behaviours with the aim of establishing predictable patterns. What are we likely to see given the same conditions, the same circumstances? Explanatory research is designed to make sense of complex issues. The aim here is to go beyond description and start looking for cause and effect. Categorical research involves sorting different objects, events or concepts. The function of this method is to highlight similarities and
differences. This can be a useful way of explaining which elements belong together and why.

Further distinctions can be made between different types of research. Categorizing falls under comparative research. Comparative research involves comparing two or more contrasting cases in order to gain a better understanding. Correlative research, on the other hand, examines the relationships between two phenomena in an effort to see if they influence each other and, if they do, to what extent. There are varying levels of association. A relationship may be incidental, there may be a loose link between the two elements or, at the other end of the spectrum, and there may be a direct cause-and-effect relationship. Once correlations are established, predictions can be made to test the relationship. If past events suggest a strong relationship between two or more characteristics, then the relationship should be seen in similar circumstances in the future. Finally, once a situation or event is fully understood and the cause-and-effect relationship has been clearly worked out, we may be able to find ways to control it. By exerting control over the ingredients, we may be able to achieve a desired outcome.

Research projects, therefore, may be carried out simply for the sake of expanding the frontiers of knowledge. However, they may also have very useful real-world applications to improve the lives we live or the conditions we live in.

(Adopted from Research Methods—the Basics, Nicholas Walliman)

APPENDIX B

Semi-structured focus-group interview

1. Please describe your own learning experiences in terms of the differences/uniqueness by using the SW-PAL, in (i) time and efforts you spent, (ii) learning outcome (effectiveness), and (iii) your preference.
2. What do you think of the SW-PAL as a tool in helping you to improve your summary writing?
3. Are there any features in SW-PAL that can be further improved?