Zhai, X., Chu, X., Meng, N., Wang, M., Spector, M., Tsai, C.-C., & Liu, H. (2022). The Effect of Multi-mode Stimuli of Feedforward and Eye Tracking on Metacognition— An Exploratory Study Using Digital Dictionaries. *Educational Technology & Society*, *25* (1), 213-227.

The Effect of Multi-mode Stimuli of Feedforward and Eye Tracking on Metacognition— An Exploratory Study Using Digital Dictionaries

Xuesong Zhai[1], Xiaoyan Chu[1], Nanxi Meng[2], Minjuan Wang[3], Michael Spector[4], Chin-Chung Tsai[5] and Hui Liu[1*]

[1]College of Education, Zhejiang University, Hangzhou, Mainland China // [2]World languages, Literatures and Cultures, University of North Texas, Denton, US // [3]Learning Design and Technology, San Diego State University, San Diego, US // [4]College of Information, University of North Texas, Denton, US // [5]Program of Learning Sciences and Institute for Research Excellence in Learning Sciences, National Taiwan Normal University, Taipei, Taiwan // xszhai@zju.edu.cn // xiaoyan_chu@zju.edu.cn // nanxi.meng@unt.edu // mwang@sdsu.edu // Mike.Spector@unt.edu // tsaicc@ntnu.edu.tw // autumnhuihui@zju.edu.cn

[*]Corresponding author

ABSTRACT: Metacognition is regarded as a retrospective skill promoting learners' learning performance, deep thinking, and academic well-being. Stimulated Recall (SR) is regarded as a reliable approach to inspiring learners' metacognition in the classroom. However, the outbreak of COVID-19, causing widespread class suspension, may impair the effect of SR on cultivating distance learners' metacognition. The current study, employing multi-mode stimuli of learners' eye movements and feedforward, aimed to develop the effect of SR on activating learners' metacognition in remote settings. Forty-eight university graduates were recruited to participate in an eye-tracking experiment using digital dictionaries. Their feedforward and eye movements were collected as multi-mode stimuli. By reviewing the consistency and discrepancies between their feedforward and eye movements, participants were invited to conduct an SR interview, which stimulated them to retrospect on their prior cognitive behaviors. The results of the metacognition scale pre-post test showed that learners' metacognitive skills were significantly improved by the stimulated recall with multi-mode stimuli. The findings theoretically enrich the metacognition strategy in the Cognitive Theories of Multimedia Learning, and practically extend the implementation of stimulated recall in distance learning contexts.

Keywords: Metacognition, Multi-mode stimuli, Stimulated recall, Eye tracking, Digital dictionary

1. Introduction

The Cognitive Theories of Multimedia Learning (CTML) emphasize the importance of metacognition for multimedia learning outcomes (Moreno & Mayer, 2007), academic well-being (Nasirzade & Nargesian, 2019), and higher-order thinking skills (Parlan & Rahayu, 2021). However, the outbreak of COVID-19 has challenged the cultivation of students' metacognition (Chakma et al., 2021). During the widespread suspension of physical classes, although the rich multimedia materials are options supporting students to conduct remote learning, the overwhelming abundance of the learning materials may distract them from their prior learning goals (Zhang & Zou, 2021). In this situation, students may thus fail to efficiently retrospect prior learning purposes and behaviors, and their metacognition may be simultaneously impaired. In addition, teachers were generally impelled to expend extra efforts to adjust to various learning techniques working online in the pandemic period, so that the retrospective strategy of inspiring students' metacognition has been less investigated (Abdullah, 2020). The above challenges lead to the need for urgent solutions to the problem of inspiring students' metacognition with a reliable approach in the widespread remote learning context.

Metacognition occurs in the condition in which learners make critical judgements on their previous learning behaviors and cognition, by which a meaning retrospection is generated (Taub & Azevedo, 2019). However, a big challenge of cultivating learners' metacognition is that students are used to recollecting the knowledge they have learned rather than retrospecting their prior behaviors and cognition (Rivers, 2020). Stimulated Recall (SR) is regarded as a reliable approach to guiding learners to implement effective retrospection of their prior behaviors, and the selection of the adaptive stimuli is a crucial factor in the successful occurrence of metacognition (Mudrick et al., 2019). Eye-tracking technology has been explored to capture online learners' behaviors as stimuli, by which some specific surveys were conducted to understand online learners' behavior (Wang et al., 2019). Regrettably, this approach may fail to stimulate learners' metacognition, since students answered the survey questions based on what they had already been told about the eye movements, so they recollected the learning process according to their eye movements without retrospection (Horská et al., 2020).

Moreno and Mayer (2007) CTML and Dunlosky's (2005) levels-of-disruption hypothesis suggested that the monitoring of feedforward and disruptions in multimedia are influential cues for metacognition. Constructing a learning setting in which students are able to compare and contrast learning materials and strategies could benefit their metacognition (Rollwage et al., 2018). Feedforward enables the prediction of learning behavior from the future concerning the desired behavior which the subject is encouraged to adopt. The current study proposed to employ stimulated recall with multi-mode stimuli including participants' feedforward and eye movements in an experiment using digital dictionaries. Students were surveyed about their feedforward first, and then their eye-media interactions were captured by eye trackers. In the third stage, by reviewing the consistency and discrepancies of the captured eye movements and learners' own prior feedforward, participants were stimulated to retrospect on what they expected when using digital dictionaries and what their real behaviors were. To examine the effectiveness of the multi-mode stimuli on improving learners' metacognitive skills, a pre- and post-test of metacognition were conducted.

2. Research background

2.1. Metacognition in multimedia learning

According to the Cognitive Theories of Multimedia Learning, a well-designed digital learning environment can significantly improve learning outcomes and perceptions when aligned with learners' cognitive processes, including essential processing, extraneous processing, and generative processing (Mayer, 2014). Essential processing is the first stage where learners get preliminary notification and classification of the presented materials. Then, in the stage of extraneous processing, learners reorganize the current orders, forms, and layout of materials according to their individualized cognitive architecture. Finally, to achieve generative processing, learners need to connect the reorganized material to their feedforward, where their metacognitive skills are aroused. Associated with the neural mechanism, metacognitive function could be examined from the frontal cortex (Frith, 2012), while surveys are regarded as a feasible measurement of metacognition in educational research (Antonietti et al., 2015).

Metacognition, inspiring learners to be aware of their cognitive process, is a higher level of thinking capacity; it is also referred to as "the thinking about thinking" (Renkl et al., 2013). Dunlosky and Metcalfe (2008) defined metacognition as a mental activity of understanding and regulating the learning process, including learners' beliefs about learning, monitoring the state of their knowledge, and controlling their learning activities. CTML emphasized that metacognition significantly impacts problem-solving, reasoning, and academic success in multimedia learning contexts (Mayer, 2014). Some researchers have proposed that metacognitive skills can help learners regulate their learning in online contexts because of their awareness of cognitive processes, and the results found that the more metacognitive skills the learners possessed, the more knowledge could be investigated from multimedia presentation to meet their needs (Antonietti et al., 2015). However, while learning during the COVID-19 outbreak, students confront less supervised environments; it is thus an urgent requirement to inspire learners to use metacognitive mentoring and conduct metacognitive control.

According to the CTML, metacognition may occur when learners connect the multimedia to their prior knowledge in generative processing (Edwards, 2010). CTML encourages comparing and contrasting students' prior knowledge and the current learning behavior in the stage of generative processing in multimedia learning (Mayer, 2014). For example, some researchers have proposed that learners' metacognitive skills were promoted when their retrospection of initial learning goals was awakened (Meuwese et al., 2014). Even though learners' actual behavior may not be consistent with their prior cognition, it also benefits the cultivation of metacognitive skills through their retrospection (Dulamă & Ilovan, 2016). The possible explanation may be due to Dunlosky's (2005) levels-of-disruption hypothesis, which states that the discrepancies within the multimedia content would stimulate learners' comprehension and metacognition when they monitor disruptions and conflicts. Some researchers have employed conflict questions to explore learners' feedback on the understanding of prime numbers, and the results showed that the learners experiencing conflict generated more metacognitive abilities (Questienne et al., 2018). Moreover, the effectiveness of monitoring disruptions relies on the relationship of the multimedia which are displayed, such as the coherence of the verbal and pictorial presentations (Mayer, 2014). Thus, we proposed that learners' feedforward on the functions of digital dictionaries and their eye movements in these areas could be employed as multi-mode stimuli.

2.2. Eye-tracking and multi-mode stimuli

The eye-tracking technique, widely employed in digital learning, is able to capture users' eye movements when they interact with learning materials, by which learners' cognitive behaviors could be observed, examined, and explained (Zhai et al., 2018b). Based on the eye-mind theory hypothesis, the eye-tracking approach allows a dynamic trace of attention to be observed via eye movements (Cortina et al., 2015). Eye movements consist of three basic evaluative criteria: fixation counts, fixation duration, and scanning paths (Lai et al., 2013; Luo et al., 2017). Firstly, fixation count was defined as the concentrations counted in certain Area of Interests (AOIs). According to Rayner (2009), a fixation count lasts over 200 milliseconds. Fixation counts could be seen as a reliable tool to gauge the level of complexity, importance, and viewing. Secondly, fixation duration was defined as the sum of duration of eye movement within certain AOIs that is examined on the time scale. Some researchers pointed out that varied learning motivations and the materials' complexity may influence learners' fixation duration (Park et al., 2015). Typically, the integration of the fixation counts and duration are employed to reflect students' focusing on certain AOIs in the media. Thirdly, the scanning path, presenting the fixations' orders, reveals the holistic logical connection of components, which is adopted to gain access to visual memories in space (Lorigo et al., 2008). Additionally, the developed visualizing technique has facilitated presentation of the eye movements, and the heatmap and scanning figures were typically illustrated to explain learners' perceptual and cognitive process (Wang et al., 2016).

Although the eye-tracking technique is regarded as an adaptable approach to obtaining objective data, it has also been suggested to integrate it with qualitative methods to investigate the driving mechanisms of cognitive processes. Previous studies have found that eye-tracking alone may lead to biased results, and the combination of eye-tracking techniques and a survey could provide a comprehensive understanding of human behaviors (Leszkowicz, 2011). Eye-tracking only tells how learners interact with digital materials from the features of their eye movements, while qualitative approaches are able to explain why the interactions occur from the perspective of learners' perceptions. The combined method could help users rethink their prior learning behaviors. For example, the eye-tracking device provides the areas of interest, but why these areas are formed remains unknown. It may be attributed to various reasons such as learner interest, confusion, and so on, which requires further investigation to connect the eye movements to the specific reasons generated. Stark et al. (2018) applied the think-aloud approach to explore gaze patterns generated by eye-tracking, which supported the reliability of combining both methods to understand the deep cognitive processes. Although the eye tracking technique has generally been utilized in some small-scale experiments, with the development of deep learning in eye-tracking recognition, the PC camera could be used as an eye tracker, facilitating the popularization of the eye tracking technique in post-pandemic learning settings.

2.3. Stimulated recall in multimedia learning

Stimulated Recall has been extensively used to help learners to retrospect their learning behavior through the stimulus, such as recorded audios and videos captured in physical classrooms (Yuan & Lee, 2014). SR was developed based on the assumption that internal activities could be verbalized from the observed external real-world events. It has considerable potential to investigate studying cognitive strategies and learning processes (Geiger et al., 2016). Mackey and Gass (2016) also suggested that SR is an effective way to recognize learners' perceptions, their interpretation of events, and their thinking at a particular point. Although SR is widely used in physical contexts (Gazdag et al., 2019), it has been less explored and employed in remote learning settings, not to mention during the outbreak of a pandemic. The successful implementation of SR in online contexts may rely on the following two factors.

One factor is the stimulus captured from online learning behaviors, and the other is the retrospective strategy adapted to remote contexts. Although students' interactive behaviors could be recorded by video or audio, their interactions with multimedia are difficult to capture and analyze in online contexts. Thus, it is an urgent requirement to explore stimuli reflecting human-computer interaction, and further to motivate learners' retrospection in massive remote learning in the post-pandemic situation. Recent research has begun to explore the comprehensive understanding of biofeedback (e.g., eye tracking and EEG) in SR in the multimedia learning context (Zhai et al., 2018a). It is suggested that eye movements are reliable stimuli to SR and that retrospection is able to explain the biofeedback in return. Besides, stimulated recall affects the reliability of the retrospective strategy; for example, some designed questions have requested students to recall their cognitive behavior, by which learners' metacognition could be aroused (Abdel Latif, 2019). However, the outbreak of the COVID-19 pandemic may hinder the implementation of the retrospective strategy. Many instructors are struggling with increased workloads online and are experiencing elevated levels of anxiety and stress, and thus may neglect conducting retrospective instruction (Schmidt-Crawford et al., 2021). Likewise, during the suspension of classroom teaching, students perceive weak interactions between the digital content, which may vitiate their intentions to recall prior cognitive behaviors (Hamdan et al., 2021).

Synthesizing the above research background, this study aimed to address the following two research questions:

- Does the employment of multi-mode stimuli of feedforward and eye movements in stimulated recall improve learners' metacognition?
- How were the multi-mode stimuli compared and contrasted to inspire learners' metacognition in using digital dictionaries?

3. Methodology

3.1. Participants

The participants recruited in this study were native Mandarin speakers who were international graduate students in a university located in the southern United States. The participants were selected based on the following three criteria: (1) all the participants were familiar with the usage of digital dictionaries, so that they were well versed in their functions. (2) Participants should have adjusted to normal visual acuity to allow the eye-tracking software to properly calibrate. (3) Participants must be willing to perform immediate stimulated recall interviews. A total of 48 international students were finally recruited, and their demographics are shown in Table 1. To show our appreciation for their participation, gifts were sent to the participants after the experiment.

Table 1. Demographic profile of the participants

Categories		Frequency	Percentages
Age(years)	1=20-25	14	29.1
	2=26-31	16	33.3
	3=>31	18	37.5
Gender	1=Male	22	45.8
	2=Female	26	54.2
Degree Program	1=Master	29	60.4
	2=PhD	19	39.6
Major	1=Science	16	33.3
	2=Social science	19	39.6
	3=Art	13	27.1
English proficiency (TOEFL Scoring)	1=70-80	1	2
	2=80-90	39	81.3
	3=90-100	7	14.6
	4>100	1	2

3.2. Selection of digital dictionaries and vocabulary

The digital dictionary was an adaptive experimental platform for this study. Firstly, different from live broadcast platforms, digital dictionaries as auxiliary learning tools are generally used for online autonomous learning, and students' metacognitive activities are especially required in this situation (Connor et al., 2019). During the epidemic, learning activities were mostly carried out in a highly self-regulated learning context, with instruction and supervision by teachers lacking. The aim of employing a digital dictionary as a representative multimedia learning platform in this study was to inspire learners' metacognition in such self-regulated learning. Secondly, information and layout overload have been observed in many digital dictionaries (Frankenberg-Garcia, 2012) which distracts users from obtaining information efficiently and achieving their learning goals (Gouws & Tarp, 2016). When facing massive amounts of content and information provided by digital dictionaries, improving learners' judgment and awareness of valid information has become an urgent concern (Niitemaa & Pietilä, 2018). Thirdly, digital dictionaries have a broad user base for varied learning purposes, such as language learning and information searching (Levy & Steel, 2015; Lew & De Schryver, 2014). Given their distinctive learning motivations and varied using habits, many users expressed their desire for individualized services from digital dictionaries (Bastos & Machado, 2016). The above concerns led to our application of a digital dictionary in this study to verify the general applicability of the developed SR approach.

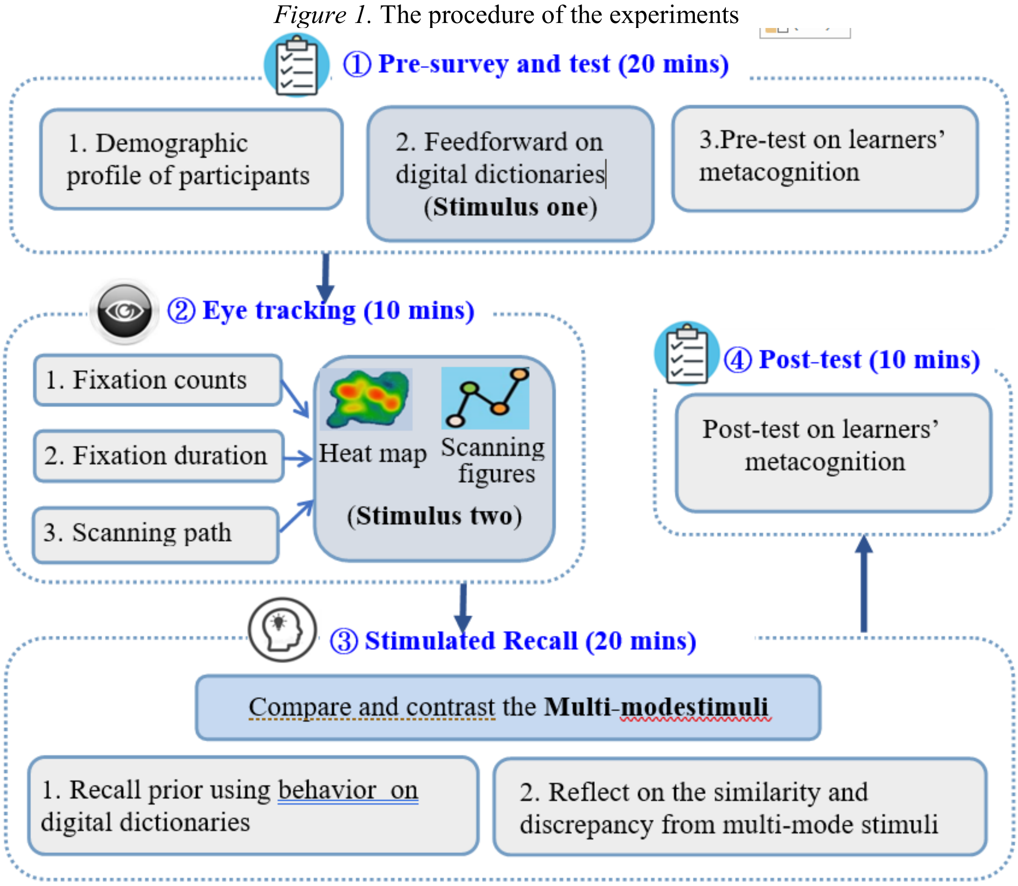
According to the criteria set by previous studies (Lew et al., 2013), the selected digital dictionaries should have similar functions and layout, including pronunciation, illustrations, definitions, phrases, synonyms/antonyms, and example sentences, so as to minimize the influence of functional distinction on users' perceptions during the experiment. Five digital dictionaries were selected according to the criteria mentioned for this study. Participants were surveyed to rank the dictionaries according to user experience. Finally, two of them were selected for this study: Youdao dictionary by Netease and Bing Dictionary by Microsoft, both of which were found to have a large user base.

Two criteria for selecting adaptive sample vocabulary in digital dictionary studies have been suggested: low frequency and polysemy (Dziemianko, 2015). By using low-frequency vocabulary, learners would focus on how to comprehensively utilize the functions of digital dictionaries to help them understand the vocabulary without being distracted by their prior knowledge of the vocabulary. According to the Corpus of Contemporary American English, low-frequency words are defined as those words that fall below the number 45,000 on a ranking of the most commonly used English words. Polysemous words were also suggested to be selected in digital dictionary studies, since polysemous words may maximize the functional display from the interface of digital dictionaries (Müller et al., 2015). According to the two criteria of vocabulary selection, two polysemous words, *tincture* and *sinew*, were selected by two professors specializing in linguistics. The researchers inputted the two words into the two digital dictionaries respectively, and took screenshots of the interfaces as source material for the eye tracking experiment.

3.3. Procedure and instruments

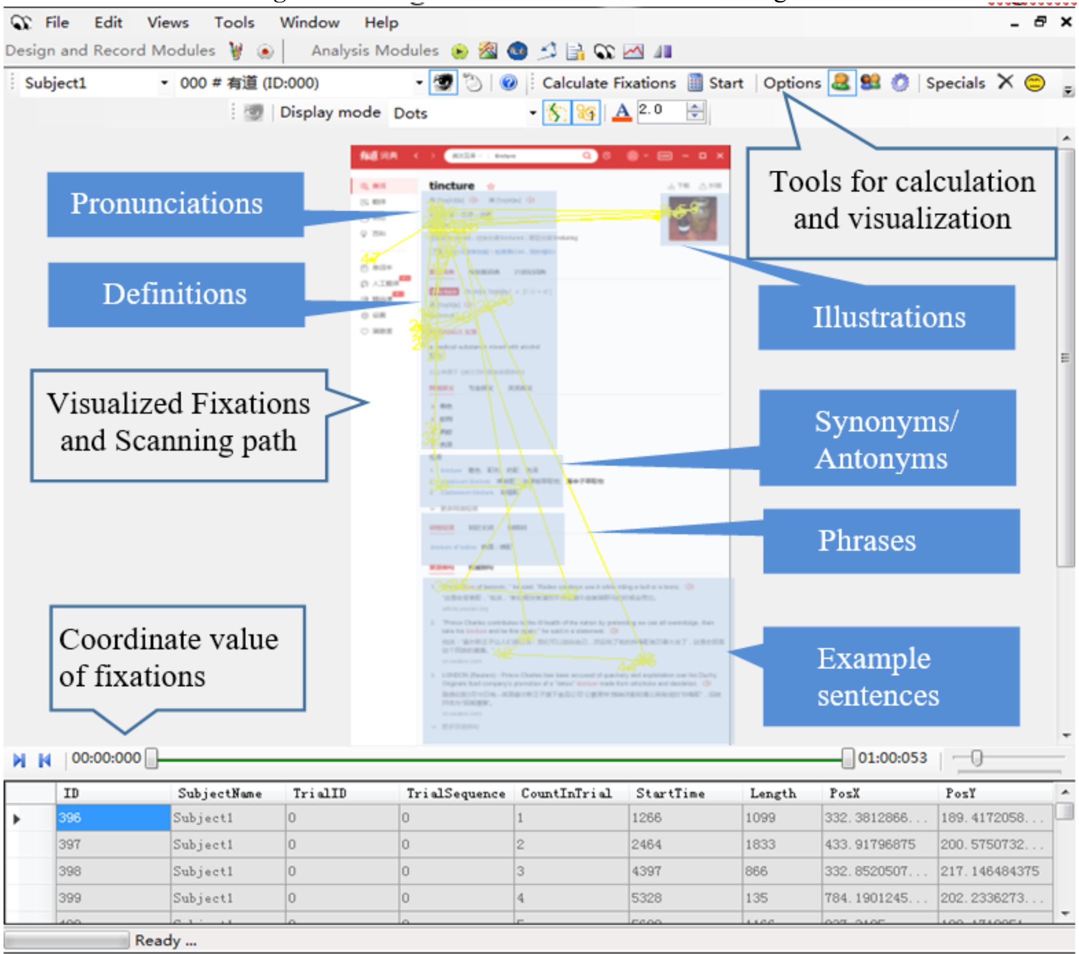
The procedure of the experiment shown in Figure 1 consists of four stages: a pre-survey and test, eye tracking, stimulated recall, and a post-test. Each participant spent around 60 minutes each time with help from an experienced teaching assistant. The experiment included a 20-minute pre-survey, 10-minute eye tracking, 20-minute stimulated recall, and a 10-minute post-test. In the first stage, a pre-survey was conducted to elicit participants' demographics, their feedforward on selected digital dictionaries, and a pre-test on metacognition. The pre-survey included demographics, feedforward, and the pre-test. According to the recommendations given in previous research on dictionaries (Collins, 2016; Frankenberg-Garcia, 2014), a 5-point scale (shown in Appendix 1), from *strongly disagree* to *strongly agree*, was employed to evaluate learners' feedforward on the perceived value of six typical functions in digital dictionaries (pronunciation, definitions, illustrations, phrases, synonyms/antonyms, and example sentences). Moreover, to examine students' metacognitive skills in the pre-post tests, a developed metacognitive scale, shown in Appendix 1, was adapted from Biasutti and Frate's research (2018).

Figure 1. The procedure of the experiments



The *Eye Tribe* eye-tracking device was employed in the second stage of the experiment. Having a reliable sampling rate from 30 Hz and 60 Hz mode, it is a reliable and adaptable tracker to capture learners' eye movements. Two open-access supporting software packages, Eyeproof and Ogama, were utilized to visualize the captured database displayed by heatmap and scanning paths (shown in Figure 2). The experiment was carried out in a laboratory with sound insulation. The participants were guided and acquainted with the equipment, procedures and the purpose of the eyetracking experiment, followed by signing the release form, granting permission to record their actions and comments. The experiment was conducted twice: once with the Youdao dictionary and once with the Bing dictionary. The Youdao dictionary was employed to display the interface of the selected vocabulary *tincture* the first time, and two days later, the Bing dictionary was used to present the interface of the other chosen vocabulary *sinew*. The experiments were conducted in one-by-one settings, since only one eye-tracker was utilized in this study. To minimize the interruption caused by the experiments, the researchers conducted the calibration of the eye movements by adjusting their head gesture, and helped participants get access to the test by a sample page, which could be completed in 5 minutes, so that the data collected in the first 5 minutes were discarded before participants got ready for the formal test.

Figure 2. The areas of interest and interface of Ogama



In the third stage, learners' feedforward and eye movements were presented to them as multi-mode stimuli in SR activities. On the basis of Cherrington and Loveridge's research (2014), the current research employed two-stepwise open questions, including recalling the original event and corrective feedback in this stage, by which learners not only reflected on how they used the dictionaries, but compared and contrasted the similarity and discrepancy between the feedforward and their displayed eye movements. Specifically, we asked (1) *What did you think according to your fixations and scanning paths;* (2) *Are there some conclusions by comparing eye movements and feedforward? What did you think of them?* To ensure the validity of recalling previous learning behaviors, previous studies have suggested that the SR interview should be conducted as soon as the experiment ends (Lyle, 2003). All the participants were interviewed by a teaching assistant approximately 20 minutes after the eye-tracking experiments were completed, and the SR interviews were recorded. After the SR, none of the participants reported any distraction due to the eye-tracker and SR interviews.

3.4. Data analysis

To test the effect of the SR with multi-mode stimuli on learners' metacognition, the current research employed the normal distribution and paired-samples *t* test by SPSS 19.0. Additionally, eye movements were evaluated by descriptive analysis and the Lag Sequential Analysis (LSA). The supporting software Ogama could generate the fixation counts and duration with adjustable criteria, and we set fixation at 200 ms in this study. Besides, to explore learners' scanning behaviors, the software GSEQ 5.1 was employed in this study to conduct the lag sequential analysis. According to the timed-event sequential data generated by Ogama, six events of the scanning paths, including *Pronunciation*, *Definitions*, *Illustrations*, *Phrases*, *Synonyms/Antonyms* and *Example Sentences*, were coded, which was followed by the implementation of the algorithms for computing inter-observer agreement.

4. Results

4.1. The pre-post tests of metacognition

In order to respond to the first research question, a *t* test was employed to measure learners' metacognitive skills according to the scoring of their pre-post tests. Firstly, the *t* value of the Kolmogorov-Smirnov (K-S) test shown in Table 2 scored 0.2 and 0.79 in the pre-test and post-test, which indicated that the data of metacognitive skills were normally distributed and qualified for the *t* test. Additionally, the reported mean value in the post-test was 4.14, while the result of the mean value in the pre-test was 3.34 ($p < .001$). The *t*-test results showed that the participants' metacognitive skills were significantly improved by the SR with multi-mode stimuli. The standard deviations are 0.26 and 0.24 in the pretest and posttest respectively, which indicated that the metacognition scoring was representative among the participants.

Table 2. The *t*-test results of learners' pre-post tests on metacognitive skill

	t-test results of metacognition					
	N	Mean	*S.D.*	K-S test	*t*-value	*p*
Pre test	48	3.34	0.26	0.20 (sig)	14.40	< .001
Post test	48	4.14	0.24	0.79 (sig)		

Table 3. The descriptive analysis of learners' feedforward and eye movements

	Feedforward				Eye-tracking on AOI	
	Perceived value (PV)		Expected Sum (ES)		Fixation counts (FC)	
	Mean	*S.D.*	Mean	*S.D.*	Mean	*S.D.*
1.Pronunciation	4.19	0.61	2.42	0.50	5.15	0.97
2.Definitions	4.69	0.55	2.92	0.61	4.90	0.90
3.Illustrations	4.60	0.57	2.73	0.71	5.50	1.07
4.Phrases	4.46	0.58	3.81	0.70	2.33	0.88
5.Synonyms/Antonyms	4.13	0.67	3.13	0.64	2.06	0.93
6.Example Sentences	4.52	0.58	4.25	0.76	15.45	5.38
(the first two)					14.33	5.36

Note. AOI refers to the area of interest.

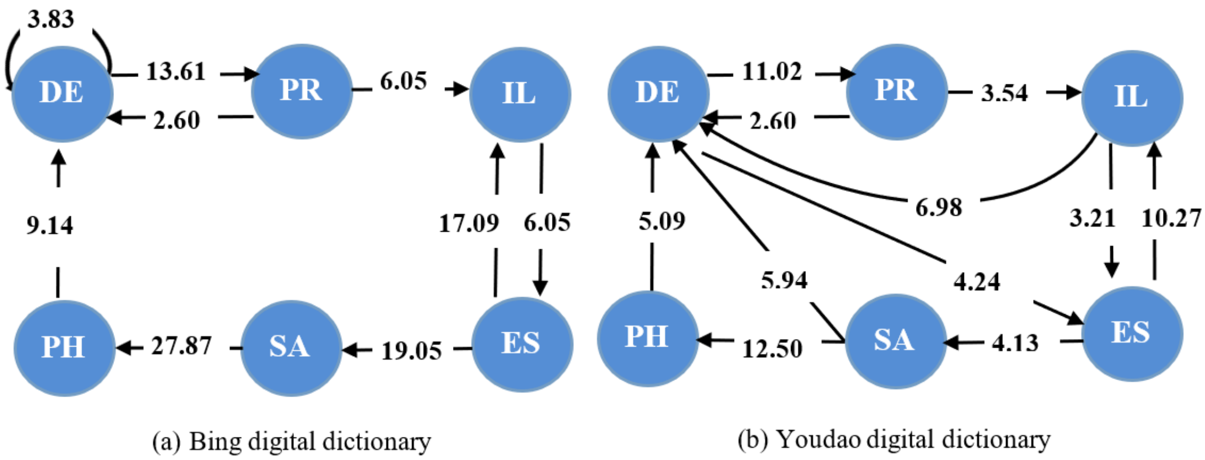
4.2. The comparison and contrast of feedforward and eye movements

In order to respond to the second research question, learners' eye movements and feedforward as multi-mode stimuli were compared and contrasted, by which learners retrospected why their eye movements were consistent or inconsistent with their prior knowledge, rather than merely recalling their prior behavior. Learners' average fixation counts (FC) and fixation duration (FD) of the two digital dictionaries were captured and generated from the supporting software Ogama. As shown in Table 3, the first three AOI on which fixation counts were mostly allocated were *Example Sentences* (FC = 15.45), *Illustrations* (FC = 5.5) and *Pronunciation* (FC = 5.15), followed by *Definitions* (FC = 4.9), *Phrases* (FC = 2.33) and *Synonyms/Antonyms* (FC = 2.06). According to the pre-survey of feedforward, the first three important functions in the digital dictionaries that the learners mostly emphasized were *Definition* (PV = 4.69), *Illustrations* (PV = 4.6), and *Example Sentences* (PV = 4.52), followed by *Phrases* (PV = 4.46), *Pronunciation* (PV = 4.19) and *Synonyms/Antonyms* (PV = 4.13). In terms of the expected sum, the first three expected functions in digital dictionaries were *Example Sentences* (ES = 4.25), *Phrases* (ES = 3.81) and *Synonyms/Antonyms* (ES = 3.13).

The above results show that learners' eye movements were partially in line with their feedforward. For example, according to the feedforward, the participants perceived a relatively higher value of using *Illustrations* and *Example Sentences* in the digital dictionaries, and their eye movements were found to be tallied with their feedforward. Likewise, *Phrases* and *Synonyms/Antonyms* were relatively less expected functionally, and the corresponding eye movements were less focused on these AOIs, which was consistent with their prior feedforward. However, as shown in Table 3 and Appendix 2, some discrepancies between learners' eye movements and their feedforward existed as well. For example, the sum of *Example Sentences* learners previously expected was 4.52 in their feedforward, while their concentrations were mainly allocated on the first two *Example Sentences* (14.33 out of 15.45 in fixation counts, and 12858.74 ms out of 13207.44 ms in fixation duration), which indicated that participants' cognitive load restricted their concentration on the rest of *Example Sentences*. Likewise, learners expected a relatively higher sum of phrases (*ES* = 3.81) and *Synonyms/ Antonyms* (*ES* = 3.13), whereas the fixation durations in these AOIs were both less than 900 ms, and the fixation counts were relatively less than that of other AOIs. (5) Besides, learners may have underestimated the value of *Pronunciation* (*PV* = 4.19, *ES* = 2.42), while relatively higher fixation counts and fixation durations were allocated in the AOI.

To further explain learners' eye movements when they used the selected digital dictionaries, Lag Sequential Analysis was employed. Six AOIs, *Pronunciation*, *Definitions*, *Illustrations*, *Phrases*, *Synonyms/Antonyms*, and *Example Sentences*, were coded in GSEQ 5.1. The eye movement data were generated from Ogama first, followed by the time sequence analysis. As shown in Figure 3, the arrows refer to the sequences learners performed from one event to another, and the coefficients predicted the correlations among the six events. The following observations in Figure 3 showed that: (1) The scan path started from *Definitions*, then moved to other events, and back to end on *Definitions*. The result indicated that the learning goal from the digital dictionaries was understanding the definitions of the selected vocabulary. (2) Their eyes moved between the *Definitions* and *Pronunciation* (z = 13.61 and 2.60 in Figure 3a. and z = 11.02 and 2.60 in Figure 3b), which suggested the assumption that users constantly tried to make connections between definitions and other multimedia information to help them understand the usage of the vocabulary. Besides, there was a set of iterative scanning behaviors between *Example Sentences* and *Illustrations* (z = 6.05 and 17.09 in Figure 3a. and z = 10.27 and 3.21 in Figure 3b), which shows that the learners tried to make illustrations connect to some specific social context.

Figure 3. Lag sequential analysis on learners' scanning behaviors



Note. DE, PR, IL, ES, SA, PH are the abbreviations of Definitions, Pronunciation, Illustrations, Example sentences, Synonyms/Antonyms, and Phrases respectively.

5. Findings and discussion

This study developed the stimulated recall approach with multi-mode stimuli to improve learners' metacognitive skills when they used digital dictionaries. Pre-post scale tests were conducted to answer the first research question, and the results showed a significant effect of multi-mode stimuli on participants' metacognitive skills. To further explain the driving mechanism, the similarities and discrepancies between feedforward and eye movements stimulated learners' metacognition as follows.

From the perspective of fixation, when learners recollected their eye movements, their cognitive behavior moved to the stage of extraneous processing and began to reorganize the material in the digital dictionaries.

Furthermore, when learners found that their eye movements were in line with their feedforward, they were encouraged to further confirm their assumptions with prior experiences, where their metacognition occurred in the stage of generative processing. For example, when students found that their fixation, consistent with feedforward, was focused on *Illustration*, they may have been involved in some real-world contexts. Two typical responses were generated from the SR interviews: (1) one is that students preferred the illustrations selected close to their previous experience and knowledge; (2) the other is that the illustrations should be located close to the definitions, which would help to spatially reduce the visual load. The findings are in line with prior research which found that learners were interested in connecting the information to their life experiences when using illustrations in digital language learning tools (Huang et al., 2012). Besides, although presented verbally, the Example Sentences provide a situational context for learners to picture activities from the sentences that help them understand the vocabulary (Huang et al., 2016). According to the SR interview, participants claimed the validity of picturing a specific contextual image in their minds, which enhanced their access to the application of the selected words. Students realized that they could activate their imagination of verbal material, which aroused their metacognitive skills of transforming media presentation. Likewise, synonyms and antonyms were relatively less noticed, which was inconsistent with their feedforward. Participants in the SR interview reported that although synonyms and antonyms of the selected words are useful to know, they were not involved in their initial learning goals. Participants realized that goal orientation is a keen factor in effective digital learning.

From the view of scanning path, definitions as an area of interest are focused. The lag sequential analysis on learners' scanning behaviors showed the significant sequential connection between definitions and some other multimedia elements, such as definitions and pronunciation. The dual coding assumption of CTML indicates that the dual channels, visual/pictorial and auditory/verbal, take effect simultaneously in the human-multimedia interaction system (Chen et al., 2017). There may be an interchange between the two channels in some situations, where users are capable of constructing their accordant psychological representations, which has been proven in the domain of vocabulary learning (Sadoski, 2005). Likewise, the significant sequential path between definitions and illustrations, shown in Figure 3(b), indicated that learners tried to recall their prior experience from the illustrations to gain access to the application of the selected words. The above retrospection may have aroused their metacognition on cross utilization of multiple elements in digital dictionaries to enhance their understanding of the unknown information. According to the SR interviews, participants concentrated on the definitions due to their confusion about which was the core and original meaning as well as the use frequency of the selected words, when polysemy was found.

Interestingly, some discrepancies between participants' feedforward and eye movements also existed, which helped learners activate their metacognition from conflict experiences (Questienne et al., 2018). Firstly, the data indicated that participants looked at the pronunciation section more often and for a longer period of time than they reported. According to the SR interviews, participants reflected on their desires for different voices, such as adult /child's voice, female/male's voice or young/older person's voice, to stimulate their auditory sense and help them remember the correct pronunciation of a word. This finding is consistent with the dual coding channels assumption. When physical representation and sensory representation are both shown, users' attention actively interacted between two modes in the multimedia learning context (Mayer, 2002). Participants' metacognition was generated by selecting their preferred pronunciation in order to meet their individualized learning requirements. Secondly, participants voiced their expectation of more example sentences. At the same time, focus was observed to be mainly on the first two example sentences, which may be due to the fact that some excessive information in the example sentences may have distracted their attention and increased their cognitive load. Some responses from the SR interview reflected that they might have overestimated their capacity, and their primary learning goal should be focused and split into specific sub-goals rather than remaining a desired but unachievable goal. Their psychological confusion may have generally occurred due to the conflicts between their feedforward and their eye movements, which awakened their metacognition that the resourceful multimedia in digital dictionaries may exceed their cognitive system's processing capacity.

6. Implications

6.1. Theoretical implications

CTML emphasizes the importance of metacognition for digital learning, and proposes finding a cuing factor as a stimulus to inspire learners' metacognitive skills (Mudrick et al., 2019), while less research has theoretically constructed a principle for it. The theoretical implication of this study contributes to exploring a principle to improve metacognition with multi-mode stimuli in digital learning settings. We recommend that multi-mode stimuli be designed and developed in SR to investigate learners' metacognition. To help learners recall their prior

behaviors, stimulated recall in previous studies employed either surveys or recorded materials such as photos, audios, and videos as stimuli. In this situation, learners most likely tend to recall the knowledge presented before, but may ignore retrospecting their previous cognitive behaviors and learning strategies.
Multi-mode stimuli are helpful for arousing learners' metacognition from two channels. For the first channel, the multi-mode stimuli mutually explain the driving mechanism of the behaviors. A single stimulus is weak in explaining the reason why behaviors happen. For example, eye tracking can tell where eyes linger but cannot explain why. The multi-mode stimuli could improve the deficiency of a single stimulus, which can enhance learners' deep understanding of the driving mechanism of learning behavior and strategies. The other channel is generating metacognition from conflict. Conflict occurs because individuals' cognitive capacity may consistently vary throughout their lives, and their expectations of cognitive capacity may be relatively hysteretic to their real behaviors. When discrepancy occurred among multi-mode stimuli, learners would instinctively retrospect the reasons caused in terms of the adaptation between prior cognitive mode and real learning behaviors.

The effect of COVID-19 has gradually impelled blend learning as an important learning channel. In the context of blend learning, the data sources are not only multivariate, but also multistage. Learners are encouraged to make comparison and contrast of these stimuli, by which their metacognition are expected to be inspired. First of all, with the assistance of information technology, the multi-mode learning behaviors and perceptions are accessible in the blend learning context. For example, the PC camera with the assistance of deep learning algorithm could be utilized as an eye tracker to capture learners' eye interaction with online material (SM et al., 2021), while their log data are also accessible from intelligent tutor system. Additionally, the longitudinal data instead of cross-section data are suggested to be employed as stimuli in this approach. Learners could much more effectively focus and reflect on a specific learning procedure between the stage of longitudinal stimuli collected, rather than recall their behaviors in general.

6.2. Practical implications

The current research has a series of practical implications for the modification and redesigning of digital dictionaries. Firstly, it is suggested that diversified pronunciation recordings, such as using a child's voice, be provided to meet learners' individual preferences. Current digital dictionaries typically only offer one or two pronunciation voices. It may be beneficial to provide a broad selection of pronunciation voices, such as the voices of male, female, older adult and young speakers to meet learners' personalized preferential treatment demands.

Second, the number of example sentences selected in digital dictionaries should be taken into consideration. According to eye-tracking data, it is suggested that cognitive overload is a central challenge in the design of digital dictionaries. Therefore, users should be able to toggle the maximum number of example sentences that appear for each entry according to their cognitive load. In addition, example sentences should be diversified to cover equivalent contexts and definitions. Even when a significant number of example sentences are presented, many of them are only for high-frequency definitions, ignoring the polysemy. The possible reason may be that example sentences were selected automatically according to search engines, which were the most frequently used but not always the most suitable. Therefore, the definition, illustration, and example should be consistent with each other.

Third, it is suggested that illustrations be spatially close to the corresponding definitions and example sentences. Based on the contiguity principle in CTML, learners' cognitive load is reduced when text and graphics are tightly spatially integrated, rather than presented separately. The illustrations selected should (1) have a relation with daily life, and (2) relate to the definitions and example sentences. Participants claimed that pictures related to their routine contributed more to their learning efficiency and drastically lowered their cognitive load. According to CTML, generative processing happens when learners actively integrate prior knowledge into working memory. Selective illustrations could facilitate learners' building of connections between words, definitions, example sentences, and pictures. They could successfully extract previous knowledge from long-term memory, then integrate the processed information with prior knowledge.

Finally, the digital dictionaries should help users to engage in the construction and modification of the dictionary content or user interface. Users showed strong interest in participating in providing feedback such as the selection criteria of the illustrations, and modifications of definition. Digital dictionaries, therefore, could be designed as an open-access or semi-open system for users' deep involvement. Digital dictionaries could not only be used as a tool for information searching, but could also help to construct a collaborative and creative learning context.

7. Limitations

Although a rigorous validation procedure was implemented to investigate learners' general using behavior and cognitive processes while using digital dictionaries, this research still suffers from some limitations. Firstly, only 48 participants from one university were recruited for this study, all of whom were native Mandarin speakers. To deepen our understanding of the individualized requirements of digital dictionaries, more consideration should be given to students' varied personalities, cultural backgrounds, language levels, technology self-efficacy, and so on. Secondly, the measurement of learners' metacognition was evaluated in general, and the results of each dimension of metacognition were not involved in this research. Thirdly, only digital dictionaries were selected as the research platform; there are, however, many other multimedia learning tools which could be involved in future studies. Finally, due to the limited function of the device and software, only the images of the final entries of digital dictionaries were included and studied, and so there will be further exploration of the synchronous influence of both visual and audio stimuli.

References

Abdel Latif, M. M. (2019). Using think-aloud protocols and interviews in investigating writers' composing processes: Combining concurrent and retrospective data. *International Journal of Research & Method in Education*, *42*(2), 111-123.

Abdullah, M. N. L. Y. (2020). The Influence of self-regulation processes on metacognition in a virtual learning environment. *Educational Studies*, *46*(1), 1-17.

Antonietti, A., Colombo, B., & Di Nuzzo, C. (2015). Metacognition in self-regulated multimedia learning: Integrating behavioural, psychophysiological and introspective measures. *Learning, Media and Technology*, *40*(2), 187-209.

Bastos, H. P., & Machado, G. P. (2016). Dictionaries on Smartphones: Learners' assessment of features and potential of dictionary apps as pedagogical tools. In *New Advances in Information Systems and Technologies* (pp. 143-156). Springer, Cham.

Biasutti, M., & Frate, S. (2018). Group metacognition in online collaborative learning: Validity and reliability of the group metacognition scale (GMS). *Educational Technology Research and Development, 66*(6), 1321-1338.

Chen, Z. H., Chen, S. Y., & Chien, C. H. (2017). Students' reactions to different levels of game scenarios: A Cognitive style approach. *Educational Technology & Society, 20*(4), 69-77.

Chakma, U., Li, B., & Kabuhung, G. (2021). Creating online metacognitive spaces: Graduate research writing during the COVID-19 pandemic. *Issues in Educational Research, 31*(1), 37-55.

Cherrington, S., & Loveridge, J. (2014). Using video to promote early childhood teachers' thinking and reflection. *Teaching and Teacher Education, 41*, 42-51.

Collins, J. B. (2016). Changes in electronic dictionary usage patterns in the age of free online dictionaries: Implications for vocabulary acquisition. *APU Journal of Language Research, 1*, 36-49.

Connor, C. M., Day, S. L., Zargar, E., Wood, T. S., Taylor, K. S., Jones, M. R., & Hwang, J. K. (2019). Building word knowledge, learning strategies, and metacognition with the Word-Knowledge e-Book. *Computers & Education, 128*, 284-311.

Cortina, K. S., Miller, K. F., McKenzie, R., & Epstein, A. (2015). Where low and high inference data converge: Validation of CLASS assessment of mathematics instruction using mobile eye tracking with expert and novice teachers. *International Journal of Science and Mathematics Education, 13*(2), 389-403.

Dulamă, M. E., & Ilovan, O. R. (2016). How powerful is feedforward in university education? A Case study in Romanian geography education on increasing learning efficiency. *Educational Sciences: Theory & Practice, 16*(3), 827-848.

Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. Sage Publications.

Dunlosky, J. (2005). Why does rereading improve metacomprehension accuracy? Evaluating the levels-of-disruption hypothesis for the rereading effect. *Discourse Processes, 40*(1), 37-55.

Dziemianko, A. (2015). Colours in online dictionaries: A Case of functional labels. *International Journal of Lexicography, 28*(1), 27-61.

Edwards, O. V. (2010). *The Effect of goal orientation of attention, learning, and metacognitive awareness* (Unpublished doctorial dissertations). University of Nevada, Las Vegas.

Frankenberg-Garcia, A. (2012). Learners' use of corpus examples. *International Journal of Lexicography, 25*(3), 273-296.

Frankenberg-Garcia, A. (2014). The Use of corpus examples for language comprehension and production. *ReCALL, 26*(2), 128-146.

Frith, C. D. (2012). The Role of metacognition in human social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1599), 2213-2223.

Gazdag, E., Nagy, K., & Sziták, J. (2019). "I Spy with My Little Eyes…" The use of video stimulated recall methodology in teacher training–The exploration of aims, goals and methodological characteristics of VSR methodology through systematic literature review. *International Journal of Educational Research, 95*, 60-75.

Geiger, V., Muir, T., & Lamb, J. (2016). Video-stimulated recall as a catalyst for teacher professional learning. *Journal of Mathematics Teacher Education, 19*(5), 457-475.

Gouws, R. H., & Tarp, S. (2016). Information overload and data overload in lexicography. *International Journal of Lexicography, 30*(4), 389-415.

Hamdan, K. M., Al-Bashaireh, A. M., Zahran, Z., Al-Daghestani, A., AL-Habashneh, S., & Shaheen, A. M. (2021). University students' interaction, Internet self-efficacy, self-regulation and satisfaction with online education during pandemic crises of COVID-19 (SARS-CoV-2). *International Journal of Educational Management, 35*(3), 713-725.

Horská, E., Nagyová, Ľ., Šedík, P., Kiełbasa, B., & Krasnodębski, A. (2020). Evaluation of cheese packaging graphics design using eye tracking and consumer survey research. *Journal of Management Cases*, *22*(2), 38-45.

Huang, C. S., Yang, S. J., Chiang, T. H., & Su, A. (2016). Effects of situated mobile learning approach on learning motivation and performance of EFL students. *Educational Technology & Society, 19*(1), 263-276.

Huang, Y.-M., Huang, Y.-M., Huang, S.-H., & Lin, Y.-T. (2012). A Ebiquitous English vocabulary learning system: Evidence of active/passive attitudes vs. usefulness/ease-of-use. *Computers & Education, 58*(1), 273-282.

Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., Lee, M.-H., Chiou, G.-L., Liang, J.-C., & Tsai, C.-C. (2013). A Review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review, 10*, 90-115.

Leszkowicz, M. (2011). "Keeping up with the Joneses." A Sociological content analysis of advertising catalogues with the eye-tracking method. *Qualitative Sociology Review, 7*(2), 85-100.

Levy, M., & Steel, C. (2015). Language learner perspectives on the functionality and use of electronic language dictionaries. *ReCALL, 27*(2), 177-196.

Lew, R., Grzelak, M., & Leszkowicz, M. (2013). How dictionary users choose senses in bilingual dictionary entries: An Eye-tracking study. *Lexikos, 23*(1), 228-254.

Lew, R., & De Schryver, G. M. (2014). Dictionary users in the digital revolution. *International Journal of Lexicography*, *27*(4), 341-359.

Lorigo, L., Haridasan, M., Brynjarsdóttir, H., Xia, L., Joachims, T., Gay, G., Granka, L., Pellacini, F., & Pan, B. (2008). Eye tracking and online search: Lessons learned and challenges ahead. *Journal of the American Society for Information Science and Technology, 59*(7), 1041-1052.

Luo, L., Kiewra, K. A., Peteranetz, M. S., & Flanigan, A. E. (2017). Using eye-tracking technology to understand how graphic organizers aid student learning. In *Eye-tracking Technology Applications in Educational Research* (pp. 220-238). IGI Global. https://doi.org/10.4018/978-1-5225-1005-5.ch011

Lyle, J. (2003). Stimulated recall: A Report on its use in naturalistic research. *British Educational Research Journal, 29*(6), 861-878.

Mackey, A., & Gass, S. M. (2016). *Stimulated recall methodology in applied linguistics and L2 research*. Routledge.

Mayer, R. E. (2002). Multimedia learning. In *Psychology of learning and motivation* (Vol. 41, pp. 85-139). Academic Press.

Mayer, R. E. (2014). Multimedia instruction. In *Handbook of Research on Educational Communications and Technology* (pp. 385-399). Springer.

Meuwese, J. D., van Loon, A. M., Lamme, V. A., & Fahrenfort, J. J. (2014). The Subjective experience of object recognition: Comparing metacognition for object detection and object categorization. *Attention, Perception, & Psychophysics, 76*(4), 1057-1068.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19*(3), 309-326.

Mudrick, N. V., Azevedo, R., & Taub, M. (2019). Integrating metacognitive judgments and eye movements using sequential pattern mining to understand processes underlying multimedia learning. *Computers in Human Behavior, 96*, 223-234.

Nasirzade, S., & Nargesian, J. (2019). The Effect of metacognitive beliefs on academic well-being mediated by perfectionist students. *Journal of School Psychology, 8*(3), 177-195.

Niitemaa, M.-L., & Pietilä, P. (2018). Vocabulary skills and online dictionaries: A Study on EFL learners' receptive vocabulary knowledge and success in searching electronic sources for information. *Journal of Language Teaching and Research, 9*(3), 453-462.

Park, B., Knörzer, L., Plass, J. L., & Brünken, R. (2015). Emotional design and positive emotions in multimedia learning: An Eyetracking study on the use of anthropomorphisms. *Computers & Education, 86*, 30-42.

Parlan, P., & Rahayu, S. (2021, March). Students' higher order thinking skills (HOTS) in metacognitive learning strategy. In *AIP Conference Proceedings* (Vol. 2330, No. 1, pp. 020035). AIP Publishing LLC. https://doi.org/10.1063/5.0043150

Questienne, L., Van Opstal, F., van Dijck, J. P., & Gevers, W. (2018). Metacognition and cognitive control: Behavioural adaptation requires conflict experience. *Quarterly Journal of Experimental Psychology, 71*(2), 411-423.

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457-1506.

Renkl, A., Berthold, K., Grosse, C. S., & Schwonke, R. (2013). Making better use of multiple representations: How fostering metacognition can help. In *International Handbook of Metacognition and Learning Technologies* (pp. 397-408). Springer. https://doi.org/10.1007/978-1-4419-5546-3_26

Rivers, M. L. (2020). Metacognition about practice testing: A Review of learners' beliefs, monitoring, and control of test-enhanced learning. *Educational Psychology Review, 33*(3), 823-862. https://doi.org/10.1007/s10648-020-09578-2

Rollwage, M., Dolan, R. J., & Fleming, S. M. (2018). Metacognitive failure as a feature of those holding radical beliefs. *Current Biology, 28*(24), 4014-4021.

Sadoski, M. (2005). A Dual coding view of vocabulary learning. *Reading & Writing Quarterly*, *21*(3), 221-238.

Schmidt-Crawford, D. A., Thompson, A. D., & Lindstrom, D. L. (2021). Condolences and congratulations: COVID-19 pressures on higher education faculty. *Journal of Digital Learning in Teacher Education*, *37*(2), 84-85.

SM, H. K., Pradyumna, G., Aishwarya, B., & Gayathri, C. (2021). Development of personal identification number authorization algorithm using real-time eye tracking & dynamic keypad generation. In *2021 6th International Conference for Convergence in Technology (I2CT)* (pp. 1-6). IEEE.

Stark, L., Brünken, R., & Park, B. (2018). Emotional text design in multimedia learning: A Mixed-methods study using eye tracking. *Computers & Education, 120*, 185-196.

Taub, M., & Azevedo, R. (2019). How does prior knowledge influence eye fixations and sequences of cognitive and metacognitive SRL processes during learning with an intelligent tutoring system*? International Journal of Artificial Intelligence in Education, 29*(1), 1-28.

Wang, C.-Y., Tsai, M.-J., & Tsai, C.-C. (2016). Multimedia recipe reading: Predicting learning outcomes and diagnosing cooking interest using eye-tracking measures. *Computers in Human Behavior, 62*, 9-18.

Wang, J., Antonenko, P., Celepkolu, M., Jimenez, Y., Fieldman, E., & Fieldman, A. (2019). Exploring relationships between eye tracking and traditional usability testing data. *International Journal of Human–Computer Interaction, 35*(6), 483-494.

Yuan, R., & Lee, I. (2014). Pre-service teachers' changing beliefs in the teaching practicum: Three cases in an EFL context. *System, 44,* 1-12.

Zhai, X., Fang, Q., Dong, Y., Wei, Z., Yuan, J., Cacciolatti, L., & Yang, Y. (2018a). The Effects of biofeedback-based stimulated recall on self-regulated online learning: A Gender and cognitive taxonomy perspective. *Journal of Computer Assisted Learning, 34*(6), 775-786.

Zhai, X., Meng, N., Yuan, J., Yang, Y., & Lin, L. (2018b). Exploring learners' cognitive behavior using e-dictionaries: An Eye-tracking approach. In *International Conference on Innovative Technologies and Learning* (pp. 165-171). Springer.

Zhang, R., & Zou, D. (2021). A State-of-the-art review of the modes and effectiveness of multimedia input for second and foreign language learning. *Computer Assisted Language Learning*, 1-27.

Appendix 1

The metacognition scale and the pre-survey on the perceptions of digital dictionaries

Items on metacognition	1	2	3	4	5
(1) When using digital dictionaries, I know my strengths as a learner.					
(2) When using digital dictionaries, I know how to select relevant information.					
(3) I know how to use the material in digital dictionaries.					
(4) I know how to organize new information in digital dictionaries.					
(5) When using digital dictionaries, I know how to connect new information with prior knowledge.					
(6) I can plan the activities when I use digital dictionaries.					
(7) When using digital dictionaries, I determine what the task requires.					
(8) When using digital dictionaries, I can select the appropriate functions.					
(9) When using digital dictionaries, I can identify the strategies depending on the task.					
(10) When using digital dictionaries, I organize my time depending on the task.					
(11) When using digital dictionaries, I modify my work according to other participants' suggestions.					
(12) I am used to asking questions to check my understanding when using digital dictionaries.					
(13) I check my approach to improve our outcomes when using digital dictionaries.					
(14) I improve my work with group processes when using digital dictionaries.					
(15) I detect and correct my errors when using digital dictionaries.					
(16) I make judgments on the difficulty of the task when using digital dictionaries.					
(17) I make judgments on the workload when using digital dictionaries.					
(18) I make judgments on the instruments when using digital dictionaries.					
(19) I make judgments on my learning outcomes when using digital dictionaries.					
(20) I make judgments on the teamwork process when using digital dictionaries.					

Items on feedforward in digital dictionaries	1	2	3	4	5
(1) I expect to use digital pronunciation to improve my learning when using digital dictionaries.					
(2) The explanatory definitions in the digital dictionaries are important to me.					
(3) The illustrations embodied in the digital dictionaries are helpful for learning.					
(4) I think the phrases are useful for learning when using digital dictionaries.					
(5) The synonyms/antonyms are important elements designed in digital dictionaries.					
(6) Example sentences benefit me a lot when I learn with digital dictionaries.					

	none	1	2-3	4-5	>5
(7) How many forms of the digital pronunciations do you suggest that digital dictionaries should offer?					
(8) How many explanatory definitions do you expect from digital dictionaries?					
(9) How many illustrations do you suggest should be presented in digital dictionaries?					
(10) How many phrases do you suggest should be provided by digital dictionaries?					
(11) How many synonyms/antonyms do you expect from digital dictionaries?					
(12) How many example sentences do you suggest that digital dictionaries should provide?					

Appendix 2

The descriptive analysis of learners' fixation duration on each AOI

	Fixation duration (ms) (FD)	
	Mean	*S.D.*
1.Pronunciation	4058.88	1206.08
2.Definitions	3687.67	1307.62
3.Illustrations	3721.90	1358.49
4.Phrases	865.79	433.90
5.Synonyms/Antonyms	688.48	382.33
6.Example Sentences	13207.44	5848.66
(the first two)	12858.74	5777.70

Note. The fixation duration was calculated in milliseconds. AOI refers to the area of interest.