Pan, W. F. (2017). The Effects of Using the Kinect Motion-sensing Interactive System to Enhance English Learning for Elementary Students. *Educational Technology & Society, 20* (2), 188–200.

The Effects of Using the Kinect Motion-sensing Interactive System to Enhance English Learning for Elementary Students

Wen Fu Pan
Department of Educational Administration and Management, National Dong-Hwa University, Taiwan // s1210@gms.ndhu.edu.tw

ABSTRACT

The objective of this study was to test whether the Kinect motion-sensing interactive system (KMIS) enhanced students' English vocabulary learning, while also comparing the system's effectiveness against a traditional computer-mouse interface. Both interfaces utilized an interactive game with a questioning strategy. One-hundred and twenty participants were chosen from an elementary school. The students were divided into three groups: Kinect, computer-mouse, and control. The participants' vocabularies were evaluated three times during a pre-test, a post-test, and a 1-month post-test. The following results were obtained: (1) there was a partially disordinal interaction relationship between the three groups and the three tests. Post-hoc comparison showed that the three tests have an order relationship. (2) The within group comparisons, both for the motion-sensing and computer-mouse groups which utilized an interactive game with a questioning strategy, displayed a relatively significant long-term retention. (3) In the between group comparison, the two interactive groups (computer-mouse and motion-sensing group) did not reach significant difference in English vocabulary learning. This means the motion-sensing interface of the KMIS was not a key-factor to affecting short-term or long-term learning retention. Therefore, our suggestion is that teachers can adopt interactive games with a questioning strategy to enhance students' long-term English vocabulary retention.

Keywords

Computer assisted learning, Human-computer interaction, Kinect sensor, Motion-sensing systems, Vocabulary learning

Introduction

Traditional one-way rote memorization method for learning English vocabulary is frequently found in schools (Smith, Li, Drobisz, Park, Kim, & Smith, 2013). However, the effects of such methods have not been found to be better than interaction (Ge, 2015). Vygotsky (1978) considered that in second language acquisition, learners need to interact with the socio-cultural environment via artifacts. These artifacts are referred to as "interfaces" between the subject and object from the viewpoint of human-computer interaction (Engeström, 2000).

Early human-computer interface (HCI) studies mostly adopted usability testing (Buur & Bødker, 2000). In the 1990s, some scholars began to cite activity theory, proposed by Leont'ev in the 1930s, as a theoretical framework of HCI design (Kaptelinin, 1996; Kuutti, 1996; Nardi, 1996). Activity theory emphasizes how to construct meaning from interaction between subject and object via artifacts (such as rules, books, etc.) (Leont'ev, 1974). Subsequently, activity theory also became one of the theoretical frameworks for language learning (Oxford, 1990). In 2003, Bedny and Karwowski divided activities into the following five levels: activity, task, action, operation, and function; they also incorporated two design types: subject-oriented and object-oriented, based on their proposed Systemic-Structural Theory of Activity (Bedny & Harris, 2005).

Subject-oriented design focuses on a subject's socio-cultural context and has often been adopted by studies of second language learning (Chapelle, 2009). On the other hand, in order to assess the usability of an emerging technology, researchers have often adopted object-oriented design (Munassar & Govardhan, 2011). This study also adopts a type of object-oriented design called "object-mental action" from activity theory. Specifically, a subject (the learner) interacts with an object (game-based animation) via the Kinect Motion-sensing Interactive System.

Edgar Dale's cone of experience theory indicates that two-way interactive learning helps learners to obtain up to 90% learning retention (Dale, 1969). Human-computer interaction also benefits learning retention (Papastergiou, 2009; Prensky, 2005); however, is this effect derived from the human-computer "interactive content," or "operating interface"? This question is worthy of further research. Therefore, in this study, we designed a game-based learning activity as the interactive "content" and a motion-sensing operation as the interactive "interface" for English vocabulary learning. The related literature is reviewed as follows.

Applying game-based learning with a questioning strategy as interactive content

Previous research has shown that game-based English learning has resulted in better retention than traditional rote memorization (Flores, 2015). Hwang, Chiu, and Chen (2015) also indicated that game-based learning is able to improve students' inquiry-based learning performance, especially in an interactive environment. Also, enjoying the game was cited as an important reason why students were willing to finish interactive tasks (Star, Chen, & Dede, 2015). The design of digital games is an important and often used method for enhancing learning motivation. A learner’s motivation to participate is enhanced through gamed-based learning (Birk, Atkins, Bowey, & Mandryk, 2016; Ronimus & Lyytinen, 2015). The goal of this study is to design an English vocabulary learning activity that integrates digitized game-based interaction.

In addition, a questioning strategy was implemented in this study to enhance the two-way interactive learning. The questioning strategy is defined as actively presenting a question and waiting for the students’ answer. Research has indicated that implementing a questioning strategy in English learning can also result in better retention (Basturkmen, 2001; Boyd & Rubin, 2006; Shomoossi, 2004; Yang, 2010).

Applying motion-sensing operation as an interactive interface

Developments in emerging technology have transformed the types of interaction between humans and computers. Past research (Chuang & Kuo, 2016; Hsiao & Chen, 2016; Sheehan & Katz, 2012) has shown that applying various motion-sensing technology to learning environments benefits students. Microsoft released the source code of Kinect (3D depth sensor) in 2012, and since then there has been a lot of development in motion-sensing applications. Currently all short-distance motion (or gesture) sensors use either infrared light emitters and sensors, ultrasonic sensors, or 3D depth sensors (Kinect) (Kumaragurubaran, 2011). The principle of the Kinect motion-sensing technique is that it employs three lenses, as well as a diffuser lens to expand or diffuse projected laser speckles. For the speckles that reach the human body, a separate camera coordinated with a light coding technique is employed to collect the 3D depth of field information regarding the human body within a 5-m tapered space (Pan, Chien, & Tu, 2012a). Pan et al. (2012a) compared the differences between infrared, ultrasonic, and Kinect sensing techniques. Applying Kinect in learning offers the following benefits: (1) it does not require a handheld controller; (2) it provides real-time feedback; (3) it is able to distinguish humans from objects; (4) it provides teachers (or developers) with a way to customize interactive content.

A lot of research has been conducted and is ongoing in applying Kinect technology to various fields (Nissimov, Goldberger, & Alchanatis, 2015; Yao, Wang, Cai, & Zhang, 2015). Kinect motion-sensors have been integrated into interactive learning, and research in this area has become a growing trend (Chuang & Kuo, 2016). For example, Sommool, Battulga, Shih, and Hwang (2013) applied Kinect motion-sensing technology to create and evaluate interactive learning classrooms; Tutwiler, Lin, and Chang (2013) applied it to multiple intelligence instruction; and Levinger, Zeina, Teshome, Skinner, Begg, and Abbott (2016) utilized Kinect in gait practice for knee replacement rehabilitation. In recent years, Pan led a team focused on the development of Kinect applications for educational situations (Pan, Tu, & Chien, 2014). Their research covered a range of applications including campus safety (Pan, Chien, Liu, & Chan, 2012b), accessible learning (Pan et al., 2012a), and interactive learning (Pan, Lin, & Wu, 2011). Pan (2013) indicated that the Kinect motion-sensing interface can better enhance students’ learning motivation compared to a more traditional computer-mouse interface. Some studies (Sommool et al., 2013; Vrellis, Moutsioulis, & Mikropoulos, 2014; Yuan, Hsieh, Chew, & Chen, 2015) have also supported the idea that the novelty of the Kinect system can attract students’ attention and increase learning motivation.

Concept framework of this study

Based on the above literature review, the concept framework of this study is shown in Figure 1.

A one-way interaction rote memorization method is frequently used in schools for learning of English vocabulary (Smith et al., 2013). In order to improve learning methods, two-way interactive learning and human-computer interfaces should be applied based on activity theory. The Kinect Motion-sensing Interactive System, or KMIS, designed in this study includes two parts: (1) applying game-based learning with a questioning strategy as interactive content; (2) applying the Kinect motion-sensing operation as an interactive interface. Past research (Pan, 2013; Sommool et al., 2013; Vrellis et al., 2014; Yuan et al., 2015) regarding Kinect motion-sensing applications in learning has usually adopted an empirical method, so an experimental design was used in this study. In addition, although some studies (Pan, 2013; Vrellis et al., 2014; Yuan et al., 2015) have also made the comparisons between Kinect motion-sensing and computer-mouse interface, these studies did not focus on the interactive relationship of the learning “content” and “interface,” or its effects on English learning. Therefore, through this experimental design, the researcher wants to assess the within-group (interactive content) and between-group (operating interface) interactive relationship, as well as compare their effects on short-term and long-term English vocabulary retention. It was expected that the findings from this KMIS design would show improvement over one-way rote memorization learning.

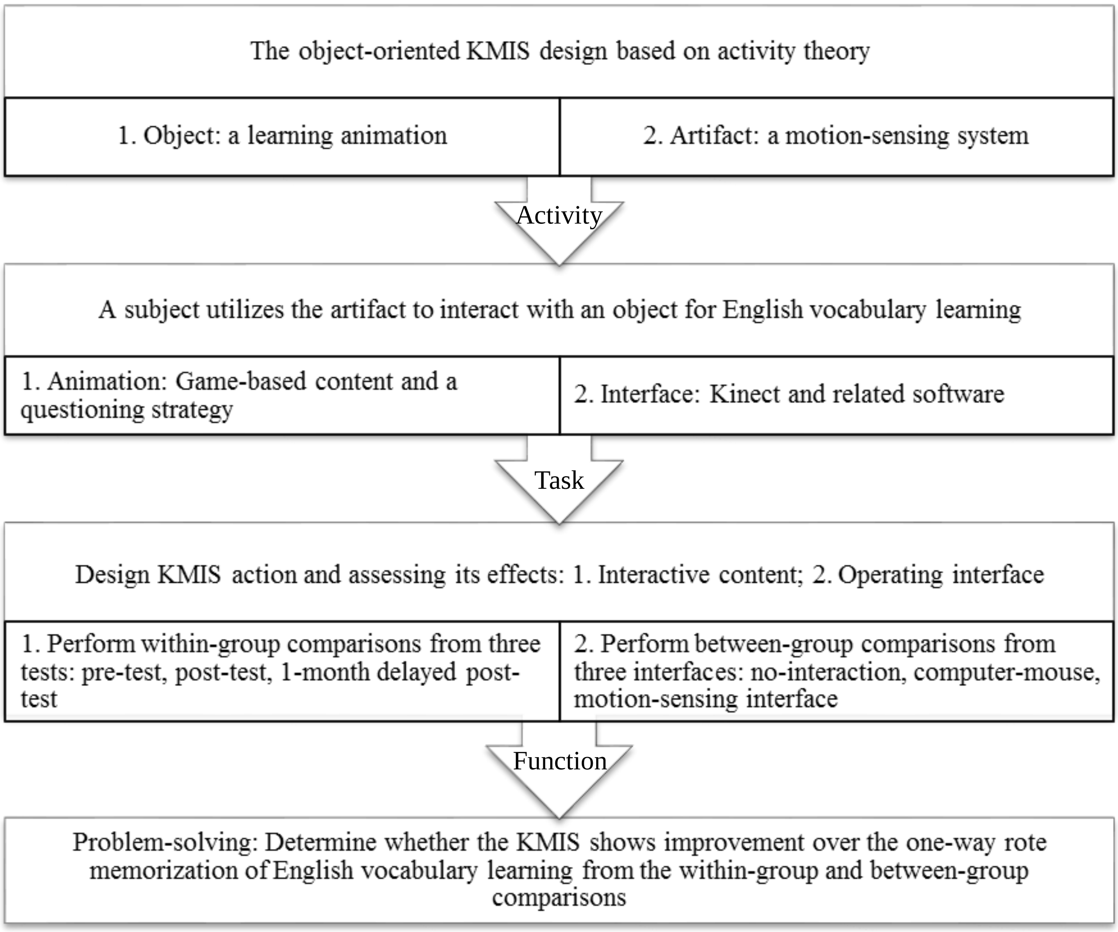

Figure 1. The concept framework of using the KMIS to improve English vocabulary learning

Objectives of the study

According to the above analysis, this study has the following three research purposes:

- To do statistical analyses on the interactive relationship between three groups (no-interaction, Kinect motion-sensing and computer-mouse) and three tests (pre-test, post-test and delayed post-test).
- To analyze the performances on the three tests (pre-test, post-test and delayed post-test) in interactive learning with an integrated questioning strategy game for English vocabulary learning.
- To compare the short-term and long-term retention effects on English vocabulary learning in the three groups.

Methods

Experimental design

The quasi-experimental design was adopted as shown in Table 1. The subjects of the experiment were divided into three groups: motion-sensing, computer-mouse, and the control group (no interaction). The motion-sensing group (X_1) was tested with the KMIS and the computer-mouse group (X_2) was tested with a traditional computer-mouse interface. Both the motion-sensing group (X_1) and computer-mouse group (X_2) had the same interactive content (a football game with a vocabulary quiz). Each group contained forty 6th grade students, who each completed the English vocabulary test three times: a pre-test (O_1), a post-test (O_2), and a 1-month delayed post-test (O_3). Types of interaction (three groups) and the three tests were the two factors in data analyses for a two-way mixed-design ANOVA. The two-way mixed-designed ANOVA was a better choice than analysis of covariance (ANCOVA) in this case because it was not only able to analyze the interactive relationships of three groups and three tests, but also able to compare the within-group and between-group differences.

Table 1. Two-way mixed-design structure (N = 120) with distinct types of interaction and tests

Experiment group	N	Pre-test	Treatment	Post-test	Delayed post-test
Control group	40	O_1	-	O_2	O_3
Computer-mouse group	40	O_1	X_1	O_2	O_3
Motion-sensing group	40	O_1	X_2	O_2	O_3

Note. O represents a 25 multiple-choice question test; the motion-sensing group (X_2) used the Kinect motion-sensing interactive system (KMIS) as the interactive interface

Research subjects

Six classes were randomly selected as research subjects with cluster sampling from 6th grade classes at a large-scale elementary school in Hualien City, Taiwan. Then two classes were randomly assigned to each of the three groups. Each group had 40 students averaging 12 years old, and the three groups had a total of 120 participants. Each group's participants had similar academic performance in school and the subjects of the three groups were examined by the homogeneity test. Statistically, the Box's test did not reach the level of significance (F = 1.522, p = .108) and the F-test for the pre-test did not reach the level of significance for difference (F = .22, p = .80, shown in table 6). Therefore, the basic background of subjects and environmental factors of the three groups were considered homogenous, and the two-way mixed-designed ANOVA was adopted. The experiment and test data were collected from October to December of 2013.

Development of the English vocabulary cognition test

The English Vocabulary Cognition Test (EVCT) for 6th grade students was created for evaluating students' learning performance. To establish the content of the test, forty words were randomly sampled from the "1200 English Vocabulary Words" endorsed by the Ministry of Education in Taiwan for elementary and junior high school students. These words were drafted into 40 multiple-choice questions for a test for 6th grade students in another elementary school in Hualien City, Taiwan. The 169 valid tests were ordered by scores, and the top and bottom 27% of the tests were grouped into high-score and low-score for item analysis of the questions. Under the significant threshold of p < = .013, twenty-five of the best (excellent) questions were selected for the formal test. Table 2 shows the Item Number, Initial Item Number in the pilot test, Item Difficulty (P), Item Discrimination (D), Critical Ration (CR), and Significance (p). The P-values appeared in between .272 and .554 (P-values = (Percentage correct $_{\text{High-score-group}}$ + Percentage correct $_{\text{Low-score-group}}$) / 2; the excellent questions' P-values ideally should be between .2 and .8), and the D-values fell between .239 and .652 (D-values = P_H -P_L; all questions' D-values ideally should be at least .2 or more). The mean P-values of the 25 questions was .407 (for excellent questions the mean of P-value ideally should be close to .5), the mean D-values was .403 (for excellent questions the mean of D-values ideally should be close to 1.0), and the mean CR-values was 4.455 (all questions reached statistical significance, p < .05).

Table 2. Item analysis of the elementary school 6th grade English vocabulary test

Item no.	Item no. in pilot test	Difficulty P	Discrimination D	Critical ration CR	Significance p
1	2	.478	.391	4.038	.000**
2	3	.380	.457	5.054	.000**
3	5	.380	.587	7.199	.000**
4	6	.402	.543	6.316	.000**
5	7	.500	.565	6.500	.000**
6	8	.467	.413	4.314	.000**
7	10	.500	.652	8.162	.000**
8	12	.522	.478	5.173	.000**
9	13	.359	.283	2.925	.004**

10	15	.489	.587	6.880	. 000**
11	16	.380	.370	3.905	. 000**
12	17	.457	.348	3.535	.001**
13	18	.413	.261	2.606	.011**
14	19	.326	.304	3.256	.002**
15	20	.554	.457	4.905	.000**
16	21	.293	.283	3.097	.003**
17	24	.304	.435	5.085	.000**
18	25	.315	.239	2.527	.013*
19	26	.272	.239	2.647	.010**
20	30	.435	.435	4.629	.000**
21	32	.424	.370	3.825	.000**
22	35	.467	.283	2.801	.006**
23	38	.478	.522	5.809	.000**
24	39	.293	.283	3.097	.003**
25	40	.293	.283	3.097	.003**
Mean:		.407	.403	4.455	.002**

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

After item analysis, the final draft of the EVCT contained 25 multiple-choice questions composed of 10 word meaning questions, 9 word form questions, and 6 word usage questions. Each question provided 4 choices, with only one correct answer. Each correct answer was scored as 4 points for a full score of 100. The EVCT was used for the pre-test, post-test, and 1-month delayed post-test. It was also used for the content of the interactive learning game for both the Kinect and computer-mouse interface.

Experimental process

The experiment was performed from October to December 2013. The treatments of the three groups for their review activity phase are outlined in Table 3. The pre-test, post-test, and delayed post-test were arranged on October 17, November 7, and December 12 of 2013. Each group received identical content both for their EVCT and review activities. The control group merely viewed the test paper with the correct answers for their review activity while the motion-sensing and computer-mouse groups reviewed using the interactive football game on their respective interfaces. The only difference between the latter two groups was the interface, point and click vs. a physical kicking motion.

Table 3. Experimental process of English vocabulary learning for three groups

Groups	Pre-test	Review activities	Post-test	Delayed
	20 min	40 min	20 min	20 min
A_1. Control group	☑	The volunteers viewed and practiced the test paper with answers.	☑	☑
A_2. Computer-mouse group	☑	The volunteers operated the game with computer mice, and the bystanders participated in watching and answering the vocabulary questions.	☑	☑
A_3. Motion-sensing group	☑	The volunteers played the game with kicking motions, and the bystanders participated in watching and answering the vocabulary questions.	☑	☑

Interactive system-design for the motion-sensing group

The subjects were divided into the following three groups for the experiment: the control group, which did not have an interactive review activity; the computer-mouse group (X_1), which utilized a traditional point and click computer interaction; and the motion-sensing group (X_2), which utilized the Kinect motion-sensing interactive system (KMIS). The structure of the KMIS is described below.

The KMIS is a system composed of both software and hardware. Figure 2 shows the Kinect Software Development Kit (SDK) from Microsoft and the Kinect Flexible Action and Articulated Skeleton Toolkit

(FAAST) offered by Suma, Krum, Lange, Koenig, Rizzo, and Bolas (2013) integrated with Windows 7, a projector, and Kinect hardware. When the participant stands at a proper distance from the Kinect sensor (about 1.5m-3.5m), the Kinect would convert the kicking motion into a computer command through FAAST and become a selection operation. The four footballs on the screen represent the four options for the multiple-choice questions. The researcher defined four kicking motions to respond to the four options: (1) the left foot kicking towards the left for the first ball (Option 1), (2) the left foot kicking forward for the second ball (Option 2), (3) the right foot kicking forward for the third ball (Option 3), and (4) the right foot kicking towards the right for the fourth ball (Option 4).

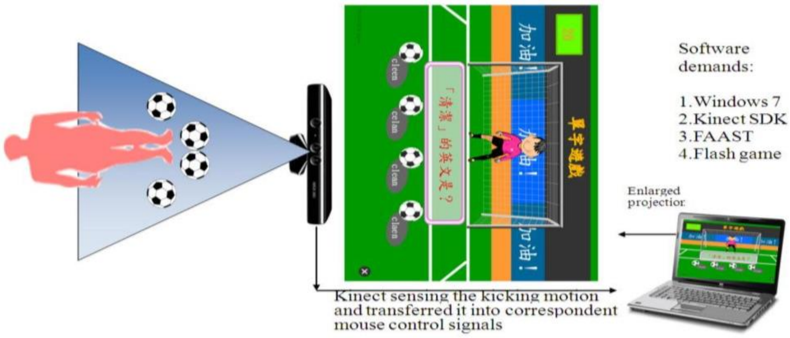

Figure 2. The KMIS interactive system design for motion-sensing group

Integrated design of game-based interactive learning and a questioning strategy

For the purposes of this research, the definition of "English vocabulary learning experience" is that learners read the vocabulary question on the review paper or screen, and then passively or actively find the correct answer from the review paper or game screen. The effects of their English vocabulary learning experience were evaluated by the three EVCT tests. The definition of "the questioning strategy" is that the questions of the English vocabulary game are actively shown on a projector screen, the learner's answer is waited for, and finally a response is given by the learner (shown in Figure 3). The questioning strategy was only utilized with the computer-mouse group and motion-sensing group.

The computer-mouse group and the motion-sensing group respectively represented traditional and novel human-computer interaction. Both utilized the football game and questioning strategy (Figure 3). In contrast, the control group only reviewed the vocabulary test paper. The football game consisted of 25 multiple-choice questions where the participants gained points by selecting the correct answers. Whether or not the correct answer was selected, the screen would still display the correct answer as feedback. This feedback was an integral part of the questioning strategy which separated the subjects in the interaction groups (computer-mouse group and motion-sensing group) with the subjects in the control group, who only reviewed the paper tests with correct answers.

Figure 3. Football game and questioning strategy integrated design for mouse and motion-sensing group

The Flash football game used in this study was free software downloaded from the internet, at http://goo.gl/dFXLKw. The software was revised by K. H. Yen, a teacher at Li Xing Primary School (Yen, 2012). It allows users (teachers) to input customized questions and implement the questioning strategy to design learning activities with its football game.

Data analysis method

A two-way mixed-design ANOVA was utilized for testing the relationship between the three types of interaction (control or no-interaction, computer-mouse, motion-sensing) and the performance on the three tests (pre-test, post-test, delayed post-test). A homogeneity test for variance was performed before the analysis, and the data analyses were presented with a descriptive statistics summary, a two-way mixed-design ANOVA, a test of simple main effects, and the LSD Method.

Results

The homogeneity test for variance was performed before the two-way mixed-design ANOVA, and the result did not reach the level of significance (Box's M = 18.972, F = 1.522, p = .108). This shows that the variance of the test scores was homogenous and that we could proceed with the successive statistical analyses. The results of the types of interaction and of the three tests, including a descriptive statistics summary, two-way mixed-design ANOVA, and a test of simple main effects, are shown in Table 4 to Table 7.

The findings regarding descriptive statistics of the types of interaction (three groups) and the three tests

Table 4 presents the population distribution, mean, and standard deviation of the types of interaction and the three tests. The overall mean of the three tests was also calculated for the post-test (M = 66.40, SD = 22.62), delayed post-test (M = 63.00, SD = 21.94), and pre-test (M = 55.33, SD = 22.06). Overall, the performance of all three groups improved, where the control group received the highest average score in the pre-test (M = 57.20), while the motion-sensing group received the highest average scores in the post-test (M = 71.30) and the delayed post-test (M = 65.60). The findings seem to reveal some variation within-group (between the three tests) or between-group. This result needs to be further tested with a mixed-design two-factor ANOVA.

Table 4. Statistics summary of types of interaction (A) and three tests (B)

B. Three tests	A. Types of interaction	*M*	*SD*	*N*
$B_{1.}$ Pre-test	$A_{1.}$ Control	57.20	22.16	40
	$A_{2.}$ Computer-mouse	54.70	22.65	40
	$A_{3.}$ Motion-sensing	54.10	21.82	40
	Total	55.33	22.06	120
$B_{2.}$ Post-test	$A_{1.}$ Control	60.70	22.22	40
	$A_{2.}$ Computer-mouse	67.20	22.52	40
	$A_{3.}$ Motion-sensing	71.30	22.41	40
	Total	66.40	22.62	120
$B_{3.}$ Delayed post-test	$A_{1.}$ Control	59.70	21.80	40
	A2. Computer-mouse	63.70	21.97	40
	A3. Motion-sensing	65.60	22.17	40
	Total	63.00	21.94	120

The findings of mixed-design two-factor ANOVA for the types of interaction and the three tests

The two-factor analysis of types of interaction and the three tests is shown in Table 5. There was interaction (F = 14.98, p < .01, η^2 = .204) between the two factors (A×B), and the three tests (B) also showed significant variation (F = 114.62, p < .01, η^2 = .495). A partially disordinal interaction relationship is shown in Figure 4, where the students' three tests before the interaction were initially ranked control group, computer-mouse group, and motion-sensing group, but the ranking was reversed in the post-test and delayed post-test. This change presents a significant interaction worth further analyses in the simple main effect test.

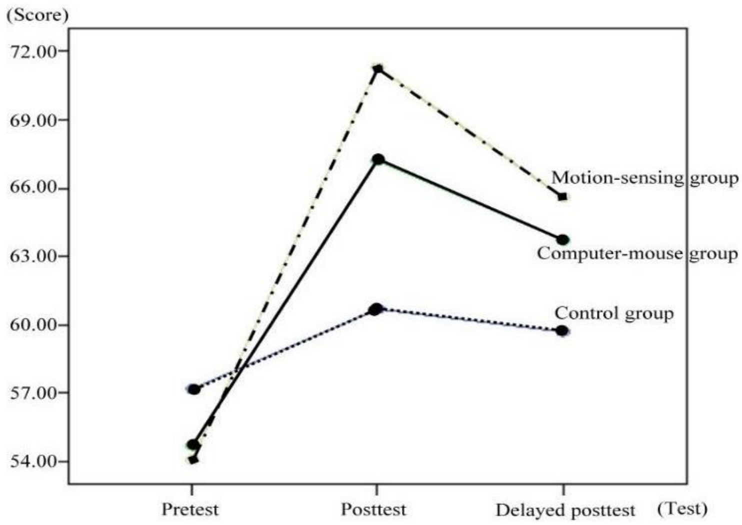

Figure 4. Statistical interaction between types of interaction and three tests

Table 5. Two-way mixed-design ANOVA of types of interaction (A) and test (B)

Variation source	*SS*	*df*	*MS*	*F*	*p*	η²
Types of interaction A	1212.09	2	606.04	.43	.652	.007
Test performance B	7712.36	1.52	5066.41	114.62**	.000	.495
Interaction A×B	2015.38	3.05	661.97	14.98**	.000	.204
Error						
Error between groups	164995.73	117	1410.22			
Residual	7872.27	178.10	44.20			
Total	183807.82	301.67	609.301			

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

The findings of the simple main effect for the types of interaction and the three tests

The simple main effect is shown in Table 6. The three types of interaction (A) did not reach a significant level of difference (F_{B1} = .22; F_{B2} = 2.28; F_{B3} = .75) in the detailed items ($B_{1.\text{pre-test}}$, $B_{2.\text{post-test}}$, and $B_{3.\text{delayed post-test}}$) of the three tests (B), revealing no between-group difference among the three types. In other words, the motion-sensing group was not superior to the computer-mouse and control groups in students' vocabulary learning. The three tests (B), on the other hand, achieved the significant difference (F_{A1} = 3.87, p < .05, η² = .090; F_{A2} = 65.00, p < .01, η² = .625; F_{A3} = 73.55, p < .01, η² = .653) in the detailed items ($A1_{.\text{control}}$, $A2_{.\text{computer-mouse}}$, and $A3_{.\text{motion-sensing}}$) of the types of interaction (A), presenting the necessity of post-hoc comparison of the three tests (B).

Table 6. Simple main effect analyses of types of interaction and three tests

Variation source	*SS*	*df*	*MS*	*F*	*p*	η²
Types of interaction(A)						
$B_{1.}$ Pre-test	216.27	2	108.13	.22	.80	.004
$B_{2.}$ Post-test	2285.60	2	1142.80	2.28	.11	.038
$B_{3.}$ Delayed Post-test	725.60	2	362.80	.75	.47	.013
Test performance(B)						
$A_{1.}$ No-interaction	260.00	1.77	146.87	3.87*	.03	.090
$A_{2.}$ Mouse	3326.67	1.49	2225.94	65.00**	.00	.625
$A_{3.}$ Body-sensing	6141.07	1.25	4912.05	73.55**	.00	.653
Error	172868.00	295.10	585.80			

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

The findings of post-hoc comparisons for the types of interactive and the three tests

The post-hoc comparison (LSD Method) findings (Table 7) showed that the within-group (three tests) comparison has an order relationship ($B_2 > B_1$, p <.05). It revealed that the post-test performance was superior to the pre-test regardless of the group. The researcher considered that the practice effect may have affected short-term learning retention. To further analyze each type of interaction, the students in the control group (A_1) presented significantly higher post-test (M_{B2} = 60.70) than pre-test (M_{B1} = 57.20) performance ($B_2 > B_1$, $p < .05$). This reveals that simply memorizing vocabulary still presented a short-term memory benefit. However, there was no significant difference in the performance between delayed post-test and pre-test, implying that over a prolonged period of time the students forgot what they had memorized and statistical significance could not be achieved. The next group, the students in the computer-mouse group (A_2) displayed an even more significant performance on the post-test (M_{B2} = 67.20) and delayed post-test (M_{B3} = 63.70) when compared to the pre-test (M_{B1} = 54.70) ($B_2 > B_1$ and $B_3 > B_1$, $p < .05$). This reveals that the traditional interaction of the computer-mouse group still led to a vocabulary learning effect. Finally, the learning performance of the motion-sensing group (A_3) was analyzed. From Table 7, the LSD method result of the motion-sensing group was identical to that of the mouse group (post-test and delayed post-test performance were higher than pre-test), revealing that the type of interaction was not a key factor affecting learning performance since the computer-mouse group also improved in the post-test and delayed post-test. However, interactive games were a key factor for long-term retention, as shown by the delayed post-test ($B_3 > B_1$, $p < .05$), since both the computer-mouse group and motion-sensing group played the same interactive game with a questioning strategy. In addition, that the post-test results are better than the delayed post-test results in both the motion-sensing and computer-mouse group ($B_2 > B_3$, $p < .05$). This reveals that some loss of retention always occurs over time.

Table 7. Simple main effect of three tests (B) with types of interaction (A)

Types of interaction	Three tests	*N*	*M*	*SD*	LSD method
$A_{1.}$ Control	$B_{1.}$ Pre-test	40	57.20	3.50	
	$B_{2.}$ Post-test	40	60.70	3.51	$B_2 > B_1^*$
	$B_{3.}$ Delayed post-test	40	59.70	3.45	
$A_{2.}$ Computer-mouse	$B_{1.}$ Pre-test	40	54.70	3.58	$B_2 > B_1^*$
	$B_{2.}$ Post-test	40	67.20	3.56	$B_3 > B_1^*$
	$B_{3.}$ Delayed post-test	40	63.70	3.47	$B_2 > B_3^*$
$A_{3.}$ Motion-sensing	$B_{1.}$ Pre-test	40	54.10	3.45	$B_2 > B_1^*$
	$B_{2.}$ Post-test	40	71.30	3.54	$B_3 > B_1^*$
	$B_{3.}$ Delayed post-test	40	65.60	3.51	$B_2 > B_3^*$

Note. $^*p < .05$.

Discussion

In agreement with past research, the interactive game with a questioning strategy used by the two groups benefited students' long-term retention

Overall, we found a partially disordinal interaction relationship between the three groups and three tests. The post-hoc comparison found the three tests to have an order relationship (the post-tests of the three groups were better than the pre-test ($B_2 > B_1$, $p < .05$), showing that the three groups all had short-term learning retention; however, the control group displayed a more apparent lack of long-term retention, as the delayed post-test was not significantly superior to the pre-test. In addition, both the computer-mouse group and motion-sensing group using interactive games not only displayed significantly better performance on the post-test than on the pre-test ($B_2 > B_1$, $p < .05$), but also displayed significantly better performance on the delayed post-test than on the pre-test (B3 > B1, $p < .05$), showing that the game-based learning with a questioning strategy used for the two interactive groups resulted in improved long-term learning retention of English vocabulary. Past studies (Dale, 1969; Papastergiou, 2009; Prensky, 2005) indicated that interactive learning should result in higher learning retention. They also indicated that a questioning strategy can be applied to English learning for better retention (Basturkmen, 2001; Boyd & Rubin, 2006; Cecil & Pfeifer, 2011; Shomoossi, 2004; Yang, 2010). Research has also showed that game-based English learning promotes retention (Flores, 2015). Therefore, affirming past research, our experimental design (the football game with a questioning strategy) also promoted long-term learning retention. Since the forty English Vocabulary Words were "randomly" sampled for the pilot test regardless of how familiar they were to the subjects, this could be a reason that the passing rate of the three tests were lower than those past tests held in case classes.

The type of interactive "interfaces" was not a key-factor affecting learning retention

The results in this study also show that there was no significant difference between the two interactive types (three group comparison). That is, the KMIS motion-sensing interface did not outperform the computer-mouse interface. Therefore it seems that the type of interactive "interfaces" were not the key-factors affecting learning retention. Nevertheless, previous research (Pan, 2013; Sommool et al., 2013; Vrellis et al., 2014; Yuan et al., 2015) indicated that the novelty of the Kinect motion-sensing interface can better increase students' attention or motivation to participate when compared to a traditional computer-mouse interface did. This reveals that the novel KMIS interface perhaps could be applied to attract students' attention or enhance motivation to participate in English vocabulary learning when compared to one-way rote memorization of English vocabulary commonly practiced in schools (Smith et al., 2013).

Since the development of applications of the Kinect in education is only just starting to expand, learners can expect much more innovation in interactive technologies. As learners adapt to motion-sensing technology, a new generation of interfaces could easily emerge for learners, similar to users' adaptation to computer-mouse operations in the 1980s (Karat, McDonald, & Anderson, 1986). The research after two decades (Forlines, Wigdor, Shen, & Balakrishnan, 2007) has discovered that, even compared to direct touch, users were still used to using mice to operate tasks on personal computers. Motion-sensing is still relatively novel compared to learners' familiarity with mouse operation. For the participants in our study, the computer-mouse group had the benefit of familiarity (familiarity results in lower cognitive load), but the operation was comparatively less novel (possibly leaving a more shallow impression in the memory) (Pan et al., 2014). As a result, we should consider that both the computer-mouse and the KMIS motion-sensing interface have their advantages in English vocabulary learning. The novel KMIS "interface" is helpful to attract students' attention and motivation; however, in this study this interface was not particularly beneficial to enhancing students' retention in English vocabulary learning.

The Kinect motion-sensing operating commands need to be intuitive in the future

In this study, participants were unfamiliar with the KMIS interface. The meanings of certain motion-sensing postures which correspond to different computer commands could differ in various research studies. This could result in unfamiliarity with operations if the definitions of postures vary across different learning environments. This also could be a hindrance in educational applications. Therefore, in further research, it would be beneficial to transform the posture definitions to make them intuitive and user friendly and avoid too complex postures to correspond to computer commands. On the other hand, further work could also consider the extent to which the KMIS approach with game-based learning can be used to learn other parts of English such as English grammar, or even other languages. This above issues are topics for future research.

Conclusion

After discussing the statistical analyses and their inferred meanings, the following conclusions were made: (1) the analyses of two-way mixed-design ANOVA reveal partially disordinal interaction relationship between the interactive types and the three tests (F = 14.98, $p < .01$, $\eta^2 = .204$). The post-hoc comparison (LSD method) shows that the three tests had an order relationship ($B_2 > B_1$, $p < .05$). This reveals that all three groups have short-term learning effects. (2) The within-group comparison, both for the motion-sensing and computer-mouse groups utilized an interactive game with a questioning strategy, displaying significant long-term (1-month) retention ($B_3 > B_1$, $p < .05$). Although the control group displayed a vocabulary learning effect in the post-test ($B_2 > B_1$, $p < .05$), there was a more apparent lack of long-term retention (B_3 not higher than B_1). (3) The between-group comparison of the two interactive groups did not reach a significant difference in English vocabulary learning, meaning that the motion-sensing interface of the KMIS was not a key-factor affecting short-term or long-term learning retention. The key-factor was the interactive content applied by the two groups.

Based on the experimental findings, our suggestion is that teachers can adopt interactive games with a questioning strategy to enhance students' long-term English vocabulary retention. Teachers also can use the novel KMIS interface for interactive operation in order to attract students' attention in English vocabulary learning. Learners are still relatively unfamiliar with using the Kinect interface for educational applications. This could be a hindrance in educational applications using Kinect. It would be beneficial to transform the posture definitions to make them more intuitive and user friendly and to avoid complex postures. In this study, the quasi-experimental design and cluster sampling were adopted for convenience; however, it could result in sampling error affecting experimental validity. Therefore, in future study, a true-experimental design should be adopted for better control of the interference factors.

Acknowledgements

The deepest gratitude goes to the Ministry of Science and Technology of Taiwan for its funding (MOST 102-2511-S-259-014), and the Review Panel for its utmost advice.

References

Basturkmen, H. (2001). Descriptions of spoken language for higher-level learners: The Example of questioning. *ELT Journal, 55*(1), 4-13.

Bedny, G. Z., & Harris, S. R. (2005). The Systemic-structural theory of activity: Applications to the study of human work. *Mind, Culture & Activity, 12*(2), 128-147.

Bedny, G. Z., & Karwowski, W. (2003). A Systemic-structural activity approach to the design of human-computer interaction tasks. *International Journal of Human-Computer Interaction, 16*, 235-260.

Birk, M. V., Atkins, C., Bowey, J. T., & Mandryk, R. L. (2016). Fostering intrinsic motivation through avatar identification in digital games. In J. Kaye & A. Druin (Chairs), *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 2982-2995). New York, NY: ACM.

Boyd, M., & Rubin, D. (2006). How contingent questioning promotes extended student talk: A Function of display questions. *Journal of Literacy Research, 38*(2), 141-169.

Buur, J., & Bødker, S. (2000). From usability lab to "design collaboratorium": Reframing usability practice. In D. Boyarski & W. A. Kellogg (Eds.), *Proceeding of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (DIS'00)* (pp. 297-307). New York, NY: ACM.

Cecil, N. L., & Pfeifer, J. (2011). *The Art of inquiry: Questioning strategies for K-6 classrooms*. Manitoba, Canada: Portage & Main Press.

Chapelle, C. A. (2009). The Relationship between second language acquisition theory and computer-assisted language learning. *The Modern Language Journal, 93*(s1), 741-753.

Chuang, T. Y., & Kuo, M. S. (2016). A Motion-sensing game-based therapy to foster the learning of children with sensory integration dysfunction. *Educational Technology & Society, 19*(1), 4-16.

Dale, E. (1969). *Audiovisual methods in teaching*. New York, NY: Dryden Press.

Engeström, Y. (2000). Activity theory as a framework for analyzing and redesigning work. *Ergonomics, 43*(7), 960-974.

Flores, J. F. F. (2015). Using gamification to enhance second language learning. *Digital Education Review, 27*, 32-54.

Forlines, C., Wigdor, D., Shen, C., & Balakrishnan, R. (2007). Direct-touch vs. mouse input for tabletop displays. In M. B. Rosson & D. J. Gilmore (Eds.), *Proceedings of the 2007 Conference on Human Factors in Computing Systems* (pp. 647-656). New York, NY: ACM.

Ge, Z. G. (2015). Enhancing vocabulary retention by embedding L2 target words in L1 stories: An Experiment with Chinese adult e-Learners. *Educational Technology & Society, 18* (3), 254-265.

Hsiao, H. S., & Chen, J. C. (2016). Using a gesture interactive game-based learning approach to improve preschool children's learning performance and motor skills. *Computers & Education, 95*, 151-162. doi:10.1016/j.compedu.2016.01.005

Hwang, G. J., Chiu, L. Y., & Chen, C. H. (2015). A Contextual game-based learning approach to improving students' inquiry-based learning performance in social studies courses. *Computers & Education, 81*, 13-25.

Kaptelinin, V. (1996). Activity theory: Implications for human-computer interaction. In B. A. Nardi (Ed.), *Context and Consciousness: Activity Theory and Human-Computer Interaction* (pp. 103-116). Cambridge, MA: MIT Press.

Karat, J., McDonald, J. E., & Anderson, M. (1986). A Comparison of menu selection techniques: Touch panel, mouse and keyboard. *International Journal of Man-Machine Studies, 25*(1), 73-88.

Kumaragurubaran, V. (2011). *Sensing, actuating and processing in the built environment: A beginner's guide to physical computing tools*. Retrieved from http://quicksilver.be.washington.edu/courses/arch498cre/2.Readings/1.Manuals/Beginners%20Guide%20to%20Physical%20Computing.pdf

Kuutti, K. (1996). Activity theory as a potential framework for human-computer interaction research. In B. Nardi (Ed.), *Context and consciousness: Activity theory and human-computer interaction* (pp.17-44). Cambridge, MA: MIT Press.

Leont'ev, A. N. (1974). The Problem of activity in psychology. *Soviet Psychology, 13*(2), 4-33.

Levinger, P., Zeina, D., Teshome, A. K., Skinner, E., Begg, R., & Abbott, J. H. (2016). A Real time biofeedback using Kinect and Wii to improve gait for post-total knee replacement rehabilitation: A Case study report. *Disability and Rehabilitation: Assistive Technology, 11*(3), 251-262.

Munassar, N. M. A., & Govardhan, A. (2011). Comparison between traditional approach and object-oriented approach in software engineering development. *International Journal of Advanced Computer Science and Applications, 2*(6), 70-76.

Nardi, B. A. (1996). Activity theory and human-computer interaction. In B. Nardi (Ed.), *Context and consciousness: Activity theory and human-computer interaction* (pp. 7-16). Cambridge, MA: MIT Press.

Nissimov, S., Goldberger, J., & Alchanatis, V. (2015). Obstacle detection in a greenhouse environment using the Kinect sensor. *Computers and Electronics in Agriculture, 113*, 104-115.

Oxford, R. (1990). *Language learning strategies: What every teacher should know*. Boston, MA: Heinle & Heinle.

Pan, W. F. (2013, July). The Different interactive interface differences on pupils' Nano learning motivation. In C. S. Lin (Ed.), *GCCIL 2013 Global Chinese Inquiry Learning and Innovative Applications* (pp. 14-17). Tainan, Taiwan: National University of Tainan.

Pan, W. F., Chien, M. Y., & Tu, S. C. (2012a). The Feasibility assessment of using a Kinect-tablet integrated system to improve electric wheelchair reversing safety. *Industrial Technology and Management, 49*, 121-125.

Pan, W. F., Chien, M. Y., Liu, C. C., & Chan, K. H. (2012b). Feasibility analysis of improving the accessibility and security of Nano-Labs via Kinect. *International Proceedings of Computer Science and Information Technology, 24*, 138-141.

Pan, W. F., Lin, H. F., & Wu, M. Y. (2011, July). Using Kinect to create active-learning situations for nanotechnology labs/classrooms. In C.-C. Chang (Ed.), *2011 3rd International Conference on Education Technology and Computer* (pp. 54-56). Singapore: IACSIT press.

Pan, W. F., Tu, S. C., & Chien, M. Y. (2014). Feasibility analysis of improving on-campus learning paths via a depth sensor. *Interactive Learning Environments, 22*(4), 514-528. doi:10.1080/10494820.2012.682585

Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers & Education, 52*(1), 1-12.

Prensky, M. (2005). Computer games and learning: Digital game-based learning. In *Handbook of computer game studies* (pp. 97-122). Cambridge, MA: MIT Press.

Ronimus, M., & Lyytinen, H. (2015). Is school a better environment than home for digital game-based learning? The Case of GraphoGame. *Human Technology, 11* (2), 123-147. doi:10.17011/ht/urn.201511113637

Sheehan, D., & Katz, L. (2012). The Impact of a six week exergaming curriculum on balance with grade three school children using the Wii FIT (TM). *International Journal of Computer Science in Sport, 11*(3), 5-22.

Shomoossi, N. (2004). The Effect of teachers' questioning behavior on EFL classroom interaction: A Classroom research study. *The Reading Matrix, 4*(2), 96-104.

Smith, G. G., Li, M., Drobisz, J., Park, H. R., Kim, D., & Smith, S. D. (2013). Play games or study? Computer games in eBooks to learn English vocabulary. *Computers & Education, 69*, 274-286.

Sommool, W., Battulga, B., Shih, T. K., & Hwang, W. Y. (2013). Using Kinect for holodeck classroom: A Framework for presentation and assessment. In J. F. Wang & R. Lau (Eds.), *Advances in Web-Based Learning-ICWL 2013* (pp. 40-49). Berlin, Heidelberg: Springer-Verlag.

Star, J. R., Chen, J., & Dede, C. (2015). Applying motivation theory to the design of game-based learning environments. In J. Torbeyns, E. Lehtinen, & J. Elen (Eds.), *Describing and Studying Domain-Specific Serious Games* (pp. 83-91). doi:10.1007/978-3-319-20276-1_6

Suma, E., Krum, D., Lange, B., Koenig, S., Rizzo, A., & Bolas, M. (2013). Adapting user interfaces for gestural interaction with the flexible action and articulated skeleton toolkit. *Computers & Graphics, 37*(3), 193-201.

Tutwiler, S., Lin, M. C., & Chang, C. Y. (2013). The Use of a gesture- based system for teaching multiple intelligences: A Pilot study. *British Journal of Educational Technology, 44*(5), 133-138.

Vrellis, I., Moutsioulis, A., & Mikropoulos, T. A. (2014, July). Primary school students' attitude towards gesture based interaction: A Comparison between Microsoft Kinect and mouse. In *Proceedings of Advanced Learning Technologies (ICALT), 2014 IEEE 14th International Conference* (pp. 678-682). doi:10.1109/ICALT.2014.199

Vygotsky, L. S. (1978). Interaction between learning and development. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind in society: The development of higher psychological processes* (pp. 79-91). Cambridge, MA: Harvard University Press.

Yang, C. F. (2010). A Discourse analysis of a native and a non-native English teacher's questioning in elementary EFL classrooms: A Case study. *Educational Linguistics Forum, 1*(1), 93-125.

Yao, N., Wang, K., Cai, Y., & Zhang, X. (2015). Application of somatosensory sensor Kinect in man-machine interaction framework of gesture recognition. *Sensor Letters, 13*(12), 1050-1054.

Yen, K. H. (2012). 雄之語文高手(足球版) [Hsiung who is a language expert (Flash: Football game)]. Retrieved from http://goo.gl/dFXLKw

Yuan, R. Q., Hsieh, S. W., Chew, S. W., & Chen, N. S. (2015). The Effects of gesture-based technology on memory training in adaptive learning environment. In *Proceedings of 2015 International Conference of Educational Innovation through Technology (EITT)* (pp. 190-193). doi:10.1109/EITT.2015.47