Chen, G.-D., Nurkhamid, Wang, C.-Y., Yang, S.-H., & Chao, P.-Y. (2014). Self-Observation Model Employing an Instinctive Interface for Classroom Active Learning. *Educational Technology & Society*, *17* (3), 14–26.

Self-Observation Model Employing an Instinctive Interface for Classroom Active Learning

Gwo-Dong Chen[1], Nurkhamid[1,2], Chin-Yeh Wang[3*], Shu-Han Yang[4] and Po-Yao Chao[5]

[1]Department of Computer Science and Information Engineering, National Central University, Jhongli City, Taiwan // [2]Faculty of Engineering, Yogyakarta State University, Yogyakarta, Indonesia // [3]Research Center for Science & Technology for Learning, National Central University, Jhongli City, Taiwan // [4]Department of Hospitality Management, Chien Hsin University, Jhongli City, Taiwan // [5]Department of Information Communication, Yuan Ze University, Jhongli City, Taiwan // chen@csie.ncu.edu.tw // nurkhamid@uny.ac.id // chinyea@db.csie.ncu.edu.tw // yoko@uch.edu.tw // poyaochao@saturn.yzu.edu.tw

[*]Corresponding author

(Submitted March 18, 2013; Revised August 1, 2013; October 7, 2013)

ABSTRACT

In a classroom, obtaining active, whole-focused, and engaging learning results from a design is often difficult. In this study, we propose a self-observation model that employs an instinctive interface for classroom active learning. Students can communicate with virtual avatars in the vertical screen and can react naturally according to the situation and tasks. Students can also immediately observe themselves mixed with a virtual environment and therefore reflect on the necessary improvements. With the designed system, potential advantages such as motivation, enjoyment, context for situated learning, engagement, social collaboration, and role-playing might arise. To examine the idea, we conducted a case study with 60 fourth grade elementary students to investigate learners' behavior as performers and peer audience and their perception of body movements and speech commands in the designed learning environment. The results show that the students' image in the vertical screen affected the students and the peer audience positively. Moreover, they positively perceived their competence and their enjoyment after they performed contextual learning activities through the body movement interface.

Keywords

Active learning, Self-observation, Instinctive interface, Body movements, Classroom learning

Introduction

Active learning involves two things for learners: the learners are actively involved in the learning process—they learn by doing, actively participating in constructing knowledge, resulting in deeper and more persistent knowledge—and are being engaged in learning (Prince, 2004). These are two core elements of active learning, i.e., activity and the promotion of engagement. John Dewey coined the phrase, "learning is doing," and later scholars identified, "knowing is doing" (Mishra, Worthington, Girod, Packard, & Thomas, 2001). More specifically, active learning may include collaborative, cooperative, and problem-based learning (Prince, 2004). Mayer (2009, pp. 21–24) categorizes active learning into two types: cognitively and physically (behaviorally) active learning. He primarily stresses utility in cognitively active learning, and this view is generally accepted in cognitivist learning theory. Indeed majority notion of active learning with technology has been of cognitive type. One example is BeyondShare (Kao, Lin, & Sun, 2008), which is an Internet-based learning environment designed to promote active learning that involved cognitive activities such as reflection, synthesis, and meaningful learning, and employed a concept map. However, when body movement is considered as an integral part of the cognition process, as it is according to the embodied cognition view, properly designed physically active learning experiences should contribute to cognition and learning. For this to occur, the learning environments need to explicitly support physical activities, and current technologies have provided diverse possibilities for achieving physically active learning.

In a typical classroom setting, students can learn by passively sitting on their chairs and using minimal body movement. However, implementing body movements, rather than simply sitting for extended periods benefits learning, facilitating effective memory, fun, and activity (Jensen, 2000). Here he focuses the discussion on physically rather than cognitively active learning. James et al. (2002) provided evidence that actively manipulating objects in a virtual reality environment improved object recognition. These indicate that a learning environment must be designed to support physically active learning that comprises body movements.

In this study, we propose a self-observation model that comprises an instinctive interface for active learning, facilitating engagement, collaboration, and physical activity.

Related work

Active learning with technology

Instead of applying passive teaching and learning practices in the classroom, contemporary educators have increasingly focused on active teaching and learning. However in traditional classrooms, students learned information, blindly followed guidelines, and learned a mass of disconnected facts, but were provided no clues or context by their teachers, who authoritatively delivered knowledge (Sawyer, 2006). This practice is called instructionism, and eventually, "scientists discovered that instructionism was deeply flawed" (Sawyer, 2006, p. 2). In the early era of computer use in classrooms, their use was based on instructionism or behaviorism. However, "the computer should take on a more facilitating role, helping learners have the kind of experiences that lead to deep learning" (Sawyer, 2006, p. 8). For example, in the electronic textbook designed as an instructivist learning tool, although it possessed certain advantages compared with traditional lectures, Herrington and Standen (2000) found it boring because students experienced difficulty to apply the knowledge in real life. According to Herrington and Oliver (2000), the tool lacks authentic contexts and activities.

Harasim (2011, p. 14) noted that epistemologically, behaviorism is in one group with cognitivism. They are of objectivism perspective assuming absoluteness of knowledge based on reality; hence, learning is acquiring knowledge. Constructivism, however, is a subjectivism paradigm, which assumes that knowledge is constructed; hence, learning is doing and creating meaning. Learning in behaviorism and cognitivism tends to be passive learning in a sense that it is more instructor-centered, while learning in constructivism promotes active learning (Harasim, 2011, pp. 58–61). Active learning facilitates knowledge acquisition, critical thinking skills, problem solving, and independent thinking. Passive learning is disadvantageous, allowing teachers to become "the sage on stage," and the lack of interactivity between teachers and students engenders quiet students (Wang, Shen, Novak, & Pan, 2009). However, obstacles arise when implementing active learning, because teachers tend to imitate the way they were taught, and most were taught as passive learners (McManus, 2001).

Technology can be applied to active learning (Shapiro, 1998; Wang et al., 2009) and may provide interactivity between learners and content knowledge (Pahl & Kenny, 2008). Jonassen et al. (2007, p.8) asserted that technology should support meaningful learning by providing a "context to support learning by doing" or active learning. Technology should no longer be used to transmit knowledge or teach. Instead, it should be used to engage students and promote critical thinking. For example, when hypertext is applied to active learning, learners can actively choose links, allowing them to effectively integrate and deeply understand knowledge, especially when the hypertext is ill-structured. In the ill-structured system, learners require cognitive effort to process the information contained in the links, understand the material, and remain oriented (Shapiro, 1998). As another example, Wang et al. (2009) designed a mobile learning environment that used real-time interaction to increase student engagement. When a video camera is used in active learning, learners can engage in meaningful activities, view their own performance, express themselves, and become creative and collaborative (Broady & Duc, 1995). Digital learning playground has also been used in active learning (Chen et al., 2013; Chen, Chuang, Nurkhamid, & Liu, 2012; Wang et al., 2010) but previous studies have not considered body movements (e.g., gesture-based interactions) or self-observations. Ozcelik and Sengul (2012) emphasized that using gestures to understand concepts can improve learning. Although active learning that leverages performance and body movements is beneficial, few researchers have explored this by using contemporary classroom technological tools.

Self-observation for learning

When an individual can see himself or herself while learning, it may positively affect corrective feedback, introspection, and the sense of self-assessment resulting in learning gains. Self-observation (seeing oneself) positively affects learning (Fireman, Kose, & Solomon, 2003). Observing self-action (Gupta & Bostrom, 2012) is called enactive learning. Observing the actions of others while they learn is called vicarious learning, because learners gain knowledge by modeling the behaviors of others. When foreign language students learn English orally, Chen (2008) noted that self-assessment promoted self-confidence, additional effort, awareness of competence, and mastery in learning, suggesting that self-assessment facilitates autonomy and lifelong learning skills.

Schunk (2011) stated, “Learning occurs either enactively through actual doing or vicariously by observing models perform (e.g., live, symbolic, portrayed electronically). Enactive learning involves learning from the consequences of one’s actions.” (p. 121). Thus, self-observation can be enactive because it displays the consequences of actions onscreen. However, self-observation can also be vicarious when students feel that the technology mirrors themselves in a learning context, and this observation feels similar to watching a television.

Instinctive interface for performing physical activity

The story of human-computer interaction evolved from command line interfaces, to graphical-user interfaces, and then to natural user interfaces (Villaroman, Rowe, & Swan, 2011). In natural user interfaces, users interact with computing devices in intuitive and instinctive ways. Learning to use the interface is expected to be quick and easy. Furthermore, this instinctive interface provides a greater level of control compared with a keyboard and mouse (Birchfield, Ciufo, & Minyard, 2006; Shiratuddin & Wong, 2011). The Kinect device provides this type of interface, enabling users to perform physical activity by using embodied interaction. Comfortable movement is possible because Kinect is wireless, requiring no attachment between the user and the device (Ozcelik & Sengul, 2012).

Role-playing activities that involve some degree of physical activity are prevalent in current pedagogies. This type of learning usually impresses learners because they perceive, act, interact with things and events in their surroundings, and their bodies link their minds to the world (Atkinson, 2010; Barsalou, 2010). Furthermore, learning consists not only of mental activities, but it also needs the learner's body to perform and interact with the learning environment; this implies embodied cognition (Atkinson, 2010). Body movements may add contextual meaning to activities that were traditionally dominated by printed materials (Gee, 2003). Kirsh (2013) maintained that performing physical activities can enable learners to acquire more knowledge compared with simply watching others; doing may result in more effective learning compared with watching.

A similar study on using body movements for learning could be that conducted by Tolentino et al. (2009), who used a horizontal ground (floor) projection as a mixed-reality science learning space. However, by using a traditional drama stage metaphor that is similar to a traditional classroom setting, our design includes two spaces: a horizontal and a vertical screen. This does not deviate much from traditional classroom culture, in that a classroom has a blackboard facing the classroom as the center of attention—which corresponds to the vertical screen in our system. Moreover, it uses current technologies in schools, such as projectors, touchable whiteboards, and desktop computers. Adopting an authentic learning environment of a Digital Learning Playground platform (Wang et al., 2010), we further enhance the platform to include self-observation feature and embodied interaction. The following section presents additional details regarding the proposed learning environment.

The learning environment

The environment implements features described in the previous section, such as active learning, self-observation, and instinctive interface. Figure 1 shows the interaction setting consisting of various parts, including a vertical display, a horizontal touch-enabled board, and an action zone for the performer. Both screens display beams from the upper and front projectors. Kinect detects commands through the performer’s body movements and voice, and renders the performer’s image on the vertical screen to place and interact with learning objects. The orchestrator of the interaction between users and the system is a desktop computer. Peer students and the teacher may stand on the left and right sides of the horizontal board (Figure 2).

The horizontal screen is a shared touch-enabled space to provide menus for a teacher to preview the lesson and to explain how to play the game. It can provide a layout of board-game items, so that a group of learners can plan when collaborating on a role-playing assignment. It can also be a virtual keyboard, for example, for one of the group members to type words with during a game session. The vertical screen provides simulative and task-situational context, for example, for learners playing a cook’s role in a virtual kitchen by using their body movements. The screen can also allow learners to see virtual events that they can authentically interact with in a situated manner.

Enabling a Kinect sensor for Xbox 360 or other motion/voice controllers can facilitate instinctive interactions. Kinect detects the learner's body movement and facilitates voice commands, for example, when choosing a type of drink. It can also produce the learner's image on the vertical screen where the ongoing task scenarios appear. The learner, as a live task performer, with his or her body movements and/or speech commands can role-play and interact with context objects.

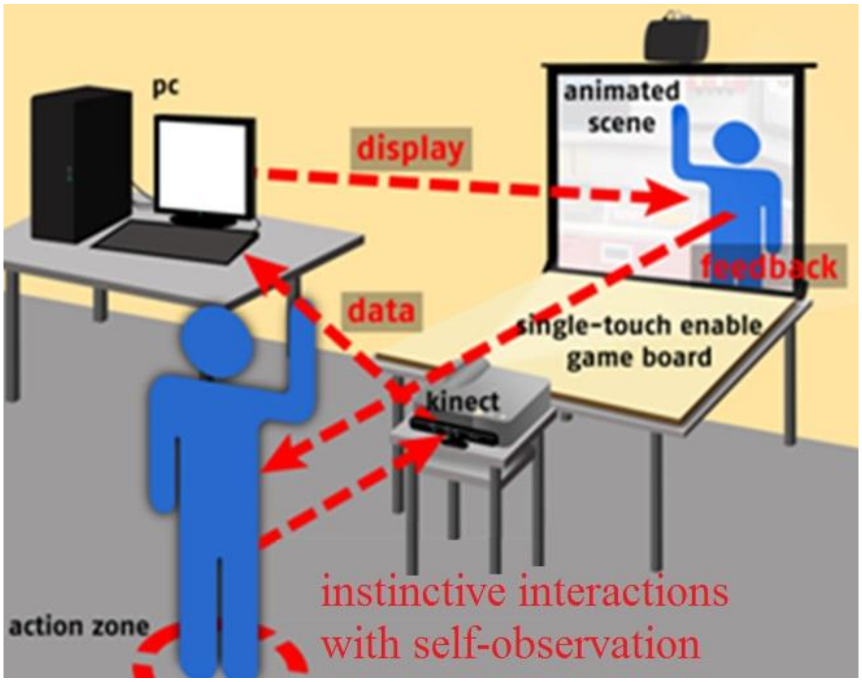

Figure 1. Self-observation model with instinctive interface for classroom learning

The potential pedagogical added values of the design may result from its built-in features and size. The vertical and horizontal screen sizes enable teachers and students to stand around the horizontal screen to conduct social collaborations and corrective feedback sessions. Including gaming and role-playing applications in the system should increase the motivation, enjoyment, and engagement of students. The group game function and animated multimedia support, which incorporates body movements, should engender "ownership and voice in the learning process," and "learning in social experience" by using "multiple modes of representation" "in realistic and relevant contexts" (Honebein, 1996, pp. 11–12). When the learning context is provided onscreen during role-playing, it combines real-world learning and the traditional classroom, engaging students as if in real locations without requiring that they travel to experience the learning context. Additional pedagogical value may result from the support of vicarious (by observing others) and enactive (by doing) learning. This should facilitate a meaningful classroom learning experience for students.

Three assumptions motivate the creation of the learning environment. First, direct embodied experiences can facilitate learning as a way of constructing knowledge. In Freiler (2008), embodied learning relied on the role of the body in knowledge construction to produce deep, experiential learning. Recent focus has been paid to incorporating body movements into interface designs for learning purposes (Johnson-Glenberg, Birchfield, Savvides, & Megowan-Romanowicz, 2010; Lee, Huang, Wu, Huang, & Chen, 2012; Ozcelik & Sengul, 2012; Tolentino et al., 2009). Second, multimodal learning optimizes knowledge construction and thus improves learning, whether for languages (Macedonia & Knösche, 2011) or science and engineering (Johnson-Glenberg et al., 2010; Tolentino et al., 2009). Multimodal learning relates to knowledge that is represented using various sensory modes such as sight, sound, movement, and touch (Birchfield et al., 2006). Multimodal learning works because cognitive processing is multimodal (Atkinson, 2010). The external representations that learners experience through various perception and knowledge processing channels improves the potential to learn (Tolentino et al., 2009). Moreover, acting out knowledge facilitates verbal learning because it involves encoding and deep processing (Macedonia & Knösche, 2011). Third, because extrapolating the findings of offline self-observation from video recordings benefits learning

(Fireman et al., 2003), observing the live actions of a person while learning should also be beneficial, facilitating instant, rather than delayed feedback from teachers and peer learners.

Following these assumptions, we designed a new learning interface. The proposed interface should be engaging and enjoyable because, rather than simply using a mouse and a keyboard, we designed the system to respond to body movements and voice commands. Consequently, the body movement and voice command interfaces should provide multimodal learning channels. Based on these three assumptions, this design is feasible. However, certain caveats exist; for example, certain learners may resist or feel uncomfortable with embodied learning (Freiler, 2008), or exhibit an initial shyness to act (Birchfield et al., 2006).

This paper may exemplify the potential educational applications of Kinect (Hsu, 2011). Kinect represents a current embodied sensor technology that is ready to catch learners' body movements, actions, and voice. To the best of our knowledge, a new feature of the proposed design is that it provides learners with the possibility to see themselves in action while they are involved in a learning scenario. The learner's animated images appear on the vertical screen, which they and their peer learners can see (Figure 1). We call this the self-observation model.

With the proposed model we want to know about the learners' behavior. Therefore, this study raised the following research questions: In the proposed design, what is the effect of the students' live performance and of their speech commands and body movements on learning behaviors? In this case, we limit our observations of the students' learning behaviors to their intrinsic learning motivation measured with the Intrinsic Motivation Inventory (IMI) survey (Schunk, Pintrich, & Meece, 2007). We chose the IMI as an instrument to understand students' learning behavior, and specifically, whether they enjoy learning by using the system we designed. Thus, we can use the results to infer whether the design is acceptable. We hope the proposed design produces a positive effect on learning behavior.

Evaluation methods

To examine the proposed model, we implemented the system as a case study for children learning English as a second language in Taiwan. Data collection includes activity videotaping, the IMI survey (Schunk et al., 2007), and interviews.

The aim of this study was to understand the effect that the students' live performance and their speech commands and body movements have on learning behavior, which their learning motivation represents. In this case, we believe that embedding students' live performance in ongoing-task scenarios not only relates the "actor" and "audience" to apprehended knowledge in a real context, but also engenders productive learning performance and a positive atmosphere. Thus, we hypothesize the following:

- Students' live performance in the ongoing-task scenarios has positive learning effects on the performer and other peers (Hypothesis 1).
- One of our assumptions was that using speech commands and body movements to act out in the task enhances learning and creates an unusual fantasy, due to near-authenticity of the role being played, not only among performers, but also among the peer audience. Therefore, we postulate that using speech commands and body movements to interact with the objects in the scenario positively affects the learning of the performer and the peers (Hypothesis 2).

Participants

The participants of this study were 60 fourth grade elementary students from two classes in Taiwan, where the experiment took place. They learn English as a second language. The English teacher of the school, who was in charge of these two classes, confirmed that the subjects' English background knowledge was similar.

Learning material

The learning objective of the class session was about knowing food names and how to order the foods; the learning material was taken from the students' English textbook. At the start, teachers explained to the participants what learning tasks and goals they were to accomplish. Each group of students, which comprised five people, needed to collaborate in planning their actions and managing their resources on the shared game table when their turn to play came. In this case, students ran a fast-food restaurant, and needed to decide their daily menu and staff arrangements (role assignment). Based on the assigned roles, each participant played serving a virtual customer who appeared on the vertical screen. The student playing the role of a counter assistant was in charge of taking orders from the virtual customers at a counter scenario. The starring student's live image appeared on the screen. Meanwhile, the student acting as a trainee would listen, watch the conversations, and then input guest orders using a virtual keyboard on the game table. Afterward, the student starring as a cook used speech commands and body movements to make food based on the guest's order in the kitchen scenario. For example, a cook had to physically flip the patty to make burgers in the virtual kitchen and respond to the beverage machine to refill a drink.

Procedure

Sixty students were divided into two 30-student groups: the experimental and control groups. The selection is based on which class they belonged to. Each group was divided into six 5-student gaming groups. Two classes took place for 80 min each for the learning activity, and were devoted to 20 min for evaluations, including interviews and questionnaires. The same teacher facilitated the process during the teaching phase. We randomly picked one class as the experimental group and another as the control group. The students of the experimental group could see their classmates' image in the scenario; conversely, the students in the control group would not appear on the screen while acting out the conversation with the characters.

To collect data, we used videotaped activities, the IMI survey, and interviews. The IMI questionnaire was used to examine subjects' intrinsic learning motivation with respect to the following aspects: interest/enjoyment, effort/importance, value/usefulness, perceived competence, and pressure/tension. The scores ranged from 1 (*lowest*) to 7 (*highest*). The interviews were also conducted to understand their perceptions of the learning experience.

Figure 2. The scenes in experimental group (left) and control group (right)

Results

We observed that the experimental group students executed their tasks smiling (75%) and made funny faces and hand gestures (Table 1). This body language indicates their curiosity toward their images in the scenario and the learning content. Regarding the control group, they were calmer while they talked to the virtual customers. The students in the control group said the following in the interview: "I would rather not to be a counter person because I don't like to speak out" and "I can learn how to talk to a customer, but it is not so fun doing that." In addition, the result from the interest dimension of IMI shows that the control group demonstrates less interest in role-playing (5.81) than does the experimental group (6.38).

To ascertain active learning through observation, the performers in the control group were more likely to perform individually, and made no significant effort to seek assistance from their teammates (0%). This phenomenon highlights the interesting relationship between the performers and the teammates. Both groups reported that they exerted considerable effort and felt competent (effort and perceived competence) after performing the tasks. In addition, all the students agreed that learning by interacting within relevant contexts helped them relate to new learned knowledge (these gave a score of over 6 on the value/usefulness dimension). One of the performers in the experimental group expressed that "seeing myself in the counter dealing with the customers really engaged me in this role."

To investigate how students perceive themselves when performing on screen and how this influences them and others, the class observation adopted two primary focuses:

- the learning behaviors of the performers, such as anxiety, active learning, enjoyment, and perception as a demonstrator, and
- the learning behaviors of others, such as attention or focus, imitative behaviors, and error correction through observation.

On average, the learners perceived that the environment was enjoyable, valuable, and essential when they used body movement and speech commands to perform contextual learning tasks. The observational data, interview, and the result of IMI were cross-analyzed in the sections that follow.

The behaviors of the performers

The interviews revealed that two students became nervous. They gave the following reasons: "I was nervous because I was afraid to give a wrong answer" and "I got nervous because I wasn't sure my response was right." We inferred that most students were unaware of their nervousness, and that this did not affect their performance. In addition, in contrast to their outward appearance in the scenario, having to give proper responses placed the performers under pressure. Regarding the students as task focus, observations indicated that the prompts and the assistance from the instructor and the peers also helped ease their tension of arising from the task focus.

Table 1 presents a summary of behaviors the students displayed when seeing themselves in the scenarios (Hypothesis 1). The five observational points are anxiety, active learning, enjoyment, and being perceived as demonstrators.

Table 1 provides the results obtained from our observation of anxiety levels produced from showing the students' images in the scenario. The observational data show that both groups of performers displayed simple nervousness. The results for the pressure/tension dimension of the IMI questionnaire show that both groups' tension is low (the highest average score is 7, whereas the lowest is 1; Table 2). Table 2 lists the results of the t-test between the CG and EG for all variables; no significant difference is indicated between the IMI variables.

Table 1. The learning behaviors of the performers (*n = 60*)

Anxiety	EG	CG
Uncomfortable swaying	8%	30%
Stiff body posture	50%	40%
Constant blink	17%	8%
Deep breathing	8%	40%
Frowning	17%	0%
Active learning	EG	CG
Active responding	58%	30%
Calling for support from teammates	17%	0%
Enjoyment	EG	CG
Smile	75%	10%
Showing curiosity:		
Making extra movement	58%	-
Making funny faces	42%	-
Hand movements indicating touching the virtual objects	33%	-

Perceived as a demonstrator	EG	CG
Standing up straight	58%	10%
Looking at the audiences	25%	0%
Speaking up	25%	10%

Table 2. The t-test analysis of IMI of performers (*n = 60*)

	EG		CG		
Variables	Avg	Std. dev.	Avg	Std. dev	Sig. (p)
Pressure/Tension	3.69	.48	3.28	.40	.570
Effort/Importance	6.50	.81	6.43	.82	.703
Perceived Competence	5.06	.37	4.97	.77	.105
Value/Usefulness	6.28	.51	6.47	.49	.700
Interest/Enjoyment	6.38	.23	5.81	.37	.059

Observations of the experimental group show notable behavior: some performers perceived themselves as knowledge demonstrators. Many EG performers were well aware they were performing, and intended to stand up straight (58%) and speak up (25%). Referring to the interview, six students voiced that they liked being called up to act as a model. Compared to the control group, the demonstrators' traits were not apparent in this experiment.

The behaviors of others

In this section, we concentrate on how the non-performers participated while their teammates were performing their tasks. Table 3 presents a summary of the observations of the behaviors of the others (Hypothesis 1). Our observational phases are their attention-focus, imitative behaviors, and active learning (supporting the performer).

Table 3. The observational points of learning behaviors of the others (*n=60*)

Attention-focus	EG	CG
Absent-minded in the task	3%	12%
Turning head toward the performer whenever performer is speaking or acting	97%	32%
When the performance is showing, the other participants are:		
Well-participated (facing the instructor when the instruction is delivered, watching the interaction shown on the screen, observing performer's performance)	70%	41%
Looking at the game table	7%	40%
Looking at other places	23%	19%
Immediate attention right after the image is appeared	97%	-
Imitative behaviors	**EG**	**CG**
Imitating performers' action	17%	12%
Simultaneously repeating the required speech	50%	18%
Active learning	**EG**	**CG**
Supporting the teammate performers (suggesting, error-correcting)	83%	35%
Supporting the performers (not aware of team competition)	43%	21%
No support	21%	70%

As the observational data of the peers' focus of attention show, having the students perform in the designed scenario successfully caught the other learners' attention (EG: 97%; CG: 32%). This is a positive learning effect because the learners were focused on the newly learned knowledge situated in an ongoing task of their classmate's role. These live performances serve as learning inputs or examples that an instructor hopes are delivered to the other participants. Moreover, this design has motivated the peer audience to imitate the performer's performance (EG: 50%; CG: 18%) and to show that they highly support the performers by correcting errors and making suggestions (EG: 83%, 43%; CG: 35%, 21%). Conversely, the students in the control group acted as passive knowledge receivers, and 70% of them did not show support. Approximately 43% of the peer audience was willing to help the performers, despite belonging to the other team. The interviews on the audience's feedback of the performance revealed these statements: "I think it was very interesting when my classmate appeared on the scene" and "I wanted to know whether he responded correctly or incorrectly, and I tried to learn from their mistakes and deliver a good performance during my turn."

The behaviors of the performers in physical learning

Regarding Hypothesis 2, Tables 4 and 5 show the IMI results and the observational data regarding engagement, active learning, and fulfillment. The results for intrinsic motivation show that the motivation of both classes through the embodied interaction is high, and that it coheres with the behaviors that the observational data reveal. Regarding the performers' behaviors, both classes displayed body expressions that demonstrated their ability to manipulate virtual objects, drawing their great curiosity to test the objects on the screen (Class A: 88%, Class B: 82%). A good number of students were highly contented and smiling (Class A: 63%, Class B: 91%) after completing the task. In relation to the non-performers, they were attracted to the task performance; some students even gave suggestions and shared the same hand movements with the performers (Figure 3). Table 4 shows a one-sample t-test that assumes a mean value of 4, the midpoint of the Likert scales from 1 (lowest) to 7 (highest), to ascertain the significance of learner responses to interacting with the system. In one sample t-test, we compared between the mean score of the sample and a known value, here the midpoint of the Likert scales, as other researchers did (for example, Jacucci et al., 2010; Mason, 2011). The results attained significance for effort/importance, value/usefulness, and interest/enjoyment, indicating that, on average, when body movement and speech commands were integrated in the system, learners found the system important, useful, and enjoyable. However, the results for the pressure/tension dimension were non-significant.

Table 4. The t-test analysis of IMI in physical learning (*n = 60*, test value = 4)

Variables	Avg	Std. dev.	Sig. (p)
Pressure/Tension	2.97	1.32	.007 ^{ns}
Effort/Importance	5.14	1.22	.002 *
Value/Usefulness	6.05	1.20	.000 ***
Interest/Enjoyment	5.58	1.29	.000 ***

^{ns}non-signficant. *significant at $p < 0.05$. ***significant at $p < 0.001$.

Table 5. The observational data of students' learning behaviors in an interactive physical learning (*n = 60*)

Active learning : Performers	Class A	Class B
Poses of readiness with both hands stretching forward	38%	36%
Trailing and exploring the kitchen tools with hands	88%	82%
Fulfillment : Performers		
Smile after tasks	63%	91%
Cheering poses (Victory poses)	38%	36%
Engagement : Non-performers		
Imitate performer's action	37%	33%
Imitate performer's speech	25%	28%

Figure 3. Imitating behavior

Discussion

We found that the involvement of body motions elicits considerable fun when performing tasks. Although some students felt stressed and shy because they saw themselves on the screen, the stress seemed not to overly affect their performance. Conversely, being in the spotlight might have encouraged students to intensify their efforts in the activities. Moreover, only students whose images appeared in the scenario could trigger the animated effects of the surrounding objects, such as touching the cash register. This "magic power" spurred students to set aside the pressure.

When learners conducted real-time observations of their self-actions while interacting with contextual objects on the vertical screen, positive learning behaviors were generated. These positive learning behaviors may result in strong learning outcomes, but those outcomes were not measured in this study. However, social cognitive theory supports that "learning interventions affect learning outcomes through reflection on observations" (Gupta & Bostrom, 2012, p. 2). Accordingly, in the proposed scheme, observing the performance of classmates might positively affect peer learners who are engaged in vicarious learning. This concurs with Gupta and Bostrom (2012), who indicated that enactive enabled and technology-mediated learning significantly and positively influence learning. However, Table 2 shows no significant difference in learner behaviors between the CG and EG.

The most interesting finding of the research was the focused attention of the peer audience while they were learning within the designed learning environment. Classmates' images in the situated scenarios were great attention getters, and the live performance kept the students focused on the learning matter. Holding the attention of a whole class on a subject matter is a difficult task, especially under game-based learning settings. We were glad to have found a way to retain the full attention and energy of learners engaged in a learning activity. The animated image on the vertical screen may provide an additional vicarious learning channel for both the performer and the peer audience.

Conclusions

We proposed a self-observation model that employs an instinctive interface, enhancing engagement and active learning by incorporating activities, role-play, and body movements. The design in self-observation model should be suitable for classroom learning because the activity does not only engage the learner who is performing the task but also the peer learners. Peer learners ought to see the performer for some reasons. The most obvious reason is that they also will have their turn to perform so they need to prepare themselves by observing their classmate performing the task.

The proposed learning environment should contribute to the study of technology-supported active learning. Thus far, active learning has been perceived primarily in terms of cognitively active learning, and this may result from the availability of previous forms of technology. However, contemporary technological tools have the ability and potential to support additional functions. These tools allow the application of physically (behaviorally) active learning under a theoretical support from, for example, embodied cognition view.

Reflecting on the experience of designing the learning space, we learned at least two lessons. First, the scenario must be designed to engage learners in the context of the roles that they must play. Relevant virtual objects appearing on the screen must be provided, and learners should be able to manipulate these objects with their body motions. The objects should respond to the movement accordingly. Second, the learning target must be realized in the objects with which learners interact. In this way, learners have the opportunity to reflect and directly experience knowledge based on the outcome of their actions. We expect the knowledge obtained to be relevant to real-world situations.

The animated image of a performer onscreen for self-observation should assist in creating deep impressions and clear experiences that may improve learning. Therefore, we believe that enabling actions through body movement, providing context that mimics a real-world situation, and seeing an individualized animated image should facilitate deep experiences and engagement in learning. We observed that this learning mechanism truly engaged students with authentic tasks. Because of the Kinect and image-processing technologies, we were able to create a near-authentic experience that embedded students' images in animated contexts and enabled learners to use language and body language to act out and present learned knowledge.

Finally, manipulating events and objects by using body movements and speech commands allows learners to engage with their assigned roles. We can conclude that students are impressed when they have the opportunity to use learned knowledge actively in a relevant context. This learning design with planned settings has great potential to generate a focused and engaging learning atmosphere that involves the whole class. Because our concern was to explore the learners' behavior in relation to our design, we did not quantitatively show the learning gains regarding language learning achievement in our system, although Macedonia and Knösche (2011) suggested the beneficial effects of body movements, for example, for word memorization.

Future research could include further exploring active and engaging learning environments by including more themed activities. Involving teachers in planning their teaching and learning activities with this system, which they can readily implement with an authoring tool, is also possible. However, current limitation is that the authoring tool for the teachers has not been available yet in our system. Besides, we chose a specific subject matter, i.e., second language learning, and created the setting for this to understand the behaviors of learners when they interacted with the system. Future study may also create the settings for other subject matters, or measure learning outcomes and learning gains from the settings. It is possible that the system can become a general purpose environment to flexibly accommodate numerous subjects for a classroom.

References

Atkinson, D. (2010). Extended, embodied cognition and second language acquisition. *Applied Linguistics*, *31*(5), 599–622. doi:10.1093/applin/amq009

Barsalou, L. W. (2010). Grounded cognition: Past, present, and future. *Topics in Cognitive Science*, *2*(4), 716–724. doi:10.1111/j.1756-8765.2010.01115.x

Birchfield, D., Ciufo, T., & Minyard, G. (2006). SMALLab: a mediated platform for education. In *ACM SIGGRAPH 2006 Educators program*. New York, NY, USA: ACM. doi:10.1145/1179295.1179329

Broady, E., & Duc, D. L. (1995). Learner autonomy and the video camera: A wider role for video recording activities? *The Language Learning Journal*, *11*(1), 74–77. doi:10.1080/09571739585200251

Chen, G. D., Chuang, C. K., Nurkhamid, N., & Liu, T. C. (2012). When a classroom is not just a classroom: Building digital playgrounds in the classroom. *Turkish Online Journal of Educational Technology*, *10*(4), 202.

Chen, G. D., Nurkhamid, Wang, C. Y., Yang, S. H., Lu, W. Y., & Chang, C. K. (2013). Digital learning playground: Supporting authentic learning experiences in the classroom. *Interactive Learning Environments*, *0*(0), 1–12. doi:10.1080/10494820.2012.705856

Chen, Y.-M. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research*, *12*(2), 235–262. doi:10.1177/1362168807086293

Fireman, G., Kose, G., & Solomon, M. J. (2003). Self-observation and learning: The effect of watching oneself on problem solving performance. *Cognitive Development*, *18*(3), 339–354. doi:10.1016/S0885-2014(03)00038-8

Freiler, T. J. (2008). Learning through the body. *New Directions for Adult and Continuing Education*, *2008*(119), 37–47. doi:10.1002/ace.304

Gee, J. P. (2003). *What video games have to teach us about learning and literacy* (1st ed.). London, UK: Palgrave Macmillan.

Gupta, S., & Bostrom, R. (2012). An investigation of the appropriation of technology-mediated training methods incorporating enactive and collaborative learning. *Information Systems Research*. doi:10.1287/isre.1120.0433

Harasim, L. M. (2011). *Learning theory and online technology*. New York, NY: Routledge.

Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational Technology Research and Development*, *48*, 23–48. doi:10.1007/BF02319856

Herrington, J., & Standen, P. (2000). Moving from an instructivist to a constructivist multimedia learning environment. *Journal of Educational Multimedia and Hypermedia*, *9*(3), 195–205.

Honebein, P. C. (1996). Seven goals for the design of constructivist learning environments. In *Constructivist learning environments: Case studies in instructional design* (pp. 11–24). Englewood Cliffs, N.J.: Educational Technology Publications.

Hsu, H. J. (2011). The potential of Kinect in education. *International Journal of Information and Education Technology*, *1*(5), 365–370.

Jacucci, G., Morrison, A., Richard, G. T., Kleimola, J., Peltonen, P., Parisi, L., & Laitinen, T. (2010). Worlds of information: designing for engagement at a public multi-touch display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2267–2276). New York, NY, USA: ACM. doi:10.1145/1753326.1753669

James, K., Humphrey, G., Vilis, T., Corrie, B., Baddour, R., & Goodale, M. (2002). "Active" and "passive" learning of three-dimensional object structure within an immersive virtual reality environment. *Behavior Research Methods*, *34*(3), 383–390. doi:10.3758/BF03195466

Jensen, E. (2000). Moving with the brain in mind. *Educational Leadership*, *58*(3), 34–37.

Johnson-Glenberg, M. C., Birchfield, D., Savvides, P., & Megowan-Romanowicz, C. (2010). Semi-virtual embodied learning-real world STEM assessment. In L. Anetta & S. Bronack (Eds.), *Serious Educational Game Assessment: Practical Methods and Models for Educational Games, Simulations and Virtual Worlds* (pp. 241–257). Rotterdam, Dutch: Sense Publishers.

Jonassen, D. H., Howland, J. L., Marra, R. M., & Crismond, D. P. (2007). Meaningful learning with technology (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Kao, G. Y.-M., Lin, S. S., & Sun, C.-T. (2008). Beyond sharing: Engaging students in cooperative and competitive active learning. *Educational Technology & Society*, *11*(3), 82–96.

Kirsh, D. (2013). Embodied cognition and the magical future of interaction design. *ACM Trans. Comput.-Hum. Interact.*, *20*(1), 3:1–3:30. doi:10.1145/2442106.2442109

Lee, W.-J., Huang, C.-W., Wu, C.-J., Huang, S.-T., & Chen, G.-D. (2012). The effects of using embodied interactions to improve learning performance. *Proceedings of the 12th International Conference on Advanced Learning Technologies (ICALT)* (pp. 557 – 559). doi:10.1109/ICALT.2012.104

Macedonia, M., & Knösche, T. R. (2011). Body in mind: How gestures empower foreign language learning. *Mind, Brain, and Education*, *5*(4), 196–211. doi:10.1111/j.1751-228X.2011.01129.x

Mason, R. B. (2011). Student engagement with, and participation in, an e-forum. *Educational Technology & Society*, *14*(2), 258–268.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York, NY: Cambridge University Press.

McManus, D. A. (2001). The two paradigms of education and the peer review of teaching. *Journal of Geoscience Education*, *49*(5), 423–434.

Mishra, P., Worthington, V., Girod, M., Packard, B., & Thomas, C. (2001). Learning science: A Deweyan perspective. *Journal of research in science teaching*, *38*(3), 317–336.

Ozcelik, E., & Sengul, G. (2012). Gesture-based interaction for learning: Time to make the dream a reality. *British Journal of Educational Technology*, *43*(3), E86–E89. doi:10.1111/j.1467-8535.2012.01288.x

Pahl, C., & Kenny, C. (2008). The future of technology enhanced active learning – A roadmap. In M. L. Lytras, D. Gasevic, P. Ordonez de Pablos, & W. Huang (Eds.), *Technology Enhanced Learning. IGI Global* (pp. 348-375). Retrieved December 19, 2012, from http://www.igi-global.com

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, *93*(3), 223–231. doi:10.1002/j.2168-9830.2004.tb00809.x

Sawyer, R. K. (2006). The new science of learning. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 1–16). New York, NY: Cambridge University Press.

Schunk, D. H. (2011). *Learning theories: An educational perspective* (6th ed.). Boston, MA: Addison Wesley.

Schunk, D. H., Pintrich, P. R., & Meece, J. (2007). *Motivation in education: Theory, research, and applications* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Shapiro, A. M. (1998). Promoting active learning: The role of system structure in learning from hypertext. *Human–Computer Interaction*, *13*(1), 1–35. doi:10.1207/s15327051hci1301_1

Shiratuddin, M. F., & Wong, K.-W. (2011). Non-contact multi-hand gestures interaction techniques for architectural design in a virtual environment. *Proceedings of the 2011 International Conference on Information Technology and Multimedia (ICIM)* (pp. 1–6). doi:10.1109/ICIMU.2011.6122761

Tolentino, L., Birchfield, D., Megowan-Romanowicz, C., Johnson-Glenberg, M., Kelliher, A., & Martinez, C. (2009). Teaching and learning in the mixed-reality science classroom. *Journal of Science Education and Technology*, *18*(6), 501–517. doi:10.1007/s10956-009-9166-2

Villaroman, N., Rowe, D., & Swan, B. (2011). Teaching natural user interaction using OpenNI and the Microsoft Kinect sensor. *Proceedings of the 2011 conference on Information technology education* (pp. 227–232). New York, NY: ACM. doi:10.1145/2047594.2047654

Wang, C. Y., Chen, C. H., Wu, C. J., Chi, Y. L., Lee, J. H., & Chen, G. D. (2010). Constructing a digital authentic learning playground by a mixed reality platform and a robot. *Proceedings of the 18th International Conference on Computers in Education* (pp. 121–128). Retrieved from the ICCE website: http://www.icce2010.upm.edu.my/papers/c2/full%20paper/C2FP71.pdf

Wang, M., Shen, R., Novak, D., & Pan, X. (2009). The impact of mobile learning on students' learning behaviours and performance: Report from a large blended classroom. *British Journal of Educational Technology*, *40*(4), 673–695. doi:10.1111/j.1467-8535.2008.00846.x