Original article

doi: 10.1111/jcal.12173

Journal of Computer Assisted Learning

Applying augmented reality to enhance learning: a study of different teaching materials

Y.-H. Hung, C.-H. Chen & S.-W. Huang

Department of Industrial Design, National Cheng Kung University, Tainan, Taiwan

Abstract The objective of this study was to determine the usefulness of augmented reality (AR) in teaching. An experiment was conducted to examine children's learning performances, which included the number of errors they made, their ability to remember the content of what they had read and their satisfaction with the three types of teaching materials, including a picture book, physical interactions and an AR graphic book. The three teaching materials were aimed to respectively demonstrate the characteristics of six bacteria with 2D graphics, 3D physical objects, and 3D virtual objects. Seventy-two fifth-grade children were randomly selected to participate in the study, and they were divided into three groups, each of which used the assigned teaching material to learn the name of the six different bacteria in intervals of 1, 2 and 3 min. Results showed that the AR graphic book offers a practical and hands-on way for children to explore and learn about the bacteria. Follow-up interviews indicated that the children liked the AR graphic book the most, and they preferred it to the other materials.

Keywords augmented reality, learning performance, physical interaction, picture book, teaching material.

Introduction

The revolution in computer interfaces has changed the way we think about computers. Computer technologies have been introduced to educational settings and have made learning a more flexible and intuitive activity. Among these technologies, augmented reality (AR) drew huge public attention because it provides a new perspective for learning by allowing learners to visualize complex spatial relationships and abstract concepts (Lin, Chen, & Chang, 2015; Phon, Ali, & Halim, 2015).

Augmented reality is defined differently by researchers. In a broad sense, AR refers to 'augmenting natural feedback to the operator with simulated cues' (p.283) (Milgram, Takemura, Utsumi, & Kishino, 1995). It could be implemented by 'any' technology blending virtual and real information in a meaningful way (Klopfer, 2008), such as mobile devices and wearable computers with Global Positioning System (GPS). On the other hand, in a narrow sense, AR could be viewed as 'a form of virtual reality where the participants' head mounted display is transparent, allowing a clear view of the real world' (p.283) (Milgram et al., 1995). This restricted view of AR reflects the fact that the early development of AR involved head mounted/hand-held displays. It augments a view of the real world with 2D images and/or 3D objects by superimposing computer graphics over real objects. Because of the nonstop advances of hardware devices and software, looking at AR in a broad sense would help designers exploit new affordances (Cheng & Tsai, 2013) by which the real world could be leveraged to enhance teaching and learning. However, for most educators, simply utilizing mature AR technology would be beneficial as

Accepted: 12 November 2016
Correspondence: Yu-Hsiu Hung, Department of Industrial Design, National Cheng Kung University, No.1, University Rd., 701 Tainan 701, Taiwan. Email: idhfhung@mail.ncku.edu.tw

they do not have to deal with unexpected technical issues in AR implementation.

Nowadays, there are two forms of AR available to educators: marker-based AR and location-aware AR (markerless-based) (Lee, 2012). Marker-based AR presents digital information to learners after pointing their mobile device camera at a marker/QR code, while location-aware AR functions without the need for labeling or supplemental reference points. The latter presents digital information (e.g., 2D/3D/audio/video content) to learners as they move geographically with location-registered mobile devices.

Delivering AR experiences via mobile devices and computers in teaching has become much more feasible than ever before. Research has shown positive effects that AR could bring to education (Bressler & Bodzin, 2013; Chen, Lee, & Lin, 2016; Yilmaz, 2016). With regards to AR in a formal learning environment, for example, school, Huang, Li, and Fong (2016b) conducted art education workshops that offered children the opportunities to practice colouring skills with an AR colouring application, Quiver (http://quivervision.com/). All the participants were asked to finish colouring sheets where 2D objects in the sheets could turn into 3D models. Results of this study demonstrated the feasibility of integrating AR into art education. In addition, Lu and Liu (2015) developed an AR application to introduce elementary school students to the ecological environments of rivers, oceans, and especially the characteristics of fish (in 3D) inhabiting coastal areas. The student participants showed learning achievements and indicated that the use of AR was fun, engaging, and interactive. With respect to AR in an informal learning environment, for example, park, garden, and zoo, Hwang, Wu, Chen, and Tu (2015) proposed a game-based AR application that allowed students to compete with each other on differing learning tasks in a butterfly ecology garden. To demonstrate that game-based learning was more effective than traditional instruction, Hwang et al. performed a comparison of their proposed application with a nongame-based AR mobile application. Results showed that the game-based AR application improved students' learning attitudes and the knowledge of butterflies.

The aforementioned empirical AR research shows the potential of AR in supporting children's learning. However, evidence of the effects of AR appeared to be shallow (Wu, Lee, Chang, & Liang, 2013). There are research gaps remaining. First, would using AR have the same educational benefits and to what extent can AR improve learning, compared with traditional teaching materials, for example, picture books, and physical interactions? Second, many studies of AR focused on development, implementation and usability evaluation of AR tools but lacked a rigorous experimental design (Ke & Hsu, 2015; Revelle et al., 2014; Vogt & Shingles, 2013). These two gaps suggest that (1) research is needed to ascertain the characteristics of AR that distinguish it from other conventional education tools (Bacca, Baldiris, Fabregat, & Graf, 2014) and (2) more empirical data from controlled evaluation studies is required to demonstrate the practical educational values of AR (Lee, 2012; Wu et al., 2013).

To address the aforementioned research needs, Di Serio, Ibáñez, and Kloos (2013) analysed the impact of teaching with AR compared with that of teaching with PowerPoint slides on the motivation of middle-school students in learning Italian Renaissance Art. They found that students were moderately motivated when they were taught within the slide-based environment but were more motivated when they were taught within the AR learning environment. In addition, Dünser, Walker, Horner, and Bentall (2012) used their developed AR book to teach ten high school students the concepts of magnetism. They compared the effectiveness of the AR book with a traditional book with text and illustrations. Their data showed that the AR book was more effective. Moreover, recently, Lin et al. (2016) demonstrated that children with disabilities had higher success rates of completing tangram puzzles (the traditional Chinese tangram and square tangram) with the assistance of AR display technology as opposed to conventional paper-based methods.

Despite the aforementioned evidence suggesting the benefits of AR in education, in the context of primary education, the efficacy and satisfaction of AR in comparison with 'the most used' classroom teaching materials – both picture books and physical interactions – (especially physical interactions) have not been explored. Most AR studies lack a rigorous experimental design that can provide evidence on the practical values of AR in education (Wu et al., 2013). Thus, the purpose of this study was to develop an experiment to assess the educational effectiveness of AR in the elementary schools, where most teachers extensively adopt picture books and physical interactions in their teaching (Ministry of Education, 2014). The research questions are as follows:

1. Can AR improve students' learning performances compared with the conventional teaching materials – picture books and physical interactions?
2. Can AR arouse more interests/motivations in students than the conventional teaching materials – picture books and physical interactions?

AR has stepped out of university labs and is now commercially available in the customer market, such as the IKEA AR furniture catalogue and the converse shoe sampler. This study is important in that (1) there has been a tendency to conduct AR studies in higher education settings, while children should be explored more for their use of AR (Bacca et al., 2014); (2) AR is ready to be massively integrated into the classroom settings (Thornton, Ernst, & Clark, 2012). Comparing the effectiveness and efficacy of AR with the most used conventional teaching materials is thus imperative. The outcomes of this study will provide insight for primary school educators regarding the adoption of AR in their teaching.

The theoretical basis and the potential of AR in education

Enhancement of spatial perception and ability

Spatial cognition is concerned with how people acquire and use knowledge about their environment (e.g., procedural/declarative/configurational knowledge) to determine where they are, how to obtain resources and how to find their way home (Denis & Loomis, 2007). The capabilities of acquiring, organizing and revising this knowledge enable humans to manage cognitive tasks in everyday life. In the context of classroom learning, children oftentimes engage in activities and cognitive processes that require them to store, recall, create and communicate about perceived spatial images (Osberg, 1997). Manipulating AR interfaces could involve the exploration and development of route/landmark/survey knowledge (Slijepcevic, 2013), as well as spatial perception and visualization (Thornton et al., 2012). It thus fosters students' spatial ability, practical skills and conceptual understanding (Ibáñez, Di Serio, Villarán, & Kloos, 2014). A good example of using AR to enhance children's spatial ability is *Cyberchase Shape Quest* released by PBS (Public Broadcasting Service) Kids. *Cyberchase Shape Quest* is a free mobile AR application that exposes children to 3D immersed puzzle worlds and engages children's spatial memory, planning and visualization skills through various levels of games (Radu, Doherty, DiQuollo, McCarthy, & Tiu, 2015). To play the game, children need to point their mobile device at a piece of paper that contains a marker image. The piece of paper becomes augmented with a 3D game world. Children can move the device around or up and down the paper, and their perspective would change and shift accordingly. Interacting with the virtual objects and solving the puzzles require children to use spatial perception and visualization skills to estimate absolute distances and scales among objects and points, and to reason about the properties of the shapes. Another good example is the AR mobile system developed by Chang, Hou, Pan, Sung, and Chang (2015). The main feature of this AR system is to provide guidance for student visitors to learn about cultural heritages through interactive texts, images and audio information. The virtual multimedia information gives users the opportunity to explore and acquire the route/landmark/survey knowledge on the geographic space of the neighbouring heritage sites. Chang et al. demonstrated that guidance using AR compared with guidance using audio was able to enhance visitors' sense of place more effectively.

Spatial ability also refers to forming and transforming visual images as an organized whole (Denis & Loomis, 2007). Spatial skills, especially the abilities to mentally transform objects, to a certain degree predict a young kid's future achievements as a designer, engineer, architect, or surgeon (Wai, Lubinski, & Bendow, 2009). Studies also found that spatial ability influences academic performance and can be increased through instructions (Potter & Van der Merwe, 2003). For example, in geometry lessons, students who studied transformations (flips, slides and turns, etc.) improved spatial ability more than those who only studied names for shapes (Clements & Battista, 1992). In engineering graphics lessons, Martín-Gutiérrez et al. (2010) developed an AR book that allowed students to see different perspectives of 3D virtual objects and complementary information for problem resolution. The problems were designed to train students' mental rotation abilities. They asked students to identify orthographic and axonometric views as well as projections of 3D virtual objects popped up in the AR book. Martín-Gutiérrez et al. conducted a validation study with 24 students and confirmed that the AR training had a positive impact on students' spatial ability improvement.

Multi-perceptual experiential learning

Augmented reality as a setting has the potential for multiperceptual experiential learning. According to the multiple resource theory (Wickens, Hollands, Banbury, & Parasuraman, 2016), humans do not have one single information processing source that can be tapped but have several different pools of resources (i.e., visual/auditory/textual/olfactory) that can be tapped simultaneously. Interacting with virtual objects in the AR environment essentially does not cause cognitive challenges as the mental resources underlying perception (i.e., visual/perceptual processing of virtual information) and those underlying the action selection/execution (i.e., psychomotor activities) are different/separate (Radu & MacIntyre, 2012).

Implementing AR in education allows children to exercise differing cognitive functions and coordinate perceptual resources with physical actions, which therefore promotes learning. Users are able to gain haptic spatial knowledge (based on touch and physical movement) through in-place physical manipulations and gain pictorial spatial knowledge (based on information in visual form) because stimuli are three-dimensionally converted to visual experiences (Shelton & Hedley, 2004; Slijepcevic, 2013). The use of both haptic and pictorial cognitive resources triggers trans-perceptual resources, which would help to develop integrated spatial knowledge and lead to more rapid and accurate perception of 3D content (Mark, 1993). The trans-perceptual knowledge initially is formed from the synthesis of the haptic and the pictorial knowledge gained in the AR environment. It is then integrated with the existing mental models and schemas to form a stable internal representation of spatial knowledge continuously updated by adding information and experiences to existing ones. Such information encoding process thus enhances the understanding of the world (Haken, 1991).

The AR interface lends itself well to task-related learning because of the exclusive connectivity between short cycles of visual perceptual activity and physical movements (Radu & MacIntyre, 2012). This provides the user with advantages for action in the world and physical processes that involve action. While multiple sensory exploration of an object is valuable for triangulating perceived information, visual perception is most effective during motion. Thus, AR provides us opportunities to experience information through visual, spatial and sensorimotor feedback in response to interface manipulations.

Method

Objective

The objective of the study was to determine the usefulness of AR in teaching and to investigate which teaching material – an AR graphic book, a picture book or physical interactions – facilitates children's learning the most. The presumption is that a teaching material that uses AR can facilitate children's learning and improve children's performance better than the most used picture books and physical interactions. In this study, children's learning performance was assessed through how much participants were able to retain the provided information in memory.

Participants

Purposeful sampling (Patton, 1990) was employed to recruit participants. The sampling criteria were as follows: (1) our study randomly selected 72 fifth-grade students from two elementary schools (Chong-Ming Elementary School & Sheng-Li Elementary School) in Tainan, Taiwan, to participate in the experiment. The age range was 12–13 years old (as mental visualization abilities develop until late childhood) (Rosser, 1994); (2) the participants had never learned the six bacteria before; (3) the participants had to have completed the basic computer skill curriculum that the school assigned for their grade. In addition, the participants were randomly divided into three groups: each of which manipulated, read or participated with the assigned teaching material, respectively. Participants in each group used only one type of teaching materials. Each group was then divided into three subgroups, each of which respectively used the assigned teaching material for the designated time interval.

The advantage of using purposeful sampling was that we were able to gain insight from a sample that (1) had a similar level of computer literacy and (2) was representative of children whose mental abilities were mature enough for absorbing knowledge from using AR designs.

Equipment/Material

The experiment was held in the natural classrooms in Chong-Ming/Sheng-Li Elementary Schools in Tainan, Taiwan, which provided comfortable environments and adequate lighting.

The experiment required the participants to be familiar with the characteristics of six bacteria, which included: myxovirus, helicobacter pylori, rhinovirus, *Saccharomyces cerevisiae*, shigella and dysenteriae. The reasons for choosing six were (1) children have a relatively limited capacity of working memory (about four items) (Rosser, 1994) when compared with adults (generally being able to hold seven plus and minus two items in working memory) (Miller, 1956; Wickens et al., 2016). Six stimuli (including spatial features of bacteria) keep the level of task difficulty a bit beyond participants' limits; (2) one reported limitation of AR for children has been cognitive overload (Dunleavy and Dede, 2014). We did not want to overwhelm/overload participants with the complexity of experimental tasks as children' learning achievements and attitudes would decrease if the amount of digital learning content and the real world environment was not carefully considered (Chu, 2014).

The participants were instructed to use the assigned teaching material within the designated time interval. Overall, four sessions were included in the experiment: preparation, error test, retention test and satisfaction investigation. The equipment settings are described as follows:

1. AR graphic book: This study used ARToolKit (ARToolkit, 2016), a software library that involves the overlay of virtual imagery on the real world, to develop the content and settings of the AR graphic book. The AR graphic book (Figure 1), *knowing bacteria*, had six pages, and each page contained a tracking marker, the name of a bacterium and a description of the bacterium. Figure 2 illustrates the AR equipment settings. The participants wore a virtual reality head mounted display (HMD) with a small video camera attached at the level of their eyes. This virtual reality HMD included a built-in colour camera that captured the image of the AR graphic book and its surroundings, as well as two micro-displays that presented the image of the AR graphic book and its surroundings to the participants. The images captured by the camera were processed by the computer, which recognized the marker in the image and overlaid the corresponding bacterium. Each bacterium had its own animation file stored in the computer. The Power Unit contained a cable for signal transmission, a socket connector for external image sources (SVGA from computer), a socket connector for USB camera signal (output), a push button for power and a connector for power supply.
2. Picture book: The picture book contained six pages, and each page included a graphic of a yellow bacterium, a name and the characteristics of the bacterium. Figure 3 shows the six bacteria and the picture book.
3. Physical interactions: The physical interactions (Figure 4) contained six physical objects, six boxes

Figure 1 An Example of the Real View and the Augmented View of the AR Graphic Book. [Colour figure can be viewed at wileyonlinelibrary.com]

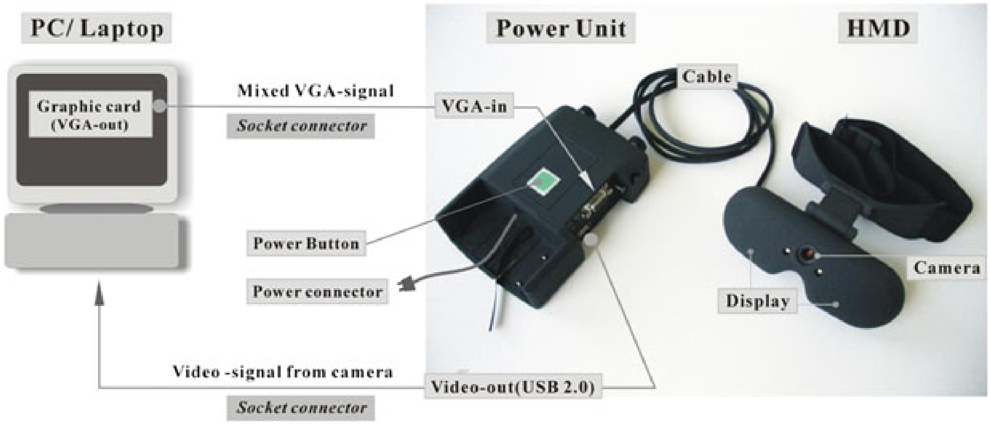

Figure 2 The Augmented Reality (AR) Equipment Settings (video-see-through USB Camera). [Colour figure can be viewed at wileyonlinelibrary.com]

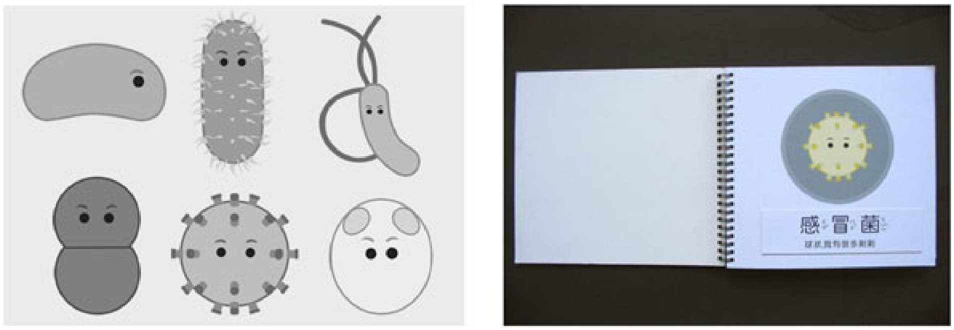

Figure 3 Six Bacteria and the Picture Book. [Colour figure can be viewed at wileyonlinelibrary.com]

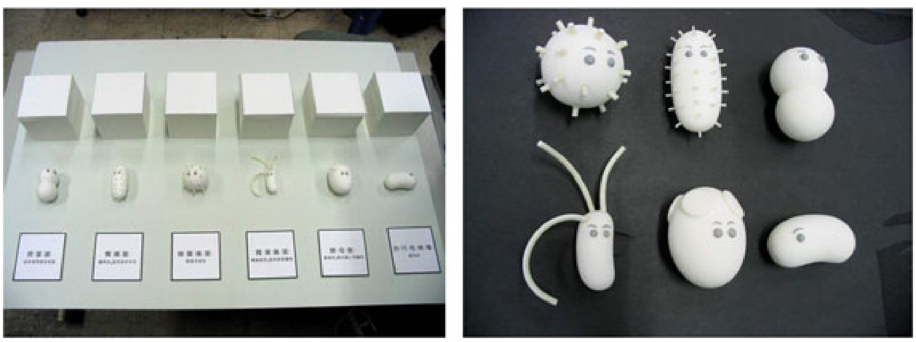

Figure 4 Six Physical Objects, Boxes, and Cards. [Colour figure can be viewed at wileyonlinelibrary.com]

to hide the physical objects and six cards to describe the characteristics of the bacteria.

Variables

Learning is an active process that involves sensory input to the brain and information transfer from sensory input to working and long-term memory (Rosser, 1994). According to Wickens et al. (2016), in general, when paying attention to the perceived information long enough (without interruption), it will be transferred to working and long-term memory, where learning potentially takes place. In this study, we posited that the effectiveness of a teaching material depends on 'how well' the teaching material facilitates the process of encoding information into the working and long-term memory. We defined students' learning performance given a teaching material as how much information is retained in mind after attempting to remember it for a certain amount of time.

Because the effectiveness/satisfaction of AR in comparison with 'the most used' classroom teaching materials (especially physical interactions) has not been explored in the literature, the aim of the experiment was to examine the usefulness of AR in teaching and to investigate which teaching material – an AR graphic book, a picture book or physical interactions – may facilitate children's learning the most. The independent variables were set to be the teaching materials and the time intervals. The dependent variables were error, satisfaction and retention. The independent and dependent variables are described as follows:

1. Independent variables

- *Teaching material*: Children's learning performance with three types of teaching materials was investigated, which included a picture book, physical interactions and an AR graphic book.
- *Time interval*: The participants were divided into groups and instructed to use the AR graphic book and the picture book and to manipulate the bacterium for intervals of 1, 2 and 3 min (because of the small number of stimuli, i.e., six bacteria). The time intervals were much longer than the decay time of working memory (typically between 5 and 10 s) (Cowan, 2008). This design allowed us to understand as time progresses 'how well' AR facilitates learning when compared with conventional teaching materials.

2. Dependent variables

- *Error*: The number of errors the children committed when attempting to identify the bacteria helped us know whether or not they had learned the characteristics of the bacteria.
- *Retention*: The number of characteristics of the bacteria the children still remembered after 1 h (much greater than working memory decay limits which suggests the occurrence of learning) helped us know which teaching material facilitated children's learning better.
- *Satisfaction*: The degree of satisfaction helped us realize each child's attitude, degree of acceptability and preference for the designated teaching material.

Procedures

Session 1 – Preparation

Before conducting the experiment, the following instructions were given: 'The teaching materials you will be using contain six tiny creatures, i.e., bacteria, which exist in everyday life. They look different from each other and have their own special characteristics. For example, one bacterium might have a circular body, while another bacterium might have an oval one. One bacterium might have short tentacles while another bacterium might have long hairs. Please try your best to familiarize yourself with all of the bacteria within the designated time interval. When you finish using the teaching material, you will be tested to see how many bacteria you can identify. If you do not have questions about the experiment and are ready to begin, please adjust your seat to the most comfortable position. We will begin now to learn about these tiny creatures.' The purpose of the instruction was to prepare the participants for the experiment and encourage them to use the teaching materials in the most natural way.

The participants were then instructed to use the assigned teaching material within the designated time interval. The participants who read the picture book or manipulated the AR graphic book were required to review the bacteria from pages one to six. The participants who engaged in the physical interactions were required to follow the same procedure as those who read the picture book or manipulated the AR graphic book. Basically, they started with the box on the left, picking it up, manipulating the bacterium and placing it back down to hide the bacterium until they had completed all of the six boxes from left to right. Figure 5 shows that the participants used different teaching materials.

There were two reasons that the experiment used the boxes to hide the physical objects: (1) The way children read the picture book or manipulated the AR graphic book was page by page and one bacterium at a time. Hence, by requiring the participants to pick up one box at a time ensured that the participants received one stimulus per bacterium; (2) the time interval from picking up

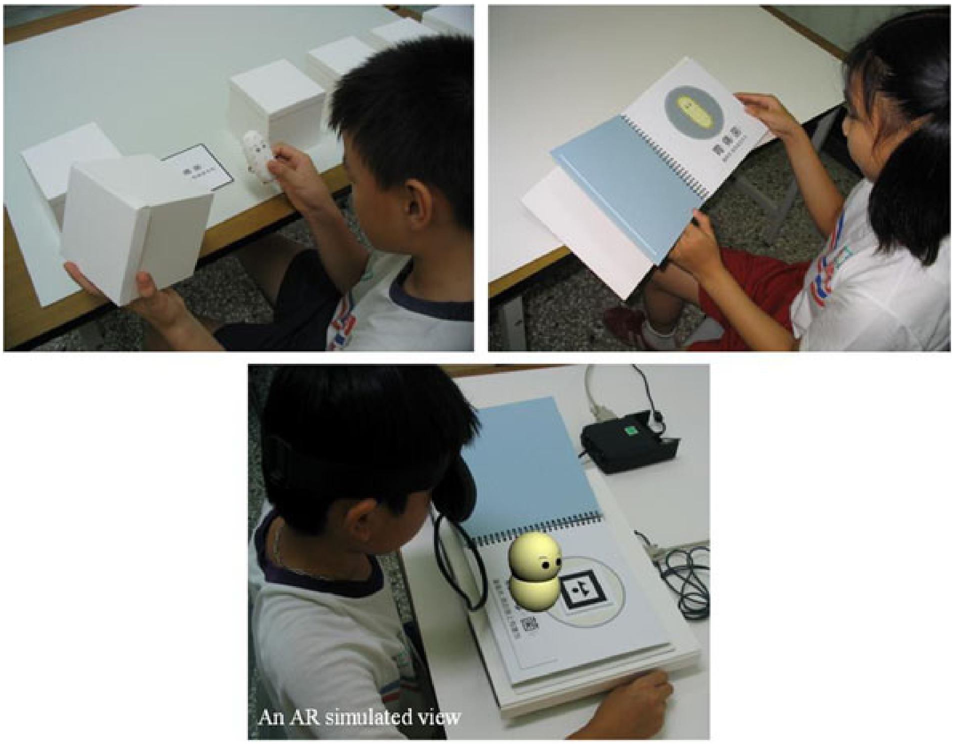

Figure 5 The Participants Used Different Teaching Materials. [Colour figure can be viewed at wileyonlinelibrary.com]

the box to placing it back over the bacterium was approximately equal to that of flipping the six book pages. The difference of the time intervals was very short and did not confound the experiment.

Session 2 – Error test

This phase examined error, one of the dependent variables. After the participants finished using the teaching materials, they were shown a screen of the six bacteria in an order different from that in which they were first presented. Then the participants received six questions, each of which included the name and description of a bacterium, and the participants were asked to identify the bacterium on the screen in front of them. The researchers recorded the participants' answers for further analysis.

Session 3 – Retention test

This phase examined retention, another dependent variable. The participants were asked to retake the error test one hour after the experiment. The participants' answers to the error test were recorded for future analysis.

Session 4 – Satisfaction investigation

This phase examined satisfaction, another dependent variable. The participants were asked to answer a five-statement satisfaction questionnaire regarding the teaching material they had used in the experiment. The five statements are listed below:

- I am interested in using this teaching material.
- This teaching material helps me learn new subjects in class.
- I like this teaching material.
- I hope my teachers can adopt this teaching material in class in the future.
- I can easily identify a bacterium given the name and description of a bacterium.

The participants were required to rate the degree of agreement with each of the five statements based on a Likert type scale (i.e., smiley face-based scale) adapted from Rambli, Matcha, and Sulaiman (2013), ranging from one to seven, from *strongly disagree* (i.e., crying face) to *strongly agree* (i.e., smiley face). Every statement was presented full-screen on a laptop. The participants could either point to the rating on the screen or say their rating aloud for the researchers to record. In addition, because gaining qualitative feedback is important for knowing the differences of the teaching materials, the researchers then conducted retrospective interviews (Hewett & Scott, 1987) of the participants. The interview was short, informal and interactive with the purpose of knowing the participants' first impression and intuitive experiences of the teaching material they used in the experiment. The interview was held in the same location as the experiment. The participants' comments about the teaching materials were recorded as well for future analysis.

Data analysis

ANOVA tests and multiple comparison analyses were employed to examine any significant difference in the participants' performances (rate of error, degree of satisfaction and retention) as a result of our using different teaching materials within different time intervals.

Results

The analyses of the error test and retention test are summarized in Table 1. With regards to the dependent variable, error, the two-way ANOVA test shows that the time interval and teaching material significantly affected the number of errors the participants made in the error test. Their interaction effect is not significant but with a *p* value close to 0.05, meaning that there is some evidence of mild interaction between the time interval and teaching material. The Tukey's studentized range test indicates that the mean errors of the participants who used the teaching materials for 1 min are significantly higher than those who used the teaching materials for 2 or 3 min. Additionally, no significant differences are found with respect to the mean errors between the participants who manipulated the AR graphic book and those who read the picture book, or between the participants who

Table 1. The Two Way ANOVA Tests for the Errors in the Recognition Test and the Retention Test (extracted from SAS)

Dependent variable	Source of variation	Treatment	Mean	DF	SS	MS	F	*p* value
Error	Time interval	1 min	2.08 [a]	2	24.53	12.26	18.79	0.000
		2 min	1.21 [b]					
		3 min	0.67 [b]					
	Teaching material	AR graphic book	1.71 [a]	2	5.78	2.89	4.43	0.016
		Picture book	1.21 [a b]					
		Physical interaction	1.04 [b]					
	Interaction			4	6.22	1.56	2.38	0.061
	Error			63	41.13	0.65		
	Total			71	77.65			
Retention	Time interval	1 min	2.29 [a]	2	23.03	11.51	7.75	0.001
		2 min	1.75 [a b]					
		3 min	0.92 [b]					
	Teaching material	AR graphic book	1.96 [a]	2	4.11	2.06	1.38	0.258
		Picture book	1.63 [a]					
		Physical interaction	1.38 [a]					
	Interaction			4	11.56	2.89	1.94	0.114
	Error			63	93.63	1.49		
	Total			71	132.32			

Note. Means with the same letter(a or b) are not significantly different using Tukey's Studentized Range Test. AR = augmented reality.

read the picture book and those who participated in the physical interactions.

With regards to the dependent variable, retention, the two-way ANOVA test shows that only the time interval significantly affected the number of errors the participants made in the retention test, with a p-value less than 0.05, and that no interaction effect exists between the teaching material and the time interval. The Tukey's studentized range test indicates that the mean errors are the same for the participants who used the teaching materials for 1 and 2 min and also the same for 2 and 3 min. In addition, no significant differences are found with respect to the mean errors among the three groups of participants subject to different teaching materials.

With regards to the dependent variable, satisfaction, we performed a reliability test on the satisfaction questionnaire. The Cronbach's alpha is 0.73, indicating that the participants' responses across the board are strongly consistent. The analysis of the satisfaction questionnaire is shown in Table 2. The one-way ANOVA test shows that the teaching material affect the ratings. The Tukey's studentized range test indicates that, in terms of Statements 2 and 5, no statistical differences exist. In terms of Statement 1, no differences are found in the ratings of the participants who manipulated the AR graphic book and those who read the picture book, or in the ratings of the participants who read the picture book and those who participated in the physical interactions. In terms of Statement 3, no differences are found in the ratings of the participants who manipulated the AR graphic book and those who participated in the physical interaction or in the ratings of the participants who participated in the physical interactions and those who read the picture book. In terms of Statement 4, it was found that the participants who manipulated the AR graphic book and those who read the picture book gave higher ratings than those who participated in the physical interactions. Overall, the degree of the participants' agreement on the questionnaire is from the lowest 4.92 to the highest 6.29, which indicates that the participants positively agreed with the statements.

Table 2. One Way ANOVA Tests for the Ratings to the five-statement Satisfaction Questionnaire (extracted from SAS)

Statement #	Teaching material	Mean	F	p value
1	AR graphic book	6.13 [a]	3.66	0.031
	Picture book	5.46 [ab]		
	Physical interaction	5.33 [b]		
2	AR graphic book	6.25 [a]	2.93	0.060
	Picture book	5.67 [a]		
	Physical interaction	6.21 [a]		
3	AR graphic book	6.29 [a]	8.61	0.000
	Picture book	5.00 [b]		
	Physical interaction	5.67 [ab]		
4	AR graphic book	5.83 [a]	8.68	0.000
	Picture book	5.79 [a]		
	Physical interaction	4.92 [b]		
5	AR graphic book	5.67 [a]	0.35	0.704
	Picture book	5.79 [a]		
	Physical interaction	5.46 [a]		

Note. Means with the same letter (a or b) are not significantly different using Tukey's Studentized Range Test. AR = augmented reality.

Discussion

Error

From Table 1, both the time interval and teaching material affected the number of errors made by the participants.

The Tukey's studentized range test for the time interval shows that, regardless of the type of the teaching materials, the participants' errors for 1 min were more than the errors made in 2 or 3 min and that the number of errors made in 2 and 3 min were the same. This suggests that the participants could have all reached the same level of learning performance had they been given enough time. In addition, the test indicates that, regardless of the different time intervals, the errors made by the participants who manipulated the AR graphic book were similar to those made by the participants who read the picture book. Even though there is no evidence showing that the errors made by the participants who read the AR graphic book were similar to those made by the participants who participated in the physical interactions, we can at least infer that the educational benefits of AR are no less than those of the traditional teaching material – a picture book.

Figure 6 assists in interpreting the results of the error test. In general, the participants who participated in the

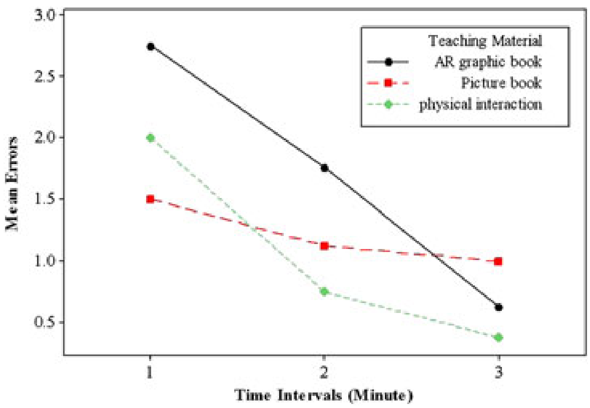

Figure 6 Time Intervals – Teaching Materials Plot for the Errors in the Error Test. [Colour figure can be viewed at wileyonlinelibrary.com]

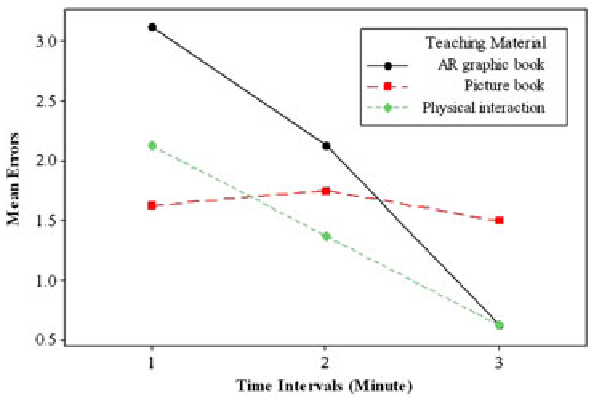

Figure 7 Time Intervals – Teaching Materials Plot for the Errors in the Retention Test. [Colour figure can be viewed at wileyonlinelibrary.com]

physical interactions seem to have the best performance. Additionally, based on the slope changes of different teaching materials, the slopes for the AR graphic book seem steeper than those of the picture book and physical interactions, which indicates that the mean errors of the participants who manipulated the AR graphic book drop most rapidly. A further study is desired to confirm this trend because only three time intervals were investigated in this study.

The error test verified the efficacy of AR in education when compared with the traditional picture book. These findings resonate with the findings in Mohammadyari and Singh (2015), who noted that integrating technology into classroom learning can lower the barrier of differentiated teaching and help overcome time and space constraints of the traditional classroom setting. It moves students from passively receiving knowledge to actively learning and gaining knowledge. From the perspective of active learning, AR does provide learners with a broader range of opportunities to acquire knowledge.

Retention

Table 1 indicates generally the participants made more errors in the retention test than those made in the error test. The two-way ANOVA test shows that the time interval had effects on the number of errors made by the participants, whereas the type of teaching material did not affect the number of errors made by the participants. Results of the Tukey's studentized range test for the time interval are similar to what had been obtained from the error test, meaning that the participants could have performed better had they been given more time. Hence, we conclude that even though AR teaching materials may not be superior to traditional teaching materials in terms of educational benefits, its effectiveness explains its real world applicability when compared with the most used teaching materials in classroom.

Figure 7 shows the results of the retention test. Again, the participants who participated in the physical interactions seem to have the best performance. Additionally, even though the teaching material did not affect the number of errors made by the participants, based on the slope changes of different teaching materials, the slopes for the AR graphic book and physical interactions drop dramatically, whereas the slopes for the picture book remain steady. A further study is desired to confirm this trend because only three time intervals were investigated in this study.

Our findings demonstrate the efficacy of AR as also suggested in Jadeja, Mehta, and Sharma (2016), who highlighted utterances of a student participant, '*Tell me and I forget. Show me and I remember, Involve me and I understand*' (p.81). Results of the analyses can answer the first research question: AR graphic book offers a practical and hands-on way for children to explore and learn about the bacteria, and AR improves students' learning performances to a similar extent as the conventional teaching materials, such as picture books and physical interactions do.

Satisfaction

In general, the participants agreed with all the statements and tended to rate them high. The one-way ANOVA test in Table 2 suggests that the type of teaching material had an impact on the participants' ratings for Statements 1, 3 and 4.

Looking at the Tukey's studentized range test for these statements, we found that Statements 1, 3 and 4 received higher ratings from the participants who manipulated the AR graphic book, suggesting that the participants liked the AR graphic book more than the other two teaching materials, and that they were more interested in using the AR graphic book in classroom than the other two teaching materials. These results support the adoption of AR in teaching. Furthermore, the participants who manipulated the AR graphic book said in the follow-up interviews that even though they had adjusted to traditional teaching materials in their classes, they would love to try to use AR because they felt that the AR graphic book was really neat and amazing, which greatly motivated them to learn. During the experiment, some of the participants even requested to use AR repeatedly or to take the AR book home.

No statistical differences were identified among the ratings for Statements 2 and 5. We conclude that the participants believed that they could choose the right answers on the error and retention tests after using their assigned teaching material. As a matter of fact, most of the participants in the follow-up interviews agreed that the teaching materials that they used facilitated learning. However, some of the participants, especially those who manipulated the AR graphic book, complained that they could not easily choose the right answers in the error and retention tests because it was hard for them to convert the 3D bacteria in their minds to the 2D bacteria on the screen, which might have led them to make more errors and this became a limitation of this study.

Interest and motivation are shown to promote learning performance (Abrantes, Seabra, & Lages, 2007). According to Baron, Vandello, and Brunsman (1996), when people experience positive emotion, they are more likely to view surrounding people and things with a pleasant state of mind. As positive emotion contributes to learning (Pekrun, 2000), the results of the satisfaction investigation answer the second research question, which states that AR arouses more interests/motivations in students than the most conventional teaching materials – picture books and physical interactions. To conclude, in order to both facilitate children's learning and increase their learning motivations in class, AR teaching material is a good alternative to conventional picture book and physical interactions.

Conclusions

The purpose of this study was to determine the usefulness of AR in teaching in comparison with two conventional teaching materials. An experiment was conducted to examine the children's learning performance, which included the number of errors they made, their ability to remember what they had learned and their satisfaction with the three types of teaching materials, including a picture book, physical interactions and an AR graphic book. Because of the age-related and individual differences in children's working memory performance (Swanson, 1996), one limitation of our study is that we provided only six bacteria for the participants and examined their performances at limited time intervals. Results showed that, after all, AR produced similar educational benefits as those traditional teaching materials did. The follow-up interviews indicated that the participants had great satisfaction and interests in using AR in class. These results answer the research questions and confirm our assumption that AR enables educators to approximate the learning performance of children using picture books or interacting with actual physical objects. Another limitation of our study is that we did not look into the effects of eye-strain and motion sickness to children's learning (Fruland, 2002). However, none of the participants reported these problems. The reason might be that the AR graphic book interested them, and they did not use the HMD long enough to cause ill physical effects.

In addition, field studies are often challenged by the sample size, owing to the realities and dynamics of data collection in a natural classroom environment. As such, statistical validity and external validity are not always as strong as what might be found in larger sample studies in laboratories. However, we were not met with this limitation in our study. First, the demarcation of a small sample size for ANOVA is about 5 per group (Norman, 2010) (smaller than 8 in this study). Second, we conducted a *post hoc* power analysis of the variables in SAS 9.4. (SAS Institute Inc.) the results showed that the power ranges for teaching material, time interval and the interaction between teaching material and time interval are from 0.69 to 0.99. The minimum power of the analyses is considered good in social sciences; the maximum is considered a high level of statistical power (Cohen, 1988). Thus, the strength of our study was the ability to draw conclusions based on a sample size that minimizes the probability of Type II error. Moreover, oursampling strategy (purposeful sampling of participants with a good level of maturity in cognition and computer literacy) allowed us to reduce the variability of the interested child population and to increase the representativeness.

Although AR did not seem superior in this study, it has been known that 3D visual/spatial aids can enhance students' understanding of a concept or process (Huang, Chen, & Chou, 2016a). AR can help students learn about not only biology (as investigated in this study) but also other subject domains in education. For instance, in ecology, AR provides an effective way to help students focus on what they need to observe in the field (Hwang et al., 2015). With respect to language learning, Chen, Su, Lee, and Wu (2007) developed an AR system that allowed children to learn Chinese characters, pronunciations and corresponding objects by putting together virtual Mandarin phonetic symbols. In the field of mathematics, Quintero, Salinas, González-Mendívil, and Ramírez (2015) used AR (instead of blackboard or paper) for the learning of Calculus. In their developed AR application, students could manipulate motion of time and parameters in algebraic expressions to see and experience how parabolas or sinusoidal curves evolve into 3D objects. Quintero showed that students not only were able to learn the concepts of Calculus, but also developed mental spatial skills. Thus, AR does have the potential to be incorporated into the classroom as it brings experiential and location-based learning to students.

AR has been widely applied in the field of education. It superimposes virtual concepts or objects over the physical world in front of users and creates new user computer interactions. Studies showed that interest increases learning (Abrantes et al., 2007), and promoting interest in classroom increases students' intrinsic motivation to learn (Hwang et al., 2015). The contribution of our study is that we confirm AR, which enables children to interact virtually with the physical world, offers a practical and hands-on way for children to learn in class. The statement we are trying to make is AR is not a technology trying to replace children's learning habits. Rather, if children feel bored about reading textbooks every day, why not give them something interesting which can be equally effective in enhancing their learning?

Currently, we are extending the framework of our study by trying to add more multimedia interactions between children and virtual content of AR teaching materials. The participants, instead of manipulating AR graphic books, use a tracking maker to trigger the virtual content of another marker. We seek to find new applications of AR in classroom environments. Nowadays, children's use of technology becomes more prevalent, and the ways educators employ technologies in education to improve learning are going to be more diverse. We believe in the future that there will be more researchers devoting themselves to the research of AR in education.

References

Abrantes, J. L., Seabra, C., & Lages, L. F. (2007). Pedagogical affect, student interest, and learning performance. *Journal of Business Research*, *60*, 960–964.

ARToolKit. (2016). Retrieved Jul, 27, 2016, from http://www.hitl.washington.edu/artoolkit/

Bacca, J., Baldiris, S., Fabregat, R., & Graf, S. (2014). Augmented reality trends in education: A systematic review of research and applications. *Journal of Educational Technology & Society*, *17*, 133–149.

Baron, R. S., Vandello, J. A., & Brunsman, B. (1996). The forgotten variable in conformity research: Impact of task importance on social influence. *Journal of Personality and Social Psychology*, *71*, 915–927.

Bressler, D. M., & Bodzin, A. M. (2013). A mixed methods assessment of students' flow experiences during a mobile augmented reality science game. *Journal of Computer Assisted Learning*, *29*, 505–517.

Chang, Y. L., Hou, H. T., Pan, C. Y., Sung, Y. T., & Chang, K. E. (2015). Apply an augmented reality in a mobile guidance to increase sense of place for heritage places. *Educational Technology & Society*, *18*, 166–178.

Chen, C. H., Su, C. C., Lee, P. Y., & Wu, F. G. (2007). Augmented interface for children Chinese learning, *Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007)*, IEEE, pp. 268–270.

Chen, C. H., Lee, I. J., & Lin, L. Y. (2016). Augmented reality-based video-modeling storybook of nonverbal facial cues for children with autism spectrum disorder to improve their perceptions and judgments of facial expressions and emotions. *Computers in Human Behavior*, *55*, 477–485.

Cheng, K. H., & Tsai, C. C. (2013). Affordances of augmented reality in science learning: Suggestions for future research. *Journal of Science Education and Technology*, *22*, 449–462.

Chu, H. C. (2014). Potential negative effects of mobile learning on students' learning achievement and cognitive load-a format assessment perspective. *Educational Technology & Society*, *17*, 332–344.

Clements, D. H., & Battista, M. T. (1992). Geometry and spatial reasoning. In D. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 420–464). New York: Macmillan.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, *169*, 323–338.

Denis, M., & Loomis, J. M. (2007). Perspectives on human spatial cognition: Memory, navigation, and environmental learning. *Psychological Research*, *71*, 235–239.

Di Serio, Á., Ibáñez, M. B., & Kloos, C. D. (2013). Impact of an augmented reality system on students' motivation for a visual art course. *Computers & Education*, *68*, 586–596.

Dunleavy, M., & Dede, C. (2014). Augmented reality teaching and learning. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 735–745). New York: Springer.

Dünser, A., Walker, L., Horner, H., & Bentall, D. (2012). Creating interactive physics education books with augmented reality. In *Proceedings of the 24th Australian computer-human interaction conference*, ACM, pp. 107–114.

Fruland, R. M. (2002). Using immersive scientific visualizations for science inquiry: Co-construction of knowledge by middle and high school students. *Annual Meeting of the American Educational Research Association*, New Orleans, USA.

Haken, H. (1991). *Synergetic computers and cognition-a top-down approach to neural nets*. Berlin etc.: Springer-Verlag.

Hewett, T. T., & Scott, S. (1987). The use of thinking-out-loud and protocol analysis in development of a process model of interactive database searching. *Human Computer Interaction–INTERACT*, *87*, 51.

Huang, T. C., Chen, C. C., & Chou, Y. W. (2016a). Animating eco-education: To see, feel, and discover in an augmented reality-based experiential learning environment. *Computers & Education*, *96*, 72–82.

Huang, Y., Li, H., & Fong, R. (2016b). Using augmented reality in early art education: A case study in Hong Kong kindergarten. *Early Child Development and Care*, *186*, 879–894.

Hwang, G. J., Wu, P. H., Chen, C. C., & Tu, N. T. (2015). Effects of an augmented reality-based educational game on students' learning achievements and attitudes in real-world observations. *Interactive Learning Environments*, 1–12.

Ibáñez, M. B., Di Serio, Á., Villarán, D., & Kloos, C. D. (2014). Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness. *Computers & Education*, *71*, 1–13.

Jadeja, A., Mehta, R., & Sharma, D. (2016). New era of teaching learning: 3D Marker based augmented reality. *International Journal of Information*, *6*, 81–88.

Ke, F., & Hsu, Y. C. (2015). Mobile augmented-reality artifact creation as a component of mobile computer-supported collaborative learning. *The Internet and Higher Education*, *26*, 33–41.

Klopfer, E. (2008). *Augmented learning: Research and design of mobile educational games*. Cambridge, MA: MIT Press.

Lee, K. (2012). Augmented reality in education and training. *TechTrends*, *56*(2), 13–21.

Lin, C. Y., Chai, H. C., Wang, J. Y., Chen, C. J., Liu, Y. H., Chen, C. W., … Huang, Y. M. (2016). Augmented reality in educational activities for children with disabilities. *Displays*, *42*, 51–54.

Lin, H. C. K., Chen, M. C., & Chang, C. K. (2015). Assessing the effectiveness of learning solid geometry by using an augmented reality-assisted learning system. *Interactive Learning Environments*, *23*, 799–810.

Lu, S. J., & Liu, Y. C. (2015). Integrating augmented reality technology to enhance children's learning in marine education. *Environmental Education Research*, *21*, 525–541.

Mark, D. M. (1993). Human spatial cognition. In D. Medychkyj-Scott & H. M. Hearnshaw (Eds.), *Human factors in geographical information systems* (pp. 51–60). London: Bellhaven Press.

Martín-Gutiérrez, J., Saorín, J. L., Contero, M., Alcañiz, M., Pérez-López, D. C., & Ortega, M. (2010). Design and validation of an augmented book for spatial abilities development in engineering students. *Computers & Graphics*, *34*, 77–91.

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995). Augmented reality: A class of displays on the reality-virtuality continuum. *Photonics for Industrial Applications*, International Society for Optics and Photonics, pp. 282–292.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Ministry of Education (2014). Grade 1–12 Curriculum Guidelines: Republic of China (Taiwan). Retrieved from http://www.naer.edu.tw/files/15-1000-7944,c639-1.php?Lang=zh-tw [Last accessed 26 November 2016].

Mohammadyari, S., & Singh, H. (2015). Understanding the effect of e-learning on individual performance: The role of digital literacy. *Computers & Education*, *82*, 11–25.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, *15*, 625–632.

Osberg, K. M. (1997). *Spatial Cognition in the Virtual Environment, Technical R-97-18*. Seattle: HIT Lab.

Patton, M. Q. (1990). *Qualitative Evaluation and Research Methods*. Newbury Park, CA: Sage.

Pekrun, R. (2000). A social cognitive, control-value theory of achievement emotions. In J. Heckhausen (Ed.), *Motivational psychology of human development*. Oxford, UK: Elsevier.

Phon, D. N. E., Ali, M. B., & Halim, N. D. A. (2015). Learning with augmented reality: Effects toward student with different spatial abilities. *Advanced Science Letters*, *21*, 2200–2204.

Potter, C., & Van der Merwe, E. (2003). Perception, imagery, visualization and engineering graphics. *European Journal of Engineering Education*, *28*(1), 117–133.

Quintero, E., Salinas, P., González-Mendívil, E., & Ramírez, H. (2015). Augmented reality app for calculus: A proposal for the development of spatial visualization. *Procedia Computer Science*, *75*, 301–305.

Radu, I., & MacIntyre, B. (2012). Using children's developmental psychology to guide augmented-reality design and usability. *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on*, IEEE, pp. 227–236.

Radu, I., Doherty, E., DiQuollo, K., McCarthy, B., & Tiu, M. (2015). Cyberchase shape quest: pushing geometry education boundaries with augmented reality. *Proceedings of the 14th International Conference on Interaction Design and Children*, ACM, pp. 430–433.

Rambli, D. R. A., Matcha, W., & Sulaiman, S. (2013). Fun learning with AR alphabet book for preschool children. *Procedia Computer Science*, *25*, 211–219.

Revelle, G., Reardon, E., Cook, K., Takeuchi, L., Ballagas, R., Mori, K., … Spasojevic, M. (2014). Electric agents: Combining collaborative mobile augmented reality and web-based video to reinvent interactive television. *Computers in Entertainment (CIE)*, *12*(3), 1–21.

Rosser, R. A. (1994). *Cognitive development: Psychological and biological perspectives*. Boston: Allyn & Bacon.

Shelton, B. E., & Hedley, N. R. (2004). Exploring a cognitive basis for learning spatial relationships with augmented reality. *Technology, Instruction, Cognition and Learning*, *1*, 323–357.

Slijepcevic, N. (2013). The effect of augmented reality treatment on learning, cognitive load, and spatial visualization abilities, Ph.D. Dissertation, University of Kentucky.

Swanson, H. L. (1996). Individual and age-related differences in children's working memory. *Memory & Cognition*, *24*, 70–82.

Thornton, T., Ernst, J. V., & Clark, A. C. (2012). Augmented reality as a visual and spatial learning tool in technology education. *Technology and Engineering Teacher*, *71*, 18–21.

Vogt, F. P., & Shingles, L. J. (2013). Augmented Reality in astrophysics. *Space Astrophysics and Science*, *347*(1), 47–60.

Wai, J., Lubinski, D., & Bendow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychologicl knowledge solidifies its importance. *Journal of Educational Psychology*, *101*, 817–835.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2016). *Engineering psychology & human performance*. Pearson, NJ: Psychology Press.

Wu, H. K., Lee, S. W. Y., Chang, H. Y., & Liang, J. C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education*, *62*, 41–49.

Yilmaz, R. M. (2016). Educational magic toys developed with augmented reality technology for early childhood education. *Computers in Human Behavior*, *54*, 240–248.