Interactive Learning Environments

ISSN: 1049-4820 (Print) 1744-5191 (Online) Journal homepage: http://www.tandfonline.com/loi/nile20

To use augmented reality or not in formative assessment: a comparative study

Kaushal Kumar Bhagat, Wei-Kai Liou, J. Michael Spector & Chun-Yen Chang

INTERACTIVE LEARNING ENVIRONMENTS
https://doi.org/10.1080/10494820.2018.1489857

To use augmented reality or not in formative assessment: a comparative study

Kaushal Kumar Bhagat [a*], Wei-Kai Liou [b*], J. Michael Spector [c] and Chun-Yen Chang [d]

[a]Smart Learning Institute, Beijing Normal University, Beijing, People's Republic of China; [b]Department of Interactive Entertainment Design, China University of Technology, Taipei, Taiwan; [c]Department of Learning Technologies, University of North Texas, Denton, TX, USA; [d]Graduate Institute of Science Education, National Taiwan Normal University, Taipei, Taiwan

ABSTRACT

Augmented reality (AR) is growing in popularity in teaching and learning due in part to powerful new technologies. What has yet to be well established is when and with which learners and learning tasks AR is an effective approach. Therefore, the aim of this study was to examine the effectiveness of using AR-based formative assessment for improving elementary students' learning achievement and motivation in a unit of instruction involving butterflies. A total of 70 students of Grade 4 were selected from an elementary school in Taiwan. The experimental group (35) underwent an AR-based intervention that involved formative assessment using iPads whereas the control group (35) followed the traditional teaching method and formative assessment. One-way Analysis of Co-variance (ANCOVA) and multivariate analysis of variance (MANOVA) were employed to analyze the data obtained. The results indicated that using the AR-based formative assessment improved not only students' learning performance but also learning motivation effectively compared with a traditional formative assessment approach. Therefore, it is recommended to conduct further studies and to consider integrating AR in formative assessments and feedback to improve learning.

ARTICLE HISTORY
Received 7 January 2018
Accepted 13 June 2018

KEYWORDS
Augmented reality (AR); elementary school; formative assessment; information and communication technology (ICT)

Introduction

Taiwan, also known as the kingdom of butterflies, is home to more than 400 species of butterflies. Identification of different butterflies based on their anatomy and physical characteristics is an important part of the elementary school curriculum in Taiwan. For the development of identification skills, students are instructed in a laboratory that uses different specimens of butterfly and pictures. While the traditional method provides detailed information about butterflies and their environments, it does not provide dynamic feedback of a student's ability to identify particular butterflies. It is well established that timely and formative feedback is a critical component of effective learning (Gagné, 1985; Spector, 2015). However, there are limited research studies based on alternative teaching methods for the identification of species (Conejo, Garcia-Viñas, Gastón, & Barros, 2016). This study is aimed at providing dynamic formative feedback using AR to improve learning and contribute an innovative use of augmented reality (AR) in formative feedback.

CONTACT Chun-Yen Chang changcy@ntnu.edu.tw Science Education Center, National Taiwan Normal University, 88 Sec. 4, Ting-Chou Rd., Taipei, Taiwan
*These authors contributed equally to this work.

The notion pursued in this study is that timely, informative feedback is known to enhance learning outcomes (Gagné, 1985; Spector, 2015; Spector & Yuen, 2016). With regard to large class sizes and complex learning tasks, providing timely and informative feedback becomes a challenge. Given the existence of powerful digital technologies, which can support visualization, such as AR systems, the general focus of this study was the degree to which using AR for formative assessment might improve learning outcomes with regard to a challenging identification and classification task with young learners. An AR system can provide real-time feedback, which addresses the timeliness of the formative feedback. In addition, an AR system, if properly configured and constrained, can also provide meaningful and informative feedback. Therefore, the focus of this study was to determine how well AR-facilitated formative assessment could improve student performance and motivation on common learning tasks involving identification and classification, which are basic intellectual skills for young learners.

Literature review

Formative assessment

Chin and Teou (2009) define formative assessment as "assessment that informs teachers about what students have learnt, indicates what students may be finding difficult, and helps teachers to adjust their teaching to maximize students' learning" (p. 1309). The general goal of formative feedback and assessment is to help learners attain the instructional goal or objective. Formative assessment is a central element of a learning activity and a core feature of many learning environments aimed at improving students' learning as they make progress in a course or program of study (Bell & Cowie, 2001; Black & Wiliam, 1998; Redecker & Johannessen, 2013). Formative assessments enable teachers to identify weaknesses and deficiencies in instructional activities, resources and approaches that can be improved in subsequent lessons (Bell & Cowie, 2001); as a result, formative assessments can also support the formative evaluation of lessons, courses and programs. More importantly, formative assessments provide learners with help in attaining intended learning outcomes (Gagné, 1985; Spector & Yuen, 2016).

From an instructional design perspective, once a specific learning goal has been identified, it is then possible to indicate how attainment of that goal will be determined (Larson & Lockee, 2014; Merrill, 2013). A final indication of the degree to which a goal has been achieved, such as at the end of a lesson or course or program, is known as summative assessment. Summative assessments can be used to improve instruction over time as part of a formative evaluation of instruction and are often used to judge individual learner performance at a point in time, such as the completion of a course (Spector & Yuen, 2016). However, summative assessments do not help a learner achieve the intended learning goal or objective, which is the purpose of formative assessments.

From a learning perspective, the purpose of a learning activity is to promote learning and successful attainment of a particular learning goal or objective (Larson & Lockee, 2014). For example, if the learning goal is to be able to identify various kinds of butterflies, then a particular learning activity might be aimed at recognizing the distinguishing features associated with different kinds of butterflies. In such a learning activity, it is relevant to provide the learner with feedback to indicate correct or incorrect identification, which is only partially informative feedback but typically the kind used in a summative assessment. More meaningful feedback could involve the reason for correct or incorrect identification, with the latter case being especially helpful in attaining the intended competence and performance (Gagné, 1985). Providing such specific informative feedback to each learner given the many possible kinds of errors and types of butterflies is a burdensome task for a teacher in a confined context with many students and with limited time for practice with feedback.

In addition to the value to timely and informative feedback, the issue of time on task is relevant to attaining the intended outcome. Learning motivation and engagement are key factors in helping a learner spend productive time on a particular learning task, in this case identifying butterflies. In that chain of reasoning, there is a place to involve visualization and an interactive digital technology to enhance motivation and ensure meaningful engagement. In this study, the added value of AR is that it supports visualization and is likely to ensure meaningful engagement in addition to being able to provide timely and informative feedback.

Augmented reality and its effectiveness

Currently, AR is an emerging technology that is becoming ever more affordable and pervasive (Johnson et al., 2012). Azuma (1997) defined AR as a system having three basic characteristics: AR (a) combines reality with a virtual world, (b) is interactive in real time, and (c) supports 3-D visualization. Chen (2013) listed some of the advantages of AR over virtual reality (VR):

λ *Multimedia and multisensory display*: AR not only provides multiple representations but also provides opportunities to the users to touch, rotate, and manipulate the virtual objects; this feature is absent in VR.

λ *Portable and cost-effective*: VR environment requires expensive and heavy devices like head-mounted display (HMD) and not easy to execute in general classroom. On the other hand, AR environment can be created using cheaper and lighter smart devices like mobile phone, tablets, etc.

λ *User friendly*: In a VR environment, users sometime experience cyber sickness that they may not experience in an AR environment.

λ *Retain user's proprioception*: In an AR based learning environment, users are not isolated from the real world, which results in proprioception (awareness of one's own body) in relation to the user's environment. However, VR environment completely takes users out of the real world situation with a resulting loss of some proprioception (e.g. when presented with a next step off a tall building in a VR environment, a user will typically hesitate or not take that next step for fear of falling even though the user is on solid ground).

Previous research studies demonstrate the benefits of AR-applications on students' learning achievement, engagement, and motivation. For example, Di Serio, Ibáñez, and Kloos (2013) examined the effectiveness of AR technology on students' motivation in visual art course. The authors found that AR technology has advantage in grasping the attention of the students because of the multimodal nature of the content. In addition, students found AR learning material more interactive compared to the content based on the PowerPoint slides, which resulted in higher motivation. Cai, Wang, and Chiang (2014) used an AR learning tool to teach chemistry course at junior high-school level. The results showed that the AR tool significantly improved the learning outcome of the students. In addition, they concluded that the AR tool was more for low-achievers compared to high-achievers. Zhang, Sung, Hou, and Chang (2014) developed a mobile digital armillary sphere (MDAS) using AR and embedded into an astronomical course. The results indicated that MDAS improved astronomical skills, learning, and had a higher impact on retention. In another study, an ARBOOK tool was developed by Ferrer-Torregrosa, Torralba, Jimenez, Garcia, and Barcia (2015) to teach anatomy. The research pointed out that AR technology is helpful to enhance students' motivation. Chang, Chung, and Huang (2016) conducted a quasi-experimental study in which they compared an AR system (*ARFlora*) with digital video learning to understand the plant growth. The results did not show any significant differences in learning achievement. However, they reported that the AR system assisted the students in retaining knowledge learned more effectively compared to video learning, and enhanced higher motivation, both of which can lead to improved learning over time. Similarly, Akcayir, Akcayir, Pektas, and Ocak (2016) investigated the effectiveness AR technology in enhancing science laboratory skills for first-year university students. The results pointed out that the AR technology not only helped the students to improve their laboratory skills but also positively affect their attitudes towards physics. In another study, Cai, Chiang, Sun, Lin, and Lee (2017) implemented AR-based motion sensing software in learning and conducted a quasi-experimental study. They found that AR helped to create active learning environment, which results into better learning outcome and motivation. Wang (2017) explored the effects of AR-based support system to assist high-school students in developing their Chinese writing skills. The research findings showed that the students in the AR-based learning system outperformed than those in the other group, especially low-achievers. The results also revealed that students who were low-achievers showed more positive attitude towards AR technology. Yilmaz and Goktas (2017) applied AR technology in story telling activities and reported that AR technology helped in enhancing the students' narrative skills, engagement, and creativity.

Based on the above literature review, it is clear that AR technology can enhance learning outcomes and motivation in different disciplines. However, a research study on the application of AR in formative assessment has yet to be explored. To fill this gap, we developed an interactive formative assessment system using augmented reality in environmental education for elementary students. The following research questions guided this study:

λ To what extent can an AR-based formative assessment improve students' learning performance?

λ To what extent can an AR-based formative assessment improve students' learning motivation?

Design and illustration of AR-based formative assessment system

In the present study, we used the Unity software platform to develop AR-based formative assessment system (see Figure 1) to improve students' learning outcome and motivation in environmental education. Figure 1(a) shows the user interface of our system. The assessment process includes two different assessments. In assessment 1, user needs to identify different butterflies whereas in assessment 2, user needs to identify different parts of the butterfly. The assessment process includes four stages. In the first stage, users need to login in the system using their username and password. In the second stage, users are provided images of different butterfly. The user has to point the in-device camera at the target image and, then the 3D AR butterfly will pop-up on the user's interface with scientific names (see Figure 1(b)). This system allows users to rotate and view the 3D butterfly from different angles. The user then needs to select the correct scientific name for the corresponding butterfly. The system records the user's answers and displays the corresponding results at the end of the assessment. In the third stage, the system displays an image of butterfly and different body part names. Users need to drag the names to the corresponding body parts of the butterfly (see Figures 1(c)) and 1(d)). In the final stage, the system displays the scores earned by the user with feedback (see Figure 1(e)).

Methodology

Research design and sample

This study used a quasi-experimental research design. A total of 70 (Males = 40, Females = 30) students of Grade 4, aged 9–10 years old were selected from an elementary school located in Taiwan and divided into an experimental group and a control group. School and class choices were based on convenience as available and willing to support the research. Groups were designed to be similar in terms of past performance and gender rather than relying on random assignment in this case due to the rather limited number of participants. The experimental group (35) underwent AR-based formative assessment using iPad whereas the control group (35) followed the traditional method of formative assessment.

Instruments

In this study, a pre-test, a post-test and the Instructional Materials Motivation Survey (IMMS) based on Keller's (2010) ARCS model were used as research instruments. A group of four subject experts with relevant years of experience developed the pre- and post-test to determine and confirm the content validity and difficulty level of the test items. The pre-test consisted of 5 multiple-choice questions

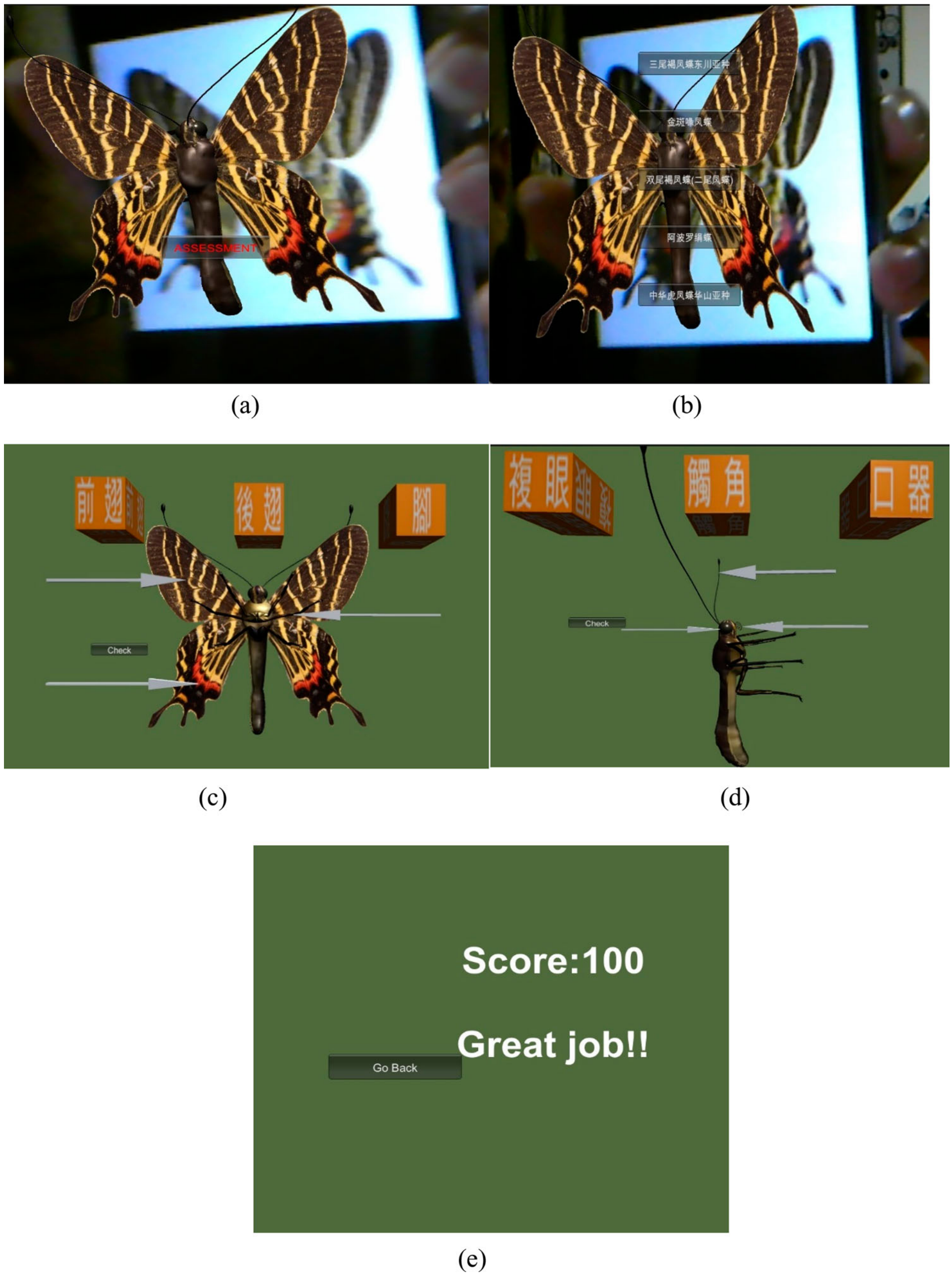

Figure 1. The AR-based formative assessment system.

(MCQ) to examine whether the students in both groups had similar achievement levels. The post-test consisted of 10 MCQ isomorphic to pre-test items. The time allotted for the pre-test was 10 min whereas for post-test 20 min. IMMS has four factors: attention, relevance, confidence, and satisfaction. It consists of 36 items with a 5-point Likert-scale. The overall Cronbach's α for the pre-test, post-test and IMMS were 0.78, 0.75, and 0.85 respectively, which are acceptable.

Procedure

Figure 2 shows the experimental procedure of this study. The experiment consisted of four stages: (a) initially, both groups took a pre-test to determine their previous knowledge of the 5 species randomly selected from those available at the laboratory; (b) the students in both groups went to the laboratory and spent the same amount of time studying 20 butterfly species and their different body parts, (c) after the lesson, the experimental group used AR-based formative assessment to strengthen their conceptual understanding whereas the control group had to complete a paper-based test used for formative assessment; in the control group, the students received delayed feedback from the teachers whereas in the experimental group real-time feedback was provided; (d) in the final stage, students from the both groups completed a post-test and IMMS motivation questionnaire. The objectives of the lesson includes: (1) identification of the different species of butterfly, and (2) identification of the different parts of the butterfly.

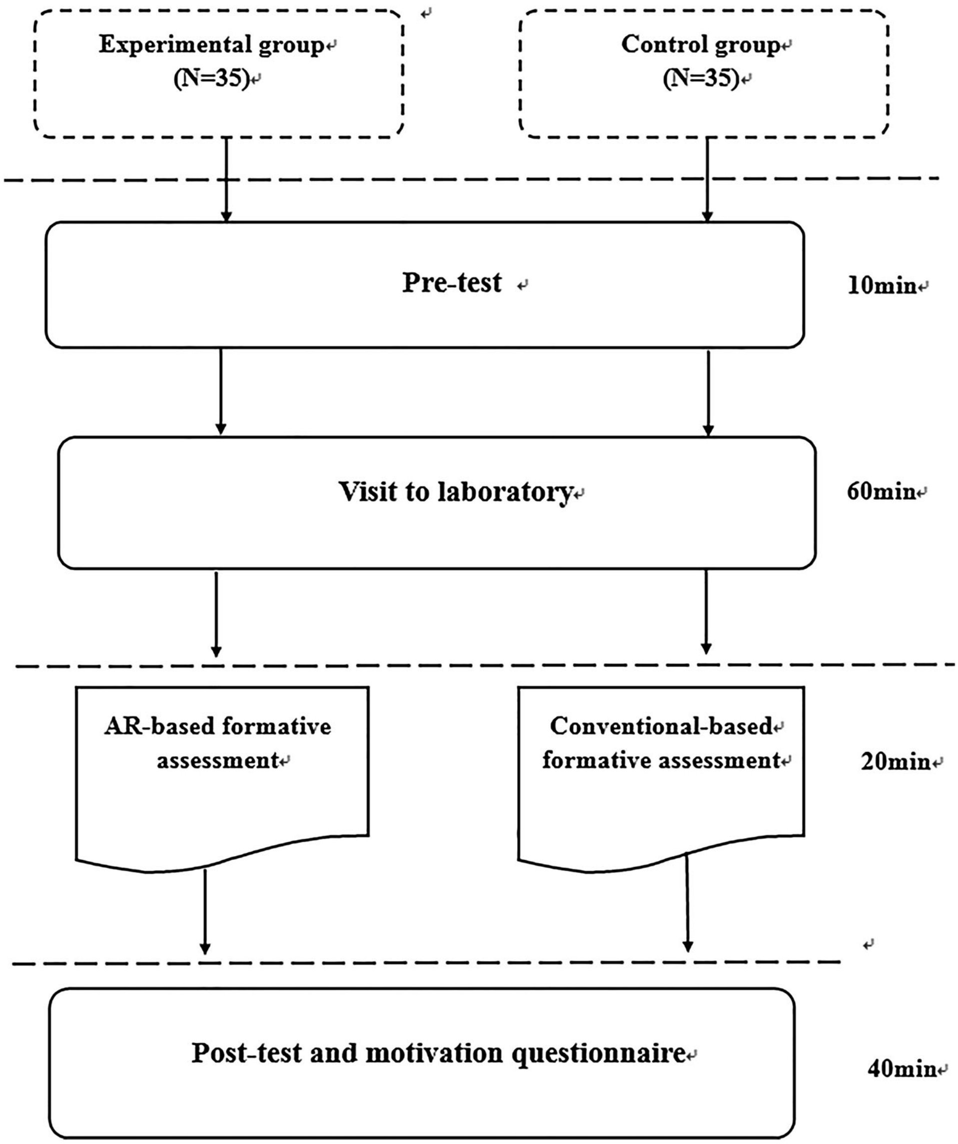

Figure 2. The experimental procedure.

Data analysis

All analyses were performed using the Statistical Package for the Social Sciences, Version 21 (SPSS 21). The statistical significance level was set at $p < 0.05$. One-way ANCOVA was applied to compare the learning performance of the students in the experimental group and the control group. With respect to the motivation, one-way MANOVA was used to test the statistical significance difference between the experimental group and the control group.

Results

Learning performance

A one-way Analysis of Co-variance (ANCOVA) was used to determine a statistically significant difference between groups on post-test scores controlling for pre-test scores. The assumption of homogeneity of regression was examined and no violation was found ($F = 0.98$, $p > .05$). Therefore, ANCOVA was employed.

The mean value and standard deviation of the post-test scores were 8.08 and 0.85 for the experimental group, and 6.77 and 0.54 for the control group, respectively. The difference was significant, $F(1, 67) = 377.72$, $p < .05$ (see Table 1). The calculated effect size (eta squared, η^2) is 0.48 which is considered to be a large effect (Cohen, 1988). This result indicated that the students in the experimental group performed better than those in the control group.

Learning motivation

Descriptive statistics (shown in Table 2), including means and standard deviation are provided for dependent variables attention, relevance, confidence, and satisfaction. The results of MANOVA revealed that there was significant difference for attention, relevance, confidence, and satisfaction between the experimental and control groups, Wilk's $\Lambda = .58$, $F = 11.48$, $p < .05$, $\eta^2 = .41$. Therefore, univariate F tests were conducted for attention, relevance, confidence, and satisfaction. The results of univariate F tests indicated a significant difference between the groups for attention ($p < .05$, $\eta^2 = .08$), relevance ($p < .05$, $\eta^2 = .14$), confidence ($p < .05$, $\eta^2 = .33$), and satisfaction ($p < .05$, $\eta^2 = .23$) (see Table 3).

Discussion and conclusion

In this study, an AR-based formative assessment system was developed and evaluated for its impact on learners' learning performance and motivation. The experimental results showed that this AR-based formative assessment is helpful in improving students' learning achievement compared to the conventional method. This result is consistent with previous research studies suggesting that AR can improve learning, although those studies did not involve AR-based formative assessment (Cai et al., 2017; Wang, 2017). The present system provides 3-D effects that simulate real-world experiences of butterflies for learners. In addition, real-time elaborated feedback is one of the important features of this system that helped students to understand the task and associated learning objectives more clearly. This type of technology-enhanced assessment also provides opportunity for the

Table 1. ANCOVA results for post-test scores.

Source	*SS*	*df*	*MS*	*F*	*p*	η^2
Pre-test	1.49	1	1.493	2.99	.08	.04
Group	31.67	1	31.67	63.50	.000	.48
Error	33.42	67	.49			
Total	3928	70				

**p < .05.*

Table 2. Descriptive statistics for IMMS motivational questionnaire.

	Group	*N*	Mean	*SD*
Attention	CG	35	2.55	.48
	EG	35	2.83	.49
Relevance	CG	35	2.08	.46
	EG	35	2.46	.45
Confidence	CG	35	2.47	.55
	EG	35	3.19	.47
Satisfaction	CG	35	2.42	.52
	EG	35	2.97	.47

students to practice more, which results in better retention of the conceptual knowledge. However, in the traditional mode of assessment, students did not receive immediate feedback with proper elaboration. Therefore, AR-based formative assessment is recommended for teaching environmental education and simple intellectual skills (e.g. discrimination and identification tasks) in elementary schools, although more research is certainly warranted.

The present study also examined the effects of AR-based formative assessment system on students' motivation. The findings revealed positive effects on motivation with students who used the AR-based system. This is in line with previous studies, such as Ferrer-Torregrosa et al. (2015), who developed an ARBOOK to teach anatomy which enhanced students' learning motivation and similar results were also found by Chang et al. (2016) when using AR for primary presentations rather than for formative assessment. This AR system provides feedback with positive reinforcement, which motivated the students. This system provides interactive and active learning environment which maybe another reason for students' higher learning motivation.

AR provides support for visualization and real-time experience, both of which are likely to be critical learning factors with other tasks. Because AR is becoming ever more affordable and more powerful, as are many other digital technologies, now is the time to be exploring how well, with which learning tasks and with which learners AR-based formative assessment is likely to be an effective learning technology.

This study shows either that any dynamic, real-time feedback can make a significant difference or that such feedback provided using AR like the one used in this case can make a significant difference in terms of learning gains. Given prior research on feedback, the latter is more likely due to the timeliness of the feedback combined with support for visualization. However, based on the Clark-Kozma media debate (see http://edutechwiki.unige.ch/en/The_media_debate) in which it was acknowledged that the design was more likely associated with outcomes while also acknowledging that new technologies make possible new designs, such as AR-based formative assessment, we conclude that AR at least indirectly contributed to improved learning. The advantage of AR-based feedback is that it can be automated for many learners whereas other forms of formative feedback require more time and effort on the part of the instructor. The visualization aspects of AR were especially suited for the butterfly species identification task but may not support other learning tasks in which visualization is not such a critical component. To sum up, the contribution of the present study is to provide empirical evidences about the effectiveness of AR in assessment process, which has remained unexplored in the research area of AR technology. We believe that the present study extends the application of AR technology in education.

Table 3. The one-way MANOVA results for IMMS motivational questionnaire.

DV	*SV*	*SS*	*df*	*MS*	*F*	η^2
Attention	Group	1.42	1	1.42	5.94*	.08
Relevance		2.50	1	2.50	11.85*	.14
Confidence		9.10	1	9.10	34.20*	.33
Satisfaction		5.29	1	5.29	21.24*	.23

$^*p < .05$.

Limitations and future directions

One limitation of this study is that it does not address the durability of learning. As Gagné (1985) and Spector (2015) have argued, learning involves stable and persistent change in what a person knows and can do. As there was no delayed post-test involved, the persistence of learning is unknown. As a result, it is recommended that future studies include a delayed post-test (e.g. three or six months after the end of the instructional sequence) to see to what extent the intervention (in this case, AR-based formative assessment) resulted in stable and persistent learning outcomes.

A second limitation is that it is not clear to what extent the technology (AR-based formative assessment) was the likely cause of improved learning or whether another kind of dynamic, real-time formative feedback would be just as good (e.g. an interactive simulation with feedback, peer-coaching and feedback, etc.). This limitation relates back to the Clark-Kozma debate mentioned earlier and to the notion that it is the use of a technology or pedagogical approach that is likely to impact learning rather than the technology or pedagogical approach itself.

A third kind of limitation goes to the issues of transfer and generalization. It is not known to what extent AR-based formative assessment would be successful in other identification and learning tasks with young learners, or whether the approach might also work with older learners or different kinds of learning tasks.

A fourth limitation involves the relatively small number of participants and lack of investigation into other identification and discrimination tasks involving young learners. Such studies will be required to determined the likelihood of transfer and as a basis for generalization to other situations.

Finally, this study did not explore initial individual differences pertaining to interest in butterflies or past performance relevant to butterfly identification. Along with timely and informative feedback and time-on-task, it is also well-established that prior learner performance is generally predictive of future performance. While this study attempted to create similar groups, the analysis did not explore the extent to which various individual differences might impact AR-based formative assessment. Another factor to examine in future studies in how a technology is being used or how any treatment is being implemented. Controlling for use is a challenge in instructor led activities, so that variations in use and implementation need to be observed and analyzed in future studies. In addition, AR-based formative assessment is worth exploring for other learning tasks and with other learners at different levels.

In short, there is certainly much more to investigate and explore with regard to AR-based formative assessment. This study is just a first, somewhat promising step in making effective use of an emerging technology to provide much improved feedback during learning activities.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was financially supported by the "Institute for Research Excellence in Learning Sciences" of National Taiwan Normal University (NTNU) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan].

Notes on contributors

Dr. ***Kaushal Kumar Bhagat*** is currently working as a post-doctoral researcher in Smart Learning Institute at Beijing Normal University. He received his Ph.D in educational technology from National Taiwan Normal University. His research area of interest includes online learning, augmented reality, virtual reality, mathematics education, flipped classroom, formative assessment and technology-enhanced learning.

Dr. ***Wei-Kai Liou*** is currently working as an assistant professor in the Department of Interactive Entertainment Design at China University of Technology. He received his Ph.D in educational technology from National Taiwan University of Science and Technology. He has published several referred journal articles and textbook. Dr. Liou was presented Taiwan Printed Circuit Association Outstanding Paper Award. His research area of interest includes augmented reality, virtual reality, mixed reality, flipped classroom, and HCI (Human–Computer Interaction).

Prof. J. Michael Spector is a professor of learning technologies at the University of North Texas. He has been editor of Educational Technology Research & Development for 15 years. He edited the third and fourth editions of the Handbook of Research on Educational Communications and Technology as well as the Encyclopedia of Educational Technology. His research focuses on intelligent support for instructional design, assessing learning in complex domains, and technology integration in education.

Prof. Chun-Yen Chang is a science education scholar in Taiwan. Currently, he serves as National Taiwan Normal University (NTNU) Chair Professor, Director of Science Education Center (NTNU), Professor of the Graduate Institute of Science Education and the Department of Earth Sciences (NTNU). Prof. Chang's major research interests include science education, e-Learning, interdisciplinary science learning and science communication.

ORCID

Kaushal Kumar Bhagat http://orcid.org/0000-0002-6861-6819

References

Akcayir, M., Akcayir, G., Pektas, H. M., & Ocak, M. A. (2016). Augmented reality in science laboratories: The effects of augmented reality on university students' laboratory skills and attitudes toward science laboratories. *Computers in Human Behavior*, *57*, 334–342. doi:10.1016/j.chb.2015.12.054

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, *6*(4), 355–385.

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education*, *85*(5), 536–553. doi:10.1002/sce.1022

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. doi:10.1080/0969595980050102

Cai, S., Chiang, F.-K., Sun, Y., Lin, C., & Lee, J. J. (2017). Applications of augmented reality-based natural interactive learning in magnetic field instruction. *Interactive Learning Environments*, *25*(6), 778–791. doi:10.1080/10494820.2016.1181094

Cai, S., Wang, X., & Chiang, F.-K. (2014). A case study of augmented reality simulation system application in a chemistry course. *Computers in Human Behavior*, *37*, 31–40. doi:10.1016/j.chb.2014.04.018

Chang, R.-C., Chung, L.-Y., & Huang, Y.-M. (2016). Developing an interactive augmented reality system as a complement to plant education and comparing its effectiveness with video learning. *Interactive Learning Environments*, *24*(6), 1245–1264. doi:10.1080/10494820.2014.982131

Chen, Y.-C. (2013). *Learning protein structure with peers in an AR-enhanced learning environment* (*Ph.D. thesis*), University of Washington. Retrieved from https://www.learntechlib.org/p/124052

Chin, C., & Teou, L. Y. (2009). Using concept cartoons in formative assessment: Scaffolding students' argumentation. *International Journal of Science Education*, *31*(10), 1307–1332. doi:10.1080/09500690801953179

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New Jersey: Lawrence Erlbaum.

Conejo, R., Garcia-Viñas, J. I., Gastón, A., & Barros, B. (2016). Technology-Enhanced formative assessment of plant identification. *Journal of Science Education and Technology*, *25*(2), 203–221. doi:10.1007/s10956-015-9586-0

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 and Education*, *68*, 586–596. doi:10.1016/j.compedu.2012.03.002

Ferrer-Torregrosa, J., Torralba, J., Jimenez, M. A., Garcia, S., & Barcia, J. M. (2015). ARBOOK: Development and assessment of a tool based on augmented reality for anatomy. *Journal of Science Education and Technology*, *24*(1), 119–124. doi:10.1007/s10956-014-9526-4

Gagné, R. M. (1985). *The conditions of learning and theory of instruction* (4th ed.). New York: Holt, Rinehart & Winston.

Johnson, L., Adams, S., & Cummins, M. (2012). *The NMC horizon report: 2012 higher education edition*. The New Media Consortium., Austin, Texas.

Keller, J. (2010). *Tools to support motivational design motivational design for learning and performance* (pp. 267–295). New York: Springer.

Larson, M. B., & Lockee, B. B. (2014). *Steamlined ID: A practical guide to instructional design*. New York: Routledge.

Merrill, M. D. (2013). *First principles of instruction. Identifying and designing effective, efficient and engaging insgruction*. San Francisco, CA: Wiley.

Redecker, C., & Johannessen, Ø. (2013). Changing assessment — towards a new assessment paradigm using ICT. *European Journal of Education*, *48*(1), 79–96. doi:10.1111/ejed.12018

Spector, J. M. (2015). *Foundations of educational technology: Integrative approaches and interdisciplinary perspectives* (2nd ed.). New York: Routledge.

Spector, J. M., & Yuen, H. K. (2016). *Educational technology program and project evaluation*. New York: Routledge.

Wang, Y.-H. (2017). Exploring the effectiveness of integrating augmented reality-based materials to support writing activities. *Computers & Education*, *113*(Supplement C), 162–176. doi:10.1016/j.compedu.2017.04.013

Yilmaz, R. M., & Goktas, Y. (2017). Using augmented reality technology in storytelling activities: Examining elementary students' narrative skill and creativity. *Virtual Reality*, *21*(2), 75–89. doi:10.1007/s10055-016-0300-1

Zhang, J., Sung, Y. T., Hou, H. T., & Chang, K. E. (2014). The development and evaluation of an augmented reality-based armillary sphere for astronomical observation instruction. *Computers and Education*, *73*, 178–188. doi:10.1016/j.compedu.2014.01.003