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

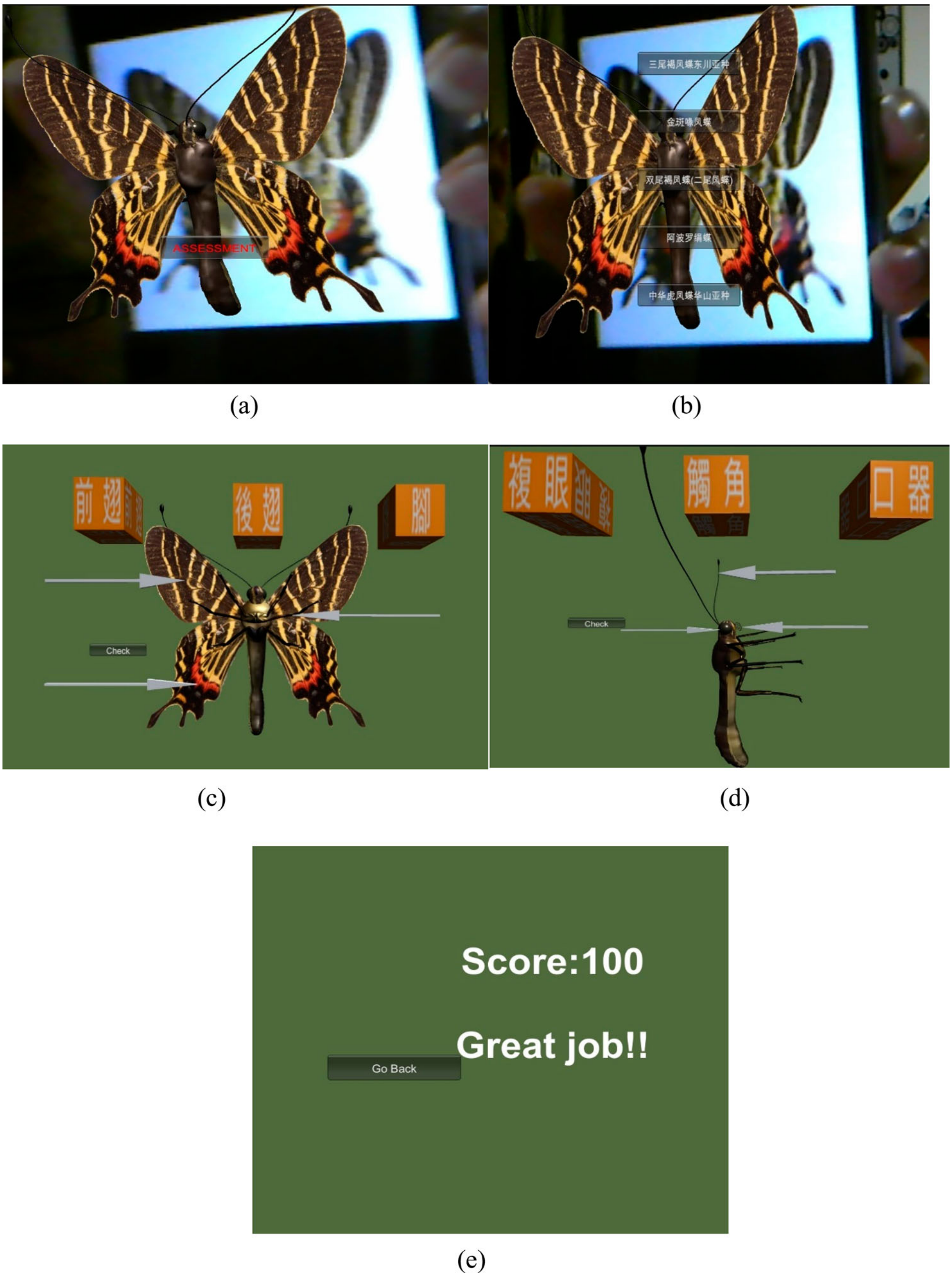

Figure 1. The AR-based formative assessment system.

(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.

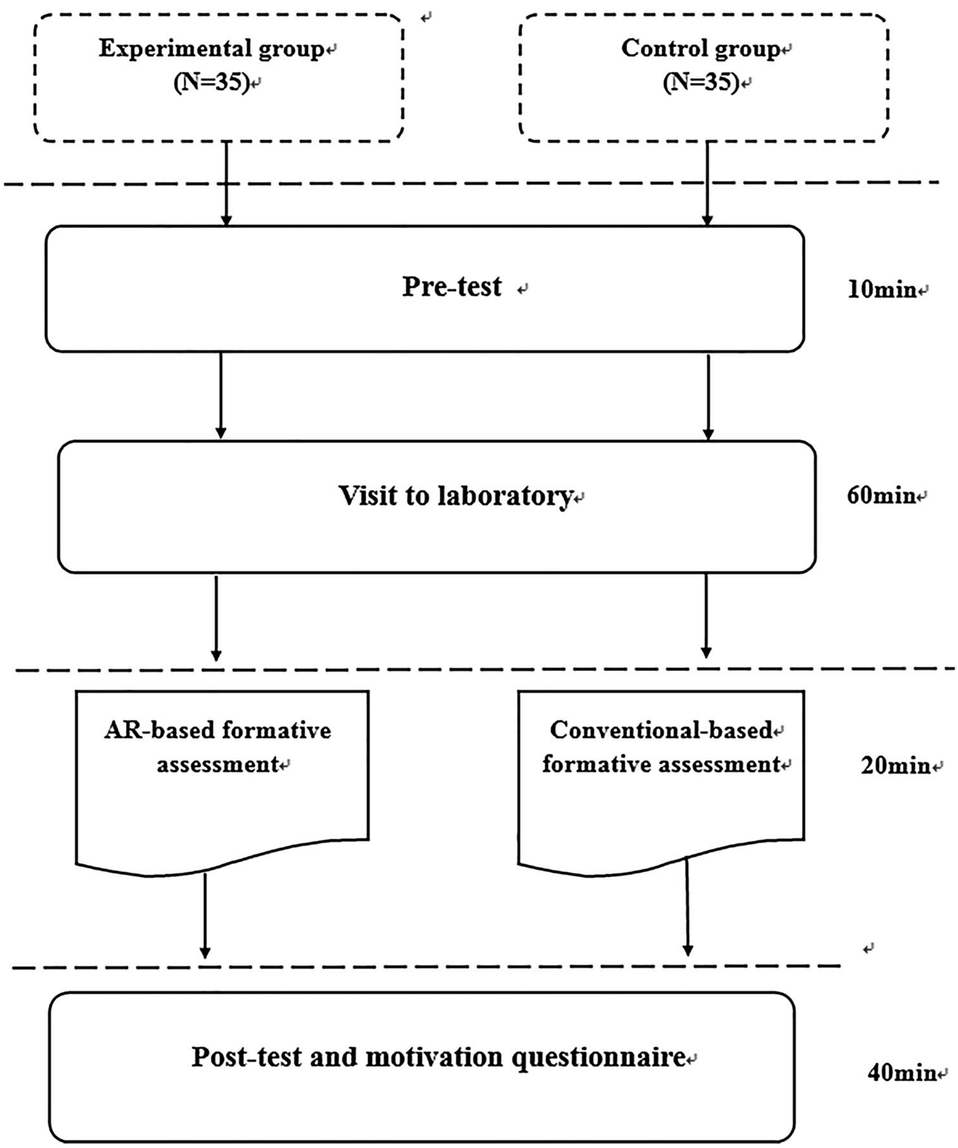

Figure 2. The experimental procedure.

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 Λ=.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 ($p < .05$, η^2 =.08), relevance ($p < .05$, η^2 =.14), confidence ($p < .05$, η^2 =.33), and satisfaction ($p < .05$, η^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

OPEN ACCESS
EURASIA Journal of Mathematics Science and Technology Education
ISSN 1305-8223 (online) 1305-8215 (print)
2017 13(1):19-34
DOI 10.12973/eurasia.2017.00602a

The Design and Development of Identification of Students' Misconceptions in Individualized Learning Environment (iSMILE) System

Kaushal Kumar Bhagat
Smart Learning Institute of Beijing Normal University, CHINA

N.P. Subheesh
Indian Institute of Technology, Chennai, INDIA

Bani Bhattacharya
Indian Institute of Technology, Kharagpur, INDIA

Chun-Yen Chang
National Taiwan Normal University, TAIWAN

Received 14 April 2016 ▪ Revised 07 June 2016 ▪ Accepted 30 June 2016

ABSTRACT

With the rapid development of technology, incorporation of Information Communication Technology (ICT) for formative assessment purpose has been increasing over the past decade. This article describes the design and development of identification of students' misconceptions in an individualized learning environment (iSMILE) system that includes accommodations for students' misconceptions in order to improve student's conceptual understanding and finally learning outcome. The assessment process is carried out by preparing two-level multiple-choice questions. Misconceptions are identified for each instructional objective to make a root question and linked questions. Depending on the answers provided by the student in first level, the linked questions are given to get more information regarding the misconception. Based on the first and second level answers of students, feedbacks are provided with the misconception identified. One of the most important differences of our iSMILE system, with respect to the other developed systems, is the inclusion of misconception based feedback. The iSMILE system has been developed by using open source ICT tools (like MySQL, Apache, etc.) which make the system free of cost and accessible and editable by any instructor from anywhere and at any time.

Keywords: Formative assessment; misconception; feedback; ICT

Correspondence: Chun-Yen Chang, *National Taiwan Normal University, 88 Sec. 4 Ting-Chou Rd., Science Education Center, 116 Taipei, Taiwan*
changcy@ntnu.edu.tw

State of the literature

- Previous research studies have indicated that formative assessment is a central element of learning environment of the 21st century to improve students' learning effectiveness.
- The main purpose of formative assessment is to provide meaningful feedback to teachers and students. Effective feedback can help to identify learner's specific errors or misconceptions related with the concept.
- We have found that there is no available system, which provides misconception based feedbacks.

Contribution of this paper to the literature

- A formative assessment system is developed to identify students' misconceptions.
- This system uses two-level multiple-choice questions for assessment purpose.
- Open source ICT tools (like MySQL, Apache, etc.) are used to design our system.

INTRODUCTION

In general, assessment plays an important role in any educational system. It identifies, collects and interprets information about the learner's conceptual understanding associated with instructional objectives (Farrell & Rushby, 2016; Rodrigues & Oliveira, 2014). Formative assessment is an integral part of assessment, which is an ongoing process of gathering data to measure the degree of attainment of course goals (Hsu, Chou, & Chang, 2011). Chin and Teou (2009, p. 1309) 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." Formative assessment is a central element of learning environment of the 21st century to improve students' learning effectiveness (Bell & Cowie, 2001; Black & Wiliam, 1998; Redecker & Johannessen, 2013). Fuchs and Fuchs (1986) conducted a meta-analysis to examine the effect of formative evaluation on students' learning achievement. They found an overall effect size of 0.70, which is moderate.

The main purpose of formative assessment is to provide meaningful feedback to teachers and students. Students can monitor their learning progress using the feedbacks (Bell & Cowie, 2001). Effective feedback can help to identify learner's specific errors or misconceptions related with the concept (Attali, 2015; Shute, 2008), improve learning and performance (Attali, 2015; Hattie & Timperley, 2007; Lefevre & Cox, 2016; Sadler, 1989; Yuan & Kim, 2015), increase learner's confidence and motivation to learn (Yuan & Kim, 2015). With the introduction of technology, the effectiveness of feedback delivery can be improved with more accuracy (Pacheco-Venegas, López, & Andrade-Aréchiga, 2015). Technology can help the assessment process to collect, and analyze the data in a shorter period of time with more accuracy (Farrell & Rushby, 2016; Redecker & Johannessen, 2013; Yuan & Kim, 2015).

Currently, many researchers have developed web-based/computer-based assessment systems and found the potential benefit in improving the learning effectiveness of the learners

(Crisp & Ward, 2008; Hsu et al., 2011; Landauer, Lochbaum, & Dooley, 2009; Pacheco-Venegas et al., 2015; Rodrigues & Oliveira, 2014; Wang, 2008, 2010). Crisp and Ward (2008) developed Pedagogical Psychology Computer Assisted Assessment (PePCAA) system to facilitate understanding of pedagogical psychology. They found that PePCAA system motivated the students and provided deeper learning. Hsu et al. (2011) developed Edu-miner system by utilizing a series of text mining techniques to evaluate learner's performance. Landauer et al. (2009) developed Write-ToLearn formative assessment tool to monitor progress in reading and writing. Pacheco-Venegas et al. (2015) developed MathDIP to provide automatic evaluation and feedback during problem solving in mathematics. Rodrigues and Oliveira (2014) developed AssiStudy as a formative assessment tool to monitor students' learning progress. Wang (2010) developed graduated prompting assessment module of the WATA (GPAM-WATA) to provide timely feedback while facing difficulties when attempting test items.

From the above, we have found that there is no available system, which provides misconception based feedbacks. To fill this research gap, we have developed iSMILE system. This paper has outlined the iSMILE system, its components and architecture, working process, and application.

COMPONENTS AND ARCHITECTURE OF ISMILE SYSTEM

The iSMILE system is composed of a database unit, a server unit, and a user interface, as shown in **Fig. 1**. In our system, MySQL and Apache Server are preferred as a database unit and a server unit respectively. PHP is used as a server-side language in order to design and code the user interface. We have used only open-source tools in our system. In an open-source system, the source code is available to all. Therefore, anyone can alter the code (to customize the software according to their needs), fix software-bugs and enhance security considerations. Thus, instructors with minimal software knowledge can also add new features and write new modules.

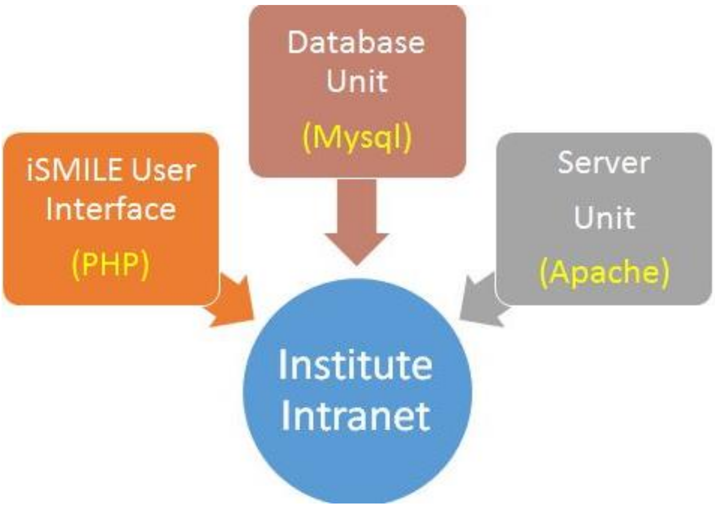

Fig. 1. The iSMILE system architecture.

iSMILE's architecture is based on the Model View Controller (MVC) design. According to the MVC design structure, the command-flow is initiated by a central controller. Upon initiation, the controller sends requests to a compatible handler application. The end user interacts with the intranet interface through this central controller. Controller handles all the inputs to the MVC model. For instance, a Graphical User Interface (GUI) controller accepts GUI inputs from the user and gives appropriate commands to the Model and View. Thus, command-actions based on respective inputs are executed. In case of an invalid-input is being given to the controller; the Model sends a command back to the controller to communicate through the View. Subsequently, the View layout displays a message through GUI like 'input-error' and 'try again'. The MVC design is illustrated in **Fig. 2**.

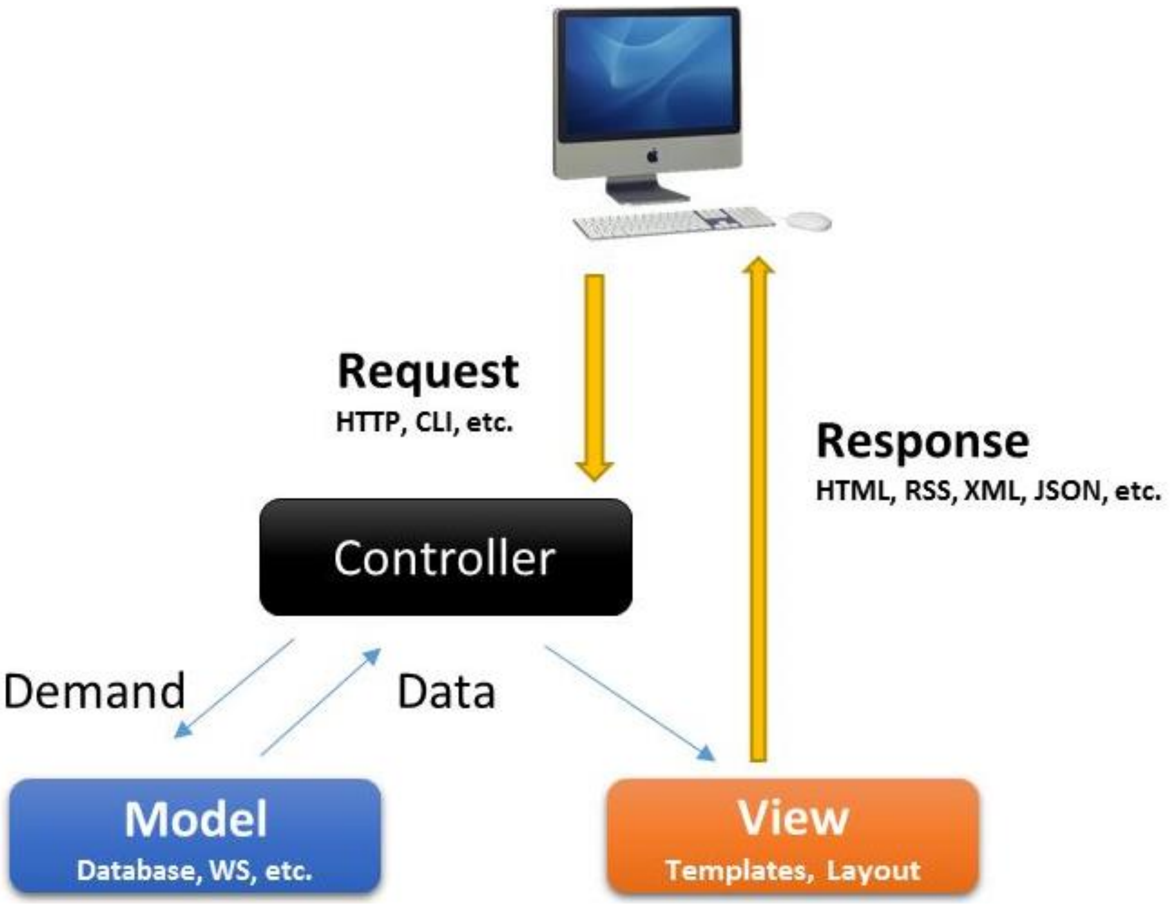

Fig. 2. MVC design.

The three components of iSMILE system architecture are described as below

Database Unit

MySQL is used as the database unit. A database design of a database system is an important stage in system development. In design phase, the database construction is coded in any formal language supported by the database management system. A database design diagram is used to show how the data is organized by forming various database tables. The database design diagram is shown as **Fig. 3**.

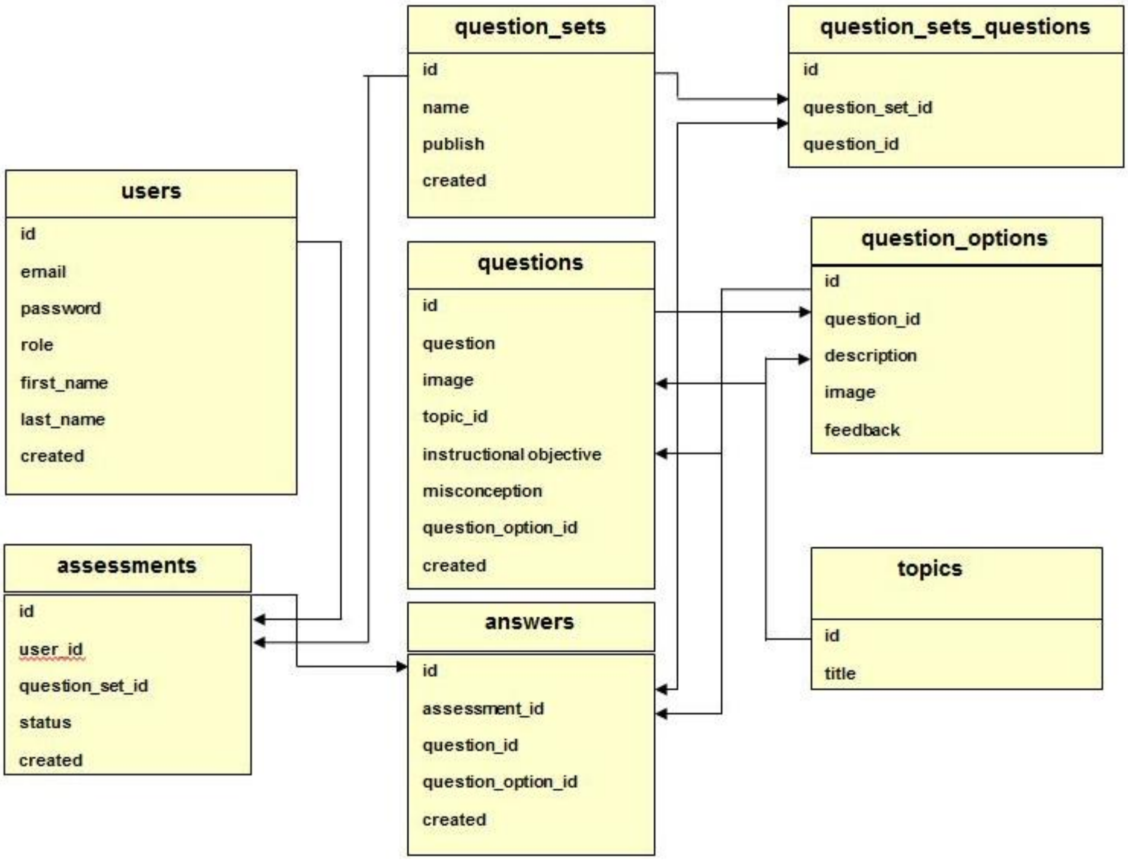

Fig. 3. iSMILE's Database Design Diagram.

Another design tool is the Data Flow Diagram (DFD). It is a graphical representation of the flow data. That is, how the data moves from the external data source or internal data storage to an external system or to an internal system. It is quite different from the Flow chart where one can determine the rate, order and circumstances of occurrence of a work but without any data input and output information, which is possible in case of DFD. iSMILE's DFD is shown in **Fig. 4**.

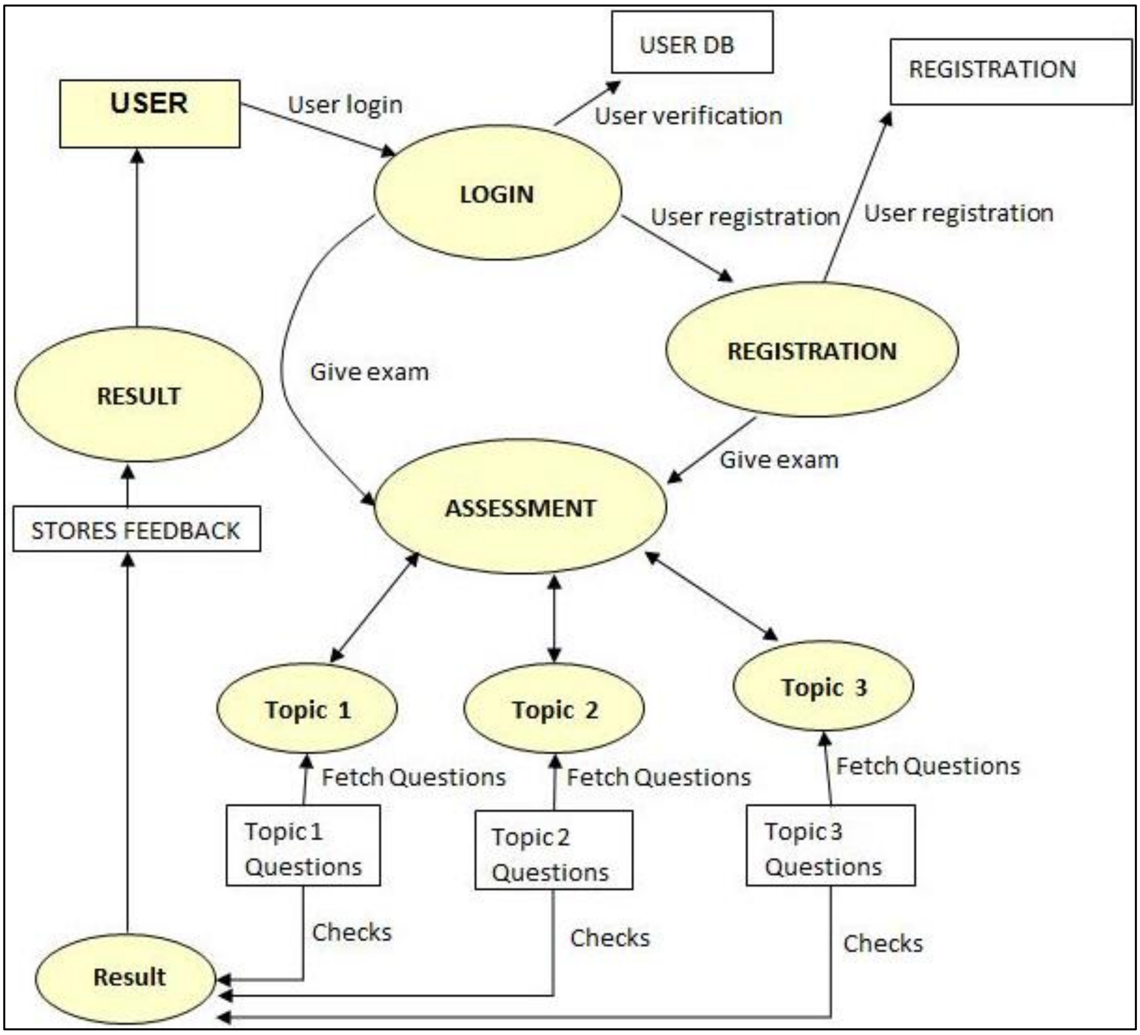

Fig. 4. iSMILE's Data Flow Diagram.

Server Unit

Apache HTTP web server software is used as a server unit in the iSMILE. We preferred Apache because it is free to download and is open-source. Hence, due to its open-source nature and add-ons support, anyone can adapt the Apache server for specific needs. In order to perform the HTTP server operations, Apache needs some other tools. Additional software tools required in the server side are:

a) Server-side programming language: PHP

b) Database tool: MySQL

c) Compatible operating system: Windows/Linux

d) Web browser supporting JavaScript: Google Chrome, Mozilla Firefox, etc.

User Interface

The user interface mainly has three stages.

Login: There is a quality login window which can be used by both instructors as well as students, as shown in **Fig. 5**. If the login credentials are valid, the user will be redirected to respective dashboard screen.

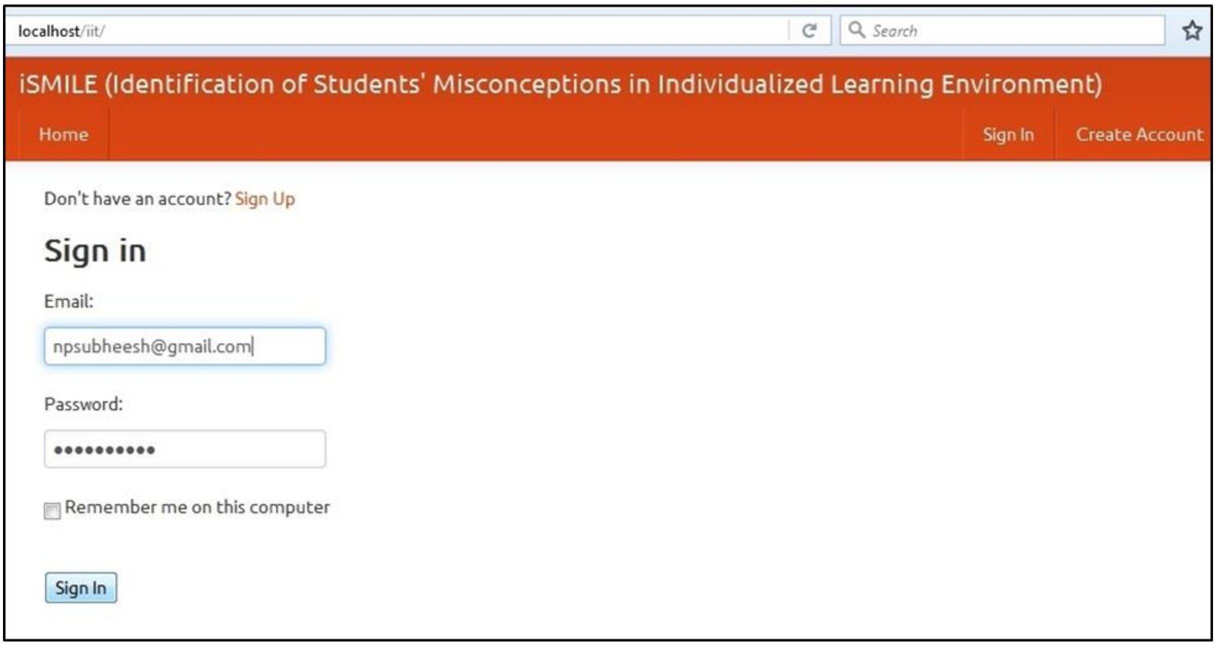

Fig. 5. Login screen of iSMILE system.

Assessment: Assessment page is the most important page in our system, as shown in **Fig. 6.** It consists of two parts namely:

(a) Assessment selection: From the given choices the student can select the assessment for taking on with the test.

(b) Appearing for the test: Here the user can answer the questions. There is also an option to skip the question. If the user has answered the first-level question, the second-level question will be provided to the user based on the selection. For each option (choice) in the first-level question will have exactly one second level question.

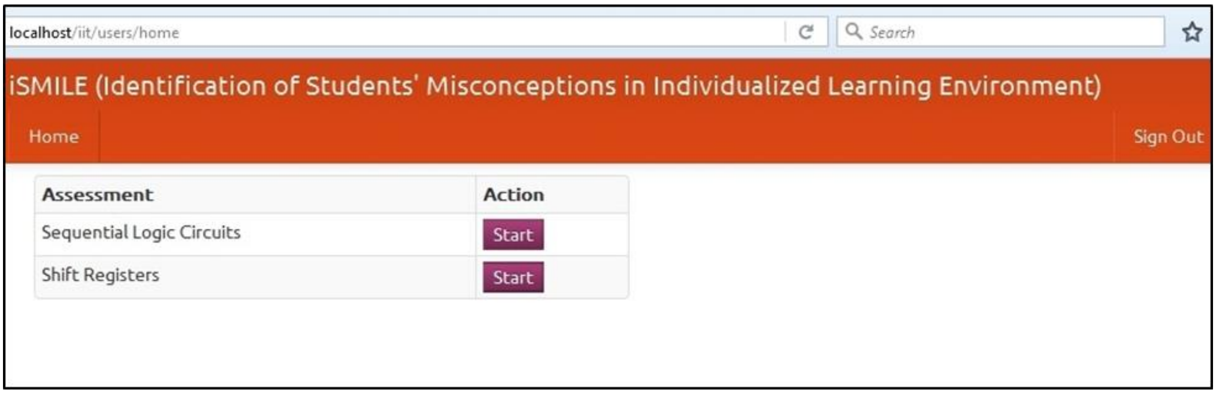

Fig. 6. Interface displaying students' assessment screen of iSMILE system.

Feedback: The student will be provided with detailed feedback upon completion of the assessment. The feedback includes the questions attended, and the answers provided by the user. There will be separate section for feedback on each question. The feedback will be stored in the application, and the user can view it at any time, as shown in Fig. 7.

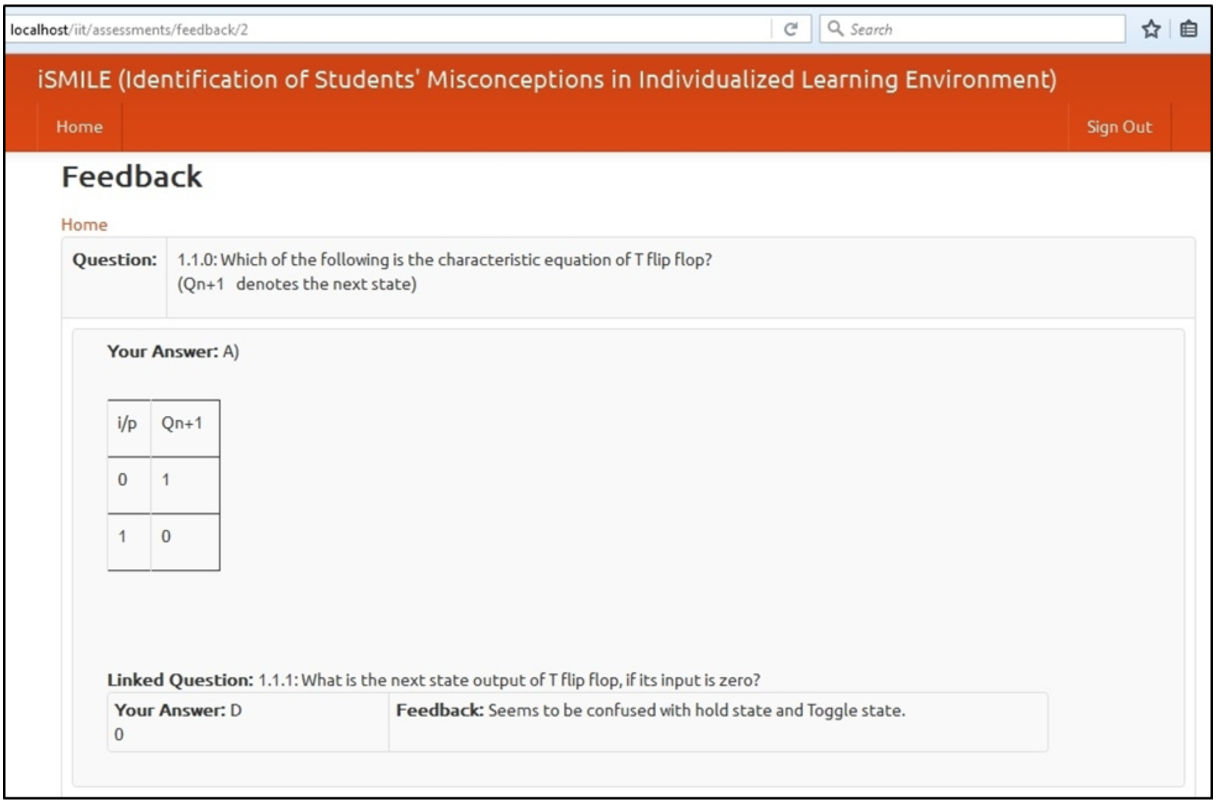

Fig. 7. Interface displaying feedback screen of iSMILE system.

WORKING PROCESS OF THE ISMILE SYSTEM

The login page of iSMILE user interface can be accessed in any web browser with a local address (e.g. http://localhost/iit/), and the option for sign-in is provided. In the first step, if the user has not registered, then he/she can register with 'create account' option. For registering a new instructor by assigning a user's email id and password, one has to get the administrative privileges. Thus, through manual editing in server computer, new instructor login is created. In the second step, after opening http://localhost/phpmyadmin/ , the system administrator can add as many instructors as required. In the third step, user has to finish the registration process, and iSMILE system will identify the student or instructor, based on the sign-in credentials they are given. We have created two different types of registration to ensure security.

When the instructor logs in the system, the application of the iSMILE provides the links (e.g., topics, questions and question-sets). If there are no question sets previously entered or if the instructor wants to create new question sets, he/she can use 'create topics', 'create question-sets' and 'create question' options to enter two-level objective questions on various topics. Each question will have two meta-data fields: instructional objective and misconception. After entering the second-level question based on the misconception, the instructor has to enter feedbacks for each possible option. The idea behind the feedback is to pinpoint the misconception. In the final step, entering all questions and feedbacks in a question-set, it has to be published by using the 'publish' button. All unpublished questions are visible to the instructor only and can be edited at any time before publishing.

When a student logs in, all the available (published) question-sets are displayed to the student. He/she can attend the assessment and can view the feedback immediately. There is a provision to skip a question also. This is to avoid random guessing. The feedbacks are stored in the system for the students, and they can view previous feedbacks at any later login. The flow charts for student's registration, login and assessment are shown in **Fig. 8**, **Fig. 9**, and **Fig. 10** respectively. The working process is further illustrated using a model set of questions from the course of Digital System Design.

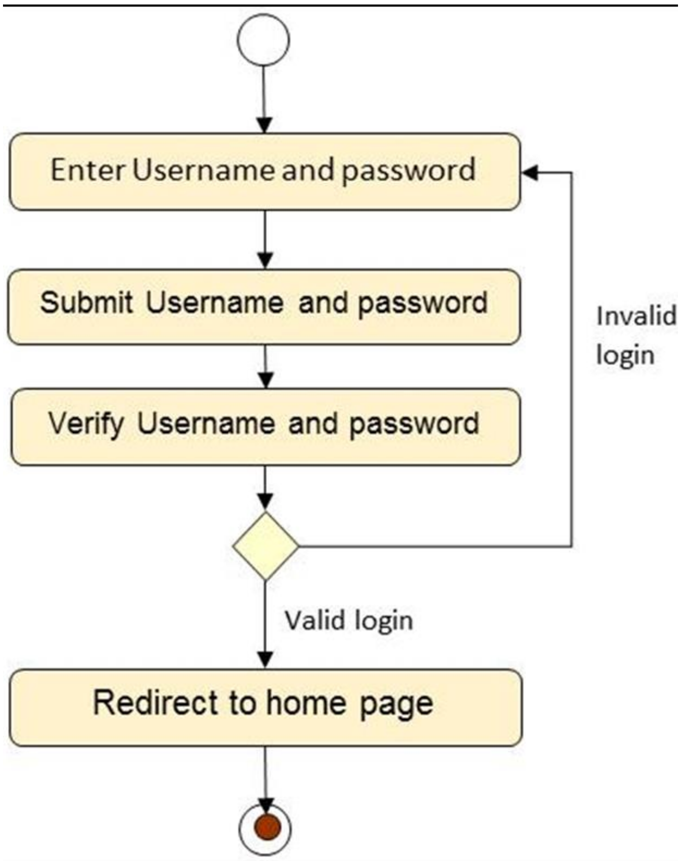

Fig. 8. Flowchart of student's registration.

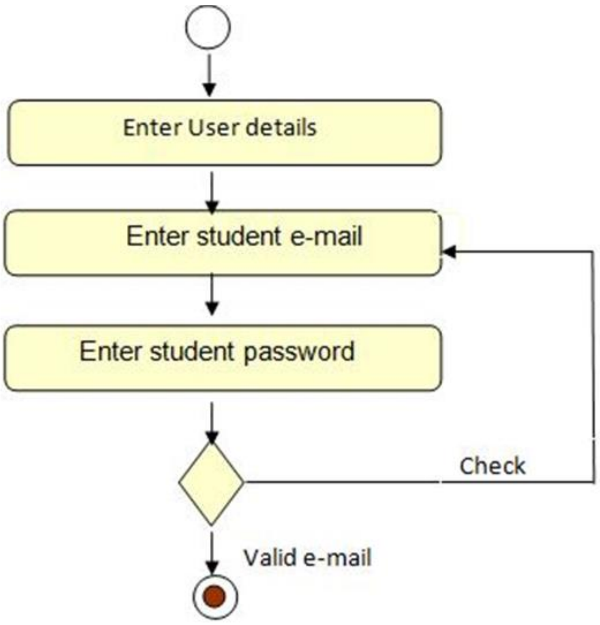

Fig. 9. Flowchart of student's login.

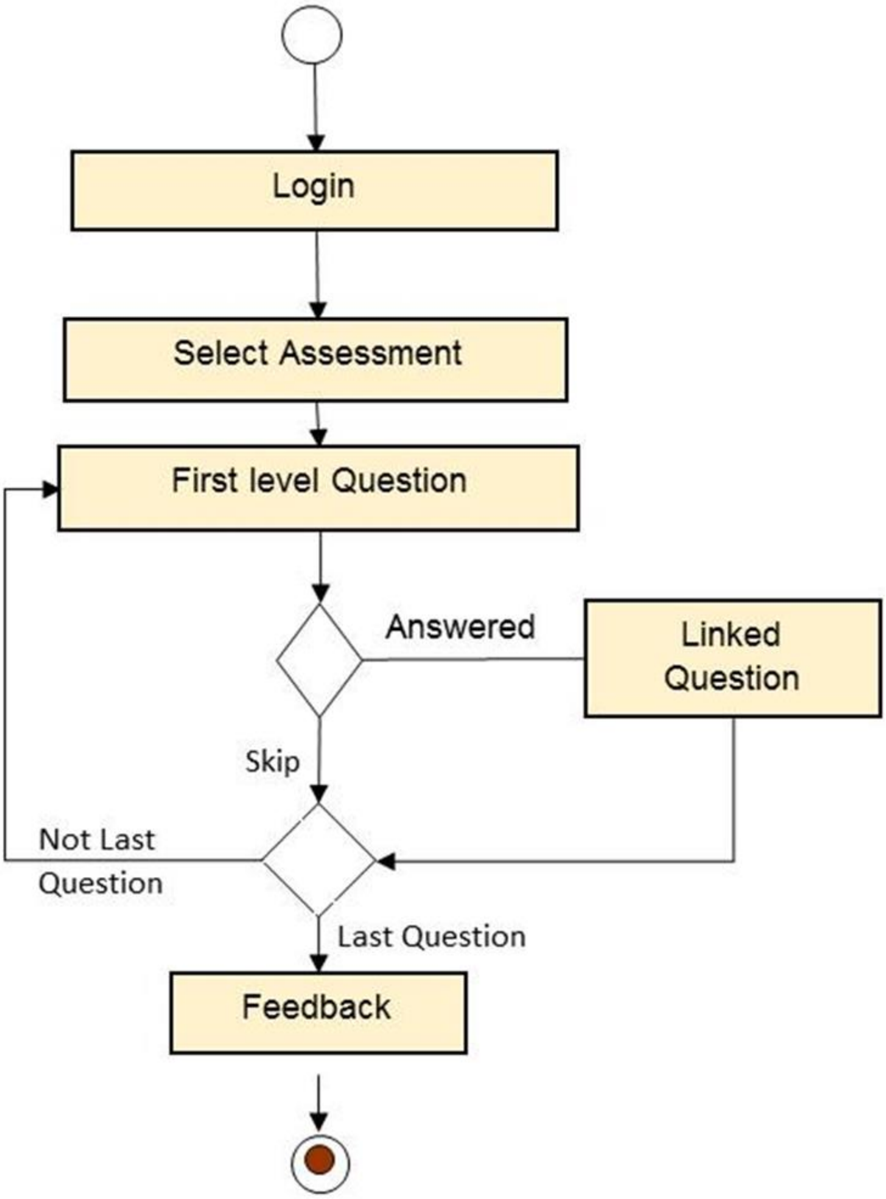

Fig. 10. Flowchart of student's assessment.

Instructional Objective: To write the characteristic (or function) tables for SR, JK, D and T flip-flops (FFs)

Misconception: Similarity in characteristic equations of T and D; SR and JK FFs.

(Root Question)

1.1.0: Which of the following is the characteristic equation of T flip flop?

(Q_{n+1} denotes the next state)

(a)

i/p	Q_{n+1}
0	1
1	0

(b)

i/p	Q_{n+1}
0	Q_n
1	$\overline{Q_{n+1}}$

(c)

i/p	Q_{n+1}
0	0
1	1

(d)

i/p	Q_{n+1}
0	$\overline{Q_{n+1}}$
1	Q_n

Correct answer is option (b).

If the student's answer is (a) for 1.1.0 (Root Question) then

(Second Level Question)

1.1.1: What is the next state output of T flip flop, if its input is zero?

(a) Hold state

(b) 1

(c) Complement of present state

(d) 0

Correct answer is option (a).

Feedbacks: For second level answer (a), and so on.

(A) Your idea regarding Toggle state of a flip-flop is improper.

(B) Your answers are contradictory. You might be guessing.

(C) Try to make a clear idea about Toggling.

(D) You seem to be confused with hold state and Toggle state.

If the student's answer is (b) for 1.1.0 (Root Question), then

(Second Level Question)

1.1.2 How a J K flip-flop is made to toggle its present state?

(a) J=0, K=0

(b) J=1, K=0

(c) J=0, K=1

(d) J=1, K=1

Correct answer is option (d).

Feedbacks: For second level answer (a), and so on.

(A) You are confused with HOLD state and TOGGLE state of JK FFs

(B) You have some problem in understanding the SET condition and TOGGLE states of JK FFS

(C) You have idea about T FFs but don't have proper idea regarding the TOGGLE and RESET state of JK FF

(D) Your Answer is excellent. The instructional objective has been met.

If the student's answer is (c) for 1.1.0 (Root Question) then

(Second Level Question)

1.1.3 What is the next state output of T flip flop if the input is 1?

(a) 1 (c) Q_n

(b) 0 (d) $\overline{Q_n}$

Correct answer is option (d).

Feedbacks: For second level answer (a), and so on.

(A) You might be confused with T FF and D FF characteristic equations.

(B) You are guessing, as your answers are contradictory.

(C) You are guessing, as your answers are contradictory.

(D) Your answers say the TOGGLE and HOLD states are confusing for you.

If the student's answer is (d) for 1.1.0 (Root Question), then

(Second Level Question)

1.1.4 If J=0, K=0 in J K flip flop, then output Q n+1 (the next state output) will be

(a) 1

(b) 0

(c) $\overline{Q_n}$

(d) Q_n

Correct answer is option (d).

Feedbacks: For second level answer (a), and so on.

(A) Your first and second level answers say you are confused with D and T FFs.

(B) It seems that you are confused with the HOLD and TOGGLE states

(C) You are guessing, as your answers are contradictory.

(D) Your idea regarding different states of T FFs is improper.

CONCLUSION

To make the formative assessment more effective, this work presents iSMILE system, which is based on students' misconceptions. This paper describes the design and features of iSMILE system and how it is better than the existing web-based formative evaluation tools. For every instructional objective, a root (first level) and four linked (second level) multiple-choice questions are prepared with respect to the identified misconceptions. At least one misconception is included in every instructional objective as a metadata field and based on this a two-level diagnostic test is carried out. Thus, based on the answers provided by the student to the first-level questions and second level questions, feedback is given to improve the learning skills of the student. In addition, feedback highlights the misconceptions associated with students' conceptual understanding. This feature of iSMILE distinguishes itself from other available web-based formative assessment systems. Our system uses the open-source ICT tools (like MySQL, Apache); therefore, it is free of cost and can be accessed from anywhere and at any time. Our system has not yet been evaluated inside real classroom settings, which is a limitation of our work. For future research, we are planning to incorporate the iSMILE system into formal classroom settings, in order to evaluate the learning effectiveness of this software.

ACKNOWLEDGEMENTS

This project is supported in part by the Ministry of Science and Technology, Taiwan, R.O.C. (previously named National Science Foundation of Taiwan) under the Grant number NSC102-2515-S-103-001-H, "Aim for the Top University Project" sponsored by the Ministry of Education, Taiwan, R.O.C., and "International Research-Intensive Center of Excellence Program" of NTNU and Ministry of Science and Technology, Taiwan, R.O.C. under Grant no. MOST 104-2911-I-003-301.

REFERENCES

Attali, Y. (2015). Effects of multiple-try feedback and question type during mathematics problem solving on performance in similar problems. *Computers & Education, 86*, 260-267. doi:10.1016/j.compedu.2015.08.011

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

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

Crisp, V., & Ward, C. (2008). The development of a formative scenario-based computer assisted assessment tool in psychology for teachers: The PePCAA project. *Computers & Education, 50*(4), 1509-1526. doi:10.1016/j.compedu.2007.02.004

Farrell, T., & Rushby, N. (2016). Assessment and learning technologies: An overview. *British Journal of Educational Technology, 47*(1), 106-120. doi:10.1111/bjet.12348

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional children, 53*(3), 199-208.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Hsu, J.-L., Chou, H.-W., & Chang, H.-H. (2011). EduMiner: Using text mining for automatic formative assessment. *Expert Systems with Applications, 38*(4), 3431-3439. doi:10.1016/j.eswa.2010.08.129

Landauer, T. K., Lochbaum, K. E., & Dooley, S. (2009). A New Formative Assessment Technology for Reading and Writing. *Theory Into Practice, 48*(1), 44-52. doi:10.1080/00405840802577593

Lefevre, D., & Cox, B. (2016). Feedback in technology-based instruction: Learner preferences. *British Journal of Educational Technology, 47*(2), 248-256. doi:10.1111/bjet.12227

Pacheco-Venegas, N. D., López, G., & Andrade-Aréchiga, M. (2015). Conceptualization, development and implementation of a web-based system for automatic evaluation of mathematical expressions. *Computers & Education, 88*, 15-28. doi:10.1016/j.compedu.2015.03.021

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

Rodrigues, F., & Oliveira, P. (2014). A system for formative assessment and monitoring of students' progress. *Computers & Education, 76*, 30-41. doi:10.1016/j.compedu.2014.03.001

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119-144. doi:10.1007/bf00117714

Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research, 78*(1), 153-189. doi:10.3102/0034654307313795

Wang, T.-H. (2008). Web-based quiz-game-like formative assessment: Development and evaluation. *Computers & Education, 51*(3), 1247-1263. doi:10.1016/j.compedu.2007.11.011

Wang, T.-H. (2010). Web-based dynamic assessment: Taking assessment as teaching and learning strategy for improving students' e-Learning effectiveness. *Computers & Education, 54*(4), 1157-1166. doi:http://dx.doi.org/10.1016/j.compedu.2009.11.001

Yuan, J., & Kim, C. (2015). Effective Feedback Design Using Free Technologies. *Journal of Educational Computing Research, 52*(3), 408-434. doi:10.1177/0735633115571929

http://iserjournals.com/journals/eurasia

Bhagat, K. K., & Spector, J. M. (2017). Formative Assessment in Complex Problem-Solving Domains: The Emerging Role of Assessment Technologies. *Educational Technology & Society, 20* (4), 312–317.

Formative Assessment in Complex Problem-Solving Domains: The Emerging Role of Assessment Technologies

Kaushal Kumar Bhagat[1*] and J. Michael Spector[2]
[1]Smart Learning Institute, Beijing Normal University, China // [2]Department of Learning Technologies, University of North Texas, USA // kkntnu@hotmail.com // mike.spector@unt.edu
[*]Corresponding author

ABSTRACT

Much of the focus on learning technologies has been on structuring innovative learning experiences and on managing distance and hybrid learning environments. This article focuses on the use of technology as an important formative assessment and feedback tool. The rationale for this focus is based on prior research findings that suggest that timely and informative feedback promotes learning. The general purpose of this article is to promote a focus on formative assessment, especially with challenging problems so as to help develop critical thinking skills. This is not a meta-analysis although we encourage such studies so as to emphasis the role that formative assessment plays in supporting learning. Much of the prior research on formative assessment has not involved advanced digital technologies and very little of that research has focused on complex and challenging problem-solving tasks. We review prior work on the use of problem conceptualizations elicited during problem solving activities and on stealth assessments of learner choices and decisions in online activities. We present a conceptual framework based on prior research and theory for conducting formative assessments in real-time with regard to complex problem-solving tasks. We then present an elaboration of how formative assessment can be used to support learning a common intellectual skill involving a discrimination task and to develop an appropriate cognitive strategy for that kind of problem. We conclude with recommendations for further research on the use of technology in support of formative assessment.

Keywords

Cognitive strategies; Complex learning; Dynamic feedback; Formative assessment; Intellectual skills; Self-regulation

Introduction

There are three major findings from research on learning in the last 50 or more years (Narciss, 2008; Spector & Yuen, 2016): (a) time on task predicts learning outcomes, (b) formative feedback tends to improve learning, and (c) prior knowledge and experience influences learning. These three findings from learning research are related and have direct implications for formative assessment. First, however, we need to define the scope and purpose of formative assessment. Formative assessment is feedback provided to the learner during an instructional sequence or learning activity that is aimed at helping the learner succeed. Timely and informative feedback is essential for formative assessment to be effective, although the amount and timing of feedback should be appropriate for a particular learner. This is how prior learning and experience is related to formative assessment. More advanced learners require less support and may regard too much feedback as intrusive according to the research on cognitive apprenticeship (Collins, Brown, & Newman, 1987). In addition, learners who spend more time on a learning activity or task are likely to gain more understanding and competence, so formative feedback that encourages continued engagement is also likely to be more effective in support of learning (Spector & Park, in press).

In general, there is sufficient research on formative assessment to support learning simple tasks in well-structured domains with outcomes targeted at simple concepts and procedures. The explosion of new technologies makes such support ever more effective. What is less well understand is how best to support learning complex and ill-structured tasks and how best to use new technologies to support formative assessment in those situations. That gap is the focus of this contribution. The good news is that there are new techniques and findings emerging that should support progress in this important domain.

Formative assessment research

Research on formative assessment presents particular challenges, even though the benefits are widely acknowledged (Black & Wiliam, 1998; Dunn & Mulvonen, 2009; Fuchs & Fuchs, 1986). One challenge is to determine the influence of formative assessment on learning. If learning is defined as stable and persistent changes in what a person knows or can do (Spector & Yuen, 2016), then there is limited research as there are few studies that examine the persistence of learning and fading of knowledge and skills over time. A second challenge involves how and when formative assessments are provided. That is to say, while a particular feedback may be intended as constructive and encouraging, it may be perceived as discouraging or disheartening. In addition, a delayed feedback may not serve well to support continued engagement in a learning activity or instructional sequence. Finally, determining whether it was a formative assessment that led to an impact on learning rather than other factors is rarely explored in a controlled manner.

One reason for the lack of randomized controlled studies involving formative assessment is an ethical about potentially disadvantaging a group of learners not being provided a preferred form of feedback. One way to address that concern is through a within subject design using repeated and alternating treatments. Such a study should be sufficiently long with sufficient alterations to potentially observe differences due to different forms of feedback and should also involve measures before, during, after and long after the instructional sequence. In such a study there is the possibility that a learner will begin to generate a form of self-assessment that is similar to the kind of formative feedback expected to be most productive in terms of learning gains. Such a meta-cognitive learning outcome can be determined and is a desirable outcome aligned with the notion of self-regulated learning (Butler & Winne, 1995).

Complex learning

Some learning tasks and problems are more complex than others. Complexity can result from there being many interacting factors, from some of those interactions being non-linear with delayed effects, from changes in the problem situation changing while a solution approach is being formulated, and from lack of learner familiarity with the general nature of the problem or problem domain (Spector, 2015; van Merriënboer, 2012). Such problems and learning tasks often have multiple acceptable solution approaches and solutions, which presents a unique challenge for formative assessment. If supportive and potentially corrective feedback is not provided early in the process and in a meaningful manner, then a learner may develop misconceptions that are difficult to overcome later in a learning progression.

One fundamental theoretical foundation that has existed for some time is based on the notions of authentic learning introduced by Dewey (1938) and more recently elaborated in the form of situated learning (Lave & Wenger, 1991). These theoretical foundations strongly support two things. First, it is important to use meaningful and realistic problems to help develop complex problem-solving skills. Of course, prerequisite knowledge cannot be assumed, although the problem-based learning community argues that much prerequisite knowledge should be introduced while working on real problems (Barrows & Tamblyn, 1980). Complex problems that occur in so many domains are typically addressed by small groups of specialists. As a consequence, developing sufficient competence to be recognized as a contributing member of a problem-solving team in an important consideration (Lave & Wenger, 1991; Milrad, Spector, & Davidsen, 2003). In short, to develop complex problem—solving skills involves the development of competence and confidence. Previous attempts that emphasis summative assessments have tended to focus on the competence aspect of that developmental process. However, legitimate peripheral participation and more recent approaches place equal emphasis on the confidence aspect involved in critical thinking and complex learning (Kelller, 2010; Lave & Wenger, 1991; Milrad, Spector & Davidsen, 2003).

One approach, consistent with the foundations just reviewed, is to follow a path of graduated complexity and help the learner develop competence in a challenging problem domain (Milrad, Spector & Davidsen, 2003). Consistent with a graduated complexity approach is to use partially worked examples beginning with examples missing only a small part of an acceptable solution and then introduce examples with more and more parts missing eventually allowing a learner to develop an entire solution approach. In such cases, digital technologies can support such a process by knowing where a learner is within a learning progression path and introducing increasingly challenges examples to complete.

Yet another approach is to identify how experts typically think about that problem and use technology to encourage a learner to think more like an expert (Spector & Koszalka, 2004). This can be done by asking experts and learners to respond to four questions: (a) What key factors influence the problem situation? (b) How would you describe each of those factors? (c) What are the relationships among those factors? and (d) How would describe each of those relationships? Those responses can be put into the form of an annotated concept map and used as the basis for formative assessment.

These approaches are consistent with Dewey (1938), Lave and Wenger (1991), as well as with mainstream instructional design experts such as Gagné and Merrill (1990), Reigeluth (1999), and van Merriënboer (2012).

Gagné and Merrill (1990) proposed focusing on enterprises, which was called a whole task somewhat later by van Merriënboer (2012).Enterprises and whole tasks are authentic (representative of actual tasks) and consistent with the notion of problem-based learning (Barrows & Tamblyn, 1980). What is changing is the ability to provide real-time, meaningful feedback during problem solving using current and emerging technologies, which has been demonstrated by a few of the assessment technologies discussed in the next section.

Current assessment technologies and systems

Researchers have already accepted the importance of formative assessment in the teaching-learning process. In the 21st century, the availability and use of technological resources like computers, mobile devices, tablets, etc. are increasing rapidly in the school and higher education (Organisation for Economic Co-operation and Development (OECD), 2015). Now the question arises, "What kind of roles can technology play to support formative assessment?" For example, Bennett and Gitomer (2009) numerated some of the important benefits of using technology, which are: (a) more informative, technology can help to track the full record of the problem-solving process adapted by the learner; (b) more efficient, saves time for scoring and error free; and (c) cost-effective, saves the expenses over the human scoring process. Spector et al. (2016) also advocated the application of current available technologies for formative assessment purposes. In this section, we discussed some of the currently available systems for formative assessment. In addition, we presented some empirical evidence supporting the key role played by different technologies in formative assessment we discussed some of the currently available systems for formative assessment.

While the gap addressed herein concerns the lack of formative assessment in complex problem-solving domains, there have been recent improvements in technologies, which can and have addressed that gap, at least in research settings. These technologies, however, have yet to see large-scale implementation in instructional settings, so the gap previously discussed remains. However, the means to systematically address that gap are now possible due to emerging technologies. Recently, many digital assessment tools have been developed, which may provide support to FA. These tools include HIMATT (Pirnay-Dummer, Ifenthaler, & Spector, 2010), AKOVIA (Ifenthaler, 2014), AssiStudy (Rodrigues & Oliveira, 2014), iSMILE (Bhagat, Subheesh, Bhattacharya, & Chang, 2017), etc.

HIMATT (Highly Integrated Model Assessment Technology and Tools) is a comprehensive tool which combines the features of DEEP (Dynamic Enhanced Evaluation of Problem Solving), MITOCAR (Model Inspection Trace of Concepts and Relations), T-MITOCAR (Text-MITOCAR), and SMD Technology (Structure, Matching, Deep Structure). HIMATT has two platforms: HIMATT Research Engine, which conducts and analyses experiments and HIMATT Subject Environment, which assigns the experiments to the individuals dynamically. The application of HIMATT includes states and changes, analysis and comparisons.

The framework of AKOVIA (Automated Knowledge Visualization and Assessment) is based on HIMATT. It is applicable for the semantic analysis of natural language (e.g., discussion forums, essay writing) and graphical knowledge representations. Automated feedback is one of the key features of AKOVIA, which can help the learners to understand their writing and improve it accordingly in an effective way.

AssiStudy is based on Service-Oriented Architectures (SOA). It creates personalized training exams based on students' profile using question from the past exams stored in the repository. These training exams provide immediate feedback explaining the mistakes made by the students. This system used various Natural Language Processing (NLP) techniques to match reference answers (RA) with student answers (SA). After the training exams, teachers use students' performance information to develop evaluation exams. Three main types of exams can be created: enumeration, specific knowledge and essay. Evaluation exams are checked by both AssiStudy and the teacher. This system tracks the performance of each student in formative as well as summative assessment and provides detailed pattern analysis to the teachers.

Identification of Students' Misconceptions in Individualized Learning Environment (iSMILE) System is developed to provide feedbacks based on misconceptions in understanding a particular concept. This system is based on Model View Controller (MVC) architecture. Assessment procedure has two levels. Firstly, student needs to answer a root question. In the next step, a linked question is provided based on the answer for the root question to evaluate the deeper understanding of the concept. After finishing both levels, students are provided elaborated feedback about their misconceptions if they make any mistakes.

Hwang and Chang (2011) developed a Formative Assessment-Based Mobile Learning (FAML) system and evaluated on a local culture course learning. FAML used to provide only hints, when the students failed to find correct answers and motivated the learners to find the answers by their own. The results showed higher learning performance, learning attitude, and learning motivation by using FAML. In another study, an online game-based formative assessment named tic-tac-toe quiz for single-player (TRIS-Q-SP) was developed by Hooshyar et al. (2016). In addition, Hooshyar and colleagues (2016) integrated TRIS-Q-SP with an Intelligent Tutoring System (ITS) to teach computer programming. Three types of feedback (delayed feedback, knowledge result and elaborated feedback) were integrated within the system. The results revealed that TRIS-Q-SP improved experimental group's problem-solving skills, which resulted into better learning achievement. More importantly, immediate elaborated feedback was one of the reasons for better learning performance. Conejo, Garcia-Viñas, Gastón, and Barros (2016) employed a web-based assessment tool called *Siette* in a botany course for higher education. This system provided elaborated immediate feedback, which means the instant correct answer with detailed explanation for the wrong answers. They found that students who used *Siette*, performed better than the students who underwent traditional method for formative assessment. More interestingly, immediate feedback helped the students for better performance, which is not possible without the use of technology. Recently, Faber, Luyten, and Visscher (2017) examined the effects of a digital formative assessment tool, *Snappet* on students' mathematics achievement and motivation. This system also provides elaborated feedback about the errors made by the students and explanation for the correct answers. They also concluded that feedback contributed for the better performance and motivation of the students who used *Snappet.*

In addition, there is important research emerging with regard to the use of advanced technologies to support stealth assessment, which involves the use of log data in an online learning environment to determine areas in which a learner may be struggling and then help guide that learner to more productive outcomes (DiCerbo, Shute, & Kim, 2017; Shute & Moore, 2017). Moreover, the ability of digital technologies to collect data and present visualizations to help a learner make progress are also becoming more prominent in support of learning (Wang, Wu, Kinshuk, Chen, & Spector, 2013; Wu, Wang, Grotzer, Liu, & Johnson, 2016; Wu, Wang, Spector, & Yang, 2013; Yuan, Wang, Kushniruk, & Peng, 2016). Such visualizations often involve extensions of concept mapping techniques previously mentioned and address both content representations as well as feedback on student performance.

Looking to the future

Overall, research indicates that technology can support formative assessment to enhance learning performance, learning attitude, and learning motivation effectively across different disciplines. There is no doubt that technology can be used to support formative assessment, although technology has more often been used to provide access to and interaction with learning resources. Given the history of emphasis on formative assessment and the potential of new technologies to extend formative assessment into complex problem-solving domains, the potential for greater impact of formative assessment on the development of competence with regard to higher order learning is high. Promising technologies include stealth assessments, automated concept map based assessments, visualizations in support of formative assessment and self-regulation skills, and tools to promote networking and collaboration.

The conceptual framework we propose is to build on the following key notions:

- Continue to introduce real-world problems, however simplified, into curricula whenever possible;
- Build on the notions of graduated complexity (Milrad, Spector & Davidsen, 2003), enterprises (Gagné & Merrill, 1990), elaboration theory (Reigeluth, 1999), and whole tasks (van Merriënboer, 2012);
- Use annotated concept maps and causal influence diagrams as a means to elicit how someone is thinking about a complex problem;
- Compare progress towards expert-like thinking based on a series of problem conceptualizations.

Concluding remarks

In closing, it is our belief that formative assessment is an important albeit neglected task of educators. The primary job of a teacher is to help students succeed in their learning activities and educational pursuits. If one accepts that last remark, then one must continue to place emphasis on formative assessment and contribute to ongoing efforts to make effective use of new and emerging technologies in support of formative assessment, especially in complex problem-solving domains.

There are a number of ways to move forward with regard to a framework for supporting the development of complex problem-solving skills. These ways include the following:

- Develop repositories of representative complex problems in a variety of domains (this is being done with regard to a number of science learning tasks by the Smithsonian Institution – see https://ssec.si.edu/);
- Develop associated repositories of how experts think about those problems using annotated problem conceptualizations of the type found in DEEP and HIMATT (Pirnay-Dummer, Ifenthaler, & Spector, 2010; Spector & Koszalka, 2004);
- Develop a version of HIMATT that can be used in classroom and online course settings to provide dynamic, real-time feedback as a learner develops a conceptualization of the problem; and,
- Track and report the development of complex problem-solving skills because of using these technologies.

We have not reported new research in this short piece. Rather, we have been urging a particular perspective to support further research on the most promising technologies and learning approaches that have evolved in the last 50 years. What would be genuinely innovative and original would be to plan, implement, deploy on a large-scale version of the proposed approach and determine the impact on the development of critical thinking and complex problem-solving skills now being emphasized in so many places.

References

Barrows, H. S., & Tamblyn, R. M. (1980). *Problem-based learning: An Approach to medical education*. New York, NY: Springer.

Bennett, R. E., & Gitomer, D. H. (2009). Transforming K–12 assessment: Integrating accountability testing, formative assessment and professional support. In C. Wyatt-Smith & J. J. Cumming (Eds.), *Educational assessment in the 21st century: Connecting theory and practice* (pp. 43-61). Dordrecht, The Netherlands: Springer.

Bhagat, K. K., Subheesh, N., Bhattacharya, B., & Chang, C.-Y. (2017). The Design and Development of Identification of Students' Misconceptions in Individualized Learning Environment (iSMILE) system. *EURASIA Journal of Mathematics Science and Technology Education, 13*(1), 19-34.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-74.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A Theoretial synthesis. *Review of Educational Research, 65*(3), 245-281.

Collins, A., Brown, J. S., & Newman, S. E. (1987). *Cognitive apprenticeship: Teaching the craft of reading, writing and mathematics* (Technical Report No. 403). Cambridge, MA: Centre for the Study of Reading, University of Illinois.

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

Dewey, J. (1938). *Experience and education.* New York, NY: Kappa Delta Pi.

DiCerbo, K., Shute, V. J. & Kim, Y. J. (2017). The Future of assessment in technology rich environments: Psychometric considerations. In J. M. Spector, B. Lockee, & M. Childress (Eds.), *Learning, design, and technology: An international compendium of theory, research, practice, and policy* (pp. 1-21). New York, NY: Springer. doi:10.1007/978-3-319-17727-4_66-1

Dunn, K. E., & Mulvenon, S. W. (2009). A Critical review of research on formative assessment: The Limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation, 14*(7), 1-11.

Faber, J. M., Luyten, H., & Visscher, A. J. (2017). The Effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment. *Computers & Education, 106*, 83-96. doi:10.1016/j.compedu.2016.12.001

Fuchs, I. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A Meta-analysis. *Exceptional Children, 53*, 199-208.

Gagné, R. M., & Merrill, M. D. (1990). Integrative goals for instructional design. *Educational Technology Research & Development, 38*(1), 23-30.

Hooshyar, D., Ahmad, R. B., Yousefi, M., Fathi, M., Horng, S.-J., & Lim, H. (2016). Applying an online game-based formative assessment in a flowchart-based intelligent tutoring system for improving problem-solving skills. *Computers & Education, 94*, 18-36. doi:10.1016/j.compedu.2015.10.013

Hwang, G.-J., & Chang, H.-F. (2011). A Formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. *Computers & Education, 56*(4), 1023-1031. doi:10.1016/j.compedu.2010.12.002

Ifenthaler, D. (2014). AKOVIA: Automated Knowledge Visualization and Assessment. *Technology, knowledge and learning, 19*(1-2), 241-248.

Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS model approach.* New York, NY: Springer.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Milrad, M., Spector, J. M., & Davidsen, P. I. (2003). Model facilitated learning. In S. Naidu (Ed.), *Learning and teaching with technology: Principles and practices* (pp. 13-27). London, UK: Kogan Page.

Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. J. G. van Merriënboer, & M. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed., pp. 125-143). New York, NY: Routledge.

Organisation for Economic Co-operation and Development (OECD). (2015). *Students, computers and Learning: Making the connection, PISA.* Paris, France: OECD Publishing. doi:10.1787/9789264239555-en

Pirnay-Dummer, P., Ifenthaler, D., & Spector, J. M. (2010). Highly integrated model assessment technology and tools. *Educational Technology Research and Development, 58*(1), 3-18.

Reigeluth, C. M. (1999). The Elaboration theory: Guidance for scope and sequence decisions. In C. M. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory* (pp. 425-453). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Rodrigues, F., & Oliveira, P. (2014). A System for formative assessment and monitoring of students' progress. *Computers & Education, 76*, 30-41. doi:10.1016/j.compedu.2014.03.001

Shute, V. J., & Moore, G. R. (2017). Consistency and validity in game-based stealth assessment. In H. Jiao & R. W. Lissitz (Eds.), *Technology enhanced innovative assessment: Development, modeling, and scoring from an interdisciplinary perspective* (pp. 31-51). Charlotte, NC: Information Age Publisher.

Spector, J. M. (2015). System dynamics modeling. In J. M. Spector (Ed.), *The SAGE Encyclopedia of educational technology* (pp. 693-697). Thousand Oaks, CA: Sage Publications.

Spector, J. M., Ifenthaler, D., Samspon, D., Yang, L., Mukama, E., Warusavitarana, A., Lokuge Dona, K., Eichhorn, K., Fluck, A., Huang, R., Bridges, S., Lu, J., Ren, Y., Gui, X., Deneen, C. C., San Diego, J., & Gibson, D. C. (2016). Technology enhanced formative assessment for 21st century learning. *Educational Technology & Society, 19*(3), 58-72.

Spector, J. M., & Koszalka, T. A. (2004). *The DEEP methodology for assessing learning in complex domains* (Final report to the National Science Foundation Evaluative Research and Evaluation Capacity Building). Syracuse, NY: Syracuse University.

Spector, J. M., & Park, S.-W. (in press). *Motivation, learning and technology: Embodied educational motivation*. New York, NY: Routledge.

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

van Merriënboer, J. J. G. (2012). Complex learning. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 681-682). New York, NY: Springer.

Wang, M., Wu, B., Kinshuk, Chen, N. S., & Spector, J. M. (2013). Connecting problem solving and knowledge-construction processes in a visualization-based learning environment. *Computers & Education, 68*, 293–306.

Wu, B., Wang, M., Grotzer, T. A., Liu, J., & Johnson, J. M. (2016). Visualizing complex processes using a cognitive-mapping tool to support the learning of clinical reasoning. *BMC Medical Education, 16*, 216. doi:10.1186/s12909-016-0734-x

Wu, B., Wang, M. Spector, J. M., & Yang, S, J. H. (2013). Design of a dual-mapping learning approach for problem solving and knowledge construction in ill-structured domains. *Educational Technology & Society, 16*(4), 71-84.

Yuan, B., Wang, M., Kushniruk, A. W., & Peng, J. (2016). Design of a computer-based learning environment to support diagnostic problem solving towards expertise development. *Knowledge Management & E-Learning, 8*(4), 540–549.

International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

Investigating the interrelationships among conceptions of, approaches to, and self-efficacy in learning science

Lanqin Zheng, Yan Dong, Ronghuai Huang, Chun-Yen Chang & Kaushal Kumar Bhagat

To cite this article: Lanqin Zheng, Yan Dong, Ronghuai Huang, Chun-Yen Chang & Kaushal Kumar Bhagat (2017): Investigating the interrelationships among conceptions of, approaches to, and self-efficacy in learning science, International Journal of Science Education

To link to this article: https://doi.org/10.1080/09500693.2017.1402142

Published online: 22 Nov 2017.

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2017
https://doi.org/10.1080/09500693.2017.1402142

Routledge
Taylor & Francis Group

Check for updates

Investigating the interrelationships among conceptions of, approaches to, and self-efficacy in learning science

Lanqin Zheng[a], Yan Dong[a], Ronghuai Huang[b], Chun-Yen Chang[c] and Kaushal Kumar Bhagat [b]

[a]School of Educational Technology, Faculty of Education, Beijing Normal University, Beijing, People's Republic of China; [b]Smart Learning Institute, Beijing Normal University, Beijing, People's Republic of China; [c]Graduate Institute of Science education, National Taiwan Normal University, Taipei, Taiwan

ABSTRACT

The purpose of this study was to examine the relations between primary school students' conceptions of, approaches to, and self-efficacy in learning science in Mainland China. A total of 1049 primary school students from Mainland China participated in this study. Three instruments were adapted to measure students' conceptions of learning science, approaches to learning science, and self-efficacy. The exploratory factor analysis and confirmatory factor analysis were adopted to validate three instruments. The path analysis was employed to understand the relationships between conceptions of learning science, approaches to learning science, and self-efficacy. The findings indicated that students' lower level conceptions of learning science positively influenced their surface approaches in learning science. Higher level conceptions of learning science had a positive influence on deep approaches and a negative influence on surface approaches to learning science. Furthermore, self-efficacy was also a hierarchical construct and can be divided into the lower level and higher level. Only students' deep approaches to learning science had a positive influence on their lower and higher level of self-efficacy in learning science. The results were discussed in the context of the implications for teachers and future studies.

ARTICLE HISTORY
Received 30 August 2017
Accepted 3 November 2017

KEYWORDS
Conception of learning; approaches to learning; self-efficacy; science education

Introduction

In recent years, there has been an increase in attention to students' perceptions and interpretations of learning science experiences in the field of science education (Lin & Tsai, 2013a; Lin, Tsai, & Liang, 2012; Shen, Lee, Tsai, & Chang, 2016). Researchers have adopted different perspectives to explore the conceptions of learning science, approaches to learning science, and self-efficacy. Several studies have investigated conceptions of learning science in high school (Lee, Johanson, & Tsai, 2008; Tsai, 2004) and undergraduate students (Chiou, Liang, & Tsai, 2012). For example, Lee et al. (2008) identified lower level conceptions of learning (i.e. memorising, testing, calculating, and practicing) and higher level conceptions of learning (i.e. increase of knowledge,

CONTACT Kaushal Kumar Bhagat kkntnu@hotmail.com Smart Learning Institute, Beijing Normal University, 12F, Block A, Jingshi Technology Building, No. 12 Xueyuan South Road, Haidian District, Beijing 100082.

applying, and understanding) in high school students. Furthermore, some researchers have started to investigate approaches to learning to examine students' experiences learning science. There were two major approaches to learning science: deep and surface approaches (Kember, Biggs, & Leung, 2004; Li, Liang, & Tsai, 2013). In addition, self-efficacy is considered a crucial aspect of interpreting students' perceptions of learning science (Lin & Tsai, 2013b). Previous studies revealed that self-efficacy was closely related to students' learning processes and outcomes (Liem, Lau, & Nie, 2008; Printrich & Schunk, 2002). Therefore, conceptions of, approaches to, and self-efficacy were considered three primary factors of interpreting students' experiences learning science.

Furthermore, the relationships between conceptions of and approaches to learning science at the high school level (Lee et al., 2008) and at the college level (Liang, Su, & Tsai, 2015) have been previously investigated. In addition, previous studies have investigated the relations between approaches to learning and self-efficacy in learning science (Lin & Tsai, 2013b; Phan, 2011). However, conceptions of learning science, approaches to learning science, and self-efficacy in learning science have not been integrated and investigated within the same model. Whether the construct of self-efficacy is a hierarchical system that can be divided into the lower level and higher level remains a puzzle. In addition, most studies focused on high school or undergraduate students' perceptions of learning science in Taiwan. Few studies have been conducted to examine primary school students' conceptions of learning science, approaches to learning science, and self-efficacy in learning science. Starting in Grade 3, science is a major subject in primary school in mainland China. Moreover, it should be noted that primary school students in mainland China are different from the samples of previous studies. Therefore, it is very interesting and valuable to investigate how primary school students in mainland China perceived and interpreted their experiences of learning science.

Literature review

Cultural impacts on learning science

Learning science means culture acquisition from the perspective of cultural anthropology (Wolcott, 1991). Numerous studies have revealed that learners' cultural backgrounds have impacts on their learning science processes (Li, 2003; Tsai, 2004). Tsai (2000) indicated that different cultural groups had different processes of acquiring scientific knowledge since students' scientific knowledge acquisition occurs in a complex social and cultural context. Li (2003) found that Chinese students' conceptions of learning put emphasis on achievement standards, the unity of knowing, and morality. While Lee et al. (2008) reported that the western students are educated in the western culture that is shaped by constructive philosophy with the notion of high individualised.

Previous studies have been conducted to examine western students' conceptions of learning (Marshall, Summer, & Woolnough, 1999; Sadi, 2017), approaches to learning (Phan, 2011), and self-efficacy (Prat-Sala & Redford, 2010; Uzuntiryaki & Capa Aydin, 2009). However, elementary school students in mainland China are educated in an eastern culture that is influenced by Confucius philosophy. Moreover, the study on elementary school students' conceptions of learning science, approaches to learning science, and self-efficacy in learning science from mainland China remains lacking.

Therefore, this investigation will be very interesting and informative for providing insights into mainland China's students' perceptions of learning science.

Research on conceptions of learning science

Richardson (1999) proposed that conceptions of learning refer to students' views about their personal experiences and learning context. Säljö (1979) conducted pioneering research on conceptions of learning via in-depth interviews and analysis. Säljö (1979) distinguished between five categories of conceptions of learning: increase of knowledge; memorisation; acquisition of facts or principles; abstraction of meaning; and an interpretive process aimed at understanding reality. Building on Säljö's study, many researchers proposed different categories of conceptions of learning (Lee et al., 2008; Marshall et al., 1999; Marton, Dall' Alba, & Beaty, 1993; Tsai, 2004; see Table 1). Lee et al.'s (2008) study indicated that students' conceptions of learning included six factors. Furthermore, numerous studies validated the effectiveness of the six factors (Chiou et al., 2012; Lin, Tsai, et al., 2012). Therefore, the present study adopted Lee et al.'s (2008) questionnaire to investigate primary school students' conceptions of learning.

Marton et al. (1993) regarded conceptions of learning as a hierarchical system. The following three categories were grouped as the lower level: increase of knowledge, memorising, and an acquisition of facts or principles level (see Table 1); the higher level categories were: understanding, an interpretive process aimed at understanding reality, and changing as a person. Similarly, Lee et al. (2008) regarded memorising, testing, calculating, and practicing as lower level categories; and, the following were higher level categories: increase of knowledge; applying; understanding, and seeing something in a new way. Therefore, lower level conceptions of learning science focused on the memorisation or practice of what was taught by teachers or read in textbooks. However, higher level conceptions of learning science represented the application of what students learned and/or their ability to see information in a new way.

Table 1. Conceptions of learning proposed by different researchers.

Level of conceptions	Säljö (1979)	Marton et al. (1993)	Marshall et al. (1999)	Tsai (2004)	Lee et al. (2008)
Lower level	Increase of knowledge	Increasing knowledge	Memorising	Memorising	Memorising
	Memorising	Memorising and reproducing	Applying equations and procedures	Preparing for test	Testing
	Acquisition of facts or principles	Applying	Making sense of concepts and procedures	Calculating and practicing tutorial problems	Calculating and practicing
Higher level	Abstraction of meaning	Understanding	Seeing in a new way	Increases of knowledge	Increase of knowledge
	An interpretive process aimed at understanding reality	Seeing things in a different way		Applying	Applying
		Changing as a person	A change as a person	Understanding	Understanding and Seeing in a new way
				Seeing in a new way	

Research on approaches to learning science

Approach to learning refers to the methods of academic learning (Biggs, 1994). Previous studies indicate that learners adopt different ways of learning when completing academic tasks (Biggs, 2001; Marton & Säljö, 2005). Typically, there are two main approaches: surface and deep approaches (Liang, Lee, & Tsai, 2010; Lin, Liang, & Tsai, 2012). The surface learning approach is elicited by surface motivations, which lead to surface strategies (Chiou & Liang, 2012). Learners who possess surface learning approaches are afraid of failure and only memorise what they are taught to pass examinations. However, the deep learning approach is stimulated by deep motivation, resulting in deep learning strategies (Chiou & Liang, 2012). Learners who have a deep learning approach are motivated by intrinsic motivation and have a deep understanding of learning content, or integrating prior knowledge with new information. The primary difference between surface and deep approaches is whether learners can produce meaning from learning materials (Chiou et al., 2012).

Research on self-efficacy in learning science

Self-efficacy refers to individuals' perceptions of their academic ability to complete learning tasks (Bandura, 1997). Self-efficacy plays a crucial role in students' academic performance (Printrich & Schunk, 2002) and motivation (Cetin-Dindar, 2016). Previous studies indicate that learners with higher levels of self-efficacy are prone to using deep learning strategies and achieve better learning performance (Liem et al., 2008). In studies examining self-efficacy, researchers used different perspectives. Initially, self-efficacy was regarded as a single scale (Glynn, Taasoobshirazi, & Brickman, 2009; Tuan, Chin, & Shieh, 2005). However, multi-dimensional self-efficacy instruments were developed with high school or undergraduate students in later studies (Capa Aydin & Uzuntiryaki, 2009; Uzuntiryaki & Capa Aydin, 2009). Lin and Tsai (2013b) developed five-factor self-efficacy scale to measure high school students' self-efficacy. The five factors were conceptual understanding, higher order cognitive skills, practical work, everyday application, and science communication. Lin and Tsai (2013b) believed that the conceptual understanding subscale measured students' self-efficacy in understanding definitions, laws, and theories. The higher order cognitive skills subscale assessed students' self-efficacy in utilising complex cognitive skills. The practical work subscale measured students' self-efficacy in completing practical activities. The everyday application subscale evaluated students' self-efficacy in solving real-life problems. The science communication subscale addressed students' self-efficacy in communicating or discussing what they have learned in science. Furthermore, Lin, Liang, and Tsai (2015) adopted four-factor self-efficacy scale to measure university students' self-efficacy in learning biology. These four factors include higher order cognitive skills, everyday application, science communication, and practical works. However, few studies investigated the construct of self-efficacy in learning science for primary school students. Thus, the present study examined the construct of self-efficacy of primary school students in mainland China. Empirical data were collected to validate the new self-efficacy constructs.

Hypothesis development

Researchers are interested in the relations between conceptions of learning science and approaches to learning science. Lee et al. (2008) found that high school students' conceptions of learning science were closely related to their approaches to learning science. Students who had higher level conceptions of learning were more likely to adopt deep approaches. Li et al. (2013) investigated college chemistry major students' conceptions of and approaches to learning chemistry. They found that juniors and seniors had higher level conceptions of learning chemistry and tended to adopt deep approaches to learning chemistry. Chiou et al. (2012) analysed the relationship between conceptions of and approaches to learning in biology. They reported that students who possessed lower level conceptions of learning tended to use a surface approach, and students who expressed higher level conceptions were more likely to adopt a deep approach. Moreover, researchers have explored the relations between self-efficacy and conceptions of learning. Tsai, Ho, Liang, and Lin (2011) reported that high school students' conceptions of learning were associated with their self-efficacy. Lower level conceptions of learning were negatively associated with self-efficacy, while higher level conceptions of learning were positively related to self-efficacy. In addition, Chiou and Liang (2012) found that conceptions of learning had an impact on self-efficacy.

Previous studies have started to focus on the relations between approaches to learning science and self-efficacy in different learning contexts. For example, Phan (2011) found that there was a positive relation between undergraduates' self-efficacy and their approaches to learning. Chiou and Liang (2012) investigated the relations between high school students' approaches to learning science and self-efficacy. They reported that approaches to learning significantly predicted self-efficacy. Therefore, according to previous studies, the following hypotheses were proposed (See Figure 1):

> Hypothesis 1 (H11): Primary school students' lower level conceptions of learning science will have a significant and positive effect on the surface approach to learning science.

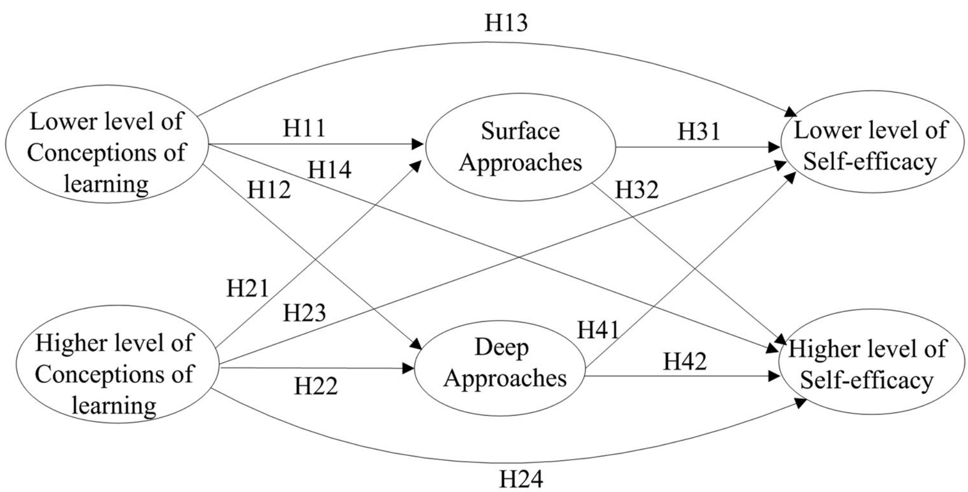

Figure 1. Proposed conceptual model.

Hypothesis 1 (H12): Primary school students' lower level conceptions of learning science will have a significant and negative effect on the deep approach to learning science.

Hypothesis 1 (H13): Primary school students' lower level of conceptions of learning science will have a significant and positive effect on lower level of self-efficacy in learning science.

Hypothesis 1 (H14): Primary school students' lower level conceptions of learning science will have a significant and positive effect on higher level of self-efficacy in learning science.

Hypothesis 2 (H21): Primary school students' higher level conceptions of learning science will have a significant and negative effect on the surface approach to learning science.

Hypothesis 2 (H22): Primary school students' higher level conceptions of learning science will have a significant and positive effect on the deep approach to learning science.

Hypothesis 2 (H23): Primary school students' higher level conceptions of learning science will have a significant and positive effect on lower level of self-efficacy in learning science.

Hypothesis 2 (H24): Primary school students' higher level conceptions of learning science will have a significant and positive effect on higher level of self-efficacy in learning science.

Hypothesis 3 (H31): Primary school students' surface approaches to learning science will have a significant and negative effect on lower level of self-efficacy in learning science.

Hypothesis 3 (H32): Primary school students' surface approaches to learning science will have a significant and negative effect on higher level of self-efficacy in learning science.

Hypothesis 4 (H41): Primary school students' deep approaches to learning science will have a significant and positive effect on lower level of self-efficacy in learning science.

Hypothesis 4 (H42): Primary school students' deep approaches to learning science will have a significant and positive effect on higher level of self-efficacy in learning science.

Method

Participants

A convenience sampling method was used to recruit participants. Participants included 1049 elementary students from Mainland China, who were enrolled in grade 4 or grade 5 elementary school classes; 550 were male and 499 were female. The participants with an average age of 11 were from Beijing, Tianjin, and Guizhou provinces. All participants were recruited from science courses, and volunteered to complete three questionnaires: conceptions of learning science, approaches to learning science, and self-efficacy of learning science. The three questionnaires are described in detail below.

Instruments

Three questionnaires were adopted to measure the conceptions of learning science (CLS), approaches to learning science (ALS), and self-efficacy of learning science (SELS) in the current study. The CLS and ALS questionnaires were adapted from a measure developed by Lee et al. (2008). The CSL scale included six subscales and 31 items (e.g. 'Learning science means memorizing the definitions and formulas in the science textbook'). The six subscales included: a 5-item memorising subscale; a 6-item testing subscale; a 5-item calculating and practicing subscale; a 5-item increasing of knowledge subscale; a 4-item applying subscale; and a 6-item understanding and seeing in a new way subscale.

The 24-item ALS scale included four subscales: an 8-item deep motive subscale; a 6-item deep strategy subscale; a 5-item surface motive subscale; and a 5-item surface strategy subscale (e.g. 'I always look forward to going to science class').

The SELS questionnaire was adapted based on measures used by Lin and Tsai (2013b) and Lin, Tan, and Tsai (2013). The SELS scale included five subscales and 32 items: a 5-item conceptual understanding subscale; a 6-item higher order cognitive skill subscale; a 7-item practical work subscale; an 8-item everyday application subscale; and a 6-item science communication subscale (e.g. 'I am certain I can master the skills in science class').

The three questionnaires were rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). This survey was administered in Mandarin Chinese. A back-translation procedure was employed to ensure the equality of the English and Chinese versions. Three experts in science education were invited to examine the content of three questionnaires to verify the validity of the questionnaires.

Data analysis

The exploratory factor analysis (EFA) was performed to examine the factor structure of three questionnaires. The confirmatory factor analysis (CFA) was further conducted to examine the construct validity of three questionnaires. In addition, 1049 participants were randomly split into two subsets for the EFA ($n = 519$) and the CFA ($n = 530$), respectively. Moreover, structural equation modeling (SEM) was conducted using Amos 18.0 to examine the relations between conceptions of, approaches to, and self-efficacy in learning science as well as to test the proposed hypotheses.

Results

EFA for CLS, ALS, and SELS

Prior to the EFA, the Kaiser-Meyer-Olkin (KMO) and Bartlett sphericity tests were adopted to examine whether the data were appropriate for factor analysis. KMO was expected to be larger than 0.5 (Field, 2000). It was found that the KMO value was 0.89 for CLS, 0.86 for ALS, and 0.92 for SELS. Furthermore, chi square was 5018.67 ($p < .01$) for CLS, 3550.292 ($p < .01$) for ALS, and 3306.448 ($p < .01$) for SELS. The findings indicated that the data had a normal multi-variable distribution (Sadi, 2017). Therefore, the data were suitable for factor analysis. Furthermore, the principle components analysis method with an oblique rotation was performed on three questionnaires of CLS, ALS, and SELS. This study adopted a factor loading larger than 0.5 for retaining the items, leading to delete 22 items.

Table 2 shows the EFA results for the CLS questionnaire. In the final version of CLS, 27 items were kept and divided into six factors, namely memorising, testing, calculating and practicing, increasing one's knowledge, applying as well as understanding and seeing in a new way. The total variance explained in the CSL was 58.29%. The Cronbach's alpha values of each scale were 0.84, 0.84, 0.66, 0.77, 0.63, and 0.78, respectively, and the

Table 2. CLS questionnaire factor analysis results.

Dimensions	CLS factors	Cronbach's alpha values	Means (SD)
Lower level	Factor 1 Memorising (CLM)	α = 0.84	Mean (SD) = 3.16 (1.02)
	Item 1	0.800	
	Item 2	0.784	
	Item 3	0.799	
	Item 4	0.822	
	Item 5	0.829	
	Factor 2 Testing (CLT)	α = 0.84	Mean (SD) = 2.42 (1.01)
	Item 1	0.816	
	Item 2	0.813	
	Item 3	0.816	
	Item 4	0.803	
	Item 5	0.800	
	Item 6	0.839	
	Factor 3 Calculating and practicing (CLC)	α = 0.66	Mean (SD) = 3.39 (0.83)
	Item 1	0.610	
	Item 2	0.592	
	Item 3	0.566	
	Item 4	0.608	
Higher level	Factor 4 Increasing one's knowledge (CLI)	α = 0.77	Mean (SD) = 3.79 (0.87)
	Item 1	0.706	
	Item 2	0.726	
	Item 3	0.715	
	Item 4	0.711	
	Factor 5 Applying (CLA)	α = 0.63	Mean (SD) = 3.46 (0.79)
	Item 1	0.512	
	Item 2	0.504	
	Item 3	0.634	
	Item 4	0.567	
	Factor 6 Understanding & Seeing something in a New Way (CLU)	α = 0.78	Mean (SD) = 3.81 (0.89)
	Item 1	0.718	
	Item 2	0.723	
	Item 3	0.717	
	Item 4	0.733	

overall reliability coefficient was 0.85. It was indicated that the scale had a good consistency if the reliability coefficient was higher than 0.60 (Nunnally & Bernstein, 1994). Therefore, the CLS had a good internal consistency.

A total of 19 items were remained in the final version of the ALS, which were grouped into four factors, including deep motive, deep strategy, surface motive, and surface strategy (see Table 3). The total variance explained was 54.88%. The Cronbach's alpha values of each scale were 0.80, 0.78, 0.77, and 0.83, respectively. The overall reliability coefficient of ALS achieved 0.81. Thus, the ALS questionnaire also had a good internal consistency.

In the final version of the SELS, there were remaining 19 items, which were divided into four factors (see Table 4). These four factors were conceptual understanding, practical work, everyday application, and science communication. The total variance explained was 54.83%. The Cronbach's alpha values of each scale were 0.78, 0.73, 0.63, and 0.82, respectively. The overall reliability coefficient of SELS achieved 0.89. Thus, the SELS questionnaire also had a good internal consistency. Furthermore, it was notable that only four factors were remained in spite of the original five factors. The main reason was that six items of higher order cognitive skills had the lower reliability coefficient and factor loading value. Thus, the subscale of higher order cognitive skills was eliminated.

Table 3. ALS questionnaire factor analysis results.

Dimensions	ALS factors	Cronbach's alpha values	Means (SD)
Deep level	Factor 1 Deep motive (DM)	$\alpha = 0.80$	Mean (SD) = 3.58 (0.85)
	Item 1	0.769	
	Item 2	0.763	
	Item 3	0.795	
	Item 4	0.774	
	Item 5	0.780	
	Item 6	0.782	
	Item 7	0.778	
	Factor 2 Deep strategy (DS)	$\alpha = 0.78$	Mean (SD) = 3.44 (0.96)
	Item 1	0.706	
	Item 2	0.721	
	Item 3	0.698	
	Item 4	0.756	
Surface level	Factor 3 Surface motive (SM)	$\alpha = 0.77$	Mean (SD) = 2.68 (1.18)
	Item 1	0.743	
	Item 2	0.638	
	Item 3	0.684	
	Factor 4 Surface strategy (SS)	$\alpha = 0.83$	Mean (SD) = 2.45 (1.08)
	Item 1	0.792	
	Item 2	0.783	
	Item 3	0.792	
	Item 4	0.794	
	Item 5	0.840	

CFA for CLS, ALS, and SELS

CFA was conducted to examine the CLS, ALS, and SELS constructs. Hair, Black, Babin, Anderson, and Tatham (2006) proposed that the relative Chi square (χ^2/df), the root mean square error of approximation (RMSEA), the goodness of fit index (GFI), the comparative fit index (CFI), the parsimony-adjusted normed fit index (PNFI), and the

Table 4. SELS questionnaire factor analysis results.

Dimensions	SELS factors	Cronbach's alpha values	Means(SD)
Lower level	Factor 1 Conceptual understanding (SEC)	$\alpha = 0.78$	Mean (SD) = 3.34 (0.86)
	Item 1	0.726	
	Item 2	0.723	
	Item 3	0.761	
	Item 4	0.735	
	Item 5	0.764	
	Factor 2 Science communication (SES)	$\alpha = 0.82$	Mean (SD) = 3.56 (0.86)
	Item 1	0.809	
	Item 2	0.801	
	Item 3	0.790	
	Item 4	0.790	
	Item 5	0.788	
	Item 6	0.799	
Higher level	Factor 3 Practical work (SEP)	$\alpha = 0.73$	Mean (SD) = 3.53 (0.83)
	Item 1	0.673	
	Item 2	0.660	
	Item 3	0.659	
	Item 4	0.674	
	Item 5	0.722	
	Factor 4 Everyday application (SEA)	$\alpha = 0.63$	Mean (SD) = 3.47 (0.88)
	Item 1	0.566	
	Item 2	0.414	
	Item 3	0.607	

parsimony-adjusted comparative fit index (PCFI) can be adopted to assess the goodness of fit. The measurement model fits the data well if χ^2/df is less than 5; RMSEA is less than .08; GFI is larger than .80; IFI, TLI, and CFI exceed .90; and both PNFI and PCFI exceed .70 (Bentler, 1990; Schumacker & Lomax, 2004). Table 5 demonstrated the results of CFA. The fitness of items for each factor of the CLS, ALS, and SELS revealed a sufficient fit and confirmed three questionnaires' structure.

Convergent and discriminant validity

In order to examine the validity of three scales, the factor loadings, composite reliability (CR), and average variance extracted (AVE) were calculated. Table 6 shows the values of factor loadings, CR, and AVE for three scales. All of the factor loading values were higher than 0.5. The CR values of all factors satisfied the threshold value of 0.60 (Bagozzi & Yi, 1988). The AVE values for some constructs were greater than 0.5, but for some constructs were less than 0.5. However, for discriminant validity, Fornel and Larcker (1981) indicated that the square root of the AVE of each construct should be greater than the correlation between the construct and other constructs and should be at least 0.50. From Tables 7, 8, and 9, it was found that all square roots of the AVE of all constructs are significantly greater than the inter-construct correlations and also were greater than 0.5. Therefore, the convergent and discriminant validities of three scales were acceptable.

Structural equation model

A structural equation model technique was used herein to examine the structural relations between the two dimensions of the CLS, the two dimensions of the ALS, and the two dimensions of the SELS. Table 10 shows the fit indices of the structural model. The results indicated that all of the fit indices exceeded the acceptable values. Therefore, the structural model fits the data well. Figure 2 shows the results of the structural model.

Table 5. The results of confirmatory factor analysis ($n = 530$).

Indices	Absolute fit index			Relative fit index			Parsimonious fit index	
	χ^2/df	RMSEA	GFI	IFI	TLI	CFI	PNFI	PCFI
CLS	1.92	0.04	0.92	0.92	0.91	0.92	0.76	0.83
ALS	2.12	0.05	0.93	0.94	0.93	0.93	0.96	0.80
SELS	1.57	0.03	0.94	0.96	0.96	0.96	0.80	0.85
Acceptance value	1–5	<.08	>.80	>.90	>.90	>.90	>.70	>.70

Table 6. The fit index of the structural model.

Indices	Absolute fit index			Relative fit index			Parsimonious fit index	
	χ^2/df	RMSEA	GFI	IFI	TLI	CFI	PNFI	PCFI
Assumed model	2.13	.03	.90	.91	.91	.91	.79	.86
Acceptance value	1–5	<.08	>.80	>.90	>.90	>.90	>.70	>.70

Table 7. The factor loadings, CR, and AVE of three scales.

Instrument	Factors	Factor loading	CR	AVE
CLS	CLM	0.60–0.74	0.81	0.46
	CLT	0.51–0.78	0.79	0.38
	CLC	0.51–0.63	0.66	0.33
	CLI	0.56–0.67	0.73	0.41
	CLA	0.50–0.64	0.60	0.31
	CLU	0.63–0.69	0.76	0.44
ALS	DM	0.52–0.69	0.78	0.34
	DS	0.62–0.70	0.76	0.52
	SM	0.62–0.78	0.76	0.51
	SS	0.72–0.80	0.83	0.57
SELS	SEC	0.56–0.61	0.73	0.35
	SES	0.60–0.70	0.81	0.41
	SEP	0.51–0.63	0.70	0.32
	SEA	0.52–0.68	0.63	0.36

Table 8. Correlations among constructs of CLS (square root of AVE in diagonal).

	CLM	CLT	CLC	CLI	CLA	CLU
CLM	**0.676**					
CLT	0.323	**0.622**				
CLC	0.391	0.163	**0.569**			
CLI	0.260	0.154	0.393	**0.639**		
CLA	0.291	0.018	0.377	0.522	**0.550**	
CLU	0.262	0.158	0.366	0.589	0.523	**0.660**

Note. Diagonal numbers (in bold) indicate the square root of the AVE.

Table 9. Correlations among constructs of ALS (square root of AVE in diagonal).

	DM	DS	SM	SS
DM	**0.579**			
DS	0.519	**0.721**		
SM	0.026	0.015	**0.716**	
SS	−0.202	−0.082	0.325	0.**756**

Note. Diagonal numbers (in bold) indicate the square root of the AVE.

Table 10. Correlations among constructs of SELS (square root of AVE in diagonal).

	SEC	SES	SEP	SEA
SEC	**0.592**			
SES	0.562	**0.643**		
SEP	0.536	0.521	**0.564**	
SEA	0.502	0.524	0.502	**0.603**

Note. Diagonal numbers (in bold) indicate the square root of the AVE.

Table 11 demonstrates the results of the path analysis. It was found that lower level conceptions of learning science had a significant effect on surface approaches to learning science ($\beta = .76$, $p < .001$). Therefore, H11 was supported. However, lower level conceptions of learning did not have a significant influence on deep approaches to learning ($\beta = .15$, $p > .05$), the lower level of self-efficacy ($\beta = -.15$, $p > .05$), and the higher level

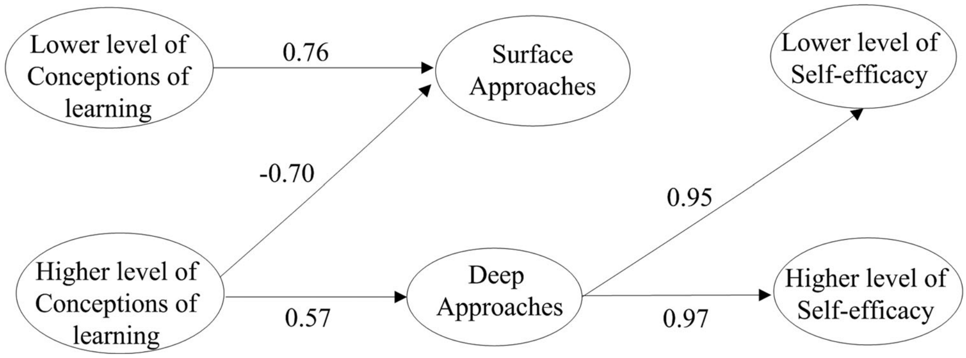

Figure 2. Structural equation model.

Table 11. The results of the path analysis.

Hypothesis	Paths	Path coefficient	Results
H11	The lower level of conceptions of learning→Surface approach to learning	.76	Support
H12	The lower level of conceptions of learning→Deep approach to learning	.15	Against
H13	The lower level of conceptions of learning→The lower level of self-efficacy	−.15	Against
H14	The lower level of conceptions of learning→The higher level of self-efficacy	−.23	Against
H21	The higher level of conceptions of learning→Surface approach to learning	−.70	Support
H22	The higher level of conceptions of learning→Deep approach to learning	.57	Support
H23	The higher level of conceptions of learning→The lower level of self-efficacy	−.24	Against
H24	The higher level of conceptions of learning→The higher level of self-efficacy	−.13	Against
H31	Surface approach to learning→The lower level of self-efficacy	.03	Against
H32	Surface approach to learning→The higher level of self-efficacy	.14	Against
H41	Deep approach to learning→The lower level of self-efficacy	.95	Support
H42	Deep approach to learning→The higher level of self-efficacy	.97	Support

of self-efficacy ($\beta = -.23$, $p > .05$). Thus, H12, H13, and H14 were not supported. Furthermore, higher level conceptions of learning science had a positive effect on deep approaches ($\beta = .57$, $p < .001$) and a negative effect on surface approaches to learning ($\beta = -.70$, $p < .001$). Hence, both H21 and H22 were supported. However, higher level conceptions of learning did not have a significant and direct influence on the lower level of self-efficacy ($\beta = -.24$, $p > .05$) and the higher level of self-efficacy ($\beta = -.13$, $p > .05$). Hence, both H23 and H24 were not supported. In addition, the surface approaches to learning did not have a significant influence on the lower level of self-efficacy ($\beta = .03$, $p > .05$) and the higher level of self-efficacy ($\beta = .14$, $p > .05$). Therefore, H31 and H32 were not supported. However, the deep approaches to learning had a significant effect on the lower level of self-efficacy ($\beta = .95$, $p < .001$) and the higher level of self-efficacy ($\beta = .97$, $p < .001$). Therefore, H41 and H42 were supported.

Discussion

The present study examined the relations between primary school students' conceptions of learning science, approaches to learning science, and self-efficacy. The EFA and CFA methods were adopted to examine the reliability and validity, and path analysis was conducted to test the structural model. The results indicated that the questionnaires about conceptions of learning science, approaches to learning science, and self-efficacy in learning science achieved a good reliability and validity. The fit indices also revealed that the structural model fits with the empirical data well. The findings indicated that primary students' conceptions of learning had a significant influence on approaches to learning, which further exerted an effect on self-efficacy in learning science.

The role of conceptions of learning

The present study revealed that primary students' conceptions of learning played a crucial role in learning science. It was found that lower level conceptions of learning science had a significant and positive effect on the surface approach to learning science. Higher level conceptions of learning had a positive effect on the deep approaches to learning science. Thus, learners with lower level conceptions of learning tended to use surface learning approaches and students with higher level conceptions of learning were inclined to adopt deep learning approaches. Better put, learners were more likely to learn science by reciting if they consider learning science as memorising what is taught in class; preparing for a test; or calculating and practicing. However, when students view learning science as increasing knowledge; applying knowledge; or understanding and seeing knowledge in a new way, they employ deep motives and meaningful learning strategies to learn science. This result was consistent with previous studies reported by Chiou, Lee, and Tsai (2013) and Li et al. (2013), who also found that conceptions of learning science were positively related to approaches to learning science. Moreover, higher level conceptions of learning science had a negative impact on surface approaches to learning science. In other words, students with higher level conceptions of learning were not likely to adopt the surface motives and strategies for learning science. Moreover, lower level conceptions of learning science did not have a significant effect on deep approaches to learning science. The possible explanation for this finding may associate with the examination-centered school culture in mainland China. Passing an examination and getting higher scores were a central theme of education in Mainland China, which was instilled in every student since primary school.

In addition, both lower and higher level conceptions of learning science did not have significant and direct effects on the lower and higher level of self-efficacy in learning science. This finding was consistent with Chiou and Liang (2012), who reported that conceptions of learning science were indirectly related to self-efficacy in learning science.

The role of approaches to learning

When the role of approaches to learning was examined, it was found that students' approaches to learning further influenced their self-efficacy. Specifically, only the deep approach to learning science had significant and positive effects on both the lower and higher level of self-efficacy in learning science. However, the surface approach to learning did not have significant effects on the lower and higher level of self-efficacy in learning science. This finding indicated that students with deep learning approaches had a sense of self-efficacy in learning science. This means that when students learned science for pursuing their inner interests, they had a higher sense of self-efficacy in learning science. Similarly, if students wanted to apply what they had learned to solve real-life problems, they tended to have a strong sense of self-efficacy in learning science. However, if students only memorised what they had learned, they did not obtain self-efficacy in learning science. In previous studies, students' deep approaches to learning were positively associated with their self-efficacy (Lin & Tsai, 2013b; Prat-Sala & Redford, 2010). A similar conclusion was also drawn from the findings in the present study.

Implications

This study had several implications for teachers and practitioners. First, conceptions of learning science had a significant, direct influence on approaches to learning and an indirect influence on self-efficacy; therefore, it was important to help students develop higher level conceptions of learning. Sophisticated conceptions of learning can encourage students to adopt deep motives and deep learning strategies. Teachers and practitioners can adopt inquiry-based learning, problem-based learning, and outdoor investigation to help students realise the purpose of learning science. The purpose of learning science was not to remember what teachers taught in class or prepare for a test; rather, it was to increase knowledge, apply knowledge, and understand knowledge in a new way. We posited that applying what students learn in class to solve real-life problems was the best way to help students obtain a better understanding of conceptions of learning science. Second, teachers and practitioners should encourage students to adopt deep motives and learning strategies for learning science; for example, when teachers guide students in obtaining a thorough understanding of what they learn in class. Students need to link prior knowledge to new information to promote meaningful learning. In addition, it was also important to motivate students to learn science via their own curiosity and interests. Third, students' self-efficacy was closely associated with their science learning performance (Capa Aydin & Uzuntiryaki, 2009); thus, it is necessary to adopt some strategies to promote self-efficacy in science, such as developing higher level of conceptions of learning, solving real-life problems, and adopting deep learning strategies.

Conclusion

In the current study, three instruments were validated to measure primary school students': conceptions of learning science, approaches to learning science, and self-efficacy in learning science. The results indicated that these three instruments had satisfactory validity and reliability. Moreover, the relations between the conceptions of, approaches to, and self-efficacy in learning science were investigated herein. The findings revealed that lower level conceptions of learning science had a positive and significant influence on surface approaches to learning science. Higher level conceptions of learning science had a positive influence on deep approaches to learning science, and had a negative influence on the surface approaches to learning science. Furthermore, the deep approach had significant and positive effects on both the lower and higher level of self-efficacy in learning science.

The present study found that primary students' conceptions of learning science and approaches to learning science were the two major components for their self-efficacy in learning science. These two components represented how students viewed the nature of and approaches to learning science. The main contribution of this study was obtaining a deeper understanding of primary school students' perceptions of learning science, which provides insights into improving pedagogical practices. In addition, the construct of self-efficacy was divided into the lower and higher level of self-efficacy in learning science.

This study was constrained by several limitations. First, all the participants in the present study were in mainland China. Since conceptions of and approaches to learning are influenced by culture, it would be interesting to examine and compare how different cultures influence students' conceptions of learning as well as approaches to learning. Second, the present study only examined primary school students' conceptions of, approaches to, and self-efficacy in learning science. It is suggested that future studies compare and investigate these three constructs in other contexts. The final limitation, likewise, deserves further investigation as students' conceptions of, approaches to, and self-efficacy in learning science were all related to their learning performance.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

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

References

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*(1), 74–94.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246.

Biggs, J. (1994). Approaches to learning: Nature and measurement of. In T. Husen & T. N. Postlethwaite (Eds.), *The international encyclopedia of education* (2nd ed., Vol. 1, pp. 319–322). Oxford: Pergamon Press.

Biggs, J. (2001). Enhancing learning: A matter of style or approach? In R. J. Stemberg & L. F. Zhang (Eds.), *Perspectives on thinking, learning, and cognitive styles* (pp. 73–102). Mahwah, NJ: Lawrence Erlbaum Associates.

Capa Aydin, Y., & Uzuntiryaki, E. (2009). Development and psychometric evaluation of the high school chemistry self-efficacy scale. *Educational and Psychological Measurement*, *69*, 868–880.

Cetin-Dindar, A. (2016). Student motivation in constructivist learning environment. *Eurasia Journal of Mathematics, Science & Technology Education*, *12*(2), 233–247.

Chiou, G.-L., Lee, M.-H., & Tsai, C.-C. (2013). High school students' approaches to learning physics with relationship to epistemic views on physics and conceptions of learning physics. *Research in Science & Technological Education*, *31*(1), 1–15.

Chiou, G. L., & Liang, J. C. (2012). Exploring the structure of science self-efficacy: A model built on high school students' conceptions of learning and approaches to learning in science. *The Asia-Pacific Education Researcher*, *21*(1), 83–91.

Chiou, G. L., Liang, J. C., & Tsai, C. C. (2012). Undergraduate students' conceptions of and approaches to learning in biology: A study of their structural models and gender differences. *International Journal of Science Education*, *34*(2), 167–195.

Field, A. (2000). *Discovering statistics using SPSS for windows: Advanced techniques for the beginner*. London: Sage.

Fornel, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement errors. *Journal of Marketing Research*, *18*(2), 39–50.

Glynn, S. M., Taasoobshirazi, G., & Brickman, P. (2009). Science motivation questionnaire: Construct validation with nonscience majors. *Journal of Research in Science Teaching*, *46*, 127–146.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). New York, NY: Prentice Hall.

Kember, D., Biggs, J., & Leung, D. Y. P. (2004). Examining the multidimensionality of approaches to learning through the development of a revised version of the Learning Process Questionnaire. *British Journal of Educational Psychology*, *74*(2), 261–279.

Lee, M. H., Johanson, R. E., & Tsai, C. C. (2008). Exploring Taiwanese high school students' conceptions of and approaches to learning science through a structural equation modeling analysis. *Science Education*, *92*(2), 191–220.

Li, J. (2003). US and Chinese cultural beliefs about learning. *Journal of Educational Psychology*, *95* (2), 258–267.

Li, W. T., Liang, J. C., & Tsai, C. C. (2013). Relational analysis of college chemistry-major students' conceptions of and approaches to learning chemistry. *Chemistry Education Research and Practice*, *14*(4), 555–565.

Liang, J. C., Lee, M. H., & Tsai, C. C. (2010). The relations between scientific epistemological beliefs and approaches to learning science among science-major undergraduates in Taiwan. *Asia-Pacific Education Researcher*, *19*, 43–59.

Liang, J. C., Su, Y. C., & Tsai, C. C. (2015). The assessment of Taiwanese college students' conceptions of and approaches to learning computer science and their relationships. *The Asia-Pacific Education Researcher*, *24*, 557–567.

Liem, A. D., Lau, S., & Nie, Y. (2008). The role of self-efficacy, task value, and achievement goals in predicting learning strategies, task disengagement, peer relationship, and achievement outcome. *Contemporary Educational Psychology*, *33*, 486–512.

Lin, Y. C., Liang, J. C., & Tsai, C. C. (2012). The relationships between epistemic beliefs in biology and approaches to learning biology among biology-major university students in Taiwan. *Journal of Science Education and Technology*, *21*, 796–807.

Lin, T. C., Liang, J. C., & Tsai, C. C. (2015). Conceptions of memorizing and understanding in learning, and self-efficacy held by university biology majors. *International Journal of Science Education*, *37*(3), 446–468.

Lin, T. J., Tan, A. L., & Tsai, C. C. (2013). A cross-cultural comparison of Singaporean and Taiwanese eighth graders' science learning self-efficacy from a multi-dimensional perspective. *International Journal of Science Education*, *35*(7), 1083–1109.

Lin, C. L., Tsai, C. C., & Liang, J. C. (2012). An investigation of two profiles within conceptions of learning science: An examination of confirmatory factor analysis. *European Journal of Psychology of Education*, *27*(4), 499–521.

Lin, T. J., & Tsai, C. C. (2013a). An investigation of Taiwanese high school students' science learning self-efficacy in relation to their conceptions of learning science. *Research in Science & Technological Education*, *31*(3), 308–323.

Lin, T. J., & Tsai, C. C. (2013b). A multi-dimensional instrument for evaluating Taiwanese high school students' science learning self-efficacy in relation to their approach to learning science. *International Journal of Science and Mathematics Education*, *11*(6), 1275–1301.

Marshall, D., Summer, M., & Woolnough, B. (1999). Students' conceptions of learning in an engineering context. *Higher Education*, *38*, 291–309.

Marton, F. M., Dall' Alba, G., & Beaty, E. (1993). Conceptions of learning. *International Journal of Educational Research*, *19*, 277–300.

Marton, F. M., & Säljö, R. (2005). Approaches to learning. In F. M. Marton, D. Hounsell, & N. J. Entwistle (Eds.), *The experience of learning: Implications for teaching and studying in higher education* (3rd ed., pp. 106–125). Edinburgh: University of Edinburgh.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Phan, H. P. (2011). Interrelations between self-efficacy and learning approaches: A developmental approach. *Educational Psychology*, *31*(2), 225–246.

Prat-Sala, M., & Redford, P. (2010). The interplay between motivation, self-efficacy, and approaches to studying. *British Journal of Educational Psychology*, *80*(2), 283–305.

Printrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research and applications*. Columbus, OH: Merrill Prentice Hall.

Richardson, J. T. E. (1999). The conceptions and methods of phenomenographic research. *Review of Educational Research*, *69*(1), 53–82.

Sadi, Ö. (2017). Relational analysis of high school students' cognitive self-regulated learning strategies and conceptions of learning biology. *EURASIA Journal of Mathematics, Science & Technology Education*, *13*(6), 1701–1722.

Säljö, R. (1979). *Learning in the learner's perspective. I: Some commonsense conceptions* (Report No. 76). Gothenburg, Sweden: University of Gothenburg.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum.

Shen, K. M., Lee, M. H., Tsai, C. C., & Chang, C. Y. (2016). Undergraduate students' earth science learning: Relationships among conceptions, approaches, and learning self-efficacy in Taiwan. *International Journal of Science Education*, *38*(9), 1527–1547.

Tsai, C. C. (2000). Relationships between student scientific epistemological beliefs and perceptions of constructivist learning environments. *Educational Research*, *42*(2), 193–205.

Tsai, C. C. (2004). Conceptions of learning science among high school students in Taiwan: A phenomenographic analysis. *International Journal of Science Education*, *26*(14), 1733–1750.

Tsai, C. C., Ho, H. N. J., Liang, J. C., & Lin, H. M. (2011). Scientific epistemic beliefs, conceptions of learning science and self-efficacy of learning science among high school students. *Learning and Instruction*, *21*(6), 757–769.

Tuan, H., Chin, C., & Shieh, S. (2005). The development of a questionnaire to measure students' motivation towards science learning. *International Journal of Science Education*, *27*, 639–654.

Uzuntiryaki, E., & Capa Aydin, Y. (2009). Development and validation of chemistry self-efficacy scale for college student. *Research in Science Education*, *39*, 539–551.

Wolcott, H. F. (1991). Propriospect and the acquisition of culture. *Anthropology & Education Quarterly*, *22*(3), 251–273.

Appendices

Appendix A: The Questionnaire Items on the COLS (Final Version)

Memorizing

(1) Learning science means memorizing the definitions, formulae, and laws found in a science textbook.
(2) Learning science means memorizing the important concepts found in a science textbook.
(3) Learning science means memorizing the proper nouns found in a science textbook that can help solve the teacher's questions.
(4) Learning science means remembering what the teacher lectures about in science class.
(5) Learning science means memorizing scientific symbols, scientific concepts, and facts.

Testing

(1) Learning science means getting high scores on examinations.
(2) If there are no tests, I will not learn science.
(3) There are no benefits to learning science other than getting high scores on examinations. In fact, I can get along well without knowing many scientific facts.
(4) The major purpose of learning science is to get more familiar with test materials.

(5) I learn science so that I can do well on science-related tests.
(6) There is a close relationship between learning science and taking tests.

Calculating and practicing

(1) Learning science involves a series of calculations and problem-solving.
(2) I think that learning calculation or problem-solving will help me improve my performance in science courses.
(3) Learning science means knowing how to use the correct formulae when solving problems.
(4) The way to learn science well is to practice constantly calculations and problem solving.

Increasing one's knowledge

(1) Learning science means acquiring knowledge that I did not know before.
(2) I am learning science when the teacher tells me scientific facts that I did not know before.
(3) Learning science means acquiring more knowledge about natural phenomena and topics related to nature.
(4) Learning science helps me acquire more facts about nature.

Applying

(1) The purpose of learning science is learning how to apply methods I already know to unknown problems.
(2) Learning science means learning how to apply knowledge and skills I already know to unknown problems.
(3) We learn science to improve the quality of our lives.
(4) Learning science means solving or explaining unknown questions and phenomena.

Understanding & seeing science in a new way

(1) Learning science helps me view natural phenomena and topics related to nature in new ways.
(2) Learning science means changing my way of viewing natural phenomena and topics related to nature.
(3) Learning science means finding a better way to view natural phenomena or topics related to nature.
(4) I can learn more ways about thinking about natural phenomena or topics related to nature by learning science.

Appendix B: The Questionnaire Items on the ALS (Final version)

Deep approach

Deep motive.

(1) I find that at times studying science makes me feel really happy and satisfied.
(2) I feel that science topics can be highly interesting once I get into them.
(3) I work hard at studying science because I find the material interesting.
(4) I always greatly look forward to going to science class.
(5) I come to science class with questions in my mind that I want to be answered.
(6) I find that I continually go over my science class work in my mind even whenever I am not in science class.

(7) I like to work on science topics by myself so that I can form my own conclusions and feel satisfied.

Deep strategy.

(1) I try to relate what I have learned in science subjects to what I learn in other subjects.
(2) I like constructing theories to fit odd things together when I am learning science topics.
(3) I try to find the relationship between the contents of what I have learned in science subjects.
(4) I try to relate new material to what I already know about the topic when I'm studying science.

Surface approach

Surface motive.

(1) I am discouraged by a poor mark on science tests and worry about how I will do on the next text.
(2) Even when I have studied hard for a science text, I worry that I may not be able to do well on it.
(3) I worry that my performance in science class may not satisfy my teacher's expectations.

Surface strategy.

(1) I see no point in learning science materials that are not likely to be on the examinations.
(2) As long as I feel I am doing well enough to pass the examination, I devote as little time as I can to studying science subjects. There are many more interesting things to do with my time.
(3) I generally will restrict my study to what is specially set as I think it is unnecessary to do anything extra in learning science topic.
(4) I find that studying each topic in depth is not helpful or necessary when I am learning science. There are too many examinations to pass and too many subjects to be learned.
(5) I find the best way to pass science examinations is to try to remember the answers to likely question.

Appendix C: The Science Learning Self-Efficacy (SLSE) Questionnaire (Final version)

Conceptual understanding

(1) I can explain scientific laws and theories to others.
(2) I can choose an appropriate formula to solve a science problem.
(3) I can link the contents among different science subjects (for example, biology, chemistry and physics) and establish the relationships between them.
(4) I know the definitions of basic scientific concepts (for example, gravity, photosynthesis, etc.) very well.
(5) I feel confident when I interpret graphs/charts related to science.

Science communication

(1) I am able to comment on presentations made by my classmates in science class.
(2) I am able to use what I have learned in science classes to discuss with others.
(3) I am able to explain clearly what I have learned to others.
(4) I feel comfortable discussing science content with my classmates.
(5) In science classes, I can clearly express my own opinions.
(6) In science classes, I can express my ideas properly.

Practical work

(1) I know how to carry out experimental procedures in the biology laboratory.
(2) I know how to use equipment (for example measuring cylinders, measuring scales, etc.) in the biology laboratory.
(3) I am able to recognize the data in the biology experiments.
(4) I know how to set up equipment for laboratory experiments.
(5) I am able to understand the relationships among variables according to research data.

Everyday application

(1) I am able to explain everyday life using scientific theories.
(2) I am able to propose solutions to everyday problems using science.
(3) I can understand the news/documentaries I watch on television related to science.

EURASIA Journal of Mathematics, Science and Technology Education
ISSN: 1305-8223 (online) 1305-8215 (print)
OPEN ACCESS
2018 14(3):987-995
DOI: 10.12973/ejmste/81151

A Cross-cultural Comparison on Students' Perceptions towards Online Learning

Kaushal Kumar Bhagat [1], Chun-Yen Chang [2,3*]

[1] Smart Learning Institute of Beijing Normal University, Beijing, CHINA
[2] Graduate Institute of Science Education, National Taiwan Normal University, Taipei, TAIWAN
[3] Science Education Center, National Taiwan Normal University, Taipei, TAIWAN

Received 20 July 2017 ▪ Revised 6 December 2017 ▪ Accepted 6 December 2017

ABSTRACT

The aim of this study was to explore cross-country (Taiwan versus India) differences in students' perceptions of online leaning by gender. The self-reported instrument, POSTOL (Perception of Students towards Online Learning), was conducted to the students from Taiwan and India. Of the total 441 respondents, there were 233 students from India and 208 from Taiwan. A 2x2 between-subjects multivariate analysis of variance was employed on the four dimensions of the POSTOL scale. This scale consists of four dimensions: instructor characteristics, social presence, instructional design, and trust. Results showed that there were significant differences, between India and Taiwan, in the perception of students towards online learning. However, there was no significant interaction effect of country by gender. Findings indicated that culture did influence students' perceptions towards online learning. There is a need to raise awareness about factors that may affect online learning experience and to provide guidance and for practice and future work.

Keywords: cross-country, gender, online learning, perception

INTRODUCTION

The portrait of online learning has moved beyond students operating through Internet-connected desktops or laptops. A new generation of technology, including Smartphones and handheld devices are today's "learning hubs". The current online learning experience exploits information on the go for the sharing and exchanging ideas, without spatial barriers (Wong, 2012). The number of online course providers and number of participants are increasing exponentially (Barak et al., 2016). Some participants in online courses are from the different cultural and linguist background (Loizzo & Ertmer, 2016; Wang, 2007). This draws the attention of many educators and researchers towards online learning.

Currently, many studies have been conducted in the area of online learning. Recent studies focused on students' course satisfaction (Lee et al., 2011), learning outcome (Horzum et al., 2014; Joo et al., 2015; Lee et al., 2011; Wang et al., 2013; You, 2016), motivation (Barak et al., 2016; Chen & Jang, 2010; de Barba et al., 2016; Joo et al., 2015; Kim et al., 2015), engagement (Barak et al., 2016; Kim et al., 2015), instructional design (Joo et al., 2015), gender differences (Ashong & Commander, 2012; Cheng et al., 2012; Kupczynski et al., 2014) and social interaction (Eryilmaz et al., 2013; Joksimović et al., 2015; Xie et al., 2013). Loizzo and Ertmer (2016) mentioned understanding the specific perceptions and experiences of online learners from different countries are important research areas. Therefore, the present study focused on cross-country differences in gender and students' perceptions towards online learning.

THEORETICAL BACKGROUND

In this study, Vygotsky's (1978) sociocultural theory is applied to understand students' perceptions towards online learning across different cultures. According to the socio-cultural theory, learning is a social process and is composed of three important themes: culture, language, and the "zone of proximal development." "Zone of

✉ kaushal.dgp@gmail.com ✉ changcy@ntnu.edu.tw **(*Correspondence)**

Contribution of this paper to the literature

- This study contributes to the existing literature by showing that the perceptions of the students towards online learning differ in the cross-cultural context.
- In addition, no study previously has investigated the students' perceptions in relation to the combined effects of country and gender.
- Cross-cultural differences have major influence on students' perceptions towards online learning. Instructional designers should apply socio-cultural theory to take into account the effect of diverse cultural background on the learner's learning behavior to enhance learning effectiveness.

proximal development" is not the central concept discussed in this study; we concentrated on the effects of culture and language to understand online learning environments. Many research studies advocated the importance of language and culture. For example, according to Mercer (2000), "Language is a tool for carrying out joint intellectual activity, a distinctive human inheritance designed to serve the practical and social needs of individuals and communities and which each child has to learn to use effectively" (p.1). Delahunty et al. (2014) describe social interaction as an important component for knowledge building process in online learning. According to (Vygotsky), learning occurs at two levels:

First, through interaction with others, and then integrated into the individual's mental structure. Every function in the child's cultural development appears twice: first, on the social level, and later, on the individual level; first, between people (interpsychological) and then inside the child (intrapsychological). This applies equally to voluntary attention, to logical memory, and to the formation of concepts. All the higher functions originate as actual relationships between individuals (Vygotsky, 1978, p.57).

Language plays a significant role in terms of learner's motivation, interaction and engagement in online learning environment (Barak et al., 2016; Mercer, 2000). However, language is one of the main barriers for the students who want to enroll in online courses from developing countries (Altbach, 2014; Kerr, 2014; Olesova et al., 2011; Osman & Herring, 2007). Participants are required to have some basic competencies, specifically digital literacy and a good level of English proficiency. If an institution offers a MOOC, it is probably in English, which is presently the international *lingua franca*. English is the native language for around 400 million people. Therefore, it still leaves around 70% of the world's total population who are unable to access the educational content published in English (Wikipedia, 2016).

Some researchers also advocate the importance of culture in online learning environment. "Culture is a set of parameters of collectives that differentiate the collectives from each other in meaningful ways. Culture is variously defined in terms of several commonly shared processes: shared ways of thinking, feeling, and reacting; shared meanings of identities; shared socially constructed environments; common ways in which technologies are used; and commonly experienced events, including the history, language, and religion of their members" (House et al., 2004, p.15). Ordóñez (2014) highlights the impact of culture in online learning and suggests instructional designers and instructor to consider culture during course development. According to McLoughlin (1999), "Culture and learning are interwoven and inseparable" (p.232). In addition, culture and language have a closed relationship. Language is an important element of cross-cultural communication that needs to be considered (Gunawardena et al., 2001). "The cyclical nature of the relationship between culture and language suggests that no complete understanding of culture can be obtained without understanding language and vice-versa" (Matsumoto, 1996, p. 266).

Barberà et al. (2014) raised an important issue that present day e-learning has expanded, and the scope for sharing academic courses between countries has increased. However, whether a course designed for the learners of a specific country for a specific discipline will be relevant to the learners of the same discipline in another country has not yet been determined. Recent literature suggests that culture plays an important role in online learning (Milheim, 2014). Some of the researchers have included culture as a dimension in their study. For example, Gunawardena et al. (2001) conducted a cross-cultural study to examine students' perceptions towards online group process and development. They found that culture influences participant's online learning behavior in terms of collectivism and context communication. Bates (2001) also concluded that culture differences influence the online teaching-learning process in terms of students' willingness to participate in the online discussion forum. Some researchers found that cultural differences create barriers in students' communication and social-interaction which affect learner's overall learning performance, motivation, and satisfaction in online learning (Hamdan, 2014; Kerr, 2014; Olaniran, 2009; Uzuner, 2009). On the other hand, some researchers advocated that students in online learning can be benefited by the international exposure and exchange of ideas among the students from different cultural backgrounds. This type of learning environments will develop and inculcate social learning and thinking of the participants (Gemmell et al., 2015; Stewart, 2004). Kim and Bonk (2002) investigated cross-cultural differences between U.S. and Finland, in online collaborative learning behavior, and found significant differences. Lim (2004) conducted a cross-cultural study to examine differences and similarities in learner's motivation in online learning between Korea and U.S. and also how cultural differences and learner's characteristics and culture affect the learning motivation of online learners. The researcher suggested that identification of learning strategy better suited for learning environment with cross-cultural differences should be the priority for the instructors and instructional designers. In another study, Zhu (2012) investigated the cultural differences between Flemish and Chinese students in online learning with respect to students' satisfaction and performance. The results revealed significant differences in students' satisfaction and academic achievement across the culture. Yang et al. (2014) found that culture plays an influencing factor in learning online. Students were interested and very positive to know other participants' cultures and backgrounds.

Previous research studies have investigated the effects of gender differences in online learning. For example, the study conducted by González-Gómez et al. (2012) revealed that females were more satisfied in e-learning environment than males. Nistor (2013) investigated the impact of gender differences in university students' attitude and participation in online courses. They found no significance difference between male's and female's attitude towards online courses. However, there was significant differences in terms of participation. Female students were more participative compared to the male students. Kimbrough et al. (2013) examined the gender differences in mediated communication. The results revealed that women used text messaging, social media, and online video calls more frequently as compared to men. In another study by Song et al. (2015), results indicated no significant gender differences in online collaboration. Liu and Young (2017) found significant gender differences in terms of learning achievement and motivation in an online community-based English reading contest.

The above literature reviews of online learning indicate that studies related to gender differences in online learning are few and findings reported are mixed and inconsistent. As a result, it remains unanswered whether there are any gender differences exist in students' perception towards online learning across different culture.

Issues and Research Questions

From the above literature review, we can draw two conclusions. First, although previous studies have focused on cross-cultural differences in online learning motivation (Lim, 2004), satisfaction, performance and knowledge construction (Zhu, 2012) and online collaboration (Kim & Bonk, 2002), so far no researchers has compared students' perception towards online learning across countries. Second, those studies do not look at gender differences across cultures, although gender differences are treated in many non-Western educational systems. The present study endeavors to fill part of the research gap by focusing on two countries in Asia with very different cultures and historical treatment of women – namely, India and Taiwan. Hence, two objectives guide this study:

1. To examine whether there are any significant differences in the student's perceptions towards online learning between India and Taiwan.
2. To explore whether there are any significant gender differences between India and Taiwan in the student's perceptions towards online learning.

METHOD

Sample and Data Collection

Indian sample

The sample was collected from different parts India using convenience sampling method. The participants only who had online learning experience responded using Google Online survey via student forums. Of the total 233 Indian students who participated in the study, there were 189 (81.1%) males and 44 (18.8%) females. Participants' ages ranged from 17 to 43 years old with a mean of 21.15 years (*SD*=3.5). In terms of the educational background, 180 (77.2%) were undergraduate, 37 (15.8%) were masters, and 16(6.8%) were Ph.D. students. **Table 1** shows the demographic statistics of the participants.

Table 1. Demographic information of the participants

		Indian sample (n=233)	Taiwanese sample (n=208)
Measure	**Category**	**Number**	**Number**
Gender	Female	44	112
	Male	189	96
Age(years)	16-20	133	38
	21-25	75	102
	26-30	19	27
	31 and above	6	40
Education	Undergraduate	180	96
	Master	37	95
	Ph.D.	16	17
No of smartphone users	Yes	135	187
	No	98	21
No of Facebook users	Yes	221	205
	No	12	03
Frequency of daily Internet use	Frequently(>5 hours)	62	123
	Normally(3-4 hours)	51	69
	Occasionally (1-2 hours)	100	15
	Rarely (<1 hours)	20	01

Taiwanese sample

The sample of Taiwanese participants from different parts of Taiwan was collected using the similar method via BBS (Bulletin Board System). Of the total Taiwanese 208 students, there were 96 (46.1%) males and 112 (53.8%) were females. Participants' ages ranged from 18 to 49 years old with a mean of 25.35 years (*SD*=6.55). In terms of the educational background, 96 (46.1%) were undergraduate, 95 (45.6%) were masters, and 17 (8.1%) were Ph.D. students.

Instrument

We employed the POSTOL instrument developed by Bhagat et al. (2016) and six items for demographic information. The POSTOL instrument is a 5-point Likert-type scale, which contains 16 items. The POSTOL instrument has four dimensions: *instructor characteristics* (5 items), *social presence* (5 items), *instructional design* (3 items), and *trust* (3 items). Bhagat et al. (2016) reported the psychometric properties of the POSTOL. Results showed that POSTOL is a reliable and valid to use. In the present study, internal consistency reliability coefficients of the four dimensions ranged between .70 to .89 for the Indian sample and .72 to .92 for the Taiwanese sample (see **Table 3**).

Data Analysis

To examine the normality of the data, skewness and kurtosis values were inspected. The skewness and kurtosis values are within the recommended values |3| and |10| for both the Indian and Taiwanese sample respectively (Kline, 2005). This proved the normality of the data. Box's M statistics was employed to examine the homogeneity of variances and was found to be 58.88 (F=1.91, p>.001). Correlation analyses and multivariate analysis of variance (MANOVA) were used to analyze the data obtained.

RESULTS

Means and standard deviations of the POSTOL subscales for Indian and Taiwanese males and females are shown in **Table 2**. For Indian sample, the highest mean was on trust subscale and the lowest mean was on instructor characteristics for both males and females. Similar results were found for the Taiwanese sample also.

Table 2. Descriptive statistics for POSTOL subscale scores on a 5-point scale

Dependent variables	Indian sample (n=233)						Taiwanese sample (n=208)					
	Gender						Gender					
	Male (n=189)			Female (n=44)			Male (n=96)			Female (n=112)		
	Mean	Skewness	Kurtosis	Mean	Skewness	Kurtosis	Mean	Skewness	Kurtosis	Mean	Skewness	Kurtosis
Instructor characteristics	1.72 (.90)	1.56	2.133	1.69 (.91)	1.28	.821	1.94 (.77)	1.22	1.65	1.84 (.81)	1.45	2.54
Social presence	1.93 (.86)	1.25	1.59	1.92 (.83)	.82	.206	2.05 (.77)	1.17	1.505	2.15 (.69)	.724	1.35
Instructional design	1.89 (.88)	1.08	.535	1.87 (.90)	1.251	.555	2.16 (.73)	.83	.707	2.24 (.75)	.646	.288
Trust	2.53 (.96)	.30	-.373	2.55 (.79)	.044	-.481	2.67 (.84)	.012	-.266	2.84 (.57)	-.298	.06

Note. Values enclosed in parentheses represent standard deviations

Table 3. Intercorrelations and internal consistency coefficients on the POSTOL subscale scores for the Indian and Taiwanese samples

Variables	Variables			
	1.	2.	3.	4.
1. Instructor characteristics				
2. Social presence	.775*(.603*)			
3. Instructional design	.697*(.509*)	.738*(.587*)		
4. Trust	.352*(.113)	.552*(.391*)	.466*(.231*)	
Internal consistency coefficients	.899(.920)	.854(.90)	.703(.727)	.706(.732)

*$p<.01$

Note. The correlation coefficients outside of parentheses belong to the Indian sample whereas values inside parentheses belong to the Taiwanese sample

Table 4. MANOVA results

Dependent Variable	Country			Gender			Country x Gender		
	F	Sig.	η^2	*F*	Sig.	η^2	*F*	Sig.	η^2
Instructor characteristics	14.16	.000	.03	1.06	.30	.002	.09	.76	.000
Social presence	9.48	.002	.02	.55	.45	.001	.90	.34	.002
Instructional design	30.12	.000	.06	.22	.63	.001	.01	.89	.000
Trust	8.41	.004	.02	1.98	.16	.005	1.81	.17	.004

Intercorrelations on the POSTOL subscales for the Indian and Taiwanese sample are shown in **Table 3**. In both the Indian and Taiwanese samples, the highest intercorrelation was between the instructor characteristics and social presence (r=.775 and r=.603 respectively). The lowest intercorrelation for the Indian sample was between the instructor characteristics and trust (r=.352) and between the instructional design and trust in the Taiwanese sample (r=.231).

MANOVA was used to examine the effects of the two independent variables (i.e., country and gender) and four dependent variables (i.e., instructor characteristics, social presence, instructional design, and trust), which were obtained by POSTOL. In addition, interaction effect of the two groups on the dependent variables was also investigated. As reported in **Table 4**, there was a significant effect of the country on students' perceptions towards online learning, Wilk's Λ =.905, F=11.40, p<.05, η^2=.09, showing a large effect. There was no significant effect of the gender on students' perceptions towards online learning: Wilk's Λ =.983, F=1.88, p>.05, η^2=.01. Interaction effects on students' perceptions were also not significant: Wilk's Λ =.985, F=1.67, p>.05, η^2=.01.

DISCUSSION AND CONCLUSIONS

Over the last decade, research in online learning has been growing exponentially due to the current advancements of information and communication technology in the education sector. Culture and language have long been found influencing factors in online learning (Barak et al., 2016; Ordóñez, 2014). Even though motivation, satisfaction, academic achievement, and instructional design have long studied, no studies have investigated students' perceptions towards online learning in the cross-cultural context. Therefore, the present study guided by socio-cultural theory, attempted to examine the differences in students' perceptions towards online learning in terms of gender and country.

An interesting similarity between the two samples was, in both Indian and Taiwanese samples and for both males and females; the highest mean was on trust. These results suggest that building trust in online learning platform should be the first priority for the course providers. This can influence students' participation in online learning and motivate them to complete the course. All the samples scored lowest on instructor characteristics. A possible reason for this result is that the students do not have the opportunity to interact with the instructor face to face.

When intercorrelations were examined, the highest correlation was between the instructor characteristics and social presence in both groups. These two variables shared 60% and 36% of variability in the Indian and Taiwanese samples, respectively. These results indicate that social presence could be an essential element for the instructors to be included in order to create an effective online learning environment to enhance and improve students' academic performance. Instructors should provide more opportunities for interaction, which can construct social presence.

The statistical results of this study showed significant differences between the Indian samples and Taiwanese samples across the fours dimensions (i.e. instructor characteristics, social presence, instructional design, and trust). India is known for multiculturalism in terms demographic, linguistic, religion and social (Bhattacharyya, 2003; Khandelwal et al., 2004). Currently, people speak in 780 different languages in India. But, English and Hindi are the official languages. On the other hand, Taiwan is a relatively monoculture country (Prowse, 2015). Taiwanese Hokkien is spoken by the majority of the population and Mandarin Chinese is used as official language. When the descriptive statistics were examined between the Indian sample and Taiwanese sample, demographic differences were found among the number of smart phone users, number of Facebook users, and frequency of daily Internet use. These results revealed that Taiwan has a better infrastructure for technology than India for providing online learning. Our results are consistent with the studies that indicated that culture and language play an influential role in online learning environment (Barak et al., 2016; Bates, 2001; Gunawardena et al., 2001; Mercer, 2000). The results of this study are also consistent with the study by Lim (2004) that highlight importance of cross-cultural settings in online learning. This study contributes to the existing literature by showing that the perceptions of the students towards online learning differ in the cross-cultural context. In addition, no study previously has investigated the students' perceptions in relation to the combined effect effects of country and gender. The results did not show a significant main effect for gender or an interaction effect between gender and country.

The discussion above leads to the conclusion that to promote active engagement of the learners and delivery of meaningful learning at a global level platform like MOOCs (Massive Open Online Courses), it is very important for the instructional designers to understand student's perceptions towards online learning in different cultural settings. Cross-cultural differences have major influence on students' perceptions towards online learning. Instructional designers should apply socio-cultural theory to take into account the effect of diverse cultural background on the learner's learning behavior to enhance learning effectiveness. Instructors and course designers need to be aware of the cultural diversity among the participants and provide more opportunities for peer to peer and instructor to peer interaction in order to solve the challenge of creating social presence. Instructors should be very careful while assigning the students from different cultural backgrounds to the same group assessment/tasks.

LIMITATIONS AND FUTURE DIRECTIONS

Our study has not violated any multivariate assumptions, but still there are some limitations, which need to be considered. First, finding the equivalent sample for comparison was the greatest challenge for us. There may be differences in the participants' socio-economic status or education system from India and Taiwan. Heterogeneity of the participants was difficult to match. Second, the effect of rural-urban digital divides was not considered. Third, we used convenience sample method for data collection, which may hinder the generalization of the results. Future studies should collect qualitative data (i.e. individual interviews, direct observations, etc.). Our study has included only two countries, which are from Asia. Including more countries from different cultural backgrounds will help to generalize the results obtained. Therefore, it is highly recommended to include Western countries for future studies to compare with Eastern countries.

REFERENCES

Altbach, P. G. (2014). MOOCs as neocolonialism: who controls knowledge? *International Higher Education, 75*, 5-7.

Ashong, C. Y., & Commander, N. E. (2012). Ethnicity, gender, and perceptions of online learning in higher education. *Journal of Online Learning and Teaching, 8*(2), 98-110.

Barak, M., Watted, A., & Haick, H. (2016). Motivation to learn in massive open online courses: Examining aspects of language and social engagement. *Computers & Education, 94*, 49-60. doi:10.1016/j.compedu.2015.11.010

Barberà, E., Layne, L., & Gunawardena, C. N. (2014). Designing online interaction to address disciplinary competencies: A cross-country comparison of faculty perspectives. *The International Review of Research in Open and Distributed Learning, 15*(2).

Bates, T. (2001). International distance education: Cultural and ethical issues. *Distance Education, 22*(1), 122-136.

Bhagat, K. K., Wu, L. Y., & Chang, C. Y. (2016). Development and validation of the perception of students towards online learning (POSTOL). *Educational Technology & Society, 19*(1), 350-359.

Bhattacharyya, H. (2003). Multiculturalism in contemporary India. *International Journal on Multicultural societies, 5*(2), 148-161.

Chen, K.-C., & Jang, S.-J. (2010). Motivation in online learning: Testing a model of self-determination theory. *Computers in Human Behavior, 26*(4), 741-752. doi:10.1016/j.chb.2010.01.011

Cheng, S.-S., Liu, E. Z.-F., Chen, N.-S., Shih, R.-C., & Chang, C.-S. (2012). Gender differences in college students' behaviors in an online question-answer discussion activity. *The Asia-Pacific Education Researcher, 21*(2), 245-256.

de Barba, P. G., Kennedy, G. E., & Ainley, M. D. (2016). The role of students' motivation and participation in predicting performance in a MOOC. *Journal of Computer Assisted Learning, 32*(3), 218-231. doi:10.1111/jcal.12130

Delahunty, J., Jones, P., & Verenikina, I. (2014). Movers and Shapers: teaching in online environments. *Linguistics and Education, 28*, 54-78.

Eryilmaz, E., van der Pol, J., Ryan, T., Clark, P. M., & Mary, J. (2013). Enhancing student knowledge acquisition from online learning conversations. *International Journal of Computer-Supported Collaborative Learning, 8*(1), 113-144. doi:10.1007/s11412-012-9163-y

Gemmell, I., Harrison, R., Clegg, J., & Reed, K. (2015). Internationalisation in online distance learning postgraduate education: a case study on student views on learning alongside students from other countries. *Innovations in Education and Teaching International, 52*(2), 137-147. doi:10.1080/14703297.2014.881264

González-Gómez, F., Guardiola, J., Martín Rodríguez, Ó., & Montero Alonso, M. Á. (2012). Gender differences in e-learning satisfaction. *Computers & Education, 58*(1), 283-290. doi:10.1016/j.compedu.2011.08.017

Gunawardena, C. N., Nolla, A. C., Wilson, P. L., Lopez-Islas, J. R., Ramirez-Angel, N., & Megchun-Alpizar, R. M. (2001). A cross-cultural study of group process and development in online conferences. *Distance Education, 22*(1), 85-121. doi:10.1080/0158791010220106

Hamdan, A. K. (2014). The reciprocal and correlative relationship between learning culture and online education: A case from Saudi Arabia. *The International Review of Research in Open and Distributed Learning, 15*(1), 309-336.

Horzum, M. B., Önder, İ., & Beşoluk, Ş. (2014). Chronotype and academic achievement among online learning students. *Learning and Individual Differences, 30*, 106-111. doi:10.1016/j.lindif.2013.10.017

House, R. J., Hanges, P. J.,Javidan, M.,Dorfman, P. W.,&Gupta, V. . (2004). *Culture, Leadership and Organizations: The GLOBE Study of 62 Nations*. Thousand Oaks, CA: Sage.

Joksimović, S., Gašević, D., Loughin, T. M., Kovanović, V., & Hatala, M. (2015). Learning at distance: Effects of interaction traces on academic achievement. *Computers & Education, 87*, 204-217. doi:10.1016/j.compedu.2015.07.002

Joo, Y. J., Oh, E., & Kim, S. M. (2015). Motivation, instructional design, flow, and academic achievement at a Korean online university: a structural equation modeling study. *Journal of Computing in Higher Education, 27*(1), 28-46. doi:10.1007/s12528-015-9090-9

Kerr, B. (2014). *Challenges in global e-learning: What does the literature tell us.* Paper presented at the *2014 International Conference on Interactive Collaborative Learning (ICL)*.

Khandelwal, S. K., Jhingan, H. P., Ramesh, S., Gupta, R. K., & Srivastava, V. K. (2004). India mental health country profile. *International Review of Psychiatry, 16*(1-2), 126-141. doi:10.1080/09540260310001635177

Kim, K.-J., & Bonk, C. J. (2002). Cross-cultural Comparisons of Online Collaboration. *Journal of Computer-Mediated Communication, 8*(1), 0-0. doi:10.1111/j.1083-6101.2002.tb00163.x

Kim, Y., Glassman, M., & Williams, M. S. (2015). Connecting agents: Engagement and motivation in online collaboration. *Computers in Human Behavior, 49*, 333-342. doi:10.1016/j.chb.2015.03.015

Kimbrough, A. M., Guadagno, R. E., Muscanell, N. L., & Dill, J. (2013). Gender differences in mediated communication: Women connect more than do men. *Computers in Human Behavior, 29*(3), 896-900. doi:10.1016/j.chb.2012.12.005

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Kupczynski, L., Brown, M., Holland, G., & Uriegas, B. (2014). The Relationship between Gender and Academic Success Online. *Journal of Educators Online, 11*(1), 1-14.

Lee, S. J., Srinivasan, S., Trail, T., Lewis, D., & Lopez, S. (2011). Examining the relationship among student perception of support, course satisfaction, and learning outcomes in online learning. *Internet and Higher Education, 14*(3), 158-163. doi:10.1016/j.iheduc.2011.04.001

Lim, D. H. (2004). Cross Cultural Differences in Online Learning Motivation. *Educational Media International, 41*(2), 163-173.

Liu, I. F., & Young, S. S. C. (2017). An exploration of participative motivations in a community-based online English extensive reading contest with respect to gender difference. *Interactive Learning Environments, 25*(1), 48-61. doi:10.1080/10494820.2015.1090457

Loizzo, J., & Ertmer, P. A. (2016). MOOCocracy: the learning culture of massive open online courses. *Educational Technology Research and Development*, 1-20. doi:10.1007/s11423-016-9444-7

Matsumoto, D. (1996). *Culture and psychology*. Pacific Grove, CA: Brooks Cole Publishing.

McLoughlin, C. (1999). Culturally responsive technology use: developing an on-line community of learners. *British Journal of Educational Technology, 30*(3), 231-243. doi:10.1111/1467-8535.00112

Mercer, N. (2000). *Words and Minds: How we use language to think together*. London, UK: Routledge.

Milheim, K. L. (2014). Cultural inclusivity in online learning *Student-Teacher Interaction in Online Learning Environments* (pp. 76-88).

Nistor, N. (2013). Stability of attitudes and participation in online university courses: Gender and location effects. *Computers & Education, 68*, 284-292. doi:10.1016/j.compedu.2013.05.016

Olaniran, B. A. (2009). Discerning culture in e-learning and in the global workplaces. *Knowledge Management & E-Learning: An International Journal (KM&EL), 1*(3), 180-195.

Olesova, L., Yang, D., & Richardson, J. C. (2011). Cross-Cultural Differences in Undergraduate Students' Perceptions of Online Barriers. *Journal of Asynchronous Learning Networks, 15*(3), 68-80.

Ordóñez, A. C. (2014). *Predicting International Critical Success Factors in e-learning: A comparison of four universities from China, Mexico, Spain and USA*. Retrieved from http://www.tdx.cat/bitstream/handle/10803/276198/Thesis%20Armando%20Cortes_24_Jul_2014_ok.pdf;jsessionid=18B3E50FA1797D31F8EB6F052E768FE2.tdx1?sequence=1

Osman, G., & Herring, S. C. (2007). Interaction, facilitation, and deep learning in cross-cultural chat: A case study. *The Internet and Higher Education, 10*(2), 125-141. doi:10.1016/j.iheduc.2007.03.004

Prowse, S. (2015). Taiwan's Growing Multiculturalism. *The Diplomat*. Retrieved from http://thediplomat.com/2015/02/taiwans-growing-multiculturalism/

Song, H.-g., Restivo, M., van de Rijt, A., Scarlatos, L., Tonjes, D., & Orlov, A. (2015). The hidden gender effect in online collaboration: An experimental study of team performance under anonymity. *Computers in Human Behavior, 50*(Supplement C), 274-282. doi:10.1016/j.chb.2015.04.013

Stewart, B. L. (2004). Online learning: a strategy for social responsibility in educational access. *The Internet and Higher Education, 7*(4), 299-310. doi:10.1016/j.iheduc.2004.09.003

Uzuner, S. (2009). Questions of culture in distance learning: A research review. *The International Review of Research in Open and Distributed Learning, 10*(3), 1-11.

Vygotsky, L. S. (1978). *Mind in society: Interaction between learning and development*. Cambridge. MA: Harvard University Press.

Wang, C.-H., Shannon, D. M., & Ross, M. E. (2013). Students' characteristics, self-regulated learning, technology self-efficacy, and course outcomes in online learning. *Distance Education, 34*(3), 302-323. doi:10.1080/01587919.2013.835779

Wang, M. (2007). Designing online courses that effectively engage learners from diverse cultural backgrounds. *British Journal of Educational Technology, 38*(2), 294-311. doi:10.1111/j.1467-8535.2006.00626.x

Wikipedia. (2016). *List of languages by total number of speakers*. Retrieved on June 19, 2016 from https://en.wikipedia.org/w/index.php?title=Special:CiteThisPage&page=List_of_languages_by_total_number_of_speakers&id=726089308

Wong, L.-H. (2012). A learner-centric view of mobile seamless learning. *British Journal of Educational Technology, 43*(1), E19-E23. doi:10.1111/j.1467-8535.2011.01245.x

Xie, K., Miller, N. C., & Allison, J. R. (2013). Toward a social conflict evolution model: Examining the adverse power of conflictual social interaction in online learning. *Computers & Education, 63*, 404-415. doi:10.1016/j.compedu.2013.01.003

Yang, J., Huiju, Y., Cen, S.-J., & Huang, R. (2014). Strategies for Smooth and Effective Cross-Cultural Online Collaborative Learning. *Educational Technology & Society, 17*(3), 208-221.

You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning. *The Internet and Higher Education, 29*, 23-30. doi:10.1016/j.iheduc.2015.11.003

Zhu, C. (2012). Student Satisfaction, Performance, and Knowledge Construction in Online Collaborative Learning. *Educational Technology & Society, 15*(1), 127-136.

http://www.ejmste.com

Education Tech Research Dev
https://doi.org/10.1007/s11423-018-9573-2

DEVELOPMENT ARTICLE

The design, implementation, and evaluation of a digital interactive globe system integrated into an Earth Science course

Wei-Kai Liou[1] · Kaushal Kumar Bhagat[2] · Chun-Yen Chang[3,4]

Abstract The aim of this study is to design and implement a digital interactive globe system (DIGS), by integrating low-cost equipment to make DIGS cost-effective. DIGS includes a data processing unit, a wireless control unit, an image-capturing unit, a laser emission unit, and a three-dimensional hemispheric body-imaging screen. A quasi-experimental study was conducted to evaluate the learning effectiveness of our system. A total of 105 junior high-school students from Taiwan participated in this 8-week experiment. The students were divided into three individual groups of 35 students each, with one control group and two experimental groups (EG1 and EG2). The results of one-way mixed design ANOVA indicated that participants in the experimental group, who used the DIGS, outperformed the other two groups, in the post-test as well as in the delayed test. These findings demonstrate that the proposed DIGS can effectively enhance the performance of the learners in an Earth Science course.

Keywords Digital Earth · Earth Science · Interactive learning environment · Learning outcomes

Wei-Kai Liou and Kaushal Kumar Bhagat have contributed equally to this work.

✉ Chun-Yen Chang
changcy@ntnu.edu.tw

1 Department of Interactive Entertainment Design, China University of Technology, Taipei, Taiwan, ROC

2 Smart Learning Institute of Beijing Normal University, Beijing, China

3 Science Education Center, National Taiwan Normal University, 88 Sec. 4, Ting-Chou Rd., Taipei, Taiwan

4 Department of Earth Sciences, National Taiwan Normal University, Taipei, Taiwan

Published online: 05 February 2018

Introduction

Various models of the Earth, such as globes showing countries, continents and other aspects of the Earth, are often used as teaching aids for formal or informal educational teaching in schools, museums as well as other education institutions. In recent years, as technologies progress, many researchers have attempted to project images utilizing Earth simulation software onto screens or as spherical representations, in the form of digital teaching aids (Ancona et al. 2002; Wu et al. 2010; Xie and Reider 2014; Zhu et al. 2008). For example, Philip M. Sadler invented the *Starlab* portable planetarium in 1977 (see https://www.cfa.harvard.edu/news/2014-24). Starlab has been used for teaching astronomy (see also http://starlab.com/), but it is expensive. Berry et al. (2007) developed a spherical display system known as *Aggie Orb*. In this system, visual effects were produced by multicolor LEDs rotating along a spherical trajectory. The heavy weight of *Aggie Orb*, which is approximately 200 lb, and its large size, are major limitations of this system. Therefore, the high costs and complexity of the equipment required, current display systems simulating the Earth are not generally available for classrooms. This article describes the development and use of a three-dimensional (3-D) simulated Earth system, utilizing existing classroom equipment to improve learning performance in the Earth Science course. Our system projects a 3-D spherical image of the Earth. In addition, this system allows the operator to use a simple handheld wireless control device to interact with the 3-D interactive globe system.

Theoretical foundations

Considerations from the cognitive theory of multimedia learning (CTML) and cognitive flexibility theory (CFT) were integrated to design digital interactive globe system (DIGS).

Mayer (2001) proposed CTML, which is based in part on dual coding theory (Paivio 1986) and cognitive load theory (Chandler and Sweller 1991). DIGS incorporated five principles of CTML to improve learning: (1) modality principle, states that learners learn better from graphics and narrations than from animation and on-screen text, (2) temporal contiguity principle, states that learners learn better when corresponding words and pictures are presented simultaneously rather than successively, (3) spatial contiguity principle, states that learners learn better when corresponding words and pictures are presented near rather than far from each other on the page or screen, (4) redundancy principle, states that learners learn better from graphics and narration than from graphics, narration and on-screen text, and (5) multimedia principle, states that learners learn better from words and pictures than from words alone. According to CTML, when symbolic information or text is presented, information is processed through the visual channel of the working memory. However, when narration text is presented, information is processed through the auditory channel. On the other hand, when multimedia information, for example, animation with spoken narration is presented, it can be processed through two channels (visual and auditory) in parallel. This type of information processing is more efficient and reduces the load placed on working memory, according to dual coding theory and cognitive load theory.

Spiro and Jehng (1990) proposed CFT, which builds upon constructivism. Constructivism refers to a naturalistic epistemology that recognizes the significance of an active learning process in which learners construct new knowledge from their current or prior knowledge (Jonassen 1994; Spector 2015). Spiro and Jehng (1990) defined cognitive flexibility as "the ability to spontaneously restructure one's knowledge, in many ways, in adaptive response to radically changing situational demands" (p. 165). CFT helps the learners to develop a deeper understanding of complex concepts and to apply that knowledge in real-world contexts (Spiro et al. 2003). This is consistent with the findings reported by Dörner (1996) in *The Logic of Failure*, which promotes rich multimedia and interactive simulations to support learning in complex domains.

CFT recommends that learning activities must provide multiple representations of the contents to discover and explore the complex problems (Spiro et al. 1988). Chieu (2007) proposed operational criteria of CFT and identified four important components of learning systems: (1) learning contents (e.g., text, images, audio, video, simulations), (2) pedagogical devices (e.g., tools provided for learners for exploring learning contents), (3) human interactions (e.g., means for engaging tutors and learners in exchanges), and (4) assessment (e.g., post-tests for determining whether learners have achieved learning objectives). Interactive features in DIGS are well aligned with CFT. In addition, changes in conceptualizations can be determined by simple tests, and it is a change in how a student thinks that is compatible with cognitive flexibility. Also, CFT is consistent with multiple forms of representations, including visualization and DIGS and its impact on learners involves multiple forms of representation.

Many empirical studies have shown the positive effects of CFT in information and communication technology-based learning environments (Dörner 1996; Fitzgerald et al. 1997; Goeze et al. 2014; Jacobson and Spiro 1995; Lima et al. 2004; Lowrey and Kim 2009; Mendes et al. 2001; Zottmann et al. 2012; Zydney 2010; Zydney and Grincewicz 2011). For example, Zydney and Grincewicz (2011) developed a software program based on CFT for socio-scientific problem solving. The results showed that the software helped the students to learn scientific processes, which promotes critical thinking and deeper learning (e.g., seeking and explaining evidence). In another study, Mendes et al. (2001) applied CFT to teach hypermedia-engineering principles. They found that the students who were exposed to CFT-based teaching performed better than those exposed only to traditional methods.

Incorporated with the above-mentioned principles of CTML and CFT, DIGS can provide learning contents (graphics, animations, simulations, etc.) with narrative descriptions simultaneously to make learning more effective and interactive.

Existing Digital Earth (DE) systems

Zhu et al. (2008) defined DE as "a virtual presentation of the planet based on geographic coordinate, and is an information system with tremendous amount of multiple resolutions and multiple scales data as shown in multiple dimensions. It can visualize the real Earth and represent historical phenomena in a digital way by using the large amount of data of the Earth, and utilizing the computer techniques, image and graphic technique, network technique, virtual reality and so on" (p. 118). Existing DE systems such as Google Earth (GE), science on a sphere (SOS) Earth system and Geo-Cosmos are new developments in visualising 2-D and 3-D representations of Earth. Demirci (2009) found that teachers are very positive towards geographic information systems (GISs) in spite of some hardware and software barriers. This positive attitude will be helpful for the development of good teaching plans in geography.

Geospatial technologies are growing in use and popularity as a result of improvements in computational power and easier access to geospatial data (Milla et al. 2005). Patterson (2007) advocated the use of GE (https://earth.google.com/) because it is a potential tool to enhance teaching methods in geography. Doering and Veletsianos (2008) found that students were motivated to use geospatial technology like GE for their class assignments. They also found there was a lack of geospatial technology-based curricula in the formal education systems.

GE maps the Earth by the superimposition of images obtained from satellite imagery, aerial photography and GIS onto a computer screen, but GE has some limitations, some of which are technical and some of which as a result of a 2-D representation. GE cannot work without an Internet connection (Patterson 2007), and it also needs high bandwidth for good performance. Wang et al. (2013) mentioned that GE required some information technology skills and therefore, it may not be suitable for the students who are not proficient in handling technology. Zhu et al. (2014), also highlighted two important limitations of the current DE systems: (1) inability to represent the whole Earth as a 3-D model clearly and (2) DE systems lack the necessary advanced functions in 3-D visualization.

The cost of current commercial DE systems is a matter of concern (Vanhoenacker 2013). Table 1 shows a comparison of the DIGS with selected commercial existing DE systems. Compared to the existing DE systems, DIGS offers instructors some additional features like real time marking and drawing functions, extensive control, including display management by walking around (MBWA), and more. Instructor led MBWA positively affects instructional practices, which results in better student learning achievement (Keruskin 2005). DIGS is an important didactic development in affordable Earth system global displays for education.

According to Roblyer (2005):

When there is a clear need for a better instructional method than those used in the past, researchers can propose that a given technology-based method is the best choice because it offers the combination of relevant symbol systems, processing capabilities, and logistical feasibility to address the need—and then do research to support that it has this relative advantage and clarify the conditions under which it works best (np).

DIGS is aligned with Roblyer's (2005) argument. Lower cost and being able to function in low bandwidth areas are the other most important features of DIGS that makes this system relatively advanced than the other existing DE systems.

Table 1 Comparison of commercial Digital Earth (DE) systems with DIGS

Features	DIGS	SOS	Geo-Cosmos
Cost (USD)	200	43,000	100,000
Dynamic information	Yes	Yes	Yes
Static information	Yes	Yes	Yes
3-D effect	Yes	Yes	Yes
Real time mark and draw function	Yes	No	No
Extensive control and MBWA	Yes	No	No
Application in the classroom	Yes	No	No
Create your own content	Yes	No	No

Need for this study

When reviewing current DE systems, three considerations need to be clarified: (1) none of these aforementioned DE systems offer didactic development in Earth Science courses for formal education, (2) the above-mentioned DE systems do not offer any opportunity for the instructors to incorporate appropriate teaching–learning materials into the systems, (3) most DE systems are too costly to be included into the mainstream education systems in developing countries, and (4) although GE is free, it lacks interactive functions and features found in the DE systems in Table 1. Based on the above considerations, the aim of this study is to design and implement a DIGS to be integrated into a general junior high-school Earth Science course.

DIGS architecture[1]

Hardware

DIGS involves a data processing unit, a wireless control unit, an image capturing unit, a laser pointing device, and a 3-D hemispheric body imaging (HBI) unit—a kind of 3-D screen (see Fig. 1). The 3-D HBI screen is designed to display the output image from the data processing unit. The laser emission unit is for emitting a laser spot on the output image.

Software

DIGS utilizes the Unity game engine. Unity is a powerful game developer engine. For independent developers, this software breaks the time, platform and cost barriers (https://unity3d.com/). We selected this cross-platform software as it supports many types of art and design resources.

Operating process with an example

In the first step, we imported real-world height maps from the website of http://terrain.party/ into Unity software to develop different Earth Science course materials (see Fig. 2). This website is available free to download the maps based on our needs. Using Unity software, real-world heights maps were converted into 3-D terrain models (see Fig. 3).

DIGS is a user-friendly system in the classroom. DIGS allows the users to navigate from one concept to another many times according to their pace. This provides an opportunity for interaction between the user and the system. Users can use the functions like *click*, *zoom in–out*, *rotate*, etc., based on their learning needs. For example, a lesson plan on Argentina terrain needs to follow the following steps:

Step 1 users need to select a country on the system to teach in the classroom.
Step 2 the system will display the selected country to the users on the interface.
Step 3 users need to double-*click* the selected country, and the system will display the 3-D terrain model of the country and need to use the *zoom in* function to view the selected country for more details (see Fig. 4).

[1] To know more information about technical description of DIGS please contact the authors.

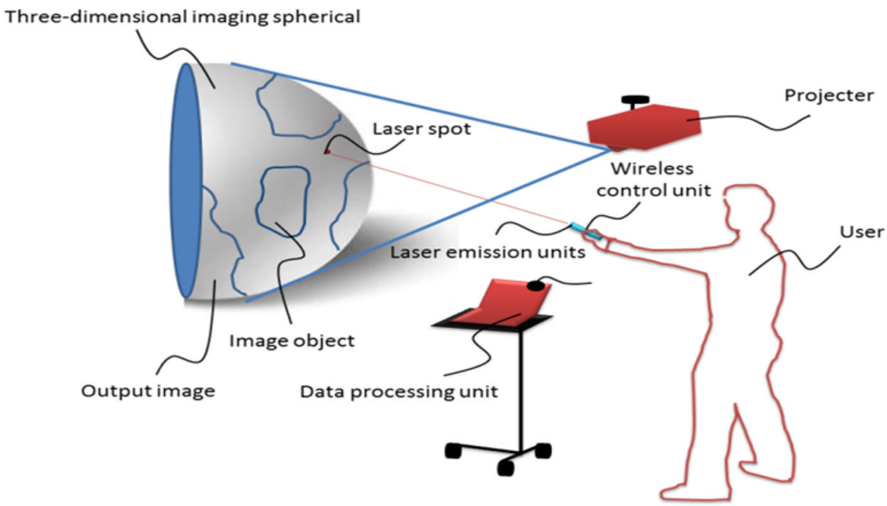

Fig. 1 DIGS architecture

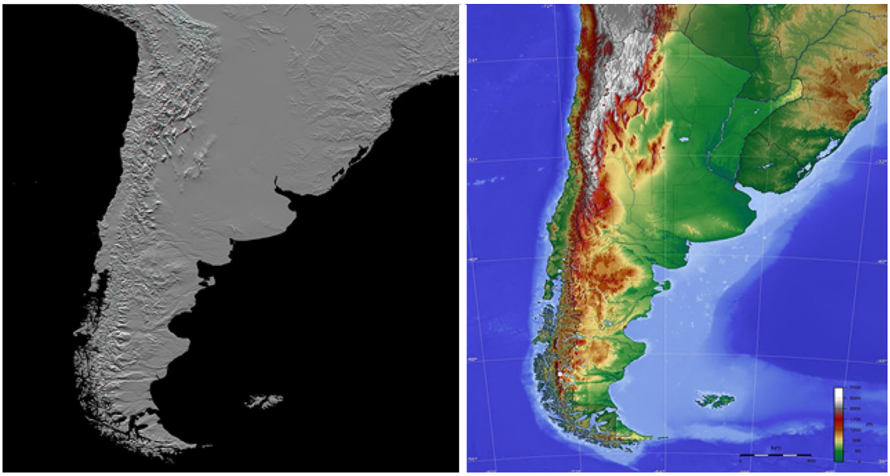

Fig. 2 Height-map downloaded from http://terrain.party/

Step 4 users can use the *rotate* function to view the 3-D terrain model from any angle (see Fig. 5).
Step 5 users can use the *zoom in* function to focus on any particular are of the 3-D terrain (see Fig. 6).

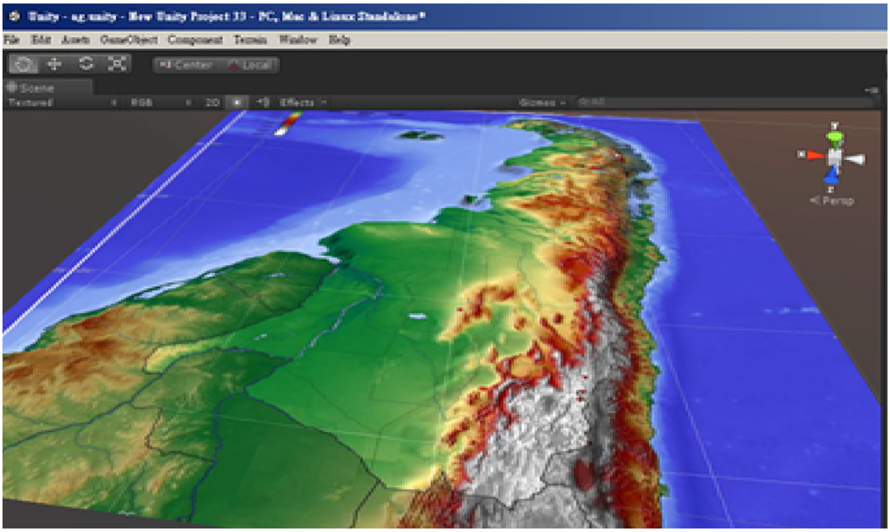

Fig. 3 Example of 3-D terrain model developed by Unity

Fig. 4 Click and zoom in functions in DIGS

Scientific contents incorporated in DIGS

This interactive system allows the users to access various Earth observation data collected from different international sources, e.g., National Aeronautics and Space Administration (NASA), US Geological Survey, National Oceanic and Atmospheric Administration. The content includes *Solar system*, *Earth tectonic system*, *Global carbon-dioxide map*, *Global ocean temperature*, *Global thundering*, *Global hurricane*, and *Global forest map*. For more details (see https://www.youtube.com/watch?v=QOUPhyvOB9o&t=140s).

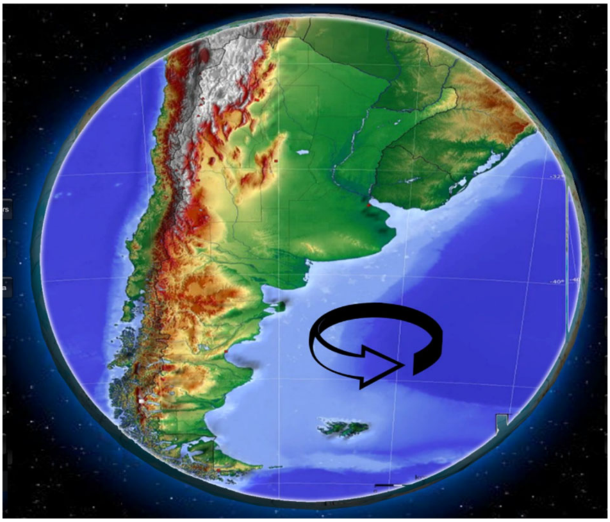

Fig. 5 The rotate function in DIGS

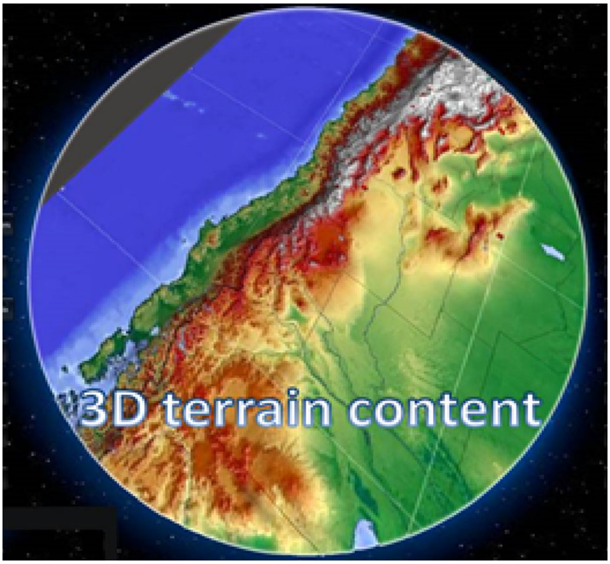

Fig. 6 The zoom in function to see more details of 3-D terrain content

Solar system

DIGS provides an interactive model of solar system (see Fig. 7). Users can point to select any planet, whereby the system provides a detailed description of the selected planet with auditory narrations.

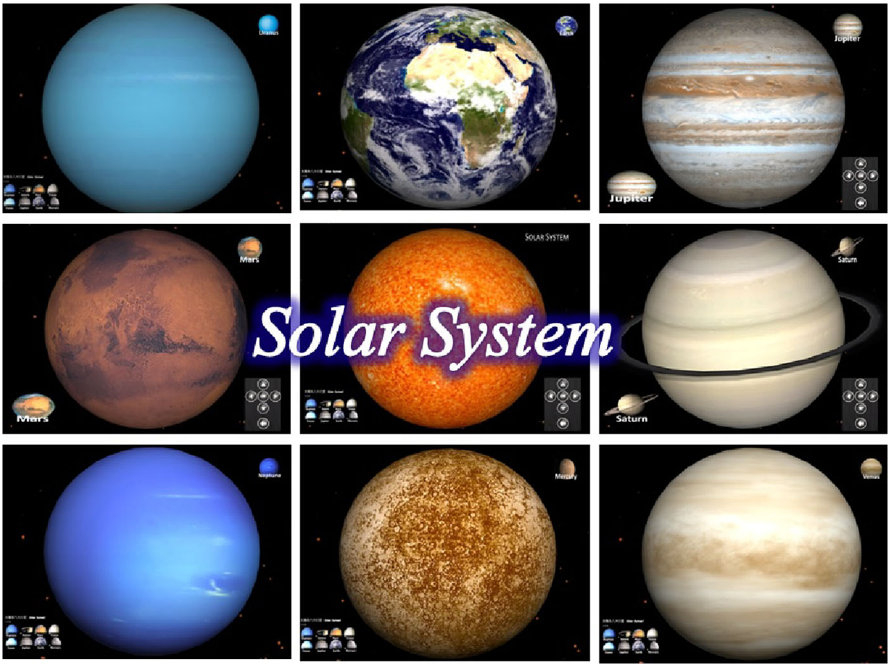

Fig. 7 Screenshot of the DIGS displaying solar system

Earth tectonic system

DIGS describes the Earth tectonic system by visualizing all the components of Earth which are responsible for the movements of the continents, formation of mountains, and ocean basins and occurrence of different natural events such as earthquakes, volcanos, etc. DIGS explains the principles of the Earth's tectonic system (see Fig. 8).

Global carbon-dioxide map

Increased emission of CO_2 is one of the major reasons for global warming (http://timeforchange.org/CO2-cause-of-global-warming). DIGS displays average global concentration of carbon dioxide in different parts of the world (see Fig. 9). This visualization helps the students to understand climate change across the globe.

Global ocean temperature

Global ocean temperature influences Earth climate and weather. In other words, ocean temperature is responsible for the natural disasters like hurricanes, typhoons, etc. DIGS provides an interactive visualization of ocean temperature across the globe (see Fig. 10).

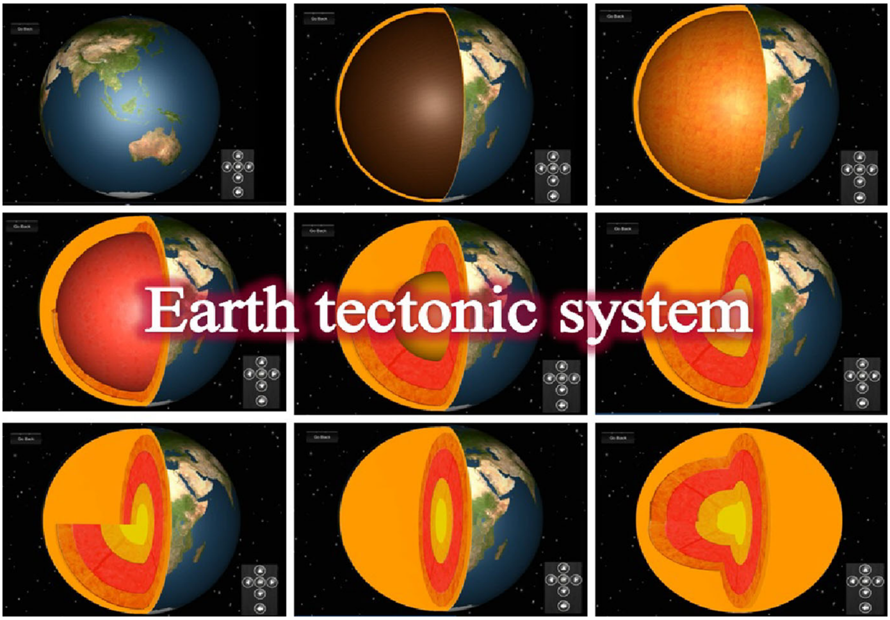

Fig. 8 Screenshot of the DIGS displaying Earth tectonic system

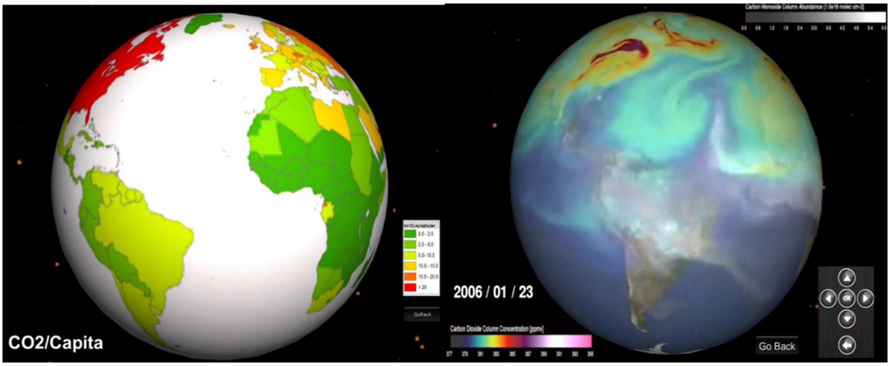

Fig. 9 Screenshot of the DIGS displaying global carbon-dioxide map

Global thundering

Thunder is caused by lightning, which is a result of intense heating and expansion of the air. DIGS shows the number of thundering across the globe based on NASA data (see Fig. 11).

Global hurricane

A hurricane is a cyclonic storm on Earth. DIGS provides interactive information for the better understanding of hurricanes across the globe. Figure 12 displays the visualization of hurricanes created by DIGS.

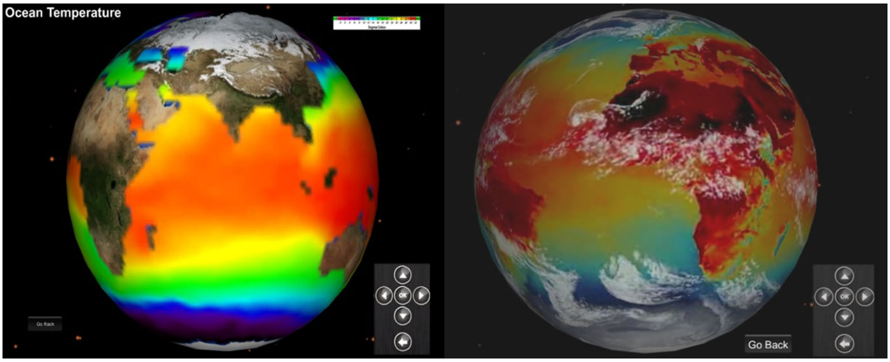

Fig. 10 Screenshot of the DIGS displaying global ocean temperature

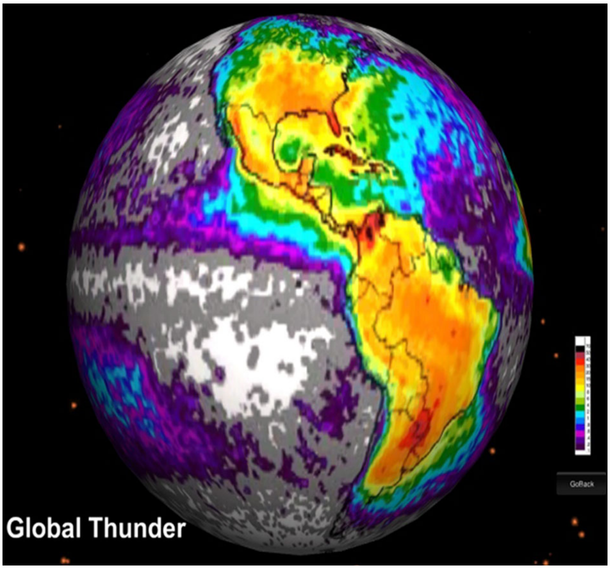

Fig. 11 Screenshot of the DIGS displaying global thundering

Global forest map

Deforestation is one of the major factors contributing to global warming. To protect our environment, understanding the current status of forests across the globe is very important. DIGS presents global forest change over a period (see Fig. 13).

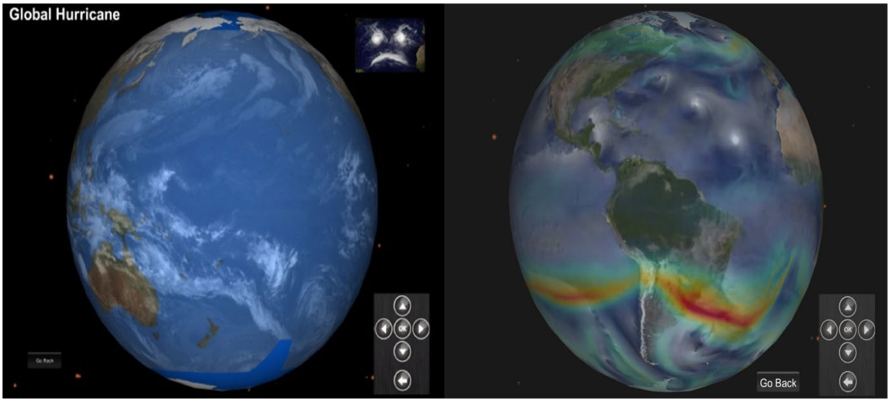

Fig. 12 Screenshot of the DIGS displaying global hurricane

Fig. 13 Screenshot of the DIGS displaying global forest map

Evaluation of DIGS

Research design and sample

The present study employed a quasi-experimental research design, directly assigning participants into groups based randomly on pre-existing classroom settings. A total of 105 (males = 56, females = 49) junior high-school students, aged 14–15 years, were selected from Taiwan and divided into three individual groups of 35 students, including: one control group (CG) and two experimental groups (EG1 and EG2). The CG was instructed using PowerPoint presentation. The other two EGs adopted digital learning system in the class. The only difference between those two experimental classes was that the first EG (EG1) adopted GE in the class while the second EG (EG2) adopted DIGS. The same teacher, using the same course content, taught all three groups. The instructor had previous experience of teaching the course content.

Procedure

This study was conducted over a period of 8 weeks, with four 40-min classes, with one class meeting per week. Therefore, the total duration of classes was 160 min. The objectives of the lesson include: (1) identification of the world distribution of volcanoes and earthquakes, and (2) understanding the Earth layers and tectonic plates.

Before the commencement of the experiment, a pre-test was conducted to assess the level of performance from the students in Earth Science. Test items in pre-test and post-test were isomorphic. The pre-test contained five objective test items and the students were given 5 min to finish the pre-test. Similarly, after the completion of the course, a post-test was administered. The post-test contained 10 objective test items and the students were given 10 min to complete the post-test. For example:

The tectonic plates of Earth are part of Earth's:

(a) Crust,
(b) Mantle,
(c) Lithosphere,
(d) Asthenosphere.

In the last week of the experiment, a delayed post-test was conducted. The content for both the post-test and delayed post-test was the same. The validity and difficulty level of the test items was done by a panel of subject experts and experienced teachers. Content validity index (CVI) developed by Lawshe (1975) was calculated for both pre-test and post-test. CVI for both pretest and post-test was greater than 0.80, which are acceptable (Polit et al. 2007).The Cronbach's α was 0.80 for pre-test and 0.78 for post-test, which are acceptable.

Data analysis

A one-way mixed design ANOVA was conducted to assess the impact of the three different interventions (CG, EG1, EG2) on participants' learning performance in Earth Science course, across three time periods (pre-test, post-test, delayed post-test). All analyses were conducted using the Statistical Package for the Social Sciences version 21 (SPSS 21). The statistical significance level was set at $p < 0.05$.

Results

As shown in Table 2, all the three groups' pre-test mean score increased in the post-test as well as delayed post-test. Therefore, a one-way mixed design ANOVA was conducted to assess the impact of the three different interventions (CG, EG1, EG2) on participants'

Table 2 Mean and standard deviation for pre-test, post-test scores and delayed post-test scores

Groups	Pre-test			Post-test		Delayed post-test	
	N	Mean	SD	Mean	SD	Mean	SD
CG	35	4.11	.47	6.54	.50	5.94	.23
EG1	35	3.97	.38	7.40	.49	6.97	.29
EG2	35	4.22	.54	8.37	.49	7.94	.33

learning performance in Earth Science course, across three time periods (pre-test, post-test, delayed post-test). Mauchy's sphericity test was used to examine the sphericity assumption. The results showed that there was no violation in sphericity assumption, W = .98, $X^2(2) = .197$, $p = .90$. Therefore, the F-value for the main effects and interaction effect did not need correction. As shown in Table 3, the main effect for time was significant, $F(2,204) = 1875.78$, $p < .05$, partial $\eta^2 = 0.94$, which is considered to be a large effect (Cohen 1988). The main effect comparing the three types of intervention was significant, $F(2,102) = 236.33$, $p < .05$, partial $\eta^2 = 0.82$, which is considered to be a large effect (Cohen 1988). In addition, there was significant interaction effect between intervention type and time, $F(2,204) = 53.27$, $p < .05$, partial $\eta^2 = 0.51$, which is considered to be a moderate effect (Cohen 1988).

These results indicated that the achievement scores did follow a different pattern through different times. It could be suggested that students in the different intervention modes performed differently. More specifically, students who were exposed to the DIGS performed better than the other two groups. This situation is further illustrated by Fig. 14.

Discussion and conclusions

In this study, a DIGS was presented for teaching Earth Science courses. DIGS was implemented into a classroom environment for evaluation purposes. This study contributes on both academic and economical view fronts. From the academic fronts of view, empirical results we gathered provide strong support that DIGS is effective and beneficial for the learners for long-term goals. DIGS provides a realistic and captivating 3-D experience to the learners. This resulted in the better performance of the students who participated in DIGS based instruction compared to the other two group. This result is in consistent with Mayer' (2001) CTML theory and Spiro and Jehng's (1990) CFT in that students' learning performance improved by using multiple modes of representations (e.g., graphics, animations, and simulations). Additionally, DIGS presents the informative terrain models that allow the users to adjust and view the displayed images from any angle. This promotes the users to engage with a hands-on-experience resulting in deeper understanding of the concepts presented.

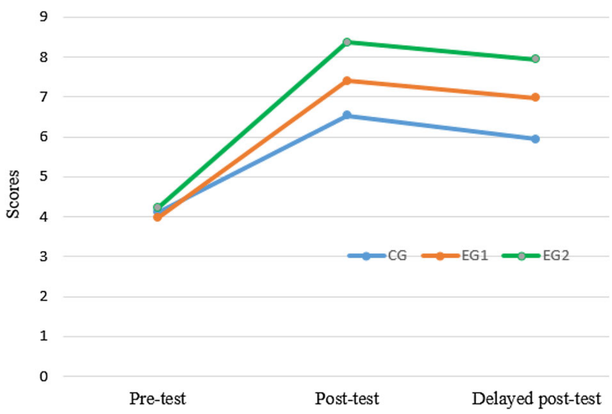

Fig. 14 Comparison of the achievement scores across different times

Table 3 Summary of mixed-design ANOVA

Sources	*SS*	*df*	*MS*	*F*	*p*	Partial η^2
Between groups						
Group	91.09	2	45.54	236.33	.000	.82
Error	19.65	102	.193			
Within groups						
Time	680.95	2	340.47	1875.78	.000	.94
Group × Time	38.67	4	9.67	53.27	.000	.51
Error	37.02	204	.182			

*$p < .05$

The low-cost of the DIGS makes it affordable for developing countries or in rural areas with limited resources. DIGS could be a solution to bridge the wide *digital divide* between countries. The most significant feature of DIGS is that it allows users to interact with its 3-D visualization system. In addition, MBWA function supports the teacher's role as a guide while placing the learner as the focus of teaching–learning process. Therefore, this system seems well suited for teaching and learning Earth Science. Teachers can develop different types of authentic tasks in combination with DIGS to enhance students' critical-thinking skills. DIGS can also be integrated as a tool in the flipped classroom model to make the classroom more interactive. In addition, DIGS can be incorporated in informal learning environments (e.g., museums) to present geographical information visually.

Limitations and future directions for research

The present study has some limitations that should be acknowledged. First, DIGS evaluation was done by using a limited number of test items focusing mainly on *understanding* as domain. Future studies should design test items covering domains like creating, analyzing, applying, etc. Second, observational data (e.g., teacher–student interactions, student–DIGS interaction, etc.) across classroom working with DIGS were not reported. Future studies should collect qualitative data with quantitative to have more in-depth conclusions. Third, we evaluated only students' learning performance for junior high-school students. Future studies should compare the effectiveness of DIGS on students' motivation, satisfaction, and attitude in addition to learning performance, across different grade levels. In addition, the training and support of teachers is worth investigating. Currently, the system is being tested in a rural school in Alabama with minimal training and support of teachers.

We recommend that educators and curriculum designers consider how to effectively integrate 3-D geospatial data lessons so as to connect the classroom with the real world. The future direction of DIGS researchers and developers is focused on creating a database of 3-D map archives as well as a variety of course materials to be used by instructors. In the next step, we are also developing social science contents in addition to science contents and will make it available free for teachers.

References

Ancona, D., Freeston, M., Smith, T., & Fabrikant, S. (2002). Visual explorations for the Alexandria Digital Earth prototype. In K. Börner & C. Chen (Eds.), *Visual interfaces to digital libraries* (pp. 199–213). Berlin: Springer.

Berry, S., Culrich, B., Kisin, M., Knezek, G., Linam-Church, W., & Liu, S. (2007). *Orb spherical display*. https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/tr-2007-94.pdf.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293–332.

Chieu, V. M. (2007). An operational approach for building learning environments supporting cognitive flexibility. *Educational Technology and Society, 10*(3), 32–46.

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

Demirci, A. (2009). How do teachers approach new technologies: Geography teachers' attitudes towards Geographic Information Systems (GIS). *European Journal of Educational Studies, 1*(1), 57–67.

Doering, A., & Veletsianos, G. (2008). An investigation of the use of real-time, authentic geospatial data in the K-12 classroom. *Journal of Geography, 106*(6), 217–225.

Dörner, D. (1996). *The logic of failure: Recognizing and avoiding error in complex situations* (R. R. Kimber, Trans.). New York: Metropolitan Books.

Fitzgerald, G. E., Wilson, B., & Semrau, L. P. (1997). An interactive multimedia program to enhance teacher problem-solving skills based on cognitive flexibility theory: Design and outcomes. *Journal of Educational Multimedia and Hypermedia, 6*(1), 47–76.

Goeze, A., Zottmann, J. M., Vogel, F., Fischer, F., & Schrader, J. (2014). Getting immersed in teacher and student perspectives? Facilitating analytical competence using video cases in teacher education. *Instructional Science, 42*(1), 91–114.

Jacobson, M. J., & Spiro, R. J. (1995). Hypertext learning environments, cognitive flexibility, and the transfer of complex knowledge: An empirical investigation. *Journal of Educational Computing Research, 12*(4), 301–333.

Jonassen, D. H. (1994). Thinking technology: Toward a constructivist design model. *Educational Technology, 34*(4), 34–37.

Keruskin, T. E. (2005). The perceptions of high school principals on student achievement by conducting walkthroughs. EdD-Doctor of Education Doctoral Dissertation, University of Pittsburgh, Pittsburgh.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x.

Lima, M., Koehler, M. J., & Spiro, R. J. (2004). Collaborative interactivity and integrated thinking in Brazilian business schools using cognitive flexibility hypertexts: The Pantheon project. *Journal of Educational Computing Research, 31*(4), 371–406.

Lowrey, W., & Kim, K. S. (2009). Online news media and advanced learning: A test of cognitive flexibility theory. *Journal of Broadcasting and Electronic Media, 53*(4), 547–566.

Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.

Mendes, E., Mosley, N., & Counsell, S. (2001). The cognitive flexibility theory: An approach for teaching Hypermedia Engineering. *SIGCSE Bulletin, 33*(3), 21–24. https://doi.org/10.1145/507758.377457.

Milla, K. A., Lorenzo, A., & Brown, C. (2005). GIS, GPS, and remote sensing technologies in extension services: Where to start, what to know. *Journal of extension, 43*(3). http://www.joe.org/joe/2005june/a6.shtml.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.

Patterson, T. C. (2007). Google Earth as a (not just) geography education tool. *Journal of Geography, 106*(4), 145–152.

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing and Health, 30*(4), 459–467. https://doi.org/10.1002/nur.20199.

Roblyer, M. D. (2005). Educational technology research that makes a difference: Series introduction. *Contemporary Issues in Technology and Teacher Education, 5*(2). http://www.citejournal.org/vol5/iss2/seminal/article1.cfm.

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

Spiro, R. J., Collins, B. P., Thota, J. J., & Feltovich, P. J. (2003). Cognitive flexibility theory: Hypermedia for complex learning, adaptive knowledge application, and experience acceleration. *Educational Technology, 45*(5), 5–10.

Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (1988). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. In V. Patel (Ed.), *Tenth annual conference of the cognitive science society* (pp. 375–383). Hillsdale, NJ: Erlbaum.

Spiro, R. J., & Jehng, J. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix & R. Spiro (Eds.), *Cognition, education, and multimedia: Exploring ideas in high technology* (pp. 163–204). Hillsdale, NJ: Erlbaum.

Vanhoenacker, M. (2013). Digital globes offer a dynamic vision. Retrieved August 10, 2016, from http://www.nytimes.com/2013/01/08/science/digital-globes-a-new-way-to-view-the-world.html.

Wang, D.-Y., Lee, M.-H., & Sun, C.-T. (2013). Effects of thinking style and spatial ability on anchoring behavior in geographic information systems. *Educational Technology and Society, 16*(3), 1–13.

Wu, H., He, Z., & Gong, J. (2010). A virtual globe-based 3D visualization and interactive framework for public participation in urban planning processes. *Computers, Environment and Urban Systems, 34*(4), 291–298. https://doi.org/10.1016/j.compenvurbsys.2009.12.001.

Xie, Y., & Reider, D. (2014). Integration of innovative technologies for enhancing students' motivation for science learning and career. *Journal of Science Education and Technology, 23*(3), 370–380.

Zhu, L., Guo, H., & Wang, C. (2008). Constructing the Digital Earth system. In *Paper presented at the image and signal processing, 2008. CISP'08*.

Zhu, L., Sun, J., Li, C., & Zhang, B. (2014). SolidEarth: a new Digital Earth system for the modeling and visualization of the whole Earth space. *Frontiers of Earth Science, 8*(4), 524–539. https://doi.org/10.1007/s11707-014-0438-7.

Zottmann, J. M., Goeze, A., Frank, C., Zentner, U., Fischer, F., & Schrader, J. (2012). Fostering the analytical competency of pre-service teachers in a computer-supported case-based learning environment: A matter of perspective? *Interactive Learning Environments, 20*(6), 513–532.

Zydney, J. M. (2010). The effect of multiple scaffolding tools on students' understanding, consideration of different perspectives, and misconceptions of a complex problem. *Computers and Education, 54*(2), 360–370. https://doi.org/10.1016/j.compedu.2009.08.017.

Zydney, J. M., & Grincewicz, A. (2011). The use of video cases in a multimedia learning environment for facilitating high school students' inquiry into a problem from varying perspectives. *Journal of Science Education and Technology, 20*(6), 715–728.

Wei-Kai Liou is currently working as a Post-doctoral Researcher in Science Education Center at National Taiwan Normal University. He received his Ph.D. in Educational Technology from National Taiwan University of Science and Technology. His research area of interest includes augmented reality, virtual reality, mixed reality, flipped classroom, and HCI (human–computer interaction).

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.

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). Professor Chang's major research interests include science education, e-Learning, interdisciplinary science learning and science communication.