A Role of Augmented Reality in Educational Contents: Intermediating Between Reality and Virtual Reality

Shohei Tsuchida and Shu Matsuura[(✉)]

Faculty of Education, Tokyo Gakugei University,
4-1-1 Nukuikita, Koganei, Tokyo 184-8501, Japan
shumats0@gmail.com

Abstract. Many science education materials exhibit simplified models of nature. This simplification is beneficial to represent the essential characteristics of nature, but it forces the learners to cognitively assign the model to reality. This paper describes a traditional content exhibiting the lunar phase, which is taught in the elementary school. The use of multiple frames of observation and the necessity of dual concept required for perceiving the lunar orbital motion and Earth's rotational motion around its axis creates an extraneous cognitive load. This study introduces a simple desktop model of the planets and a multi-view display of the model planets using augmented reality (AR) and virtual reality (VR). Eye-tracking experiments are performed to examine the role of AR to intermediate between the spatial arrangement of real objects and the VR display observed from a fixed position on an object that represents the Earth. The results indicate that participants who experimented with the desktop model took more time to check and move their eye gaze between AR, VR, and the real model, in the beginning phase of the trials. Therefore, it is suggested that AR intermediates cognition of a view outside of the orbit and a surface of the Earth.

Keywords: Cognitive load · Learning · Desktop model

1 Introduction

Traditional learning materials of science education are usually based on simplified models that illustrate the essential character of nature or the essence of theories. Although this simplification or idealization enables clear scientific understanding of the complex natural phenomena, 2D illustrations on paper sometimes generate a high cognitive load for less-experienced learners.

In this study, the description of the lunar phases is used as an example to investigate the effect of augmented reality (AR) on learners' cognition load. Elementary school students learn the basic principles of planetary movements theoretically and through observations, which are often conducted at home. Although students are capable of changing their understanding of the lunar phases in a scientifically accurate way, they might continue to hold views inconsistent with the scientific perspective [1, 2]. Learners have to relate the lunar phases to the orbital motion of the moon. Here, multi-frame of observation brings about extraneous cognitive loads for the learners [3].

One frame of observation is the view from outside the Earth's orbit, where the sun looks still (universe view). The other frame of observation is the view from a fixed position (usually the location of the learner's country) on the surface of the Earth (earth view). The Shape of the moon that is observed from the earth view is determined by the relative positions of the sun, the Earth, and the moon. Although this is not taught at elementary schools, the angle between the Earth's axis and the orbital plane also has to be considered.

To tackle the cognitive load produced by using multiple frames, constructivistic activities using balls and a light are often conducted in classrooms [4–7]. In addition, digital learning materials provide a synchronized animation of the planet motions on the display, and the multi-view display enables learners to observe from both the universe and earth views. A virtual observation using AR proves effective for junior high school students [8].

In this study, an AR system is used, which includes a tabletop model that represents the sun, the Earth, and the moon displayed in both the universe view (AR) and the earth view (virtual reality, VR). It is hypothesized that the AR display plays a role in supporting the learner to relate the arrangement of the real tabletop blocks with the moon phases exhibited in the VR display. We discuss the gaze time on AR display, VR display, and real objects in a simple problem-solving process using eye-tracking equipment.

2   Method

2.1   Tabletop Toy-Model of the Sun, the Earth, and the Moon

Two of 50 mm × 50 mm square blocks with a thickness of 10 mm representing the Earth and the moon, on a worksheet with a sun marker, were used to show the eight location of the moon around the Earth, which result in the different phases of the moon (Fig. 1). The AR application was created using Vuforia (Qualcomm) AR package on the development environment Unity v.5.3.2.f1 (Unity Cooperation). The Vuforia
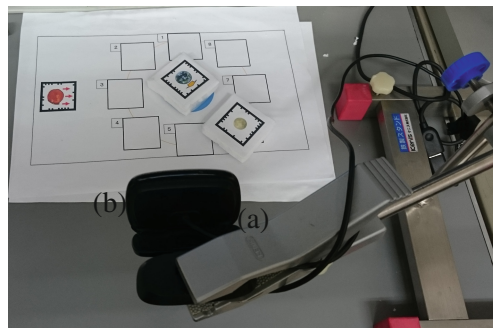

Fig. 1. The worksheet and blocks with the planet markers. (a) the blocks with the earth and the moon markers; (b) the worksheet with a sun marker; (c) the web camera to capture the scene on the worksheet. The view is of the participant.

"frame markers" (35 mm × 35 mm) for AR detection were pasted on the block surface. The size of the markers was set to be smaller than the block size to avoid occlusion by the participants' fingers. The worksheet where the blocks were placed, was the size A4. A web camera (Logicool C920R) captured the images of the markers. The participants watched the AR and VR view displayed on a 27' monitor (Apple Thunderbolt Display) set at around 70–100 cm from the face of the participant.

2.2   The Moon Phase Question

The participants of the experiments were students at a teacher's training university aged from 20 to 22. Nine students participated in the sessions. They were majoring in science education for elementary school. Thus, they have experience in establishing understanding of the moon phases.

The participants were asked to arrange the blocks in order to reproduce an indicated moon phase. The moon phases requested were the waxing crescent moon in the east, the waning crescent moon in the west, or the waxing gibbous moon in the west.
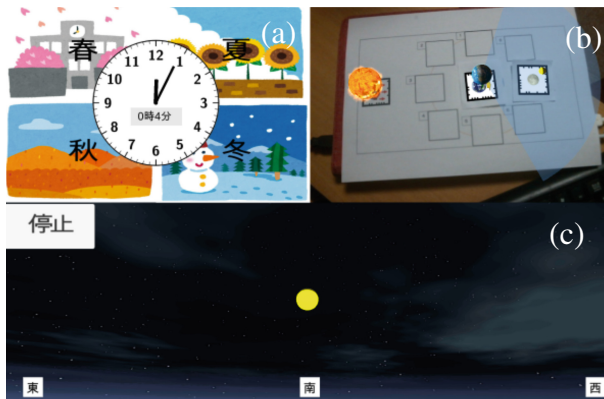
Fig. 2. A screen shot of the display. (a) the time corresponding to the observations in (b) and (c). (b) AR (RB mode) display. A gray blue zone in the right side of the earth image shows the visible area from the view position. VR display from the view position on the earth (Color figure online).

The display consists of three windows, i.e., AR, VR, and the time presentation as shown in Fig. 2. The AR window displays the graphics of the sun, the Earth, and the moon from the universe view superimposed on the block markers. Here, we examined two types of representations. First, the video image of the environment was exhibited in the "real background video (RB)" mode. The AR display consists of two types of images, namely, the image of a real scene captured by the web camera, and the AR graphics superimposed on the real scene. The RB consists of these components. Second, the environment image was not projected in the "blank background (BB)" mode, because the BB mode lacks a visual cue for three-dimensional cognition. The RB mode provides three-dimensional perspective in the real scene, together with

the planet configurations in the universe view. This mode augments the real model, and it is expected to stimulate a three-dimensional sensation.

The VR window shows the moon lit by the sun from the earth view. The corresponding viewing position was predetermined on the marker of the earth block. The right side of the window corresponds to the east, and the left side indicates the west. The shape of the moon and its position correspond to the arranged planet model. Finally, the time presentation shows a clock to indicate the time corresponding to the arrangement of the model blocks.

The participants were asked to sit in front of the display to solve the problem by manipulating the blocks. The eye-tracking measurement was performed using "ViewTracker" (DITECT Corporation), a head-mounted-type eye tracker. Eye tracking continued for fewer than 100 s, and the participants solved the problems within this period.
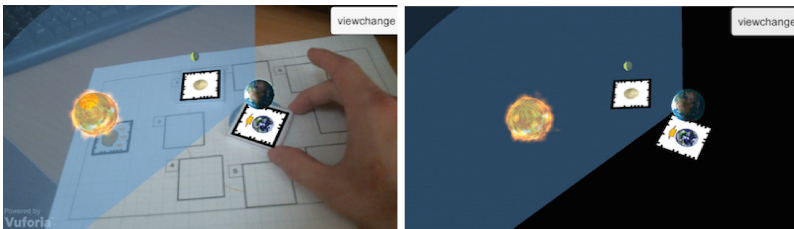
Fig. 3. Comparison of the RB (left) and BB (right) modes of the equivalent scene.

Each participant engaged in two successive sessions with different problems. In these sessions, the types of AR windows, RB, and BB, were changed. Participants had another set of successive sessions, with the order of RB-BB converted, on another day. This conversion was done to avoid a memory effect, which would result in the participants' solving the new problem too fast. After each session, the participants were interviewed on how they solved the problem.

3   Results and Discussion

3.1   Types of Problem Solving

Approach Types. The solution to the questions involves two phenomena, namely the lunar orbital motion (LOM) and the Earth's rotational motion around its axis (ERM). Based on the interview, the participants' approaches to this problem are grouped into three types. In the first type, the participant tries to solve the lunar orbital motion (LOM), namely the phase of the moon. In the second approach, the participant attempts to solve the Earth's rotational motion (ERM), namely the direction of the moon seen from the Earth. The explanations of the third type were vague in solving ERM and LOM separately.

Behavior Types. In addition, there are two types of behaviors. Students of the first type immediately manipulate the model, experimenting until he/she finds the solution (experiment-oriented type). Another type of student tries to solve the problem conceptually, reaches a satisfactory answer, which he/she tests by manipulating the model (think-first type).

Table 1. Number of cases for three approach and two behavior types. Each cell is divided into two orders of real background (RB) and blank background (BB) conditions.

	ERM			Vague			LOM		
Experiment oriented	8	RB → BB	4	0	RB → BB	0	9	RB → BB	4
		BB → RB	4		BB → RB	0		BB → RB	5
Think first	7	RB → BB	4	3	RB → BB	2	3	RB → BB	2
		BB → RB	3		BB → RB	1		BB → RB	1
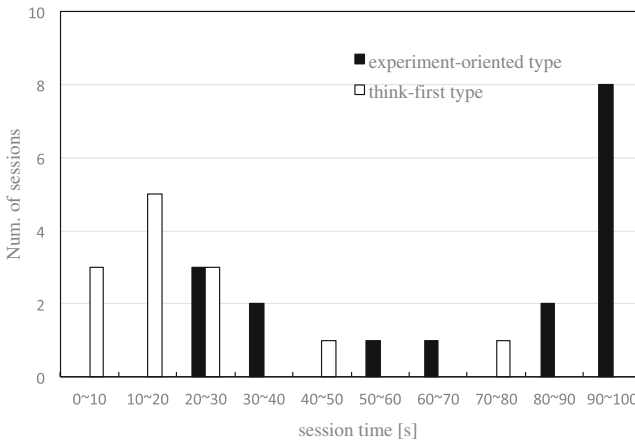
Fig. 4. Histogram of session times for experiment-oriented and think-first type participants

The students' sessions are grouped according to the types described above. Table 1 shows the number of sessions of these six groups, and further, the table shows the number of sessions under two types of AR viewing. Accordingly, the ERM and LOM numbers are almost equivalent for experiment-oriented students. However, ERM is more frequent than LOM in think-first type students. Since ERM has only two possible choices of east or west, the think-first type participants might feel ERM consideration is an easier task to begin with and therefore chose to perform this one first.

Figure 4 shows a comparison of the session times for experiment-oriented and think-first type participants. The figure shows that think-first type participants finish earlier, while many experiment-oriented participants require a longer session time. This result suggests that experiment-oriented participants make full use of the visualization to solve the problem.

3.2 Gaze Time on the Visual Components

The times of the participants' sessions were divided into dwell time on the AR display component (AR-dwell), the VR component (DR-dwell), and the real model blocks (real-dwell), and motion time gazing point from one component to the other (motion).

Comparison Between the Real Background (RB) and the Blank Background (BB). As seen in Fig. 3, the background scene display seems effective for the spatial cognition of planet CG. Here, we compare the total dwell times on the AR display with and without the environment image in Fig. 5. The figure includes the sessions of both experiment-oriented and think-first types. The dwell time on AR with RB exhibits a peek within a small time-range of 10–20 s. On the other hand, the dwell time on BB exhibits a high peak at a long time-range of 90–100 s. This result implies that the AR display without a background video requires a larger load for spatial cognition as compared with AR with the real scene.
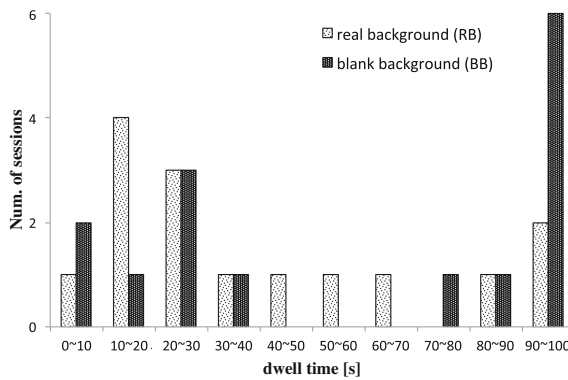

Fig. 5. Histogram of total dwell time of each session on AR display with the real and the blank background (RB and BB, respectively).

We now discuss the effect of the change of RB and BB in the successive two sessions, i.e., we compare the change from RB to BB, and the change from BB to RB. Let us denote the total dwell time on the AR display in the first session as T_{AR1}, and the time in the second session as T_{AR2}. Then, the difference in dwell time between the two sessions is $\Delta T = T_{AR1} - T_{AR2}$.

Figure 6 shows the comparison of the time difference ΔT for successive RB-to-BB and BB-to-RB sessions. The BB-to-RB sessions show a positive ΔT, decreasing AR viewing. This implies that less time is required for cognition using the RB type AR, as compared with the BB type, in addition to an effect of practice. On the other hand, the RB-to-BB sessions show positive and negative ΔT. Negative or small ΔT implies that the cognition of BB requires more gaze, regardless of the effect of practice.

Comparison of Dwell Times in Three Time Zones. To characterize the time change in the display and object model, each session is divided into three time zones. The first one-third of a session is called "time zone 1," the middle one-third is "time zone 2," and the final period is referred to as "time zone 3."
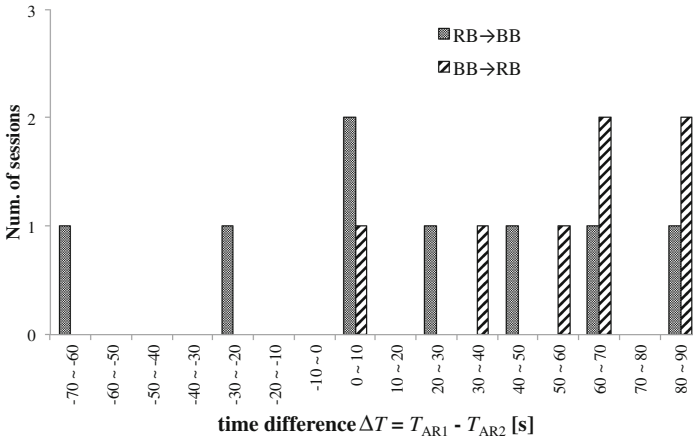
Fig. 6. Histogram of the difference of dwell time on the AR view between two successive two sessions. In successive sessions, the real background image is switched on from BB to RB, or switched off, from RB to BB.

Figure 7 shows the histograms of dwell times in the real object model, AR and VR display in the three time zones. Gazes for around 0.2 s are most frequent on the real objects. Although gazes on AR and VR are less frequent than on real objects in a small time-range, $\sim$2-s gazes are more frequent than gazes on real objects. Gazes on the AR for more than 2 s are particularly frequent in time zone 2. Moreover, gazes on the VR for more than 1 s are observed in time zone 3. This suggests an intermediating role of AR between real objects and the VR image from a view of the object.

Furthermore, gazes on the AR area often concentrate on the image of the participant's finger image taken from the camera, as seen in Fig. 8. Seeing their fingers implies that they try to collect visual feedback that will supplement muscular sensory feedback. The participants rotate the block, watching the AR display, and getting a visual feedback from the image of their hands. Then, from time to time, they confirm the resultant moon phase from the earth view represented in the VR display. In this way, the AR display intermediates between the universe view and earth views through the visual feedback of their manipulation.

The aim of the original question is to get a correct VR image by arranging the blocks. Thus, it is natural to begin with seeing and manipulating the blocks and to certify the result in the VR window at the end. AR in this study provides pictorial cues to help understand both the moon phase lit by the sun and the earth view at short notice.

The percentages of dwell time in the cells were shown in the top middle of the cells. Participant's fingers are inside the center and center bottom cells, which show high percentages.
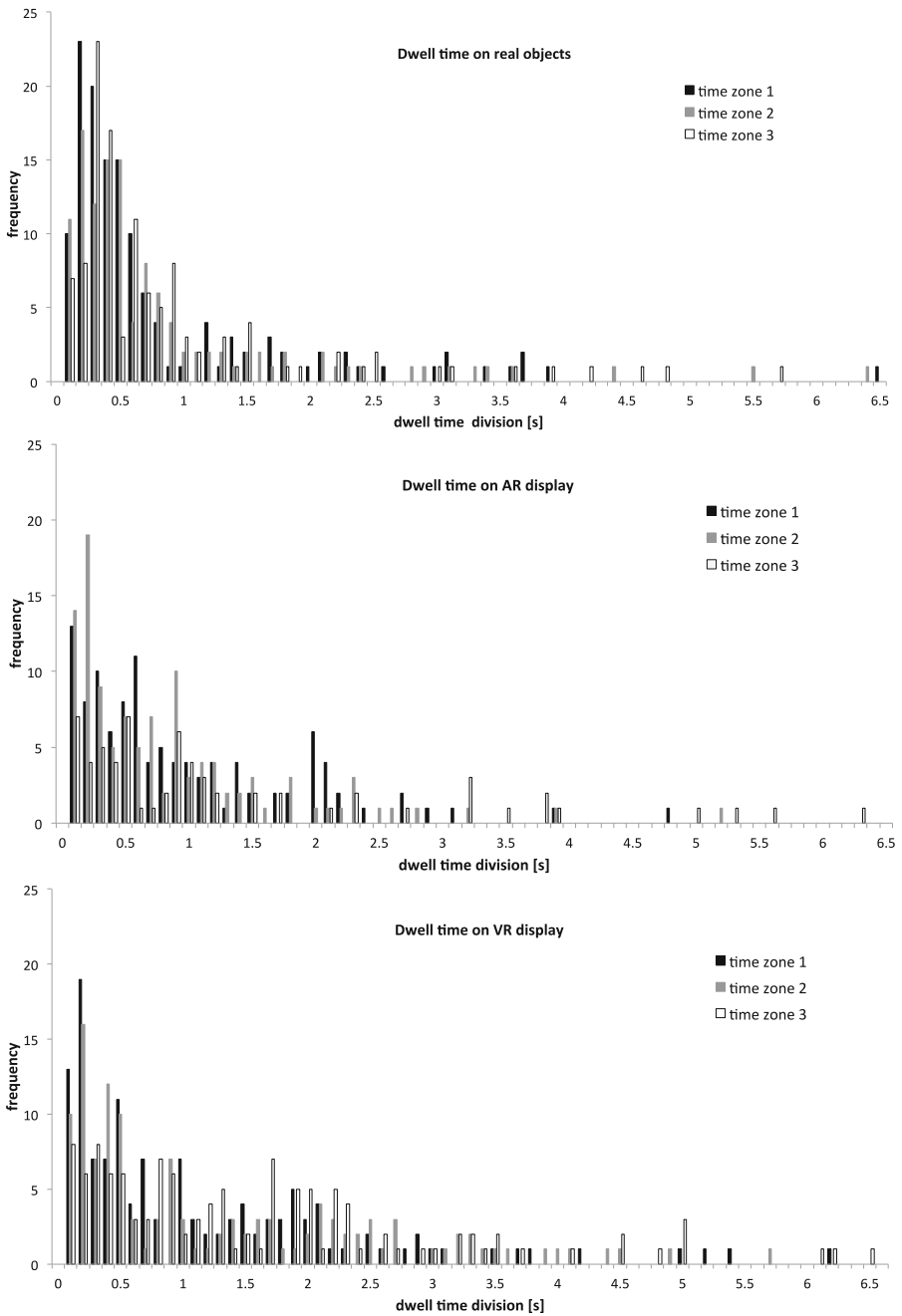
Fig. 7. Histogram of the dwell time on real objects (top), and AR (middle) and VR (bottom) displays for three successive time zones of the sessions.
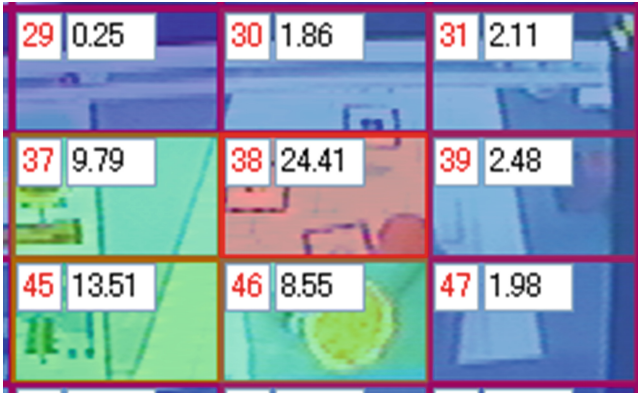
Fig. 8. AR display area divided into 9 cells with the rate of dwell.

4   Conclusions

An AR learning material with a real object model was created to investigate the effect of AR on problem solving. Differences appeared between the experiment-oriented and think-first type of approach, and the AR and VR were more utilized by the former type of participants. In AR, it was suggested that spatial cognition was achieved in shorter times by superimposing 3D graphics on the image of the 3D environmental scene. A temporal change was found in the frequency of viewing real objects, AR, and VR. It was suggested that observing the AR image intermediates the manipulation of real objects from the universe view and the VR image from the earth view.

These results indicate that the combination of real objects, AR images, and VR images is potentially beneficial for experimental usage in a collaborative, educational environment, if it activates the discussion on intermediating different views or concepts.

References

1. Stahly, L.L., Krockover, G.H., Shepardson, D.P.: Third grade students' ideas about the lunar phases. J. Res. Sci. Teach. **36**(2), 159–177 (1999)
2. Bailey, J.M., Slater, T.F.: A review of astronomy education research. Astron. Edu. Rev. **2**(2), 20–45 (2003)
3. Sweller, J.: Cognitive Science **12**(2), 257–288 (1988)
4. Suzuki, M.: Conversations about the moon with prospective teachers in Japan. Sci. Educ. **87**(6), 892–910 (2003)

5. Foster, G.W.: Look to the moon, students learn about the phases of the moon from an "Earth-centered" viewpoint. Sci. Child. **34**(3), 30–33 (1996)
6. Abell, S., George, M., Martini, M.: The moon investigation: instructional strategies for elementary science methods. J. Sci. Teach. Educ. **13**(2), 85–100 (2002)
7. Kavanagh, C., Agan, L., Sneider, C.: Learning about phases of the moon and eclipses: a guide for teachers and curriculum developers. Astron. Educ. Rev. **4**(1), 19–52 (2005)
8. Tian, K., Endo, M., Urata, M., Mouri, K., Yasuda, T.: M-VSARL system in secondary school science education: lunar phase class case study. In: IEEE 3rd Global Conference on Consumer Electronics, pp. 321–322 (2014)