3D learning spaces and activities fostering users' learning, acceptance, and creativity

Donghee Shin[1] · Seyoung Park[2]

Abstract

As an instructional material, 3D instructions afford people to learn procedural-manipulative tasks. Observing and emulating motions presented in 3D animations is important in learning contexts. This study examined the effects of visual cueing in an effort to identify the optimal way to present information in a 3D virtual environment. While previous studies have found that animations are more effective than static images for learning procedural-manipulative tasks, the transient nature of dynamic visualizations might create an unnecessary cognitive load on learners. To compensate strategically for the lack of permanency, this study suggests combining dynamic and static visualizations into one medium by adding transparent static images (visual cueing) to an animation. A between-subjects experiment was conducted to examine the effects of visual cueing on cognitive load and learning outcomes. The study found that the hybrid of dynamic and static visualizations was beneficial for reducing cognitive load, although it did not improve learning outcomes. The results suggest a design strategy for improving the effectiveness of 3D instructional animations. The findings are broadly applicable to numerous learning contexts, such as virtual reality storytelling, augmented reality games, and diverse gamification services.

Keywords Visual cueing · 3D animation · Procedural-manipulative task · Animation versus static images · Cognitive load theory · Multimedia learning

✉ Donghee Shin
  Donghee.Shin@zu.ac.ae

  Seyoung Park
  sypark@skku.edu

[1] College of Communication and Media Sciences, Zayed University, Kalifa City-B, PO Box 144534 Abu Dhabi, United Arab Emirates

[2] School of Communication, Sungkyunkwan University, Seoul, Korea

Introduction

Learning environments are becoming more diverse as a response to the development of multimedia technologies (Shin 2017). In particular, online distance learning (e.g., the Massive Open Online Course) has become increasingly common, and instructional animation is used steadily and increasingly as an efficient and effective learning tool, regardless of discipline. In addition, teaching tools and technologies have evolved gradually, and dynamic types of instructional materials have emerged (Hwang and Shin 2018).

The 3D virtual environment offers exciting possibilities for application in many educational fields. 3D instructional media are useful when realistic and sophisticated visualizations are needed in, for example, the areas of human anatomy, mechanics, or architecture (Shin 2012). 3D instructional tools demonstrate their highest value in teaching and learning procedural-manipulative tasks that carry a high risk of accidents (e.g., brain surgery). The probability of failure can be minimized by observing, imitating, and repeating a manipulation procedure in a virtual environment similar to the real world. Based on the progress in this area thus far, the question about the most effective way to present information for learning procedural-manipulative tasks in a 3D virtual environment has arisen.

As educational multimedia tools have developed, the contrast in the use of animated or static images is becoming a highly controversial topic. Numerous studies have found that dynamic visualizations are superior to static ones in terms of format, but previous research has produced mixed results (see Höffler and Leutner 2007). Even so, with respect to learning procedural-manipulative tasks (e.g., assembling a machine or using a tool), many previous studies have found that animations are more effective than static images (Ayres and Paas 2007; Van Gog et al. 2009; Wong et al. 2009; Paas and Sweller 2012; Garland and Sanchez 2013; Castro-Alonso et al. 2015). This is because animations reduce the extraneous cognitive load by activating the mirror neuron system and providing mental representations of motions.

However, instructional animations could be limited in terms of their transient information effects, which means that the lack of permanency in dynamic media leads to generating additional cognitive loads (Hwang and Shin 2018). Instructional animations can also be too complex or fast for learners to grasp the contents (e.g., the apprehension principle; Tversky et al. 2002). The animations' intrinsic limitations generate an unessential cognitive load, which disturbs the learning of complex and advanced skills. Embedding static images in instructional animations could alleviate the cognitive load by compensating for the lack of permanency in the dynamic visualizations.

This study examines the effects of adding static images strategically (i.e., visual cueing) to increase the effectiveness of instructional animations. It compares animation with static images, and suggests strategies for combining the advantages of dynamic and static visualizations into one medium. The users in this study are required to learn a Japanese crystal 3D puzzle for the experiment, testing whether adding transparent images to an animation improves user

performance in the 3D virtual environment. Based on cognitive load theory and multimedia learning principles, this study aims to develop an optimal design for instruction in procedural-manipulative tasks in a 3D virtual environment.

Literature review

User and cognitive load

Cognitive load theory regards the manner in which cognitive resources are focused and used during learning (Chandler and Sweller 1991). Cognitive load refers to the amount of mental effort required to process information in the working memory (Mayer and Moreno 2003). Because the human cognitive processing capacity (working memory) required for schema acquisition is limited, it is crucial to manage the cognitive load carefully through the presentation of information. The theory is useful for explaining ways to present information to accomplish the learning of a problem-solving process and for appropriate instructional designs. Although changing the instructional format can control germane cognitive load, extraneous cognitive load differs in people in the way it contributes to the formation of schemas while disturbing information processing. In other words, when designing instructional materials, considering the ways to decrease extraneous and increase germane cognitive load is important (Hwang and Shin 2018).

Procedure, manipulation, and tasks

In the term "procedural-manipulative tasks," the word "procedural" implies that the task requires a procedure (process) as opposed to declarative knowledge. Procedural knowledge concerns "knowing how" to perform a certain action, whereas declarative knowledge relates to "knowing that" something is (Ryle 2009). Procedural knowledge is proposed to "underlie the performance of actions, particularly skilled actions, like riding a bicycle or playing the piano, and cognitive (as well as motor) actions, like language production or skilled mathematical problem solving" (Berry and Dienes 1993, p. 153). Another characteristic of a procedural task is its certain goal to be accomplished and its series of steps to achieve such a goal. Moreover, the order of steps is important because, if disordered, execution or comprehension of the overall procedure could be disabled (Arguel and Jamet 2009).

The word "manipulative" is used in this study to refer to tasks requiring certain actions performed on objects (i.e., object manipulation) using the human body (mostly the hands) to accomplish a specific goal, such as assembling a machine or using a tool. In many cases, to perform these types of tasks successfully, learning specific procedural knowledge is required. Therefore, the term "procedural-manipulative task" is used in this study, as opposed to the term "manipulative task." Moreover, previous studies have used related terms, such as "hand-manipulative task" (Ayres et al. 2009), "procedural-motor knowledge" (Höffler and Leutner 2007), and

"human motor skill" (Van Gog et al. 2009; Wong et al. 2009), because most manipulative tasks involve a specific sequence of motions and actions.

In sum, the term "procedural" emphasizes that a series of actions is needed to accomplish a particular task, and "manipulative" stresses that a task requires the control or handling of an object using the human body. Therefore, in this study, "procedural-manipulative task" is defined as going through a sequence of motion steps to complete a purposed object manipulation.

## Combining animations with static images using transparent ones

The compelling question concerns the ways through which one can present an animation and static images together in one medium. Few studies have attempted to combine animation with static images. However, among them, Arguel and Jamet (2009) seemed to be the only one, thus far, that examined the effect of adding static images to an animation. They found that for learning procedural contents (i.e., how to give first aid), a combination of video and static pictures produced better learning outcomes than either of the two produced individually. Arguel and Jamet (2009) observed that the number of static pictures, which were presented using video, is an important moderating factor, because showing fewer pictures was more beneficial. This observation might be a result of the relationship between showing too many pictures and increased extraneous cognitive load, which is necessary to the integration of the pictures with the video.

Although Arguel and Jamet (2009) found a positive effect of combined video and static pictures on learning, the technique of presenting a series of static pictures (three or four snapshots of important steps from the video) beneath the video, as shown in Fig. 1, might not be effective for learning procedural-manipulative tasks. A major problem is the spatial split-attention effect that is a result of presenting multiple objects in a medium (Ayres and Sweller 2005). To solve this problem, this study

Fig. 1 Example of adding static images to video (Arguel and Jamet 2009)

Fig. 2 Example of adding static
images to animation
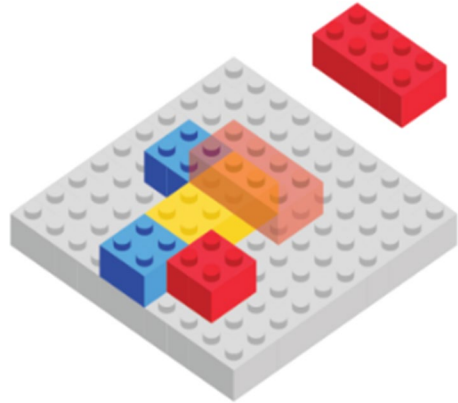
suggests overlaying transparent images on an animation (Fig. 2); such approach
would certainly have broad applications across learning contexts. However, for
the purposes of this study, it is considered useful for learning various manipula-
tion tasks. Although this alternative approach cannot present all the manipulation
phases simultaneously, additional information might not be necessary because the
outcomes of the previous phases are cumulated into the present object (e.g., imagine
the Lego manipulation shown in Fig. 2).

Using transparent images has, among others, advantages regarding learning pro-
cedural-manipulative tasks. The most critical advantage is that this approach can
present the goal and the motions of each step simultaneously. Static visualization
requires integrating a succession of images and imagining the necessary motions
from them, an approach that is likely to cause a spatial split-attention effect (see
Ayres and Sweller 2005). Dynamic visualization could also be demanding because it
requires the learner to comprehend, remember, and integrate simultaneously a series
of informational elements, a process that can cause a temporal split-attention effect
(see Ayres and Sweller 2005). The effect could be aggravated when the observer
(learner) has been unaware of why the motion is needed (goal of the motion) until
the motion is complete. Adding transparent images to animation can compensate for
the effects of spatial and temporal split-attention by presenting the motion and the
goal almost simultaneously (Hwang and Shin 2018).

Rather than presenting static images outside the animation (Fig. 1), adding
transparent images to animation can focus learners' attention on one figure in one
medium; this approach corresponds to the spatial contiguity principle (Ayres and
Sweller 2005). This type of visualization is clearly more appropriate to virtual or
augmented reality than an actual physical environment.

It is important to recognize that the addition of transparent images must be distin-
guished from other types of visual cues, such as using arrows or accent colors. Most
visual cues in animations have merely an attentional function to guide learners to the
essential information, and these cues do not function to highlight the relationships
between or among learning elements (De Koning et al. 2009). However, compared
with types of visual cues that emphasize only the key elements, transparent images

could deliver additional precision and informational elements about the actions. In other words, adding transparent images as visual cues could spotlight the relationships between or among sequential motions, as well as function as an attentional cueing.

Based on the discussion, the following research question guides this study.

RQ: Does adding transparent images to an animation improve learning material designs in 3D virtual environment?

With this overarching RQ in place, the following three hypotheses are tested.

When learning a procedural-manipulative task for the first time by watching and imitating an instructional animation,

Hypothesis 1 Cognitive load is lower for participants exposed to animation with visual cues than for those exposed to animation without visual cues.

When learning a procedural-manipulative task for the second time,

Hypothesis 2a Participants exposed to animation with visual cues spend a shorter time on the task than those exposed to animation without visual cues.

Hypothesis 2b Participants exposed to animation with visual cues have a lower cognitive load than those exposed to animation without visual cues.

When recalling what they learned from the instructional animation,

Hypothesis 3 Participants exposed to animation with visual cues demonstrate better learning performance than those exposed to animation without visual cues.

Methods

To examine the effects of visual cueing on cognitive load (Hypotheses 1 and 2) and learning performance (Hypothesis 3), a between-subjects experiment was conducted. The experimental treatment (independent variable) was the presence of visual cues in the instructional animation. There were two conditions: (1) animation with visual cues and (2) animation without visual cues (see Fig. 3a, b). The dependent variables were: (1) participants' cognitive loads while watching the instructional animation, measured by (a) the amount of time it takes for them to imitate the manipulation and (b) self-report score; and (2) learning performance, measured by (a) the amount of time it takes for them to accomplish the recall task and (b) completion rate.
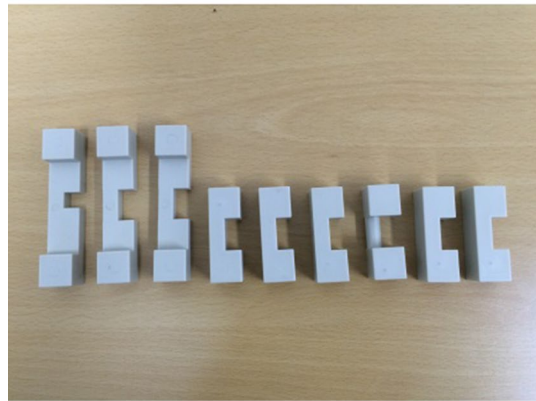
Participants

The experiment was conducted over 2 days. The sample comprised 32 undergraduate students (15 male and 17 female) at a comprehensive university in Seoul, Korea.

Fig. 3 a Assembled Japanese
crystal 3D puzzle. **b** Disassem-
bled Japanese crystal 3D puzzle
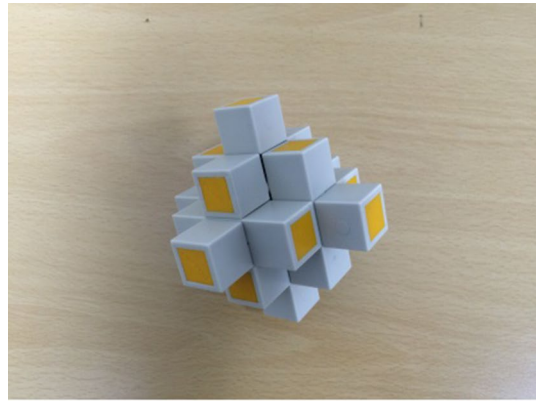
(a)

(b)

The participants were recruited via an online bulletin board. The participants were
assigned randomly to one of two groups of 16 subjects. The first group was the con-
trol group and the second group was the experimental group. The average age of the
participants was 22.43.

Research design

Task material: Japanese crystal 3D puzzle

The task material, "Japanese crystal 3D puzzle," was chosen to meet the definition
of a procedural-manipulative task, which is "going through a sequence of motion
steps to complete a purposed object manipulation" (see Sect. 2.2) (Fig. 3a, b). This
3D puzzle is a type of "burr puzzle" that requires the user to combine notched blocks
into an interlocking unit that ultimately resembles a crystal. The units are composed
of about 30 units with diverse colors. Each unit is about 5–7 cm-long. The puzzles
are made of plastics and are precision-made for easy sliding and accurate fitting of

the pieces. This is an appropriate example of a procedural-manipulative task because it requires the learner to understand and remember a sequence of certain motions to assemble the parts into a unit. Instead of the many other possible object manipulative tasks, such as origami models (Wong et al. 2009) or Lego blocks (Castro-Alonso et al. 2015), the 3D puzzle was adopted as this study's task material because it was more likely to be unfamiliar to the participants. In fact, in the screening process, all participants responded that they had no experience with that type of puzzle.

Thus, it attempted to control the potential bias possibly related to prior experience.

When the participants were given the puzzle, the blocks were disassembled and arranged in the order of the assembly steps, as shown in Fig. 3b. The image of completely assembled puzzle blocks (Fig. 3a) was presented in the introductory part of the instructional animation.

Stimulus material: instructional animation

A short instructional animation was the experimental stimulus material. The video was full-screen size (1920*1080) and lasted for about 2 min. The exact duration of the instructional video was 120 s. In the introductory part of the video, a brief description of the imitating task was provided that explained how to arrange the nine disassembled blocks in their order of assembly steps (Fig. 4b), which needed to be performed to create the unit shown in Fig. 4a.

After the introduction, the animation demonstrated how to assemble the blocks. The demonstration was presented in the 3D virtual environment. Previous studies have found no significant differences between the physical and virtual environments in learning performance and outcome (Castro-Alonso et al. 2015). This study establishes two learning phases to see if there are any different results. In this study's animation, the image of hands manipulating the blocks was not shown because that image has not be found to help learning (Castro-Alonso et al. 2015), and the goal of the observed action is more important than the presence of virtual hands (Van Gog et al. 2009).

Two different versions of assembly instructions were provided, one with visual cues and another without visual cues. In the animation without visual cues, the blocks were assembled step-by-step without any added visual cues (Fig. 5a). In the animation with visual cues, a transparent static image was present, which referred to the goal of each step (Fig. 5b).

Design: learning performance (time on task and completion rate)

The testing phase was a recall task that measured how much the participants remembered about the assembly procedure. Learning performance was evaluated by two variables: (1) time required to accomplish the recall task and (2) the completion rate. Immediately after each participant completed the two learning phases (and the questionnaires), the experimenter placed the disassembled puzzle blocks on the desk, directed each participant to start the task, and began to video record the participant. By analyzing the video-recorded data, the researcher evaluated the amount of time
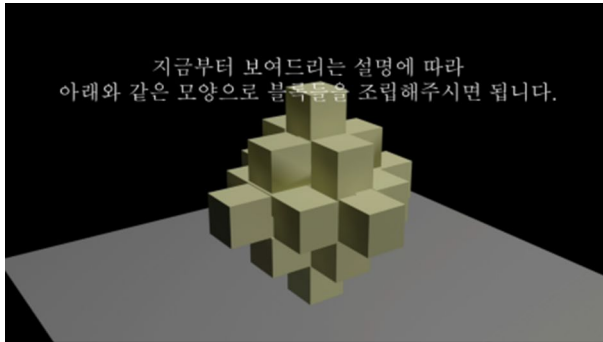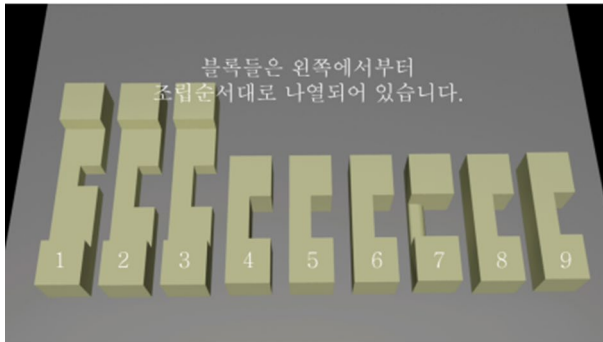
Fig. 4 **a** Assembled Japanese crystal 3D puzzle in the instructional animation. **b** Disassembled Japanese crystal 3D puzzle in the instructional animation

it took to accomplish the task and the completion rate of each participant. Because there was a time limit of 5 min, the completion times ranged from 0 to 300 s, and the completion rates ranged from .00 (0 of 12 steps = completely failed) to 1.00 (12 of 12 steps = succeeded perfectly).

Measurement

Cognitive load (time spent on task and self-reported score)

There are numerous ways to measure cognitive load (e.g., dual-task method, self-reporting method, and brain activity measures), and the best way to measure it accurately or effectively remains controversial (see Brunken et al. 2003). Per the Paas et al. suggestion (2003), this study used two methods to measure cognitive load subjectively and objectively, and directly and indirectly. Cognitive load while watching the instructional animation was measured by the amount of time a participant took to imitate the manipulation (or the time it took a participant

(a)

(b)
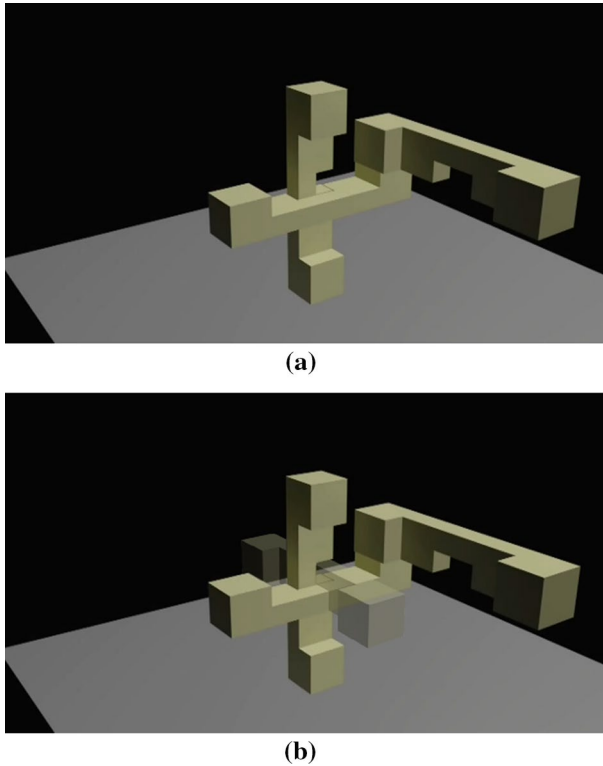
Fig. 5 **a** Animation without a visual cue. **b** Animation with a visual cue

to follow and complete the whole manipulation procedure), and it was measured after the imitation manipulation by the participant's self-report score.

As an indirect and objective measurement of cognitive load, the amount of time required to imitate the assembly actions completely was assessed (Brunken et al. 2003). While it has various meanings, the amount of time required for certain tasks has been advocated and used widely as well (Paas et al. 2003). While watching the animation and imitating the actions, many of the participants had instances when they hesitated, two to three times on average in the first learning phase, and one time on average in the second learning phase. These hesitations were analyzed from the recorded videos. The hesitation behavior suggested indirectly that the participants were thinking of where and how to locate the blocks. As previous studies have shown (e.g., Shin and Chung 2017; Shin et al. 2016a) and in accordance with common sense, the more difficult the task was for a participant, the longer and more frequent were the hesitations. Therefore, the time required to complete the imitating task could indicate the extent of a participant's cognitive load during the task (Brunken et al. 2003). This point is concurred by Shin and Chung (2017) and Shin et al. (2016a, b). While time spent on a task can mean different things than learners' cognitive loads, time spent to complete the

task can be one aspect of measurement. As there are numerous studies using time spent on tasks to measure cognitive load, time spent on a task can be justified in this study.

Instruments

As a direct and subjective method for measuring cognitive load, a participant's self-report score was obtained from responses to questions on a paper-and-pencil questionnaire taken immediately after each learning phase. The items in the questionnaire were modified from the NASA-TLX (Task Load Index; Hart and Staveland 1988) and Paas' mental effort scale (2003). Developed by NASA, TLX is a widely used, subjective, multidimensional assessment tool that rates perceived workload in order to assess a task, system, or team's effectiveness or other aspects of performance.

The NASA-TLX was modified in its wording and in the nuances of its terminologies. Four instruments were used from the NASA-TLX, whereas two instruments were adopted from Paas' scale (2003). Combining the two instruments may not be the best way to examine these measures because of their inherent differences. Although the two were combined, they are not mixed up randomly. The NASA-TLX used base fundamental measures, whereas Paas's measurement (2003) used an adjunct tool.

The instrument comprised six items that employed seven-point Likert-type scales. The same questions were asked after the first and second learning phases. The reliability analysis of the items after the first and second learning phases were Cronbach's alpha of .694 and .743, respectively. Following are the measurement items used.

1. The activity covered formulas that I perceived as very complex.
2. The activity covered concepts and definitions that I perceived as very complex.
3. The instructions and/or explanations during the activity were very unclear.
4. The instructions and/or explanations were, in terms of learning, very ineffective.
5. The instructions and/or explanations were full of unclear language.
6. The activity really enhanced my understanding of the topic(s) covered.

Procedure

Each experimental session was divided into three phases: (1) first learning, (2) second learning, and (3) testing. In the first learning phase, the participants' cognitive loads while learning the task (through watching and imitating the instructional animation) was measured to test Hypothesis 1. In the next phase, the participants went through the same learning process again, and their cognitive loads while learning the task were measured to test Hypothesis 2. In the last phase, learning performances (the amount of time it took to complete the task and the recall rate) were measured to test Hypothesis 3.

In the first learning phase, the participants watched a short instructional animation describing how to solve the Japanese crystal 3D puzzle, and they imitated the object manipulation simultaneously. During the video presentation, the participants were permitted to pause or rewind the video whenever they wanted for whatever reason. This was designed intentionally to examine the cognitive load indirectly. That is, if the participants were to use pause or rewind often, we assume that their cognitive loads are higher than those who do not. To further measure the subjective cognitive load, the participants completed a paper-and-pencil questionnaire immediately after watching and imitating the assembly task. The procedure was identical in the second learning phase.

After the two learning phases, the participants entered the testing phase. They attempted the manipulative task of assembling the blocks without watching the videos. The experimenter notified them that there was a 5-min time limit for the assembly. While the participants assembled the puzzle, the experimenter video recorded their activities (of which the participants had been informed and to which they had consented).

A priori power analysis

A power analysis was conducted at the planning stage to determine the proper sample size that would enable accurate and reliable statistical judgments regarding the effects. For the power analysis, GPower and F test MANOVA were used: repeated measures and within-between interaction (two groups and measured for two times). Two kinds of power analysis were conducted: a priori and sensitivity. The result of a priori power analysis provided information on the required sample size for a given α (.05), power (.95), and effect size. This study expected a medium effect size based on the results of previous study, which was .61. The power analysis estimated that the required sample size was 35. Because around 40 people would participate in this study, we conducted a sensitivity power analysis, which predicts the smallest possible effect size this study may acquire for a given α (.05), power (.95), and sample size (35). The result was .38. This indicates that this study can detect a medium level effect size around .38 for the effect of the treatment. This required about three times more participants than the crossover design. The sensitivity analysis showed that the smallest possible effect size the study can detect when there were 80 participants would be .74, which was much less sensitive than what the crossover design could detect (.41).

Results

Design: learning performance (time on task and completion rate)

The variables of learning phases are measured and the results of learning performance are as follow (Table 1).

Table 1 Summary of the variables

Dependent variables	Condition	Mean	Standard deviation	Standard error of the mean
First learning phase				
Time spent on task	Control	194.000	99.195	24.799
	Experimental	178.688	76.940	19.235
Self-reported cognitive load	Control	4.083	.991	.248
	Experimental	3.990	.908	.227
Second learning phase				
Time spent on task	Control	130.500	35.948	8.987
	Experimental	107.313	25.804	6.451
Self-reported cognitive load	Control	3.750	1.058	.264
	Experimental	3.063	.875	.219
Testing phase				
Time spent on task	Control	159.813	93.136	23.284
	Experimental	165.063	100.353	25.088
Completion rate	Control	.729	.411	.103
	Experimental	.828	.302	.076

Dependent variables	$t(df = 30)$	p value	Mean difference	SE
First learning phase				
Time spent on task	.488	.629	15.313	31.384
Self-reported cognitive load	.279	.782	.094	.336
Second learning phase				
Time spent on task	2.096	.045	23.188	11.063
Self-reported cognitive load	2.003	.054	.688	.343
Testing phase				
Time spent on task	− .153	.879	− 5.250	34.228
Completion rate	− .774	.445	− .099	.128

Dependent variables	Condition	Mean	Standard deviation	Standard error of the mean
Procedure				
First learning phase				
Time spent on task	Control	194.000	99.195	24.799
	Experimental	178.688	76.940	19.235
Self-reported cognitive load	Control	4.083	.991	.248
	Experimental	3.990	.908	.227
Second learning phase				
Time spent on task	Control	130.500	35.948	8.987
	Experimental	107.313	25.804	6.451
Self-reported cognitive load	Control	3.750	1.058	.264
	Experimental	3.063	.875	.219

Table 1 (continued)

Dependent variables	Condition	Mean	Standard deviation	Standard error of the mean
Testing phase				
Time spent on task	Control	159.813	93.136	23.284
	Experimental	165.063	100.353	25.088
Completion rate	Control	.729	.411	.103
	Experimental	.828	.302	.076

Table 2 Summary of the experimental statistics ($n = 16$ per group)

Dependent variables	Condition	Mean	Standard deviation	Standard error of the mean
First learning phase				
Time spent on task	Control	194.000	99.195	24.799
	Experimental	178.688	76.940	19.235
Self-reported cognitive load	Control	4.083	.991	.248
	Experimental	3.990	.908	.227
Second learning phase				
Time spent on task	Control	130.500	35.948	8.987
	Experimental	107.313	25.804	6.451
Self-reported cognitive load	Control	3.750	1.058	.264
	Experimental	3.063	.875	.219
Testing phase				
Time spent on task	Control	159.813	93.136	23.284
	Experimental	165.063	100.353	25.088
Completion rate	Control	.729	.411	.103
	Experimental	.828	.302	.076

Control = Animation without visual cues; Experimental = Animation with visual cues

An independent samples *t* test was performed to compare the two groups' means of cognitive load (time to complete the task and self-report score) and learning performance (time to complete the task and completion rate). The descriptive statistics are presented in Table 2.

No significant effect of visual cueing on cognitive load during the first learning phase was found. There was no significant difference between two conditions regarding the time spent on the task (watching and imitating the animation) or in the self-reported cognitive load scores. Therefore, Hypothesis 1 was not supported.

However, there was a significant difference between two conditions in the second learning phase regarding the time spent on the task ($t(30) = 2.096$, $p = .05$ rounded from .045). The time spent on the task for the animation with visual cues ($M = 107.31$, $SD = 25.804$) was shorter, on average, than for the group with animation without visual cues ($M = 130.50$, $SD = 35.948$); this supports Hypthesis 2a. Similarly, the effect of visual cueing on self-reported cognitive load was marginally

Table 3 Summary of *t* test results (*n* = 16 per group)

Dependent variables	*t*(*df* = 30)	*p* value	Mean difference	SE
First learning phase				
Time spent on task	.488	.629	15.313	31.384
Self-reported cognitive load	.279	.782	.094	.336
Second learning phase				
Time spent on task	2.096	.045	23.188	11.063
Self-reported cognitive load	2.003	.054	.688	.343
Testing phase				
Time spent on task	− .153	.879	− 5.250	34.228
Completion rate	− .774	.445	− .099	.128

SE standard error of the mean difference

significant ($t(30) = 2.003$, $p = .054$). The self-reported cognitive load scores of the group with visual cues ($M = 3.063$, $SD = .875$) was lower, on average, than for the group without visual cues ($M = 3.750$, $SD = 1.058$); this supports Hypothesis 2b. Therefore, Hypothesis 2 was fully supported.

Last, the effects of visual cueing on learning performance were not statistically significant. The effect of visual cueing on the accuracy score was not significant, and no significant difference was found between the two conditions in the testing phase regarding the time it took to complete the task (assembling the blocks without instructions). Therefore, Hypothesis 3 was not supported.

The results of the *t* tests are shown in Table 3.

Discussion

This study investigated the effects of visual cueing using a between-subjects experimental design. Unlike the similar study by Arguel and Jamet (2009) that examined the effect of combining static images with animation, this study found no significant effects of visual cues on the participants' learning performance, on average. The results found no significant influence of visual cueing on cognitive load during the first learning phase or on recall performance during the testing phase. Adding static images to an animation was beneficial only for decreasing the cognitive load in the repeated (second) learning phase.

It is notable that the results were different in the first and second learning phases, although all of the procedures and materials were identical. One plausible explanation for this difference is that the presentation format of the animation with the visual cues might have been somewhat unfamiliar to the participants. The experiment was conducted without a tutorial phase, and there was no explanation of the visual cues (transparent static images) in the animation. Although most of the participants quickly understood the meaning and purpose of the visual cues, they might have needed additional time to familiarize themselves with the instructional material. However, when they had become accustomed to the procedure of the learning phase

(in the second phase), they could understand easily what the visual cues meant, which could account for the differences in effects between the first and second learning phases. This phenomenon is similar to carryover effects, since the participants learned the process through the first phase. That is, the first phase may inadvertently affect performance in the second one. This is similarly found in other studies of user research. For example, Shin and Chung (2017) found in their experiment on user game research that second phase game performance is improved from the initial phase testing.

A second reason for the different outcomes between the first and second phases is that the visual cues in the second phase could have functioned as recall cues. A number of the participants in the group with visual cues tended to begin working on the puzzle as soon as they saw the transparent image on the animation, before watching the full demonstration. They reproduced the motions needed to assemble the blocks using only the visual cues. Therefore, in the second phase, these participants might have been able to perform the imitating task more quickly and easily than those in the group without visual cues because of their experience in the first learning phase. This explanation is similarly found in previous studies such as Shin et al. (2016a, b). Previous findings have shown that visual cues play different roles in different contexts, such as recall, behavioral, and cognitive cues. Shin et al. (2016a, b) found that visual cues have functioned as engaging factors in single-handed interaction.

In addition, the difference in cognitive load was not related to learning performance in the testing phase. Neither visual cueing nor the mediating effect of cognitive load significantly influenced recall memory. According to De Koning et al.'s (2009) meta-analysis, most visual cues in animations function only with respect to attention by guiding the learner's attention to the essential information, and they do not function to spotlight the relationships between or among the learning elements or emphasize the instructional organization. They also found that visual cues in animations benefit comprehension and transfer the learning contents, but they do not benefit retention. Similarly, in this study, visual cueing apparently helped to reduce cognitive load during the viewing of the instructions by guiding attention toward the key information (i.e., the goal of each step), but it was not helpful for the participants' recall.

Overall, the results found that the hybrid format of dynamic and static visualization is beneficial for reducing cognitive load, although it did not lead to improving learning outcomes in the experimental sample. Arguel and Jamet (2009) argued that static pictures could compensate for the transient nature of animations, and that static images in an animation could help learners to remember and integrate key information. Their results verified that permanent information works as a type of external memory that helps to relieve the burden on working memory (Hegarty 2004). However, this and other related studies suggest that there could be numerous moderators, such as learners' visuospatial abilities, number of static pictures (Arguel and Jamet 2009; Castro-Alonso et al. 2015), and the pace of the animation (De Koning et al. 2011). In particular, the pace of animation could be an important moderator because it relates fundamentally to the transience of animation. The pace of animation also closely relates to learner control. In this study, learner control was permitted while viewing the animation as a way to objectively measure the cognitive

load. However, concerning the positive effect of learner control on cognitive load (see Ayres and Paas 2007), the results might differ under the circumstance that the effects of the factor are controlled. In that case, the pace of animation should play an important role.

Despite these results, certain doubts remain regarding the negative influences of the hybrid format. Because dynamic movements in an animation could capture attention more than non-dynamic visual cues (De Koning et al. 2009), static cues in an animation could be less effective, or even become distractions, by splitting attention to accommodate the moving and static images. To investigate the influence of visual cueing with precision, follow-up studies such as those that employ eye tracking (see Mayer 2010) would be useful. In addition, a detailed examination of the negative aspects of combining dynamic and static visualization could help to explain the complexities of a hybrid approach.

Although this study's results regarding visual cueing and cognitive load are interesting, they are limited. First, the between-subjects experimental design found substantial individual differences in the participants' abilities to perceive the positions and motions of the blocks in virtual space (visuospatial ability). Perhaps the sample size was so small that the influences of visual cueing were not statistically evident. In other words, the differences between the participants could have obscured the effects of visual cueing on learning performance. Therefore, an experiment might be more successful as a within-subjects experimental design.

This study also could not address a possible interaction effect of visual cueing and verbal instruction on learning. In practical learning contexts, an instructional animation usually includes verbal and visual instructions and explanations. Certain multimedia learning principles, such as the temporal contiguity principle (Mayer and Moreno 2002; Shin et al. 2016a, b), emphasize that both types of information should be presented in instructional animations. Therefore, the relationship between visual cueing and verbal instruction should be considered in future studies.

Conclusion

This study investigated the effects of visual cueing in 3D animations regarding learning procedural-manipulative tasks. It attempted to combine the advantages of dynamic visuals with static pictures in one medium, which was a step beyond a simple comparison of animation to static images. In a between-subjects experimental design, transparent static images were added to an animation to assess whether cognitive load during the viewing and imitating of the instructions would be lower than when an animation was without visual cues. Based on a review of previous relevant studies and the results of this empirical investigation, the results suggest certain visual design strategies to improve the effectiveness of instructional animations in the 3D virtual environment. Moreover, the findings of this experimental study are broadly applicable to numerous learning contexts, such as virtual or augmented reality environments.

Compliance with ethical standards

Conflict of interest The author declares that there is no conflict of interest regarding the publication of this article.

References

Arguel, A., & Jamet, E. (2009). Using video and static pictures to improve learning of procedural contents. *Computers in Human Behavior, 25*(2), 354–359.

Ayres, P., Marcus, N., Chan, C., & Qian, N. (2009). Learning hand manipulative tasks: When instructional animations are superior to equivalent static representations. *Computers in Human Behavior, 25*(2), 348–353.

Ayres, P., & Paas, F. (2007). Making instructional animations more effective: A cognitive load approach. *Applied Cognitive Psychology, 21*(6), 695–700.

Ayres, P., & Sweller, J. (2005). The split-attention principle in multimedia learning. *The Cambridge Handbook of Multimedia Learning, 2,* 135–146.

Berry, D., & Dienes, Z. P. (1993). *Implicit learning: Theoretical and empirical issues*. New York, NY: Psychology Press.

Brunken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 53–61.

Castro-Alonso, J. C., Ayres, P., & Paas, F. (2015). Animations showing Lego manipulative tasks: Three potential moderators of effectiveness. *Computers & Education, 85,* 1–13.

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

De Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review, 21*(2), 113–140.

De Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2011). Attention cueing in an instructional animation: The role of presentation speed. *Computers in Human Behavior, 27*(1), 41–45.

Garland, T. B., & Sanchez, C. A. (2013). Rotational perspective and learning procedural tasks from dynamic media. *Computers & Education, 69,* 31–37.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index). *Advances in Psychology, 52,* 139–183.

Hegarty, M. (2004). Dynamic visualizations and learning: Getting to the difficult questions. *Learning and Instruction, 14*(3), 343–351.

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17*(6), 722–738.

Hwang, Y., & Shin, D. (2018). The roles of visual experience in enhancing user performance in virtual environments. *Social Behavior and Personality, 46*(1), 11–24. https://doi.org/10.2224/sbp.6500.

Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and Instruction, 20*(2), 167–171.

Mayer, R. E., & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational Psychology Review, 14*(1), 87–99.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52.

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review, 24*(1), 27–45.

Paas, F., Tuovinen, J., Tabbers, H., & van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38,* 63–71.

Ryle, G. (2009). *The concept of mind*. New York, NY: Routledge.

Shin, D. (2012). 3DTV as a social platform for communication and interaction. *Information Technology and People, 25*(1), 55–80.

Shin, D. (2017). The role of affordance in the experience of virtual reality learning. *Telematics and Informatics, 34*(8), 1826–1836. https://doi.org/10.1016/j.tele.2017.05.013.

Shin, D., An, H., & Kim, J. (2016a). How the second screens change the way people interact and learn: The effects of second screen use on information processing. *Interactive Learning Environment, 24*(8), 2058–2079. https://doi.org/10.1080/10494820.2015.1076851.

Shin, D., Choi, M., Kim, J., & Lee, J. (2016b). Interaction, engagement, and perceived interactivity in single-handed interaction. *Internet Research, 26*(5), 1134–1157. https://doi.org/10.1108/IntR-12-2014-0312.

Shin, D., & Chung, K. (2017). The effects of input modality and story-based knowledge on users' game experience. *Computers in Human Behavior, 68,* 180–189. https://doi.org/10.1016/j.chb.2016.11.030.

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human–Computer Studies, 57*(4), 247–262.

Van Gog, T., Paas, F., Marcus, N., Ayres, P., & Sweller, J. (2009). The mirror neuron system and observational learning: Implications for the effectiveness of dynamic visualizations. *Educational Psychology Review, 21*(1), 21–30.

Wong, A., Marcus, N., Ayres, P., Smith, L., Cooper, G. A., Paas, F., et al. (2009). Instructional animations can be superior to statics when learning human motor skills. *Computers in Human Behavior, 25*(2), 339–347.

Donghee Shin is a Professor of College of Communication and Media Sciences at Zayed University. Prior to Zayed University, Dr. Shin was a professor in Korea and US including Sungkyunkwan University and Penn State University. Dr. Shin was awarded Distinguished Endowed Professor of Ministry of Education in Korea. He also served as a Principal Investigator of BK21 Plus, a national research project hosted by the Ministry of Education in Korea.

Seyoung Park is a master student at the Department of Interaction Science at Sungkyunkwan University. She majored in human computer interaction and education.