

The Effect of Stereoscopic Three-Dimensional Images on Vocabulary Learning

Regina Kaplan-Rakowski
University of North Texas, United States
ORCID: 0000-0002-6769-7784

Received: 25.12.2018

Accepted: 18.07.2019

Published: 17.10.2019

Abstract

Vocabulary is crucial for language learning. Therefore, searching for the most effective methods to master vocabulary is an important quest. This study investigated the effect of stereoscopic three-dimensional (S3D) images on recall and retention of foreign language vocabulary. S3D images, frequently used in virtual reality visualizations, are highly realistic and differ from non-stereoscopic three-dimensional (NS3D) images in that they provide an impression of depth. In this within-subject study, American subjects ($N = 82$) were exposed to 16 Polish words accompanied by corresponding images. Half of these images were in NS3D format, and half were in S3D format. The vocabulary was counterbalanced throughout four randomly distributed versions of the experiment. Quantitative data were drawn from immediately administered and one-week-delayed productive and receptive vocabulary tests. Qualitative data were gathered through background questionnaires and follow-up surveys, which included a mixture of open-ended and Likert scale questions. Two-tailed, paired t-tests showed no significant difference on vocabulary recall and retention of test scores for words represented with NS3D and S3D images. Post hoc tests revealed that lower performance on S3D images occurred only for those subgroups of subjects who exhibited discomfort or lack of experience with S3D technology, or both.

Keywords: *Vocabulary learning; Stereoscopic 3D images; Multimedia; Virtual reality; Immersive environments*

Introduction

Educators have long used visual aids to make learning more meaningful and effective (Bush, 2007). Some forms of teaching with visual aids have included physical objects (i.e., realia), which showed benefits for learning (Rule, Barrera, & Steward, 2004). Physical objects are often impractical in the classroom and, therefore, teachers rely on digital substitutes, such as images displayed on computer screens.

Digital images can vary in their complexity and degree of realism (Alessi & Trollip, 2001). One of the most important factors contributing to the realism of an image is the extent to which depth is depicted. In this context, the simplest form of an image is a two-dimensional (2D) image, with no depiction of depth (e.g., a line drawing). A more complex form is a 2.5D image, with a pseudo depiction of the third dimension of depth (e.g., an image with texture or shading). In this study, this type of image is referred to as a non-stereoscopic three-dimensional (NS3D) image. An even more realistic looking image, with a technologically enhanced sense of depth, is referred to as a stereoscopic three-dimensional (S3D) image.

Based on the concept of binocular disparity, the perception of depth is achieved in S3D images with two nearly identical pictures taken from two slightly different angles, eyes-width apart. This action mimics the way human eyes view the world and creates the illusion of depth (Price, Lee, & Malatesta, 2014). Figure 1 presents an example of three varying levels of realism. The main purpose of this study was to test the difference in effectiveness between the NS3D and S3D images to see which format is more effective for vocabulary learning.

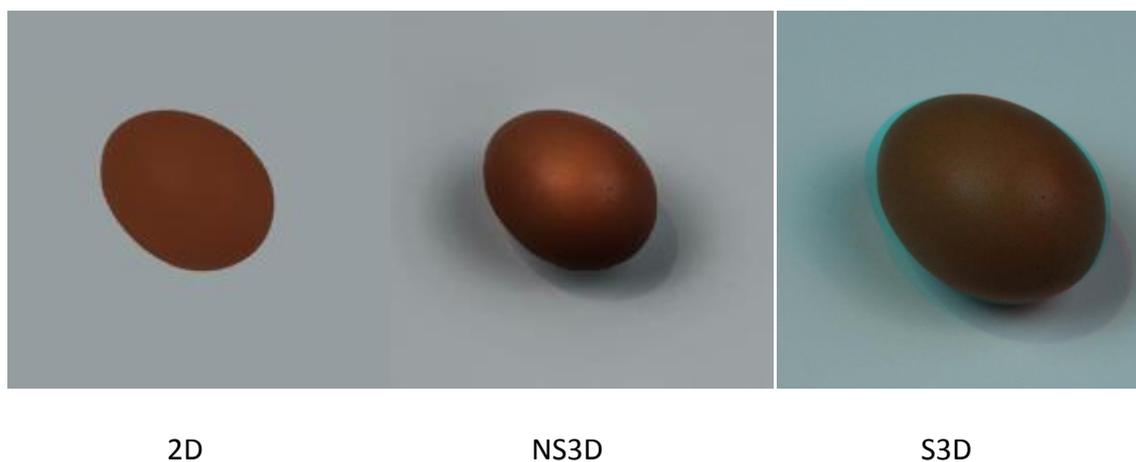


Figure 1. Images of Increasing Levels of Realism.

In the pictures above, the two-dimensional (2D) egg has no cues of depth. The non-stereoscopic three-dimensional (NS3D) egg uses shade and illumination to cue depth. The stereoscopic three-dimensional (S3D) image requires anaglyph glasses for viewing.

While testing the effectiveness of alternative visuals, such as line drawings (Dwyer, 1969), animated images (Lowe, 1999), and photographs (Mayer & Sims, 1994), has received considerable attention, empirical research on the use of NS3D and S3D images in education is limited to studies involving spatial learning (McIntire, Havig, & Geiselman, 2014). Some examples of domains where research has been conducted on depth-related tasks using S3D technology are surgery (Lewis, Zaritsky, Heinrichs, & Nezhat, 2006), imaging (Wen, Chang, Reinhold, Lo, & Markey, 2018), and teleoperation (Drascic, 1991). Studies using S3D images for nonspatial learning tasks are much less common.

In the past, learners were often limited to non-immersive NS3D images. A large body of evidence shows a decisive advantage of implementing images together with text for language learning (Chun & Plass, 1996; Kim & Gilman, 2008; Yeh & Wang, 2003). One explanation for why images lead to better learning is dual-coding theory (Paivio, 1986), which claims that the visual and the textual channels together help encode information better than either channel on its own. Another reason is that images provide greater immersion than text alone, which is sometimes associated with increased learning. S3D images are an example of a format in which increased realism is known to trigger greater immersion relative to (standard) NS3D images.

The contemporary cohort of learners may be even more susceptible to immersion-based learning approaches than previous generations because a large and growing portion of students' time is spent in environments with rich visual stimuli (Anderson & Jiang, 2018), such as digital games, virtual reality, movies, and virtual environments. As McGraw (2004, p. 154) pointed out:

“With the increasing prevalence of three-dimensional instructional media and immersive environments, it is imperative that the effects of three-dimensional stimuli on learners be considered.” Moreover, virtual reality, which often uses S3D visualizations, is increasingly influencing education. An important empirical question thus arises whether S3D images promote learning or not (Price, Lee, & Malatesta, 2014).

Because S3D images enhance immersion by providing a spatial sense of depth, a challenge in researching the effectiveness of S3D images has been figuring out how to disentangle the spatial characteristics of learning content from the delivery format of that content. A general conclusion is that when the learning content is depth-related, then S3D images may provide benefits relative to NS3D images (McIntire, Havig, & Geiselman, 2014). Whether or not the increased realism and immersion of S3D images provide benefits for nonspatial content remains an open question. This study examines the question in the context of foreign language (FL) vocabulary.

Research on the effectiveness of S3D images for learning has yielded various results, ranging from positive (Drascic, 1991; Ferdig, Blank, Kratcoski, & Clements, 2015), to mixed (Barfield & Rosenberg, 1995; Hansen, Barnett, MaKinster, & Keating, 2004), to nonsignificant (Cid & Lopez, 2010; Cliburn & Krantz, 2008; Keebler, 2011). McIntire, Havig, and Geiselman (2014) performed a review of studies that compared the effectiveness of NS3D and S3D displays on various types of learning. Their assessment covered 162 articles and considered the outcomes of 184 experiments. A comprehensive measure showed that in 60% of the experiments, S3D viewing was more advantageous for performance as compared with NS3D viewing. Mixed results were found in 15% of the experiments, and the remaining 25% of experiments showed no advantage of S3D.

When learning content is spatial or operational, several studies showed S3D images to be no different from, or better than, NS3D images for learning. For example, Price, Lee, Subbarao, Kasal, and Aguilera (2015) conducted an experiment showing planetarium visitors (N = 498) a movie in NS3D and S3D formats, to ascertain which format could better help the viewers determine the shape of the galaxy. Immediate tests showed no significant difference, but delayed tests revealed some positive impact of the S3D format. Price, Lee, and Malatesta (2014) exposed 261 subjects to twelve NS3D and S3D images to test how well they could perceive spatial elements in each image format. The results showed no significant difference between the formats. In another study, Cid and Lopez (2010) failed to find supporting evidence that showing learning content in S3D was beneficial in making students understand lunar phases. More recently, a dissertation by Wuerzburger (2017) explored whether learning information technology content can benefit from using an S3D display. However, such a display showed no added learning value compared to an NS3D display.

This present study is the first to examine the effectiveness of S3D images for FL vocabulary learning. Within the field of vocabulary learning, the most closely related previous work was by Sato and Suzuki (2010), who embedded a 2D static gloss and an NS3D animation gloss in a multimedia dictionary. The learning content consisted of spatial prepositions, such as “above,” “on,” and “over.” Therefore, the content had an inherent aspect of dimensionality involving visual perception. However, the small sample restricted the power of the test, leading to insignificant results. This study builds on the work of Sato and Suzuki with an expanded sample, the use of nonspatial content, and, most important, the use of S3D images instead of NS3D images.

The multimedia learning theory (Mayer, 2009) and the claims that humans naturally favor three-dimensionality (Tavanti & Lind, 2001) suggest that increased dimensionality should activate germane cognitive load, which is advantageous for learning (Sweller, 1994). Meanwhile, researchers warn about possible pitfalls of increased dimensionality, such as the possibility that more complex images will provoke extraneous cognitive load, which is disadvantageous for learning (Sweller, 1994). Most current research shows examples in which visually rich environments, such as virtual reality, may impede learning (Makransky, Terkildsen, & Mayer, 2019).

An inconsistency and a gap in the literature on learning exist that have not previously been reconciled, that is, whether S3D or NS3D images are better for learning when immersion is important, but content has no spatial component. In the case of foreign language vocabulary learning, prior research has repeatedly shown that immersion is particularly important. Meanwhile, analysis should be done on content that is inherently nonspatial. The existing literature, practice, and theories thus motivate the main purpose of this study and form the nondirectional research question: Is there a significant difference between NS3D and S3D images in their effect on immediate and delayed, productive and receptive recall of FL vocabulary?

Methodology

This study utilized a within-subjects experimental design with random assignment to treatments. The independent variable was the image type with two values: NS3D or S3D. The dependent variables were the average recall scores on the immediate, delayed, productive, and receptive tests. Two-tailed, paired *t*-tests measured whether average recall test scores of the NS3D and S3D treatments differed significantly from each other. Each *t*-test used scores from 82 participants for the immediate and 53 for the delayed tests. The number of subjects decreased due to attrition.

Participants

The participants were volunteers from a large public university in the Midwest region of the USA. Participants were required to be at least 18 years old and native speakers of English who could not speak Polish or any cognate language to Polish. The sample was 70% female and 30% male. The mean age was 20.5 years.

On average, subjects reported using electronics for 6.34 hours a day. A large percentage (97%) of subjects reported having seen a movie in 3D but only 28% of them had played a 3D video game. Only 22% of the subjects reported owning a system that displays 3D content, with Nintendo 3DS being the most popular system to own (12%).

Research Procedure and Content

The research procedure consisted of eight steps: (1) consent form and background information survey, (2) warm-up tutorial, (3) experiment, (4) immediate productive recall test, (5) immediate receptive recall test, (6) follow-up questions, (7) delayed productive recall test, and (8) delayed receptive recall test. Prior to the experiment, the participants viewed a warm-up tutorial with

experiment instructions and a slideshow containing NS3D and S3D images, which was intended to make the subjects more accustomed to the S3D technology.

In the experiment, the participants viewed 30 new Polish concrete nouns accompanied by corresponding images and English translations (see Figure 2). Half (15) of the images were NS3D, and the other half (15) were S3D. Each subject was exposed to each slide for 15 seconds. Participants wore anaglyph 3D glasses to view S3D images.



Figure 2. A Non-stereoscopic Three-dimensional (NS3D) Image and Its Stereoscopic Three-dimensional (S3D) Counterpart (The NS3D image is on the left; the S3D image, on the right. Anaglyph 3D glasses must be worn for stereoscopic perception of the S3D image).

To ensure an equal exposure of the words in the NS3D and the S3D treatments to participants, four versions of the experiment content were devised. While the words in each version were the same, the format (NS3D versus S3D) of individual images and the sequence of their exposure varied from version to version and thus from participant to participant.

Instrumentation and Reliability

The immediate and the one-week-delayed tests followed identical procedures. All instruments were administered right after the experiment and were completed with pen and paper. Students had five minutes to finish each test (ten seconds per word). In addition to immediate, delayed, productive, and receptive performance tests, the participants filled in a background information survey and a follow-up questionnaire.

Each performance test covered all 30 FL words encountered in the experiment. The rationale behind the inclusion of the whole list of words was driven by Nation's (2001) advice regarding increasing test reliability. Cronbach's alpha ranged from $\alpha = 0.73$ to $\alpha = 0.58$ across conditions and experiment version.

In the productive tests, the participants were assessed on their ability to recall English translations of the Polish words. Each correct translation received one point. Missing or incorrect translations received zero points. Possible total scores ranged from zero to 30. The receptive

recall tests consisted of 30 multiple-choice items. The participants were asked to choose the appropriate translation of the target words from among four choices. One point was granted for each correct choice.

The sequence of productive tests preceding the receptive tests was important because the receptive recall test was in a multiple-choice format. Had the sequence of those tests been reversed, the receptive test would have served as a learning experience. As such, it could have influenced the answers on the productive tests. Consequently, the risk of a cross-test effect would have arisen, causing an internal validity flaw.

Verifying the answers on three occasions strengthened the content validity of the assessment. The raters were highly educated native Polish speakers, grounded in English studies. The scoring was straightforward: Answers were either correct or incorrect, with no room for negotiation, which allowed for unbiased ratings.

The background information survey contained fill-in-the-blank and multiple-choice questions on participants' gender, age, school rank, native language, FL knowledge, and prior knowledge of Polish. Further, the participants answered questions on their ownership and use of electronic devices, especially those capable of displaying S3D images.

Finally, the participants completed an eight-item follow-up questionnaire. The questionnaire was administered to elicit participants' perceived engagement with NS3D and S3D images, self-reflection on their performance tests, opinions on the usefulness of S3D images for learning vocabulary, and discomfort while using S3D technology. Participants also could provide additional comments.

Post Hoc Analyses

The analyses of the emerging data motivated several additional tests, which sought to answer exploratory, not primary research, questions. The topics covered the impact of discomfort on the effectiveness of S3D images and the impact of experience with S3D technology. For each of these issues, an identical procedure was used. The full sample of subjects ($N = 82$) was divided into two sets of subgroups: (1) the subgroup reporting discomfort and the subgroup with no such discomfort and (2) the subgroup reporting experience with S3D technology and the subgroup with no such experience.

The subjects' performance between NS3D and S3D words was analyzed separately for each set of subgroups. This procedure involved the computation of paired *t*-tests measuring whether average recall test scores of the NS3D treatment were significantly better than average recall test scores of the S3D treatment within each subgroup. Because each test had an expected outcome, each hypothesis in the post hoc analyses was one-directional.

Results

Immediate, Delayed, Productive, and Receptive Test Scores

The primary research question was whether a significant difference exists between NS3D and S3D images in their effect on immediate, delayed, productive, and receptive recall of FL

vocabulary. Table 1 displays the mean scores (number of correct items) from the productive recall test, as well as associated hypothesis test statistics. The mean score for words displayed with NS3D images was 4.87 (out of 15) for immediate testing and 2.59 for delayed testing. The mean score associated with S3D images was 4.61 for immediate testing and 2.25 for delayed testing. This is, when the subjects learned through NS3D images, they scored 0.26 higher, on average, than when the subjects used S3D images—a difference that is of no statistical significance ($p = 0.40$). For delayed testing, the subjects scored 0.33 higher, on average, on NS3D images, which resulted in an insignificant difference ($p = 0.08$). The effect size calculated using Cohen's d was 0.50 for immediate testing and 0.54 for delayed testing.

Table 1. Results of Paired t-tests on Immediate, Delayed, Productive, and Receptive Test Scores

	Mean (Percent, Minimum, Maximum)		Mean Difference	<i>df</i>	<i>p</i> (Two- Tailed)	<i>t</i>	Cohen's <i>d</i>
	NS3D	S3D					
Immediate Testing							
Productive	4.87 (32%, 1, 15)	4.61 (31%, 1, 10)	0.26	81	0.40	0.84	0.19
Receptive	9.72 (65%, 3, 15)	9.43 (63%, 3, 15)	0.29	81	0.23	1.20	0.27
Delayed Testing							
Productive	2.59 (17%, 0, 8)	2.25 (15%, 0, 7)	0.33	50	0.08	1.76	0.50
Receptive	7.22 (48%, 5, 12)	6.72 (45%, 4, 10)	0.50	45	0.07	1.82	0.54

Therefore, for the full sample of subjects, the null hypothesis of no significant difference between the effect of S3D and NS3D images on FL recall in the productive recall test cannot be rejected. As with the productive test, the subjects showed no significant difference on the receptive test between NS3D and S3D words. These results pertain to both immediate and delayed test scores.

Post Hoc Analyses

The post hoc analyses of the data involved additional tests pertaining to the impact of discomfort on the effectiveness of S3D images and the impact of experience with S3D technology.

The Impact of Discomfort on the Effectiveness of S3D Images

The subjects gave feedback on discomfort on two separate occasions. First, the follow-up questionnaire revealed that 33 subjects (40%) reported experiencing discomfort when viewing S3D images. The reports of discomfort were also made in open-ended comments. Because these percentages are a large proportion of the sample, the researcher conducted additional tests to establish how discomfort, or the lack of it, influenced the performance scores.

The full sample of subjects ($N = 82$) was divided into two subgroups based on whether or not they reported experiencing discomfort when viewing S3D images. Thirty-three participants

(40%) reported discomfort and 49 (60%) reported no discomfort. To evaluate the difference in performance test scores between NS3D and S3D images, the researcher conducted four additional paired *t*-tests, for the discomfort and no-discomfort subgroups.

Table 2 summarizes the results of paired *t*-tests for the subgroup reporting discomfort while viewing S3D images. For productive tests, the mean NS3D score was 5.58 and the mean S3D score was 4.52, a difference of 1.06. Because the researcher assumed that discomfort would impede subjects from learning the words presented with S3D images, a one-tailed *p*-value was taken into consideration. The difference of mean scores between NS3D and S3D words was significant ($t_{(33)} = 1.96$, $p = 0.03$) with a large Cohen's *d* effect size of 0.69.

Table 2. The Results of Paired *t*-tests for Discomfort and No-Discomfort Subgroups

Subgroup	Measure	Mean Score		<i>df</i>	<i>p</i> (One-Tailed)	<i>t</i>	Cohen's <i>d</i>
		NS3D	S3D				
Discomfort	Productive	5.58	4.52	32	0.03	1.96	0.69
	Receptive	10.48	9.76	32	0.02	2.10	0.74
No-discomfort	Productive	4.39	4.67	48	0.20	0.84	0.24
	Receptive	9.20	9.20	48	1.00	0.00	0.00

For receptive tests, the mean NS3D score was 10.48 and the mean S3D score was 9.76. The difference of 0.72 was significant ($t_{(33)} = 2.10$, $p = 0.02$). Again, the effect size was large with a Cohen's *d* of 0.74. These results suggest that the overall lack of benefits from using S3D technology is driven by the subgroup of learners who reported experiencing discomfort.

As Table 2 shows, for the productive recall tests conducted on the learners who reported no discomfort while viewing S3D images, the mean NS3D score was 4.39 and the mean S3D score was 4.67. Even though the mean S3D score was higher (by 0.28), the difference in means was not significant ($t_{(33)} = 0.84$, $p = 0.20$) and Cohen's *d* was low, at 0.24. Similarly, no significant difference ($t_{(33)} = 0.00$, $p = 1.00$) was detected in the receptive tests for the learners who reported no discomfort while viewing S3D images. In fact, the mean NS3D score and the mean S3D score were identical (9.20) in this case.

The Impact of Experience with S3D Technology

One of the questions in the background information survey asked, "Have you ever played a video game in 3D?" Based on the answers to that question, the full sample of subjects ($N = 82$) was divided into two subgroups: the subjects who reported having played a video game in S3D ($n_1 = 23$; 28%) and the subjects who reported not having played a video game in S3D ($n_2 = 59$; 72%). To evaluate the difference in performance test scores between NS3D and S3D images, four additional paired *t*-tests were conducted, for the subgroup of the subjects who reported having played a video game in S3D and the subgroup without such an experience.

Table 3 reports the results of paired *t*-tests for the subgroup that had played a video game in S3D. For the productive test, the mean NS3D score was 5.52 and the mean S3D score was 5.13.

Because the assumption was that experience playing a video game in S3D would help the subjects recall the words in the S3D format, a one-tailed p -value was taken into consideration. The difference of 0.39 between NS3D and S3D words was not significant ($t_{(23)} = 0.58, p = 0.28$).

Table 3. Results of t-test Scores of Subjects with and without Stereoscopic Three-Dimensional (S3D) Experience

Subgroup	Measure	Mean Score		df	p (One-Tailed)	t	Cohen's d
		NS3D	S3D				
S3D Experience	Productive	5.52	5.13	22	0.28	0.58	0.25
	Receptive	9.61	9.83	22	0.31	0.49	0.21
No-S3D Experience	Productive	4.61	4.41	58	0.27	0.60	0.16
	Receptive	9.76	9.27	58	0.05	1.71	0.45

For the receptive test, the mean NS3D score was 9.61 and the mean S3D score was 9.83. The difference of 0.22 points was not significant ($t_{(23)} = 0.49, p = 0.31$) on a one-tailed test.

As Table 3 shows, for the productive recall tests conducted on the learners who reported not having experience playing a video game in S3D, the mean NS3D score was 4.61 and the mean S3D score was 4.41. Because the assumption was that the learners without experience playing a video game in S3D would do worse on S3D images, a one-tailed test was used for this subgroup. Even though the mean S3D score was higher by 0.20 points, the difference in means was not significant ($t_{(23)} = 0.60, p = 0.27$).

A significant difference ($t_{(23)} = 1.71, p = 0.05$) was detected in the receptive test for the learners who reported not having played a S3D video game. The mean NS3D score was 9.76 and the mean S3D score was 9.27, a difference of 0.49 points. The effect size was also large for this test, with a Cohen's d value of 0.45. These results suggest that the overall lack of benefits from using S3D technology is driven by the subgroup of learners who are inexperienced with S3D technology.

Discussion and Recommendations

The findings of this study were framed through the lens of Merrill (2002), who claims that sound instruction is built on aspects of effectiveness, engagement, and efficiency. The previous literature found that 60% of reviewed experiments showed that S3D viewing was more effective than alternative formats for learning and 15% of the studies had mixed results. The remaining 25% of the experiments revealed that S3D and NS3D viewing are similarly effective for learning (McIntire, Havig, & Geiselman, 2014). This study showed no statistically significant difference in the effectiveness between S3D and NS3D images, even though the engagement with S3D images was significantly higher.

This study is aligned with the studies that found no overall difference in the effectiveness of the two viewing formats (Cid & Lopez, 2010; Price, Lee, & Malatesta, 2014; Price et al., 2015; Wuerzburger, 2017). Meanwhile, those studies differed from this one. For instance, Price et al. (2015) study found no significant benefit to either of the formats. It tested dynamic (film) visualizations, while this study tested static images. Price et al. also found a positive effect of S3D film on delayed tests. This study did not show such an effect. On the contrary, the NS3D effect, even though nonsignificant, was stronger compared with S3D.

This study also supports the finding of no significant benefit to using S3D representations when learning tasks are not explicitly depth-related (McIntire, Havig, & Geiselman, 2014). The original contribution of this study includes further analysis of the factors beyond learners' overall performance that allow for better understanding of when and how NS3D and S3D images differ in their effectiveness.

Regarding efficiency, technological progress tends to decrease the cost of acquiring and using S3D technology and, therefore, increases S3D efficiency, relative to NS3D efficiency. Currently, producing and viewing NS3D images requires only basic technology skills and minimal resources. The production of S3D images requires more advanced photographic editing skills. Moreover, equipment for creating and viewing S3D images, such as S3D cameras and glasses or goggles, are necessary. The use of NS3D images thus arguably is more efficient than S3D images. The advancement of technology should decrease the relative cost and consequently improve the level of efficiency of S3D images. The 3D technology industry has been actively promoting S3D-related content and technologies to the educational sectors.

The feedback on the follow-up questionnaire revealed that the students were highly engaged with S3D images. This outcome is well aligned with a recent dissertation by Wuerzburger (2017) that showed subjects being more attracted to S3D visualizations, but without added learning value relative to the NS3D format. Increased engagement is one of the characteristics of S3D technology that may, in turn, drive greater effectiveness. The researcher of the present study performed further analysis of the interaction between engagement and S3D technology on learning. Engagement has long been identified as an important component of successful learning. Engaged learners attain their learning goals faster and with more depth. They tend to stay more focused and motivated throughout the learning process (Hannafin & Hooper, 1993). Supplemental statistical analysis measured the perceived levels of engagement with NS3D and S3D images. The mean engagement score for NS3D images was 3.22 (out of 5); for S3D images, 3.66. The perceived engagement with S3D images was significantly greater than with NS3D images ($p = 0.01$), with a moderate effect size (Cohen's d of 0.58). In addition to formal statistical tests, multiple comments ($n = 18$, which is 22% of the sample) in the follow-up questionnaire fortified the idea that the S3D images are "more engaging," "more stimulating," and "more interactive." In sum, both the quantitative and the qualitative analyses suggest that learning with S3D images promotes enhanced engagement. Consequently, learning with S3D images could lead to more successful learning, especially in settings such as foreign language learning where engagement and context are of particular importance.

Several limitations of the study motivate suggestions for improvement in future research. For example, this study examined learning only of concrete, static nouns. Future studies should test abstract nouns, other parts of speech, or other learning content. Dynamic visualizations should also be taken into consideration, with the involvement of movies, video games, or more sophisticated virtual reality platforms, such as *Oculus Rift*.

The subjects in the current study were exposed to vocabulary in a semi-contextualized manner (e.g., contextualization through relevant images). Such an approach was done purposefully, to reduce the possibility of extraneous variables. Nevertheless, future studies should introduce words in a more contextualized way, for example, by embedding them in a story.

While this study attempted to diminish the novelty effect by introducing the S3D images in the warm-up tutorial, prior to the experiment, it is likely that the novelty effected was not fully eliminated. Therefore, subjects should experience S3D images for extended periods before experiments are conducted in future studies.

Contributions and Implications

Previous investigations on the effect of NS3D and S3D images on learning are limited to studies encompassing spatial learning (McIntire, Havig, & Geiselman, 2014) and do not focus on vocabulary learning. This study is the first one concerning S3D technology to pay explicit attention to vocabulary learning and the first to incorporate content that does not have an inherent spatial component. These characteristics of the research design allow for findings that can be better generalized across educational settings and are not restricted to the previously studied spatial topics, such as geometry, navigation, or architecture.

An additional area of long-standing concern for instructional designers is the extent to which discomfort (Bando, Iijima, & Yano, 2012; Kim, Yoo, & Seo, 2013) or cognitive overload (Sweller, 1994) can mitigate the otherwise beneficial aspects of a new technology. This study includes post hoc analyses of how discomfort or lack of experience, or both, with S3D technology interacts with the use of S3D images. These analyses revealed that when discomfort or lack of experience is present, the effectiveness of learning with richer-stimuli content (S3D images) is diminished. For the first time, the analyses of discomfort and lack of experience with S3D technology have been included in an instructional unit that tests the impact of S3D images on vocabulary learning.

Historically, academics in the field of instructional design and technology have presented their work through the perspective of Merrill (2002) by evaluating the extent to which instruction is effective, efficient, and engaging. One of the findings of this study was that S3D images can be as effective for learning as NS3D images. Therefore, instructional designers should be advised to implement those visual representations that are more efficient (i.e., faster and cheaper to develop) or more engaging, or both. As of 2019, the general assumption is that NS3D images are more efficient. With S3D technology becoming more affordable, the differences in efficiency are diminishing. Together, these interpretations are an application of Merrill's perspectives to a new fashioning of instructional design.

The findings of this study lead to several practical implications for instructional designers, instructional technologists, and FL instructors. This study allows for better understanding about how to facilitate explicit vocabulary learning by taking effectiveness, efficiency, and engagement into account. While no evidence exists that the unconditional use of S3D images is always superior for vocabulary learning, this study does provide some insights into the circumstances when S3D images could be more effective than NS3D images, as well as the concerns that should be addressed to design effective lessons by incorporating S3D images.

Acknowledgements

This article is derived from my doctoral dissertation entitled: "The Effect of Stereoscopic Three-Dimensional Images on Recall of Second Language Vocabulary" (2016), Southern Illinois University Carbondale, USA. I am deeply grateful for all the constructive feedback received from my Ph.D. committee members: Christian Sebastian Loh (Chair), Peter Fadde, Krassimira Charkova, Lingguo Bu, and Lourdes Albuixech.

References

- Alessi, S. M. & Trollip, S. R. (2001). *Multimedia for learning methods and development* (3rd ed.). Boston, MA: Allyn and Bacon.
- Anderson, M., & Jiang, J. (2018). *Teens, social media & technology 2018*. Pew Research Center's Internet and American Life Project, 31. Retrieved on 24 December 2018 from www.pewinternet.org/2018/05/31/teens-social-media-technology-2018/
- Bando, T., Iijima, A., & Yano, S. (2012). Visual fatigue caused by stereoscopic images and the search for the requirement to prevent them: A review. *Displays*, 33, 76-83.
- Barfield, W. & Rosenberg, C. (1995). Judgments of azimuth and elevation as a function of monoscopic and binocular depth cues using a perspective display. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), 173-181.
- Bush, M. D. (2007). Facilitating the integration of culture and vocabulary learning: The categorization and use of pictures in the classroom. *Foreign Language Annals*, 40(4), 727-745.
- Chun, D. M. & Plass, J. L. (1996). Effects of multimedia annotation on vocabulary acquisition. *The Modern Language Journal*, 80(2), 183-198.
- Cid, X. C. & Lopez, R. E. (2010). The impact of stereo display on student understanding of phases of the moon. *Astronomy Education Review*, 9(1), DOI: 10.3847/AER2009044.
- Cliburn, D. & Krantz, J. (2008). Towards an effective low-cost virtual reality display system for education. *Journal of Computing Sciences in Colleges*, 23(3), 147-153.
- Drascic, D. (1991). Skill acquisition and task performance in teleoperation using monoscopic and stereoscopic video remote viewing. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 35(19). San Francisco, CA: Sage.
- Dwyer, F. M. (1969). An experimental analysis of the use of realism in illustrations used to complement programmed instruction. *Programmed Learning and Educational Technology*, 6, 147-153.
- Ferdig, R., Blank, J., Kratcoski, A., & Clements, R. (2015). Using stereoscopy to teach complex biological concepts. *Advances in Physiology Education*, 39(3), 205-208.
- Hannafin, M. J. & Hooper, S. R. (1993). Learning principles. In M. L. Fleming & W. H. Levie (Eds.), *Instructional Message Design: Principles from the Behavioral and Cognitive Sciences* (2nd ed.), pp. 191-227. Englewood Cliffs, NJ: Educational Technology Publications.

- Hansen, J. A., Barnett, M., MaKinster, J. G., & Keating, T. (2004). The impact of three-dimensional computational modeling on student understanding of astronomical concepts: A quantitative analysis. *International Journal of Science Education, 26*(11), 1365-1378.
- Keebler, J. R. (2011). *Effects of 3D stereoscopy, visuo-spatial working memory, and perceptions of simulation experience on the memorization of confusable objects* (Unpublished doctoral dissertation). University of Central Florida, Orlando, FL.
- Kim, D. & Gilman, D. A. (2008). Effects of text, audio, and graphic aids in multimedia instruction for vocabulary learning. *Educational Technology & Society, 11*(3), 114-126.
- Kim, D., Yoo, J., & Seo, Y. (2013). Qualitative analysis of individual and composite content factors of stereoscopic 3D video causing visual discomfort. *Displays, 34*(3), 223-240.
- Lewis, L. A., Zaritsky, E., Heinrichs, W. L., & Nezhat, C. (2006). Comparison of two-dimensional and three-dimensional camera systems in laparoscopic surgery. *Obstetrics & Gynecology, 107*(4), 78S.
- Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education, 14*(2), 225-244.
- Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction, 60*, 225-236.
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Mayer, R. E. & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 86*(3), 389-401.
- McGraw, T. M. (2004). The effects of two-dimensional stimuli and three-dimensional stereoptic stimuli on spatial representation in drawings. *Studies in Art Education, 45*(2), 153-169.
- McIntire, J. P., Havig, P. R., & Geiselman, E. E. (2014). Stereoscopic 3D displays and human performance: A comprehensive review. *Displays, 35*, 18-26.
- Merrill, M. D. (2002). First principles of instruction. *Journal of Structural Learning and Intelligent Systems, 14*(4), 459-466.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- Paivio, A. (1986). *Mental representations: A dual-coding approach*. New York: Oxford University Press.
- Price, C. A., Lee, H. S., & Malatesta, K. (2014). Stereoscopy in static scientific imagery in an informal education setting: Does it matter? *Journal of Science Education and Technology, 23*(6), 721-734.
- Price, C. A., Lee, H. S., Subbarao, M., Kasal, E., & Aguilera, J. (2015). Comparing short- and long-term learning effects between stereoscopic and two-dimensional film at a planetarium. *Science Education, 99*(6), 1118-1142.

- Rule, A. C., Barrera, M. T., & Steward, R. A. (2004). Using descriptive adjective object boxes to improve science vocabulary. *Montessori LIFE, Spring*, 28-33.
- Sato, T. & Suzuki, A. (2010). Do multimedia-oriented visual glosses really facilitate EFL vocabulary learning? A comparison of planar images with three-dimensional images. *Asian EFL Journal*, 12(4), 160-172.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4, 295-312.
- Tavanti, M. & Lind, M. (2001). 2D vs. 3D: Implications on spatial memory. *Proceedings of the IEEE Symposium on Information Visualization, INFOVIS'01* (pp. 139-145). Washington, DC: IEEE Computer Society Press.
- Wen, G., Chang, H. C., Reinhold, J., Lo, J. Y., & Markey, M. K. (2018). Virtual assessment of stereoscopic viewing of digital breast tomosynthesis projection images. *Journal of Medical Imaging*, 5(1). DOI: <https://doi.org/10.1117/1.JMI.5.1.015501>
- Wuerzburger, J. (2017). *The impact of stereoscopic three-dimensional learning strategy on undergraduate technology students' procedural learning in information technology* (Unpublished doctoral dissertation). Indiana State University, Terre Haute, Indiana, USA.
- Yeh, Y. & Wang, C.-W. (2003). Effects of multimedia vocabulary annotations and learning styles on vocabulary learning. *CALICO Journal*, 21(1), 131-144.

Correspondence: Regina Kaplan-Rakowski, Ph.D., Department of Learning Technologies,
College of Information, University of North Texas, Denton, Texas, United States
